

RENSIT:

RadioElectronics. NanoSystems. Information Technologies

Journal "Radioelectronics. Nanosystems. Information Technologies" (abbr. RENSIT) publishes original articles, reviews and brief reports, not previously published, on topical problems in **radioelectronics (including biomedical) and fundamentals of information, nano- and biotechnologies and adjacent areas of physics and mathematics.**

Designed for **researchers, graduate students, physics students of senior courses and teachers.**

It turns out **4 times a year** (that includes 4 issues)

Authors of journal are academicians, corresponding members and foreign members of Russian Academy of Natural Sciences (RANS) and their colleagues,

as well as other russian and foreign authors on presentation of their manuscripts by the members of RANS, which can be obtained by authors before sending articles to editors.

And also after its receiving - on recommendation of a member of editorial board of journal, or another member of Academy of Natural Sciences, that gave her opinion on article at request of editor.

The editors will accept articles in both **Russian and English** languages.

Articles are internally peer reviewed (**double-blind peer review**) by members of the Editorial Board.

Some articles undergo external review, if necessary.

Journal RENSIT is included in the **DB SCOPUS (21.10.2017, Elsevier, Amsterdam, Netherlands), EBSCO Publishing**, in the international abstracts database - **Ulrich's International Periodicals Directory** (USA, New York, <http://www.ulrichsweb.com>), in the **AJ and DB VINITI RAS** (<http://www.viniti.ru>), and DB **Russian Science Citation Index (RSCI)** (http://elibrary.ru/project_risc.asp). Included in the **List of Higher Attestation Commission of the Ministry of Education and Science of the Russian Federation (15.01.2018, No. 757).**

Full-text content is posted in the DB of the **Russian Scientific Electronic Library**

- information resource on the Internet <http://elibrary.ru> and is available for registered users.

And also - in Open Access **CyberLeninka NEB** of Russian Federation <http://cyberleninka.ru>.

On journal's website <http://www.rensit.ru> posted metadata publications

and **RENSIT: Radioelectronics. Nanosystems. Information Technologies - english version** (cover-to-cover translation) of journal, which is a party to **CrossRef**.

The founder - the **Russian Academy of Natural Sciences**

Publisher - Publishing Center of the Russian Academy of Natural Sciences

Publisher Address: 119002 Moscow, per. Sivtsev Vrazhek 29/16

CONTENTS

RADIOELECTRONICS

FRactal Applications in Radio Electronics as Fractal Engineering

Alexander A. Potapov 215

Numerical Simulation of Low-Voltage Logic Gates Based on Field-Effect Nanotransistors with a Variable Working Area Diameter

Nikolae V. Masalsky 233

NANOSYSTEMS

Heating Dynamics of Metal-Dielectric Structures with Nanometer-Thin Conductive Films under the Influence of Microwave Fields

Vladimir V. Starostenko, Evgeniy V. Grigoriev, Alim S. Mazinov, Sergey P. Arsenichev, Evgeniy P. Taran, Aleksandr V. Starosek, Ibraim Sh. Fitaev ... 243

CONDENSED MATTER NUCLEAR PHYSICS

Strange Traces of a "Strange" Radiation

Anatoly I. Nikitin, Vadim A. Nikitin, Alexander M. Velichko, Tamara F. Nikitina 249

INFORMATION TECHNOLOGIES

Information Technologies Based on Noise-Like Signals: III. Spectrum of Sequences Periods Forming by a Discrete Chaotic Algorithm

Nikita A. Ageykin, Vladimir I. Grachev, Nadezhda G. Petrova, Alexander M. Popov, Viktor I. Ryabenkov, Vladimir V. Kolesov 269

Problems of Message Routing Algorithms in Streaming Data Processing Systems

George G. Bulychev, Alexey V. Chernykh 279

QR-Decomposition Based Receivers for Downlink MIMO-SCMA

Islam Kanatbekuli, Dmitriy A. Pokamestov, Yakov V. Kryukov, Evgeniy V. Rogozhnikov, Renat R. Abenov 291

Holographic Processing of Moving Sources in a Shallow Sea with Intense Internal Waves

Venedikt M. Kuz'kin, Sergey A. Pereselkov, Elena S. Kaznacheeva, Vladimir I. Grachev, Sergey A. Tkachenko, Pavel V. Rybyanets 301

FRACTALS IN PHYSICS

Light Diffraction by Flat Geometric Bifractals

Galina V. Arzamastseva, Mikhail G. Evtikhov, Fedor V. Lisovsky, Ekaterina G. Mansvetova 309

MEDICAL PHYSICS

Method of Brain Wave Train Electrical Activity Analysis – The Theoretical Basis and Application

Olga S. Sushkova, Alexei A. Morozov, Nadezhda G. Petrova, Margarita N. Khokhlova, Alexandra V. Gabova, Alexei V. Karabanov, Larisa A. Chigaleychik, Karine Yu. Sarkisova 317

SCIENCE METHODOLOGY

Computational Experiment – Nondimensionalization of Equations, Computational Stability and Program Testing

Mikhail G. Evtikhov, Vladimir G. Evtikhov 331

Triad Matter - Substance - Material

Sergey P. Gubin 341



RUSSIAN ACADEMY
OF NATURAL SCIENCES

DEPARTMENT OF
RADIOELECTRONICS,
NANOPHYSICS AND
INFORMATION TECHNOLOGIES
PROBLEMS

RENSIT:

**RADIOELECTRONICS.
NANOSYSTEMS.
INFORMATION
TECHNOLOGIES.**

2022, Vol. 14, no. 3

FOUNDED IN 2009

4 ISSUES PER YEAR

MOSCOW

Editor-in-Chief

VLADIMIR I. GRACHEV

grachev@cplire.ru

Deputy Chief Editor

Vladimir A. Makarov, DrSci, MSU

Deputy Chief Editor

Sergey P. Gubin, DrSci, IGIC RAS

Executive Secretary

Rostislav V. Belyaev, PhD, IRE RAS

belyaev@cplire.ru

EDITORIAL BOARD

Anatoly V. Andreev, DrSci, MSU

Vladimir A. Bushuev, DrSci, MSU

Vladimir A. Cherpenin, DrSci, c-mRAS, IRE

Alexander S. Dmitriev, DrSci, IRE

Yuri K. Fetisov, DrSci, MIREA

Yuri V. Gulyaev, DrSci, acad.RAS, IRE

Yaroslav A. Ilyushin, DrSci, MSU

Anatoly V. Kozar, DrSci, MSU

Vladimir V. Kolesov, PhD, IRE

Albina A. Kornilova, PhD, MSU

Alexander V. Okotrub, DrSci, SB RAS

Aleksey P. Oreshko, DrSci, MSU

Igor B. Petrov, DrSci, c-m RAS, MIPT

Alexander A. Potapov, DrSci, IRE

Vyacheslav S. Rusakov, DrSci, MSU

Alexander S. Sigov, DrSci, acad.RAS, MIREA

Eugeny S. Soldatov, DrSci, MSU

Ikhamasuren Enkhdor, DrSci (Mongolia)

Yoshiyuki Kawazoe, DrSci (Japan)

Kayrat K. Kadyrzhanov, DrSci (Kazakhstan)

Peter Paul Mac Kenn, DrSci (USA)

Jiao Licheng, Ph.D. (China)

Deleg Sangaa, DrSci (Mongolia)

Andre Skirtach, DrSci (Belgium)

Enrico Verona, DrSci (Italy)

ISSN 2414-1267

The journal on-line is registered by the Ministry of Telecom and Mass Communications of the Russian Federation. Certificate El. no. FS77-60275 on 19.12.2014

All rights reserved. No part of this publication may be reproduced in any form or by any means without permission in writing from the publisher.

©RANS 2022

EDITORIAL BOARD ADDRESS

218-219 of, 7 b., 11, Mokhovaya str.,

125009 MOSCOW, RUSSIAN FEDERATION,

TEL. +7 495 629 3368

FAX +7 495 629 3678 FOR GRACHEV

DOI: 10.17725/rensit.2022.14.215

Fractal applications in radio electronics as fractal engineering

^{1,2}Alexander A. Potapov¹Kotelnikov Institute of Radioengineering and Electronics of Russian Academy of Sciences, <http://www.cplire.ru/>

Moscow 125009, Russian Federation

²Jinan University, JNU-IREE RAS Joint Laboratory of Information Technology and Fractal Processing of Signals, <https://jnu.edu.cn/>

Guangzhou 510632, China

E-mail: potapov@cplire.ru

Received July 15, 2022, peer-reviewed July 22, 2022, accepted July 29, 2022.

Abstract: The use of the fractal paradigm is presented - the main directions for introducing textures, fractals, fractional operators, dynamic chaos and methods of nonlinear dynamics for the design and creation of real technical projects in radio electronics - fractal radio systems, taking into account the hereditary, non-Gaussianity and scaling of physical signals and fields. The substantiation of the use of fractal-scaling and texture methods for the synthesis of fundamentally new topological texture-fractal methods for detecting signals in the space-time channel of scattering waves (a new type of radar) is discussed. It is shown that the use of fractal systems, sensors and nodes is a fundamentally new solution that significantly changes the principles of constructing intelligent radio engineering systems and devices. It is shown that the use of computational dielectric metasurfaces brings to a new level all the functional characteristics of a multifunctional system of topological texture-fractal processing of signals and fields in solving classical problems of detection, measurement, recognition and classification by intelligent radio engineering systems and devices. The concept of "fractal engineering" is introduced, the methodology of its use is discussed.

Keywords: fractal, texture, scaling, signature, target detection and recognition, radar, fractal aggregation, computational metasurface, nonlinear dynamics, radiophysics, fractal engineering

UDC 519.24 + 537.86 + 621.396.96

For citation: Alexander A. Potapov. Fractal applications in radio electronics as fractal engineering. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2022, 14(3):215-232. DOI: 10.17725/rensit.2022.14.215.

*"Scientists study the world as it is;
engineers create the world,
that didn't exist before."
Theodor von Karman*

CONTENTS

1. INTRODUCTION (216)
2. REVIEW OF THE MAIN RESULTS OF THE IRE RAS IN THE FIELD OF FRACTAL RADIO ELECTRONICS (216)

- 2.1. PRIMARY INFORMATION RADIO SYSTEMS (216)
- 2.2. ENSEMBLES OF TOPOLOGICAL TEXTURE-FRACTAL FEATURES (217)
- 2.3. DEVELOPMENT OF FRACTAL TECHNOLOGIES (217)

- 2.4. A NEW CLASS OF TOPOLOGICAL TEXTURAL-MULTIFRACTAL FEATURES (218)
- 2.5. FRACTAL FUSION OF MULTIBAND RADAR IMAGES (219)
- 2.6. SELECTED CLASSICAL EXAMPLES OF FRACTAL PROCESSING (220)
- 2.7. RADAR EQUATION FOR A FRACTAL TARGET (222)
- 3. COMPUTATIONAL DIELECTRIC METASURFACES AND TOPOLOGICAL TEXTURE-FRACTAL PROCESSING (223)
 - 3.1. PRINCIPLES (223)
 - 3.2. DESIGN (224)
 - 3.3. RESULTS (224)
- 4. FRACTAL ENGINEERING AND PHILOSOPHY OF FRACTAL ENGINEERING (225)
- 5. CONCLUSION (227)
- REFERENCES (227)

1. INTRODUCTION

The description of real processes occurring in modern radio-physical and radio-technical systems currently involves taking into account the hereditary (memory), non-Gaussianity and scaling (self-similarity, self-similarity) of physical signals and fields. All these concepts are included in the definition of fractal sets or fractals, first proposed by B. Mandelbrot in 1975 [1,2].

At the Kotelnikov Institute of Radioengineering and Electronics of the Russian Academy of Sciences (IRE RAS) from 1979 to the present, research has been initiated and is being developed in the fundamental scientific direction "Fractal radio physics and fractal radio electronics: design of fractal radio systems". The obtained results allow us to speak about the Russian school of fractal engineering and fractal engineering in the field of radio electronics. The purpose of this work is to review the formation of this school on the examples of the main results of the research carried out at the IRE RAS.

2. REVIEW OF THE MAIN RESULTS OF THE IRE RAS IN THE FIELD OF FRACTAL RADIO ELECTRONICS

2.1. PRIMARY INFORMATION RADIO SYSTEMS

Primary radiophysical information comes from various modern radio systems in the form of a one-dimensional signal and/or radar image (RLI) - **Fig. 1**. Here, SAR is a synthetic aperture radar; UAV is an unmanned aerial vehicle. MIMO (Multiple input-multiple output) spatial signal coding technology generally implies that each radio device participating in data exchange will have several spatially distributed weakly correlated receiving and transmitting antennas. The main idea of fractal MIMO radars is the use of fractal antennas and fractal detectors [3-6,16,17,24-30].

The ability of fractal antennas to operate simultaneously at several frequencies or emit a broadband probing signal gives a sharp increase in the number of degrees of freedom, which determines many important advantages of this type of radar and greatly expands the possibilities of adaptation.

To reflect these features, a new term "fractal-frequency MIMO systems (FF

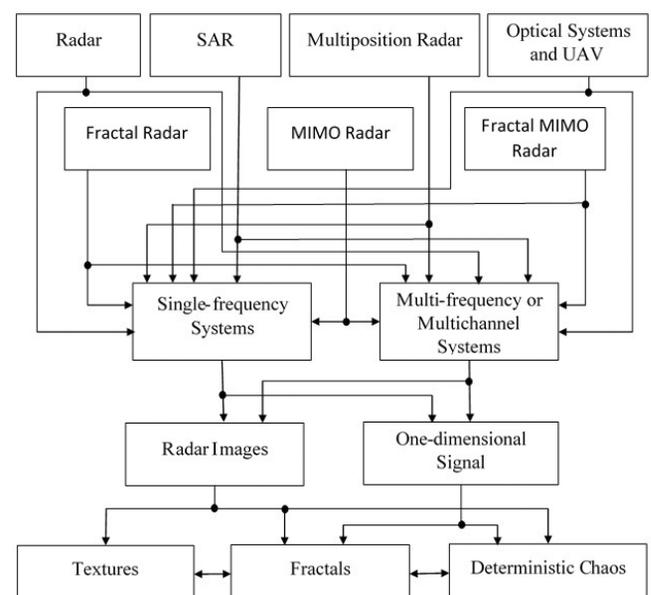


Fig. 1. Radio systems of initial primary information (fractal radar and fractal MIMO radar included).

MIMO)" has been introduced, which more fully reflects their physical capabilities.

2.2. ENSEMBLES OF TOPOLOGICAL TEXTURE-FRACTAL FEATURES

All currently existing new dimensional and topological (and not energy!) features or invariants and methods for detecting subtle objects against the background of intense reflections from the sea, land and meteorological precipitation are compactly presented in Fig. 2. Functional relationships between various features and methods are also noted here. Data Fig. 2 logically continue the data of Fig. one.

Thus, the introduction of the concepts of "deterministic chaos", "texture", "fractal", "fractal dimension D " and "fractal signature

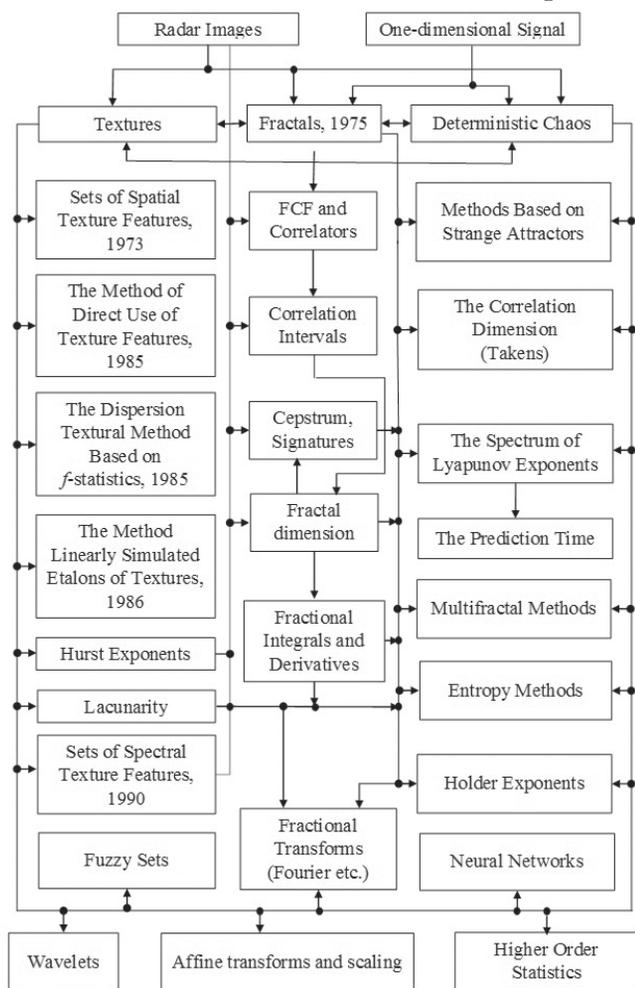


Fig. 2. Topological features and methods for detecting low-contrast (inconspicuous) objects against the background of intense noise and interference.

$D(t,f)$ " into the scientific use of radar made it possible to propose and then apply new non-energetic (!) signs or invariants (Fig. 2), which are combined under the generalized concept of "sample topology" ~ "fractal signature". Their application is widely and in detail presented in articles and monographs [3-17,21-30]. Data Fig. 2 are based on long-term full-scale experiments that were carried out jointly with the Central Design Bureau "Almaz" and other leading industrial organizations of the USSR, starting from the now distant 80s. XX century [31].

2.3. DEVELOPMENT OF FRACTAL TECHNOLOGIES

On Fig. 3 and Fig. 4 schematically shows the main stages of fundamental research in texture and fractal areas. Here, compared with the data for 2021 [23], in the diagram below Fig. 3, two additional and very important sections appeared – "Fractal complexing

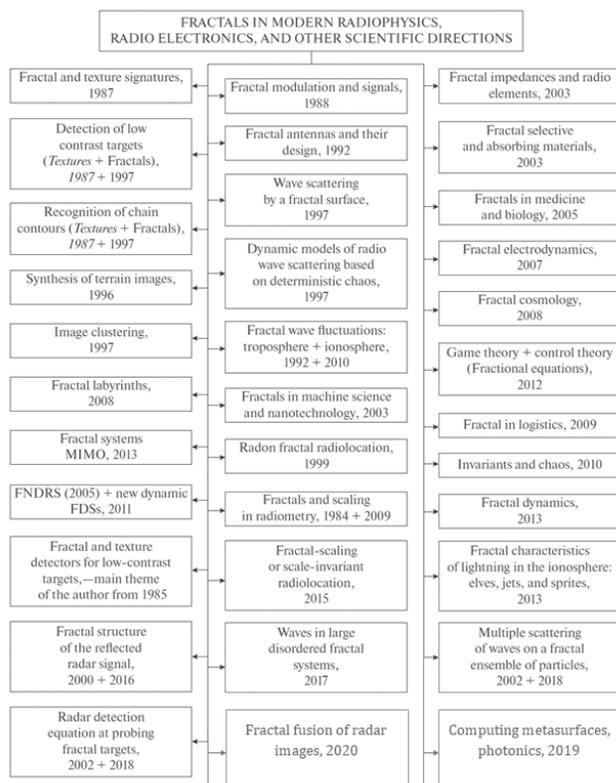


Fig. 3. A sketch of the development of breakthrough fractal technologies (this includes results related to texture): FNORS – fractal non-parametric radar signal detector, FOS – fractal signal detector.

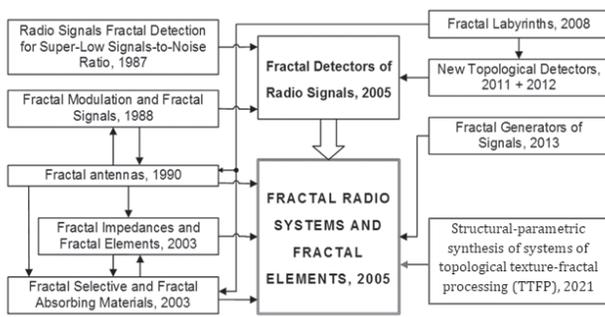


Fig. 4. The concept of fractal radio systems, sensors, devices and radio elements.

of radar images" and "Computational metasurfaces, photonics", and in Fig. 4 – section "Structural-parametric synthesis of systems of topological texture-fractal processing (TTFP)". On the first topic (Fig. 3), together with co-authors-students from the Air Force Acad. Professor N.E. Zhukovsky and Yu.A. Gagarin received a patent of the Russian Federation [32] and published a number of articles [33-37].

On Fig. 4, one additional and very significant section appeared – "Structural-parametric synthesis of topological texture-fractal processing systems (TTFP), 2021." [35]. The presented scheme of structural-parametric synthesis allows us to speak about *the optimality* of TTFP of multidimensional images, since the efficiency criterion is the maximum of one or several probabilistic characteristics simultaneously – classification (segmentation), detection or recognition, depending on the tasks solved by the system.

2.4. A NEW CLASS OF TOPOLOGICAL TEXTURE-MULTIFRACTAL FEATURES

For the proposed structural-parametric synthesis, we for the first time and specifically proposed in [37] a fundamentally new class of topological textural-multifractal features that allow us to jointly evaluate various fractal texture properties. In this case, all issues of joint estimation of scaling, singular, multifractal and anisotropic properties of

the texture of images of any kind are solved. We called this class of features a *directed morphological multifractal signature* (DMMFS) [37].

The calculation of the DMMFS begins with the calculation of local morphological multifractal indices (LMMFI) $L_q(\varepsilon, r)$ for the required number of angular directions of analysis r_{\max} of the processed image $I(m, n)$, here $m = 1, 2, \dots, M$, $n = 1, 2, \dots, N$ are the number of rows and columns, respectively. In this case, an array of corresponding values is formed in the "direction-scale" coordinates for each order of the scaling moment q , where $-\infty \leq q \leq \infty$, $q \neq 0$.

The calculation of the generalized statistical sum $Z(q, \varepsilon, r)$ for each direction of analysis r ($r = 1, 2, \dots, r_{\max}$) is carried out according to the generated set of "upper" $\{U_{\varepsilon, r}(m, n)\}$ and "lower" $\{X_{\varepsilon, r}(m, n)\}$ of "coverings" obtained as a result of morphological processing (dilation and erosion, respectively) of the array $\{B_r(m, n)\}$, rotated by the required number r_{\max} of the angular positions of copies of the original image using the set "flat" horizontally oriented structural elements $\{Y_\varepsilon, \varepsilon = 1, 2, \dots, E\}$, whose length w ($w = 2\varepsilon + 1$) corresponds to the analyzed scale. The generalized partition function $Z(q, \varepsilon, r)$ of the q_{th} order on each analyzed scale ε for each rotated image $B_r(m, n)$ is determined by the following relationship

$$Z(q, \varepsilon, r) = ((2\varepsilon)^{-1} V(\varepsilon, r)) V^{-q}(\varepsilon, r) \times \left(\sum_{m=1}^M \sum_{n=1}^N (|U_{\varepsilon, r}(m, n) - X_{\varepsilon, r}(m, n)|)^q \right), \quad (1)$$

where

$$V(\varepsilon, r) = \sum_{m=1}^M \sum_{n=1}^N (U_{\varepsilon, r}(m, n) - X_{\varepsilon, r}(m, n)) \quad (2)$$

is the "volume" of the image surface $B_r(m, n)$ on the scale ε enclosed between the corresponding coatings.

The determination of the multifractal signature (MFS) $\mathbf{S}_{q,r}$ is carried out by calculating the values of the LMMFP, measured between adjacent analysis scales by the expression

$$L_q(\varepsilon, r) = \left(\log \frac{\varepsilon}{\varepsilon + 1} \right)^{-1} \log \frac{Z(q, \varepsilon + 1, r)}{Z(q, \varepsilon)}, \quad (3)$$

followed by the formation of an array

$$\mathbf{S}_{q,r} = \begin{bmatrix} \mathbf{L}_{-\infty}(\varepsilon_1) & \mathbf{L}_{-\infty}(\varepsilon_2) & \cdots & \mathbf{L}_{-\infty}(\varepsilon_{\max} - 1) \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{L}_{-1}(\varepsilon_1) & \mathbf{L}_{-1}(\varepsilon_2) & \cdots & \mathbf{L}_{-1}(\varepsilon_{\max} - 1) \\ \mathbf{L}_1(\varepsilon_1) & \mathbf{L}_1(\varepsilon_2) & \cdots & \mathbf{L}_1(\varepsilon_{\max} - 1) \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{L}_{\infty}(\varepsilon_1) & \mathbf{L}_{\infty}(\varepsilon_2) & \cdots & \mathbf{L}_{\infty}(\varepsilon_{\max} - 1) \end{bmatrix}, \quad (4)$$

where $\mathbf{L}_q(\varepsilon) = [L_q(\varepsilon, r_1) \ L_q(\varepsilon, r_2) \ \cdots \ L_q(\varepsilon, r_{\max})]^T$ is the column vector of the LMMFP for r_{\max} , the number of rotations of the q_{th} order for the analysis scale ε and $[\cdot]^T$ is the transposition operator. Next, the prevailing orientation directions of the texture elements are determined on the corresponding analysis scales based on the approximation by ellipses of the set of LMMFP values $\{L_q(\varepsilon, r)\}$ formed for a given index q in the polar coordinate system, and the determination of the ellipticity parameters $k_{el}(q, \varepsilon)$ and the angle $\psi(q, \varepsilon)$ ellipse tilt. All further mathematical operations and notation are given in [37].

The results of image processing (Fig. 5) from an array of real radar images showed that the segmentation accuracy using NMMFS turned out to be 24.8–63.5% higher compared to the accuracy achieved using MMFS.

The use of the developed class of topological features of DMMFS in segmentation problems provides an increase in the differentiation accuracy when processing anisotropic images up to 86.5% and improves the segmentation accuracy by 35.6% when processing images

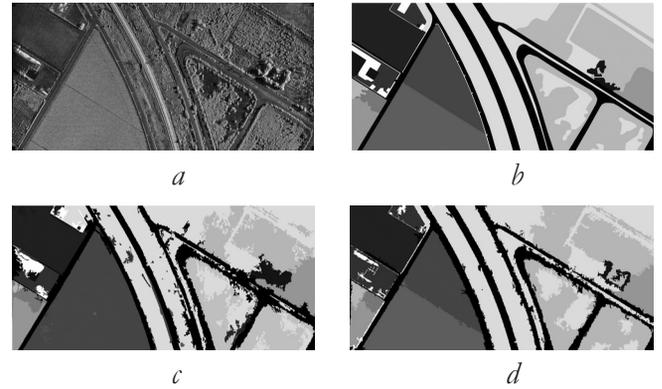


Fig. 5. Results of fractal segmentation of rural radar data (a), performed by an expert manually (b) and using the (Fuzzy C-Means) algorithm using the features of MMFS (c) and NMMFS (d) [37].

with multifractal properties. In the case when the only differentiating characteristic for segmented images is information about the angular dependence of their texture elements, the segmentation accuracy increases by 78%. The use of a new class of topological features in RLI segmentation problems improves the accuracy of their differentiation up to 63.5%.

2.5. FRACTAL AGGREGATION OF RADAR IMAGES OBTAINED BY MULTIBAND SAR

Despite the current existence of methods, methods and algorithms for texture-fractal processing of two-dimensional images, the problem of efficient integration of multidimensional radar images has not yet been solved (see Fig. 6a,b and Fig. 7).

Taking into account the results of a statistical analysis of the brightness values and fractal dimension D of two-dimensional radar images in [32–35], we proposed a new method for complexing radar images obtained by multiband SAR. Multifractal aggregation is based, in contrast to the known ones, on the simultaneous calculation of local multifractal dimensions by the sliding window over all initial radar images by the method of iterative coverage, and can significantly increase the information content of images, estimated by entropy.

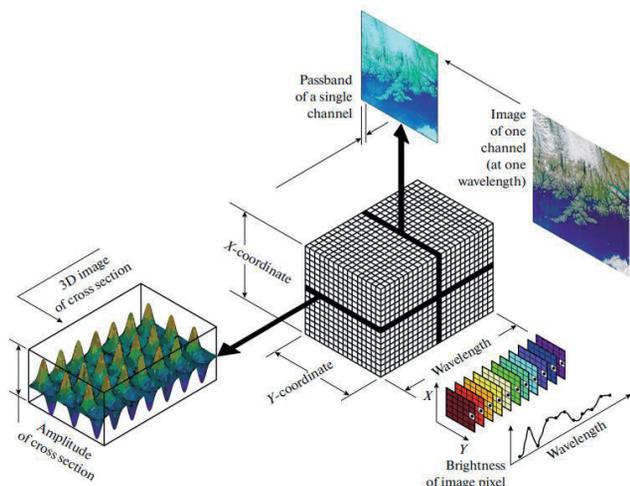


Fig. 6a. Model for storage and analysis of multi- and hyperspectral data [35].

The use of this method can significantly increase the information content of radio systems, especially in the case of jamming and masking objects in some or all frequency ranges.

In general, the scheme of structural-parametric synthesis of systems for optimal texture-fractal processing of multidimensional images is shown in Fig. 8 [35]. Here R_{psegm} , R_{po} , R_{psp} are the probabilities of correct segmentation, detection and recognition.

The presented scheme of structural-parametric synthesis allows us to speak about the optimality of texture-fractal processing of multidimensional images, since the efficiency criterion is the maximum of one or several probabilistic characteristics simultaneously – classification (segmentation), detection or recognition, depending on the tasks solved by the system.

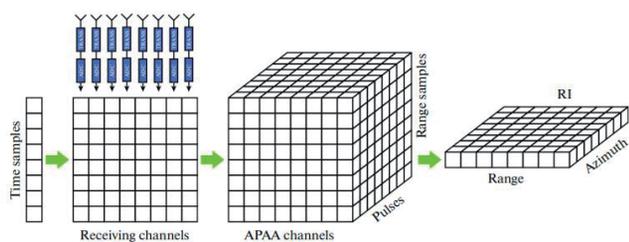


Fig. 6b. Model of Radar Data Cube Formation [35].

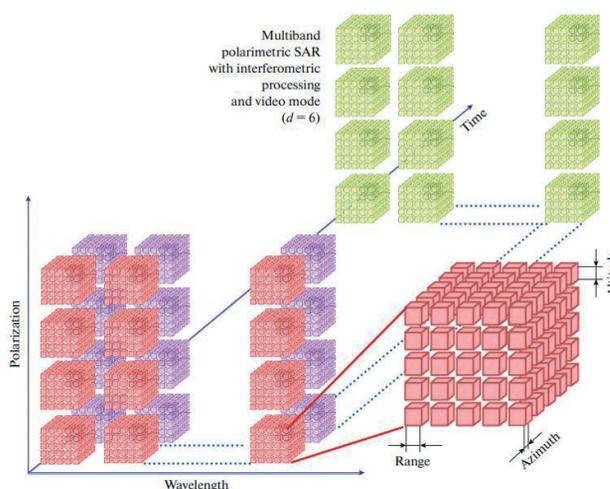


Fig. 7. Model for the formation of a multidimensional cube of radar data [35].

2.6. SELECTED CLASSICAL EXAMPLES OF FRACTAL PROCESSING

Let us present selected experimental results of TTFI of multidimensional signals from objects of different physical nature (Fig. 9-14).

In particular, in Fig. 14 shows the unique multifractal characteristics of high-altitude discharges in the ionosphere. Every day, 4 million lightning strikes the sky, about 50 every second. And above the lead thunderstorm fronts, in the upper atmosphere, a light show of "ghost lightning" unfolds: blue jets, red-violet sprites, red rings of elves soaring in the sky. These are discharges of very high energy, which strike not into the ground, but into the ionosphere!

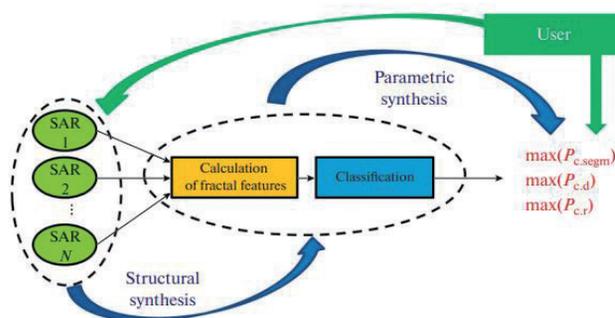


Fig. 8. Scheme of structural-parametric synthesis of the system of texture-fractal processing of multidimensional radar images [35].

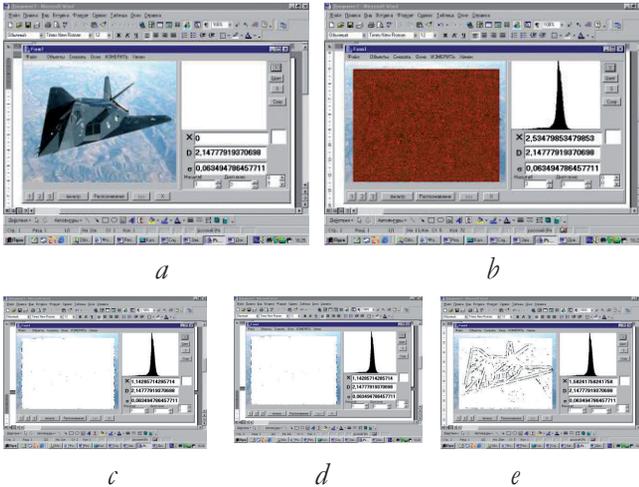


Fig. 9. Fractal image processing of the F117 aircraft image – (a): (b) image of the F117 aircraft in noise at $q_{02} = -3$ dB, (c) example of fractal nonparametric filtering (FPF) of the aircraft image at the current value $D^{(1)} \propto X$, (d) example of the image FPF of the image aircraft at the current value $D(2) > D(1)$, (e) – an example of the FIF of an aircraft image at the value $D(3) > D(2) > D(1)$; everywhere on the right is the current paretian D .

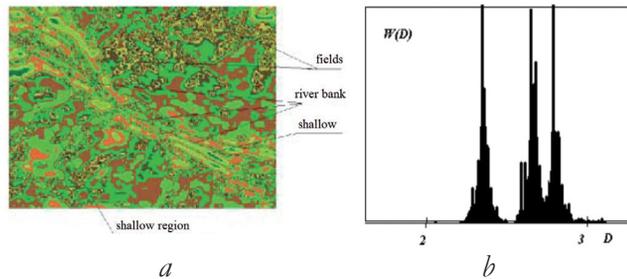


Fig. 10. An example of distinguishing types of earth surfaces by the field of fractal signatures D (left) and the empirical distribution D (right) during segmentation by D of land cover textures on radar imagery.

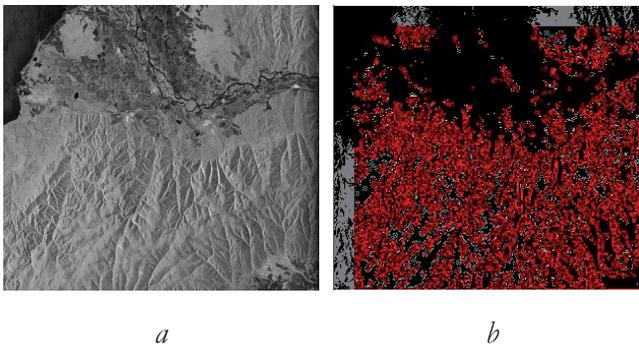


Fig. 11. Delta of the Selenga River in the PALSAR SAR image (left), the result of fractal processing (right); wavelength 23 cm, spatial resolution about 7 m.

The history of their discovery is very interesting. Sprites, for example, were

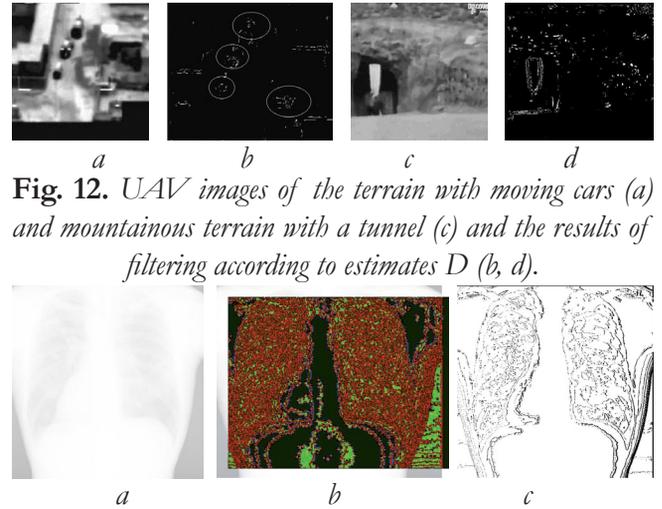


Fig. 12. UAV images of the terrain with moving cars (a) and mountainous terrain with a tunnel (c) and the results of filtering according to estimates D (b, d).

Fig. 13. An example of solving the problem of fractal clustering of an X-ray image (a) by the value of the estimate of the fractal dimension D (b) and fractal contour detection (c).

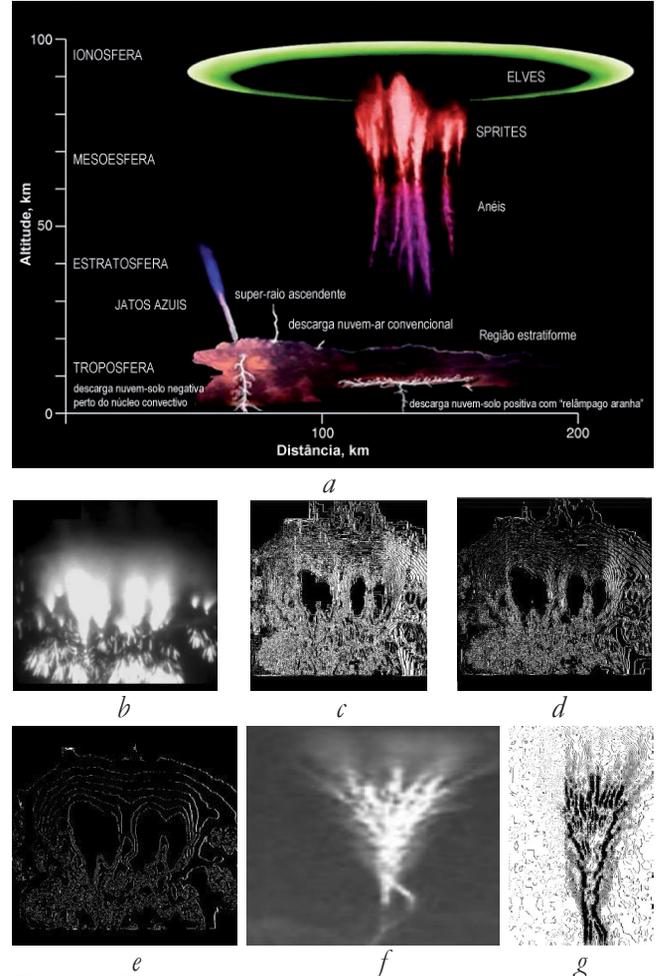


Fig. 14. Dynamic fractal structures in the atmosphere - (a), sprite (shooting from a spacecraft - USA, NASA) - (b); jet (IR - survey from spacecraft - China) - (f); multifractal filtering of the sprite image: map $D = 2.3$ - (c); map $D = 2.6$ (d); map $D = 2.8$ (e); fractal filtering of the jet image: map D - (g); KA is a spacecraft.

discovered by accident on the night of July 5-6, 1989 in the United States.

High-altitude electric discharges (20-100 km) are divided into several main types: elves, jets, sprites, halos, etc. They clearly confirmed the existence of a global electric circuit (GEC) on our planet and provided new opportunities for its study. This literally shocked specialists in astronomy, atmospheric electricity, radiophysics, atmospheric acoustics, gas discharge physics and aerospace safety. On maps of fractal dimension (Fig. 14), external, basic and hyperfine structures are clearly distinguished. Dynamic spatiotemporal features and sprite morphology can be explained, in particular, by the multifractal geometry of discharges and percolation. Here, modeling based on fractal labyrinths is also applicable, which well reflect the physics and morphology of such ionospheric structures. By the way, the data in Fig. 14 are the world's first results of fractal processing of such structures, which at one time caused a stir at radar conferences in the USA and China.

Conclusion: numerous results (SAR, UAV(unmanned aerial vehicle), medicine, space, mechanical engineering, etc.) show that fractal processing methods improve the quality and detail of objects and targets in active and passive modes by about several times. These methods can be successfully used to process information from space, aviation systems, low-observable high-altitude pseudo-satellites (HAPS) or detect HAPS and UAV clusters, synthesized space antenna clusters, space debris, etc.

2.7. RADAR EQUATION FOR A FRACTAL TARGET

In [14], questions of the general theory of multiple scattering of electromagnetic waves in fractal discrete randomly inhomogeneous media are considered in detail on the basis of modifications of the

classical Foldy-Tverskoy theory. An integral equation is obtained for a coherent field and the second moment of the field for a fractal scattering medium.

Based on the developed theory, the value of the backscattered signal from the fractal medium [4,14] was calculated using the classical radar equation. The received signal power P_s is determined by the radar equation.

Here we end up with two cases:

- for the far zone and a flat fractal target (Euclidean dimension $\mathbf{E} = 2$); then

$$P_s \propto \frac{1}{r^{4-D}}, \quad (5)$$

- for the far zone and volumetric fractal target (Euclidean dimension $\mathbf{E} = 3$); then

$$P_s \propto \frac{1}{r^{5-D}}. \quad (6)$$

Here r is the distance to the target. The results (5) and (6) show that the reflected radar signal can be used to estimate the fractal dimension D of the probed fractal medium or fractal target (such as a dynamic layer of snow, rain, etc.).

Similarly, on the basis of (5) and (6), one can obtain a solution for anisotropic disordered large fractal systems: cascades of fractals nested into each other, graphs from chains of fractals, percolation systems, nanosystems, space debris, clusters of UAVs, or small-sized space satellites (SSCs), including mini- and microclasses, dynamic synthesized space antenna constellations (cluster apertures), low-observable high-altitude pseudo-satellites (HAPS), spatially distributed space systems (clusters) from small SSCs for solving problems of monitoring emergency situations, etc.

3. COMPUTATIONAL DIELECTRIC METASURFACES AND TTFS

Continuous improvement of topological texture-fractal processing (TTFP) of signals and fields in modern radio physics and radio electronics implies a constant improvement in the speed of information processing and the search for new physical principles for its implementation. Here, undoubtedly, the future belongs to photonic and radio-photonic technologies. Below are selected results in the field of photonics, radiophotonics, computational meta-optics and 2D dielectric *metamaterials* (MM) or *computational metasurfaces* (MS), which were obtained by us with Chinese scientists at the Joint Laboratory of Information Technology and Fractal Signal Processing of Jinan University in Guangzhou, China for period 2019-2021. The results have been published in leading international scientific journals [38-45] in the USA and Switzerland. It should be noted that China has a special state program, and in 2015 China became the world leader in terms of production of photonics devices.

3.1. PRINCIPLES

The concept of *computational metamaterials* was first introduced in 2014: computational MMs are metamaterials that can perform desired mathematical operations on arbitrary wave signals as they propagate through them. By designing the geometry of dielectric metasurfaces (MS), optical analog computing devices with various functions, such as spatial differentiators, integrators, equation solvers, etc., can be obtained. The most common are MMs, in which the structure of the elements (*meta-atoms* a few nanometers in size), the size, and the distance between the elements are much smaller than the wavelength of the exciting field. In this case, the result of the

interaction of individual elements leads to the fact that the properties of the MM are determined not so much by the properties of its constituent elements, but by an artificially created periodic structure.

MSs as a typical device prototype consist of artificial subwave structures over flat surfaces of a dielectric material, which facilitate flexible control of the amplitude, phase, and polarization of electromagnetic waves. The use of diffuse and transmissive MMs for optical analog computing paves the way for the realization of fully integrated spatial filtering devices. It can be said that MT is *the root of many fascinating topological phenomena in physics and exotic manipulations with waves*.

Despite the huge interest in MS and the presence of a large number of English-language reviews, this issue is practically not covered in the Russian-language literature, with rare exceptions. Multipole resonances of dielectric nanoparticles provide a promising way to tune the optical transfer function (OTF) of an MS. Compared to plasmonic structures, dielectric nanostructures can not only solve the problem of losses and increase the efficiency of the MS, but also facilitate the control of light scattering and propagation. In addition to resonant schemes for realizing spatial differentiation and image edge detection, non-resonant schemes such as spin-orbit interaction in Pancharatnam-Berry phase MSs have been investigated for image edge detection. The use of dielectric MMs instead of classical plasmonic structures reduces ohmic (thermal) losses and makes it possible to control not only the electric, but also the magnetic component of the light wave. In other words, due to the free penetration of the electromagnetic field into dielectrics,

such nanostructures have both electric and magnetic Mie resonance.

3.2. DESIGN

Next, we present the design of a polarization-insensitive MS. Our papers present a way to implement high-performance optical MSs that perform isotropic and polarization-insensitive edge detection on an arbitrary 2D image [39,40,43,44].

The implementation of the Laplace operation as optical analog computing has recently attracted attention, and a compact device with high spatial resolution has not yet been invented. The spatial differentiation operation is important in image processing and its applications such as sharpening and edge-based segmentation. In this case, the Laplacian, the simplest operator of the isotropic derivative in two dimensions, is of particular importance. Here we would like to emphasize that our device operates in transmission mode, while almost all previous ones usually worked in reflection mode. Transfer mode operation is important for image processing because this device can be used directly as the first layer of a recognition system and in image processing applications.

We have proposed a *Laplace metasurface* that can perform almost perfectly the Laplace operation for different configurations of the incident light field – **Fig. 15**. The proposed Laplace MS is based *on the excitation of a bound state in a continuum*, which has demonstrated exotic optical properties. The unit cell (metaatom) consists of a silicon brick (blue) with a thickness $b = 163$ nm and a bottom glass substrate (gray). On the right in Fig. 15 is a top view of the unit cell. The period is $a = 331$ nm, the width of the silicon brick is $d = 251$ nm. Four square voids $s = 33$ nm wide are located in the center of all edges.

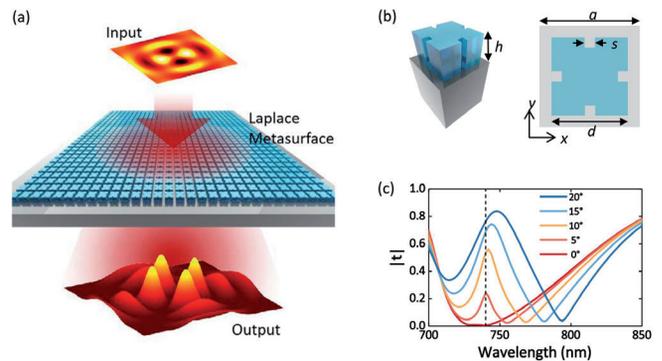


Fig. 15. (a) Dielectric Laplace MS transforming the input two-dimensional space function into another function as the Laplace operator; (b) Unit cell of the Laplace dielectric MS; (c) Transmission spectra of the Laplace MS at various angles of incidence along the x direction for the p -wave.

The highly symmetric mode profile provides an almost isotropic Laplace OTF. The proposed Laplace MS can be configured to operate at different wavelengths in transmit mode, which is advantageous for applications in optical computing, medical diagnostics, machine vision, and more. We emphasize that although we are everywhere talking about the scattering of electromagnetic waves, the results presented can be easily transferred to the case of scattering of waves of a different nature, for example, acoustic waves.

3.3. RESULTS

One of the applications of the Laplace operation is to detect the edges of problematic targets, etc. on the image. We demonstrate that Laplace's proposed MS [44] can be used for traffic sign recognition, which is critical for automated driving (**Fig. 16**).

We also used a typical QR-code as an input 2D image since QR-codes are now important in our daily life and edge detection for them plays a critical role in QR-code region detection (**Fig. 17a**). The QR-code we have chosen carries information about the Chinese hieroglyph meaning "Light". Through the processing procedure discussed

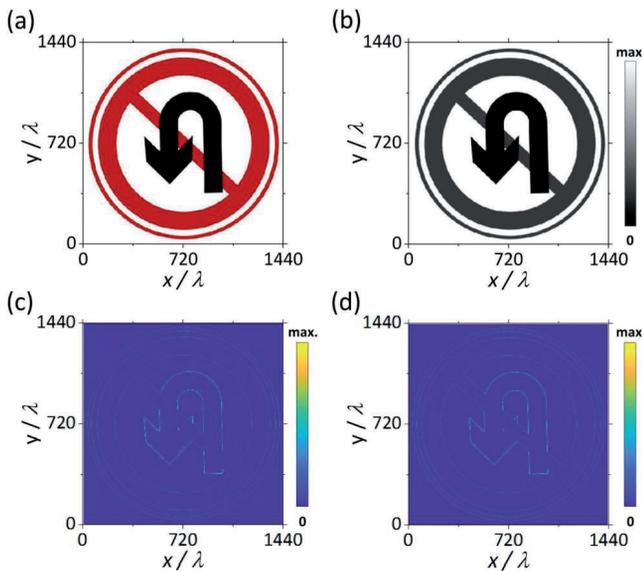


Fig. 16. (a) Color image of a road sign; (b) The corresponding grayscale image as input; (c) and (d) Output image from the ideal Laplace operator and Laplace MP for (b), respectively.

in this context, we can obtain the results from the ideal Laplace operation and from the Laplace MS, which are shown in Fig. 17b and Fig. 17c.

Due to the limited volume of this article, here we only briefly list other areas of photonics, which are also studied in our joint papers [38-45] and are presented in detail there. These are: control of light scattering by nanoparticles using magnetoelectric coupling and zero backscattering (the theory of light scattering by nanoparticles and electromagnetic multipoles, numerical simulation, verification experiments in

the frequency range from 4 to 7.5 GHz) [38,41,42]; strong optomechanical coupling in chain waveguides made of silicon nanoparticles with quasi-bound states in a continuum (photon-phonon interaction with microstructures) [45], etc.

The use of computational dielectric MSs as a whole brings to a new level all the functional characteristics of the multifunctional system of *topological texture-fractal processing* (TTFP) of signals and fields proposed by us at the end of the 20th century in solving classical problems of detection, measurement, recognition and classification by intelligent radio engineering systems and devices. So, the continuum of all the data in Fig. 1-17 shows our use of the fractal paradigm to create real technical projects in radio electronics.

4. FRACTAL ENGINEERING AND PHILOSOPHY OF FRACTAL ENGINEERING

As you know, the use of scientific principles for the design and creation of real technical objects and structures is a field of human activity called engineering. Currently, on the basis of the results of modern science, technology (technologies) and engineering, the world is showing an active interest in posing problems of the philosophy of

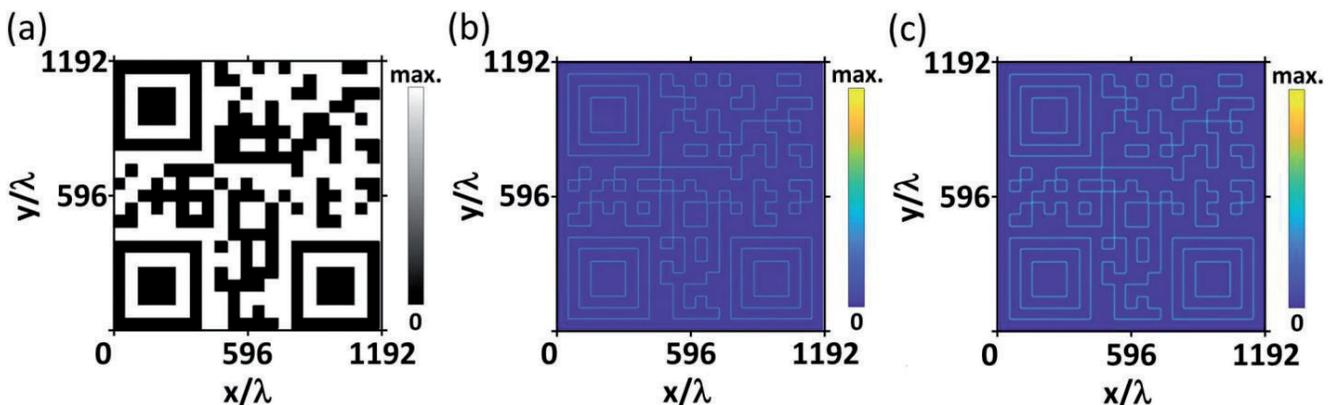


Fig. 17. (a) Input image consisting of a QR code; (b) the output image of the ideal Laplace operation; (c) exit from the Laplace MS. All images are a light intensity profile; the pixel sizes are set to 2.88λ .

engineering. In 2002, the Chinese scientist Li Bocun published the book "Introduction to the Philosophy of Engineering" [18], then in 2003 the American scientist Luis Buchiarelli published the book "Philosophy of Engineering" [19]. In [18], the concept of "three principles" (science, technology, engineering) was introduced, in which they were clearly separated. More than fifty categories of philosophy of engineering (plan, decision-making, goal, drawing up a plan, etc.) were considered, and the problems of philosophy of engineering were analyzed [18,20].

This article proposes the concepts of "fractal engineering" in the hope of contributing to the problems of the philosophy of engineering. The triad of science, technology and engineering (of course, using fractals as an example), regardless of the aforementioned Chinese book on the philosophy of engineering [18], was presented by the author 10 years ago in Issue No. 1 of the RENSIT journal for 2012 – all 142 pages of the issue were given to two large articles [46,47] (essentially a book; thanks to RENSIT Editor-in-Chief for that).

Here it seems appropriate to also present the results on fractal antennas obtained by us at the end of the 20th century. In 1988, the author, together with the Almaz Central Design Bureau, carried out the first development and design of such unusual (for that time) fractal antenna structures (in particular, an operating model of a fractal slotted antenna array was made in the millimeter and centimeter wave range (MMW and SMW) for a portable solid-state dual-frequency coherent radar based on parametrons with a complex phase-shift keyed signal of an extra-large base (patent [54]). This digital radar (the size of a small

case) was installed on a helicopter, and the author worked with it for a long time and received the first radar images of land covers and objects (**Fig. 18** [3-5,7,31,49-54]. And before that it was also necessary to be able to strictly calculate the parameters of a unique two-frequency fractal receiving-transmitting antenna for two bands and then to make several almost industrial samples.

It should be noted that it was on this radar that the author first studied the fractal properties of code M-sequences with a period of up to $2^{20} - 1$ (the base of a complex signal on MMW is up to 1048576). The quantization of the input signal in the radar took place in the stochastic number system. The signal represented by such a code exhibits its fractal properties. Like a hologram, any fragment of which carries

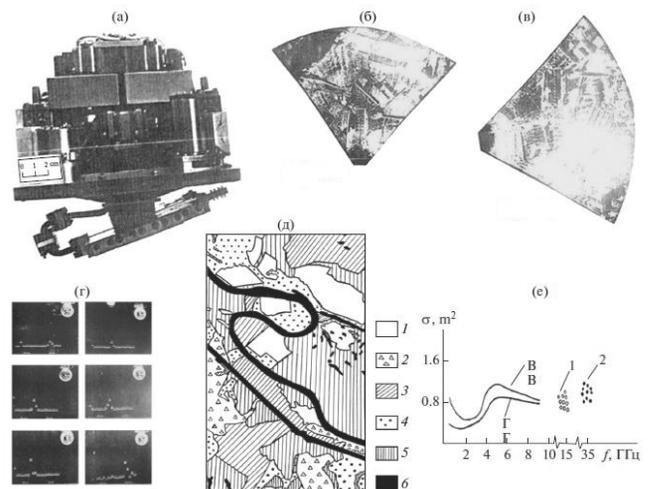


Fig. 18. Portable digital solid-state dual-frequency coherent radar based on parametrons with a complex phase-shift keyed signal of an extra-large base and with a fractal slit grating in the MMW and CMW ranges (a) and some results of field tests of the radar: (b and c) – the first radar images at a wavelength of 8.6 mm; (d) are the characteristic shapes of the envelope of the signal reflected by the ground cover; (e) – an example of a reference synthesized map of an inhomogeneous terrain in terms of energy, textural and fractal features; (f) is the average effective scattering area as a function of frequency for horizontal (H) and vertical (V) polarizations, 1 and 2 are the data of the author, who participated in the field experiment as a "located object" [31,49-54].

information about the complete object, any fragment of a stochastic code contains information about the amplitude of the quantized signal [31]. Then, on this module, a new radar method based on the Radon transform was also implemented for the first time [5,7,31,49-52,54].

All this is actually fractal engineering and fractal engineering with elements of the philosophy of engineering. We can say that that time, namely, the 80s of the XX century, was the beginning of the birth of the Russian philosophy of fractal engineering and fractal engineering! And it was a serious and advanced project in the great USSR.

5. CONCLUSION

The global fractal method was created and demonstrated in many ways in the published works of the author and directly in the references to this work. As a result, a new semantic space has been formed in the scientific world with its properties and tasks unusual for classical radiophysics, radio engineering and radar. I consider the problem of "sample topology" [3-5] to be one of the most important in all radio electronics, and at the same time I am convinced that without fractality and scaling, the entire classical theory of detection and recognition of multidimensional signals in the future will lose its causal significance for the fundamental concepts of signal and noise.

The performed studies are a priority in the world and serve as a basis for further development and substantiation of the practical application of topological fractal-scaling and texture methods in modern radiophysics, radar, nanotechnology and photonics, as well as in the improvement of fundamentally new and more accurate topological texture-fractal optimal methods

for detecting and measurement of signal parameters in the space-time radar and navigation channel of propagation of waves with scattering [3-17,21-37,46-59].

The use of fractal systems, sensors and nodes is a fundamentally new solution that significantly changes the principles of building intelligent radio engineering systems and devices. Fractal processing methods improve the quality and detail of objects and targets by several times.

Computational meta-optics is revolutionizing hardware with the advantages of ultra-fast speed, ultra-low power consumption, and parallel information processing in general-purpose, including fractal, applications. The recent advent of metasurfaces has made it possible to fully manipulate electromagnetic waves at subwavelength scales, promising feature-rich, high-performance, compact, and flat optical processors.

The author expresses his sincere gratitude to the Chinese scientists with whom they jointly completed and published articles on photonics and computational metaoptics in leading international scientific journals [38–45].

REFERENCES

1. Mandelbrot B. *Les Objects Fractal: forme, hasard et dimension*. Paris, Flammarion, 1975, 190 p.
2. Mandelbrot B. *The Fractal Geometry of Nature*. N.Y., Freeman and Comp., 1982, 468 p.
3. *Professor Alexander Alekseevich Potapov. Fractals in Action: A Bio-Bibliographic Index*. Ed. Acad. Yu.V. Gulyaev. Moscow, Raduga Publ., 2019, 256 p.
4. Potapov Alexander A., Wu Hao, Xiong Shan. *Fractality of Wave Fields and Processes in Radar and Control*. Guangzhou: South

- China University of Technology Press, 2020. 280 p.
5. Potapov AA. *Fractals in radiophysics and radar*. Moscow, Logos Publ., 2002, 664 p.; Potapov AA. *Fractals in Radiophysics and Radar: Sampling Topology*. Ed. 2nd, revised and additional, Moscow, Universitetskaya kniga Publ., 2005, 848 p.
 6. Potapov AA. Fractals and chaos as the basis of new breakthrough technologies in modern radio systems. Supplement to the book: Kronover R. *Fractals and chaos in dynamical systems*. Moscow, Technosphaera Publ., 2006, p. 374-479.
 7. Bunkin BV, Reutov AP, Potapov AA et al. *Issues of advanced radar*. Moscow, Radiotekhnika Publ., 2003, 512 p.
 8. Podosenov SA, Potapov AA, Sokolov AA. *Pulse electrodynamics of broadband radio systems and fields of coupled structures*. Moscow, Radiotekhnika Publ., 2003, 720 p.
 9. Potapov AA, Gulyaev YuV, Nikitov SA, Pakhomov AA, German VA. *The latest methods of image processing*. Ed. A.A. Potapov. Moscow, Fizmatlit Publ., 2008, 496 p.
 10. Potapov AA, Bystrov RP, Gvozdev AE, Nemtsov AV. *Selected issues of theory and practice of linear and non-linear radar*. In 3 parts. Ed. A.A. Potapov. Moscow, Central Research Institute of the Ministry of Defense of the Russian Federation Publ., 2014, 654 p.
 11. Podosenov SA, Potapov AA, Fowzon J, Menkova ER. *Nonholonomic, fractal and coupled structures in relativistic continuous media, electrodynamics, quantum mechanics and cosmology*. In 3 vols. Ed. A.A. Potapov. Moscow, URSS Publ., 2016, 1128 p.
 12. Margolin VI, Potapov AA, Farmakovskiy BV, Kuznetsov PA. *Development of nanotechnologies based on nanocomposites*. St. Petersburg, SPbGETU "LETI" Publ., 2016, 190 p.
 13. Potapov Alexander. On the Indicatrices of Waves Scattering from the Random Fractal Anisotropic Surface. ГЛАВА 9 В КН.: *Fractal Analysis - Applications in Physics, Engineering and Technology*. Ed. Fernando Brambila. Rijeka, InTech, 2017, pp. 187-248.
 14. Potapov AA. Multiple scattering of waves by a fractal ensemble of particles and in large disordered fractal systems. In: *"Turbulence, Atmospheric and Climate Dynamics"*. Moscow, Fizmatkniga Publ., 2018, p. 564-573.
 15. Akinshin RN, Rummyantsev VL, Peteshov AV, Pafikov EA, Potapov AA. *Physical foundations of the device of rocket and artillery weapons. Algorithms and devices for the operation of airborne artillery reconnaissance airborne radio equipment*. Penza, Branch of VA MTO, Penz. art. eng. in-t Publ., 2018, 400 p.
 16. Potapov AA. Application of the principles of fractal-scaling or scale-invariant radar in SAR, UAV and MIMO systems. In: *Radar. Results of theoretical and experimental studies*. In 2 books. Book. 2. Moscow, Radiotekhnika Publ., 2019, p. 15-39.
 17. Beliaeva VS, Klyuev DS, Neshcheret AM, Potapov AA, Sokolova YV. Fractal Antenna Systems with Chiral Metamaterials Substrates for MIMO Systems. In: Hu Z., Petoukhov S., He M. (eds). *Advances in Artificial Systems for Medicine and Education V. AIMEE 2021*. (Lecture Notes on Data Engineering and Communications Technologies, V. 107). Cham, Switzerland, Springer Int. Publ., 2022. P. 329-345.
 18. Li B. *Gongcheng zhexue yintun: Wo zao wu gu wo zai* [An introduction to philosophy of engineering: I create, therefore I am]. Zhengzhou, Daxiang chubanshe, 2002, 452 p. (In Chinese).

19. Bucciarelli LL. *Engineering philosophy*. Delft, Delft Univ. Press, 2003, 102 p.
20. Bao Ou. The formation of the philosophy of engineering in China. *Problems of the Far East*, 2013, 4:135-147.
21. Potapov AA. *Fractal method and fractal paradigm in modern natural science*. Voronezh, Nauchnaya kniga Pibl., 2012, 108 p.
22. Potapov AA. Fractal method, fractal paradigm and the method of fractional derivatives in natural science. *Bulletin of the Nizhny Novgorod University. N.I. Lobachevsky. Series: Mathematical Modeling. Optimal control*, 2012, 5(2):172-180.
23. Potapov AA. Mathematical foundations of the fractal-scaling method in statistical radiophysics and applications. *RENSIT: Radioelectronics. Nanosystems. Information Technology*, 2021, 13(3):245-296. DOI: 10.17725.rensit.2021.13.245.
24. Potapov AA. Fractals, scaling and fractional operators in information processing (Moscow Scientific School of Fractal Methods at V.A. Kotelnikov IRE RAS, 1981–2011). *Sat. scientific tr. "Irreversible processes in nature and technology"*. Moscow, MSTU im. N.E. Bauman and the Lebedev Physical Institute of RAS, 2012, IV:5-117.
25. Potapov AA. On strategic directions in the synthesis of new types of radar texture-fractal detectors of low-contrast objects with the selection of their contours and the localization of coordinates against the background of intense interference from the surface of the earth, sea and precipitation. *Proceedings of the IV All-Russian Scientific and Technical Complex "RTI Systems VKO-2016"*, devoted to the 100th anniversary of NIIDAR and the 70th anniversary of the RTI them. acad. A.L. Mints (Moscow, Mints RTI, June 02-03, 2016). Moscow, Bauman MSTU Publ., 2017, p. 438-448.
26. Potapov Alexander A. Chaos Theory, Fractals and Scaling in the Radar: A Look from 2015. ГЛАВА 12 В КН.: *The Foundations of Chaos Revisited: From Poincaré to Recent Advancements*. Ed. C. Skiadas. Switzerland, Basel, Springer Int. Publ., 2016, p. 195-218.
27. Potapov Alexander A. Postulate "The Topology Maximum at the Energy Minimum" for Textural and Fractal-and-Scaling Processing of Multidimensional Super Weak Signals against a Background of Noises. ГЛАВА 3 В КН.: *Nonlinearity: Problems, Solutions and Applications. Vol. 2*. Ed. Ludmila A. Uvarova, Alexey B. Nadykto, and Anatoly V. Latyshev. New York, Nova Science Publ., 2017, p. 35-94.
28. Potapov AA. Fractal-scaling or scale-invariant radar: discovery, justification and ways of development. Topical issues of the state, operation and development of aircraft avionics systems, problems of training specialists. *Sat. scientific Art. Based on the materials of the II All-Russia. scientific-practical. conf. "Avionics"* (March 16-17, 2017). Voronezh, Military Educational and Scientific Center of the Air Force "Air Force Academy named after Professor N. E. Zhukovsky and Yu. A. Gagarin" Publ., 2017, p. 143-152.
29. Potapov AA. Fractal Scaling or Scale-invariant Radar: A Breakthrough into the Future. *Universal Journal of Physics and Application*, 2017, 11(1):13-32.
30. Potapov AA. Fractal-scaling or scale-invariant radar and fractal processing of signals and images. *Sat. scientific works for the 65th anniversary of the creation of the Kotelnikov IRE of RAS and the 110th anniversary of Academician V.A.*

- Kotelnikov. Moscow, Kotelnikov IRE of RAS Publ., 2018, p. 99-104.
31. Potapov AA. Synthesis of images of the earth's covers in the optical and millimeter wave bands (01.04.03 - Radiophysics). *Abstract dis. ... Dr. phys.-math. nauk*, Moscow, IRE RAN Publ., 1994, 44 p.
32. Kuznetsov VA, Potapov AA, Alikulov EA. Method for fractal complexing of multifrequency radar images. *Patent RU 2746038 C1 G06T 5/50* (2020.09). (Invention priority 09/05/2020, State registration date 04/06/2021). Bull. No. 10.
33. Potapov AA, Kuznetsov VA, Alikulov EA. Analysis of methods for complexing images generated by multi-band radar stations with a synthetic aperture. *Izv. Russian universities. Radioelectronics*, 2021, 24(3):6-21.
34. Potapov AA, Kuznetsov VA, Alikulov EA. Multifractal fusion of multiband radar images. *TeX report Int. conf. "XXII Kharitonov thematic scientific readings. Supercomputer modeling and artificial intelligence"* (Sarov, May 24-27, 2021). Sarov, RFNC-VNIIEF Publishing House, 2021, p. 108-109.
35. Potapov AA, Kuznetsov VA, Alikulov EA. Structural-parametric synthesis of systems for optimal texture-fractal processing of multidimensional radar images. *Radiotekhnika i elektronika*, 2022, 67(1):51-67.
36. Potapov AA, Kuznetsov VA, Alikulov EA. Simulation model for the formation of images of radar stations with a synthetic aperture based on a multi-beam approach, taking into account the presence of radar shadows. *Sat. scientific tr. "Microwave engineering and telecommunication technologies"*, Issue. 3:460-461. Sevastopol, 2021.
37. Potapov AA, Kuznetsov VA, Pototsky AN. A new class of topological textural-multifractal features and their application for processing radar and optical low-contrast images. *Radiotekhnika i elektronika*, 2021, 66(5):457-467.
38. Danping Pan, Tianhua Feng, Wei Zhang, and Alexander A. Potapov. Unidirectional light scattering by electric dipoles induced in plasmonic nanoparticles. *Optics Letters*, 2019, 44(11):2943-2946.
39. Wan Lei, Pan Danping, Yang Shuaifeng, Zhang Wei, Potapov Alexander A, Wu Xia, Liu Weiping, Feng Tianhua, and Li Zhaohui. Optical analog computing of spatial differentiation and edge detection with dielectric metasurfaces. *Optics Letters*, 2020, 45(7):2070-2073.
40. Pan Danping, Wan Lei, Potapov Alexander A., and Feng Tianhua. Performing Spatial Differentiation and Edge Detection with Dielectric metasurfaces. *QELS_Fundamental Science "OSA Technical Digest Conf. on Lasers and Electro-Optics (CLEO) (San Jose, California, USA, 10-15 May 2020)"*. Washington, Optical Society of America, 2020, Paper FW4B.2.pdf. 2 pp. (From the session "Inverse Design and Computation (FW4B)").
41. Feng Tianhua, Potapov Alexander A., Liang Zixian, and Xu Yi. Huygens Metasurfaces Based on Congener Dipole Excitations. *Physical Review Applied*, 2020, 13(2):021002.
42. Feng Tianhua, Yang Shuaifeng, Lai Ning, Chen Weilian, Pan Danping, Zhang Wei, Potapov Alexander A., Liang Zixian, and Xu Yi. Manipulating light scattering by nanoparticles with magnetoelectric coupling. *Physical Review B.*, 2020, 102(20), 205428.

43. Wan Lei, Pan Danping, Feng Tianhua, Liu Weiping, Potapov A.A. A review of dielectric optical metasurfaces for spatial differentiation and edge detection. *Frontiers of Optoelectronics*, 2021, 14(2):187-200.
44. Wan Lei, Pan Danping, Ouyang Min, Zhang Wei, Potapov Alexander A., Liu Weiping, Liang Zixian, Feng Tianhua, Li Zhaohui. Laplace metasurfaces for optical analog computing based on quasi-bound states in the continuum. *Photonics Research*, 2021, 9(9):1758-1766.
45. Yang Shuairfeng, Wan Lei, Wang Fugen, Potapov Alexander A., and Feng Tianhua. Strong optomechanical coupling in chain-like waveguides of silicon nanoparticles with quasi-bound states in the continuum. *Optics Letters*, 2021, 46(18):4466-4469.
46. Potapov AA. Essays on the development of fractional calculus in the works of A.V. Letnikov. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2012, 4(1):3-102.
47. Potapov AA. Reflections on the fractal method, the method of fractional integro derivatives and the fractal paradigm in modern natural science (From the author's notebooks). *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2012, 4(1):103-142.
48. Zasovin EA, Borzov AB, Bystrov RP, Ilyasov EP, Potapov AA, Sokolov AV, Titov AN. *Radio engineering and radio-optical systems*. Moscow, Krugly God Publishing House, 2001, 752 p.
49. Opalenov YuV, Potapov AA, Fedyunin SYu. Formation of complex phase-shift keyed signals in remote sensing problems. *Tez. report STC "Formation of complex signals"* (Suzdal, 28.11-I.12.88). Moscow, Publishing house of the USSR NIO Union, 1988, p. 49.
50. Opalenov YuV, Potapov AA, Fedyunin SYu. Radiophysical measuring complex with a complex phase-shift keyed signal for the study of natural resources: principles of construction and analysis of experimental results. *Tez. report II All. STC "Theory and Technique of Spatio-Temporal Signal Processing"* (Sverdlovsk, 6-10.06.89). Sverdlovsk, UPI Publishing House, 1989, p. 110-111.
51. Opalenov YuV, Potapov AA, Fedyunin SYu. Radiophysical measuring complex with a complex FM signal in the millimeter wave range. *Radiotekhnika*, 1991, 11:67-70.
52. Opalenov YuV, Potapov AA, Strukov AZ. Experimental studies of radars with full digital polarization sounding and fractal-radon filtering. *Tez. report II STC "Radio-optical technologies in instrumentation"* (Sochi, 13-17.09.2004). Moscow, Publishing house of Bauman MSTU, 2004, p. 26-27.
53. Potapov AA. Fractal electrodynamics. Numerical modeling of small fractal antenna devices and fractal 3D microstrip resonators for modern ultra-wideband or multiband radio engineering systems. *Radiotekhnika i elektronika*, 2019, 64(7):629-665.
54. Opalenov YuV, Potapov AA. A method for synthesizing a radar image and a device for its implementation (Registered in the State Register of Inventions of the Russian Federation on August 27, 2003). *Patent for invention No. 2211461* dated August 27, 2003. Priority from 18.06.2001.
55. Akinshin NS, Potapov AA, Minakov EI, Timoshenko AV, Perlov AYu. A method for evaluating the performance of a computer complex in monitoring the technical condition of radar stations and sensor platforms. *Radiotekhnika i elektronika*, 2022, 67(5):493-499.

56. Potapov Alexander Alekseevich. Brief scientific biography. In: "*International New Materials Industrial Development Forum*" (Jining, China, 11-13 December, 2019). Jining: Jining National High-tech Industrial Development Zone, 2019, p. 8 (Chinese, Japanese, Russian).
57. Potapov AA. Fractal engineering and fractal engineering are new concepts in the theory and practice of fractals and dynamic chaos. Proc. of the 12th Int. Scientific Conf. "Chaos and Structures in Nonlinear Systems. Theory and Experiment" (Kazakhstan, Pavlodar-Bayanaul, June 17-19, 2022). Pavlodar-Bayanaul, Toraighyrov University Publ., 2022.
58. Potapov AA. Fractal technologies: problems and prospects. *Sat. abstract XXVII Baikal All-Russian. conf. with international participation "Information and mathematical technologies in science and management"*, Baikal session (Irkutsk, Baikal, Olkhon Island, July 1-8, 2022). Irkutsk, ISEM SB RAS, 2022, <https://conf.isem.irk.ru/event/14/abstracts/407/>.
59. Potapov Alexander A. Author's Approaches to Fractal Engineering and the Philosophy of Fractal Engineering: Fractal Radio Systems and International Priorities in the Study of Fractal Applications in Radio Electronics. *Proc. IEEE Conf. Antenna Measurements and Applications (IEEE CAMA)* (China, Guangzhou, 14-17 November, 2022). (In press).

DOI: 10.17725/rensit.2022.14.233

Numerical simulation of low-voltage logic gates based on field-effect nanotransistors with a variable working area diameter

Nikolae V. Masalsky

Research Institute for System Research of the Russian Academy of Sciences, <http://www.niisi.ru/>
Moscow 117218, Russian Federation

E-mail: volkov@niisi.ras.ru

Received May 27, 2022, peer-reviewed June 1, 2022, accepted June 8, 2022

Abstract: The results of numerical simulation of low-voltage logic gates based on silicon CMOS nanotransistors with a cylindrical geometry of a fully enclosing gate with a variable diameter of the working area are presented. Changing the geometry of the working area from a conventional cylindrical shape to a truncated conical shape improves the electrophysical characteristics and allows you to compensate for the limitations that arise due to scaling. Numerical studies of conical prototypes were carried out using mathematical simulation performed using the TCAD instrument technological simulation program based on the n- and p-type nanotransistor models developed by TCAD. The simulation results demonstrate improved electrostatic characteristics. The conical structure in the control voltage range from 0 to 0.6 V is characterized by a higher transistor current, a maximum I_{on}/I_{off} current ratio, a low leakage current and a slope of the subthreshold characteristic close to the theoretical value. The dynamic characteristics of the developed physical models of the inverter and the chain of 11 inverters are numerically investigated for the optimized ratio of the diameters of the working area of 8.1/10 nm and the gate length of 25 nm. At control voltages of 0.6 V and a frequency of 25 GHz, the inverter model predicts a maximum switching delay of 1.5 ps, the limit level of active power of 0.21 μ W, static 4.4 pW. The active power consumed by the circuit consisting of 11 inverters is 2.34 μ W at control voltages of 0.6 V.

Keywords: silicon-on-insulator technology, cylindrical CMOS nanotransistor, logic gate, low supply voltage, TCAD simulation

UDC 621.382.323

Acknowledgments: The work was carried out within the framework of the state task of the Federal State Institution FNTs NIISI RAS "Conducting fundamental scientific research (47 GP)" on topic No. FNEF-2022-0022 "Mathematical support and tools for modeling, designing and developing elements of complex technical systems, software systems and telecommunication networks in various problem-oriented areas".

For citation: Nikolae V. Masalsky. Numerical simulation of low-voltage logic gates based on field-effect nanotransistors with a variable working area diameter. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2022, 14(3):233-242. DOI: 10.17725/rensit.2022.14.233.

CONTENTS

1. INTRODUCTION (233)
 2. TRANSISTOR PROTOTYPE STRUCTURE (234)
 3. PROTOTYPE MODEL (235)
 4. TRANSISTOR PROTOTYPE SIMULATION (236)
 5. INVERTER SYNTHESIS (238)
 6. CONCLUSION (240)
- REFERENCES (240)

1. INTRODUCTION

Modern trends in scaling CMOS nanotransistors a priori lead to a decrease in the distance between the source and drain, reducing the controllability of the gate and violating the classical potential distribution, which thereby worsens the carrier transfer in

the transistor's working area [1-3]. Despite this, further scaling and improvement of the design of CMOS transistors continues to improve the performance of integrated circuits [4]. Various technological and topological solutions are used to level the degradation consequences of scaling [5-9]. For example, heterogeneous alloying or the use of gate materials with different output operation. The undoubted success of these solutions is always associated with a great complication (a sharp rise in price) of the technological process of manufacturing chips [4]. In this paper, we consider a new approach to overcome the effects of scaling associated with changing the geometry of the transistor's working area. As a result, its performance may increase, as well as compensate for scaling limitations, in particular due to short-channel effects (SCE).

Silicon-on-insulator (SOI) structures have been recognized as the basic material for the development of nanoelectronics. In order to effectively use the area occupied by the SOI transistor, the transistor channel can be induced over the entire surface of the working area – a 3D channel [10,11]. In this paper, the cylindrical architecture of the CMOS nanotransistors with a surrounding gate and with a variable diameter of the working area in the form of a truncated cone is investigated. In this case, from the drain side, the diameter of the working area is smaller than its diameter from the source side. This technique achieves the screening of the drain, which reduces the influence of hot media. As a result, the transistor current in the open state increases [10,12]. In addition to this, this design is characterized by significant SCE suppression and a low capacitance value.

In this paper, the potential application of the proposed architecture of the SOI CMOS

nanotransistor for the development of low-voltage logic gates are investigated. To solve this problem, their electrophysical characteristics are optimized using mathematical modeling carried out in the TCAD instrument-technological modeling software environment [13] based on the TCAD models of n- and p-types of nanotransistors developed here, and the dynamic characteristics of logic gates with low supply voltage are analyzed.

2. TRANSISTOR PROTOTYPE STRUCTURE

On **Fig. 1** shows the 3D architecture of a silicon cylindrical CMOS nanotransistor with a conical working area and a surrounding

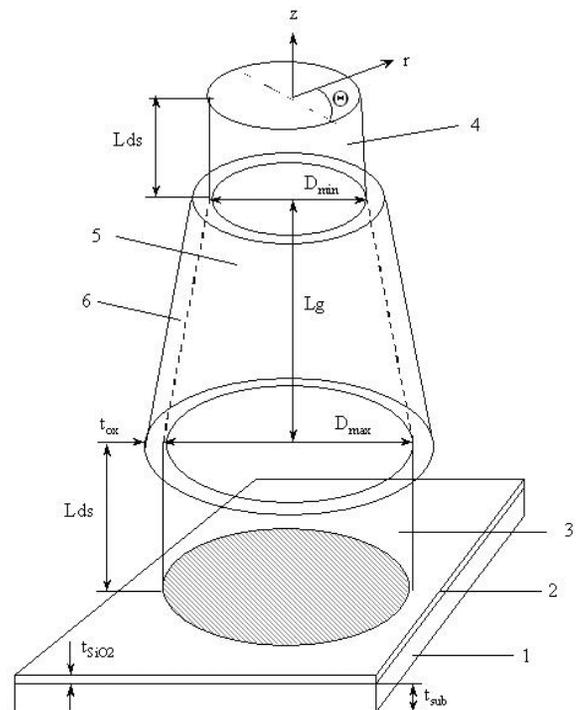


Fig. 1. A sketch of a silicon cylindrical surrounding gate SOI CMOS nanotransistor with a working area in the form of a truncated cone. Here 1 is a silicon substrate, 2 is a silicon oxide film, 3 is a source, 4 is a drain, 5 is a working area, 6 is a gate dielectric, L_{ds} is the length of the drain (source), D_{max} is the diameter of the working area areas from the source side, D_{min} is the diameter of the working area from the drain side, the z axis is directed from the source to the drain, the origin is on the upper face of the source.

gate. The working area of the transistor is a truncated cone with a height of L_g . It is completely covered by a silicon oxide film (a gate dielectric) with a thickness of t_{ox} . The working area is limited to the source and drain areas. The entire structure is placed on a SOI substrate with a thickness of t_{sub} and a thickness of silicon oxide t_{SiO_2} .

It is well known that when scaling cylindrical transistor structures, only a decrease in the diameter of the working area (D) leads to a decrease in the transistor current [14]. At the same time, in order to suppress SCE in a cylindrical architecture, the connecting geometric parameters must be met: $L_g \leq 8.6\sqrt{Dt_{ox}}$ [11,15]. At the same time, the diameter of the working area is scaled more easily than the thickness t_{ox} , which determines both the breakdown voltage at the gate and the level of the tunnel current of the gate [11]. The competitiveness of these factors was determined by the development of the design of the working area of a cylindrical SOI CMOS nanotransistor with a working area in the form of a truncated cone, with the help of which the restriction on parameter D is removed. In this case, from the source side for a large diameter (or D_{max}), the condition for suppressing the SCE is not fulfilled, and from the drain side for a small diameter (D_{min}), it is fulfilled. At the same time $D(z) = D_{min} - \frac{\Delta D}{L_g}z$ и $\Delta D < L_g$, where $\Delta D = D_{max} - D_{min}$. This task is solved for the minimum possible L_g , D_{max} , D_{min} and low-alloy workspace. The requirement for the doping level is important, because a low level ensures approximately uniform distribution of carriers over the entire cross-section of the working area. As the channel doping increases, the carrier potential on its surface will be greater than in the center of the working area. This will pull the electrons to the surface and lead to a deterioration in

the flow of current in the narrowing part of the working area at the drain.

3. PROTOTYPE MODEL

Computer simulation of the electrophysical characteristics of transistor structures was carried out using the DESSIS program of the ISE TCAD package [13]. A TCAD model has been developed for a silicon cylindrical surrounding gate SOI CMOS nanotransistor with a working area in the form of a truncated cone. The "vertical" was considered (see Fig. 1) a design variant taking into account the surface recombination of charge carriers by the Shockley-Reed-Hall mechanism, high degradation of field mobility and without taking into account quantum effects [16,17]. In our calculations, the diameter D varied, the gate length L_g is fixed. The floating base effect is compensated by the choice of the D range. It should be borne in mind that with a decrease in diameter, the capacitance of the drain-source junction decreases, the threshold voltage (U_{th}) increases accordingly and, consequently, the response time of the transistor [8]. The level of doping of the source and drain is limited by an exponential increase in the direct tunnel current between the source and drain and a decrease in the breakdown voltage [16,18]. The boundaries of the working area and the source/drain are sharp. There is no overlap between the gate and the source and drain areas. The thickness of t_{ox} is set in such a way as to exclude the influence of the constant tunnel current of the gate

In the course of numerical experiments, prototypes with different D_{max}/D_{min} ratios were analyzed. The main model parameters of the prototypes are given in **Table 1**.

The following designations are accepted in Table 1. Here N_{ds} is the source and drain doping concentration, t_g is the thickness of the polysilicon gate, v_{SRH} is the rate of

Table 1

Main parameters of prototypes

Parameter	Value	Parameter	Value
L_g, nm	25	N_A, cm^{-3}	1.5×10^{15}
D, nm	6.5-10	L_{ds}, nm	50
t_{ox}, nm	1.2	N_{ds}, cm^{-3}	5×10^{19}
t_{SiO_2}, nm	50	$v_{SRH}, \text{cm/sec}$	3×10^5
$t_{sub}, \mu\text{m}$	0.6	$t_{SRH}, \mu\text{s}$	10
t_g, nm	8	T, K	300

surface recombination by the Shockley-Reed-Hall mechanism, t_{SRH} is the lifetime of non-basic charge carriers by the Shockley-Reed-Hall mechanism.

4. TRANSISTOR PROTOTYPES SIMULATION

Simulation of the electrophysical characteristics of prototypes of SOI CMOS nanotransistors with cylindrical geometry, a surrounding gate and a conical working area was carried out in the range of control voltages from 0 to 0.6 V.

On Fig. 2 shows the longitudinal profiles of the surface potential of prototypes of conical CMOS transistors with different diameter ratios. They are compared with the potential of classical cylindrical geometry at $D_{min}/D_{max} = 1$. It is clearly seen that the profiles of the surface potential of conical prototypes on the drain side are shifted upwards (has a large curvature) due to the narrowing of the silicon

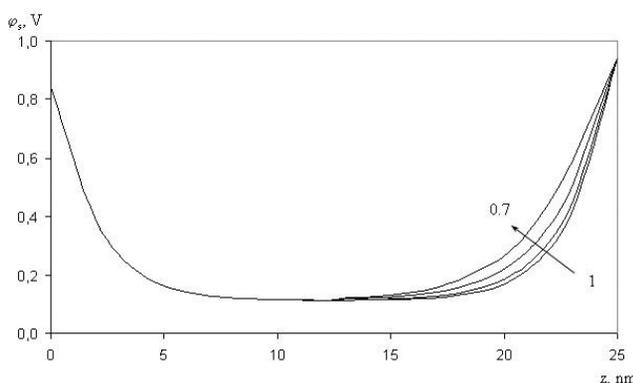


Fig. 2. Longitudinal distributions of the surface potential at $U_{ds} = 0.1 \text{ V}$ in the working area of prototypes with different D_{min}/D_{max} ratios from 1 (the lowest curve) to 0.7 (the uppermost curve) in increments of 0.1.

working area. Then, a kind of drain shielding is implemented, which reduces the electric field strength in this area [5,8].

It also follows from the simulation results that for all conical prototypes, the peak of the electric field is less than for cylindrical geometry. The magnitude of the electric field peak is minimal for $D_{min}/D_{max} = 0.65$ and maximal for $D_{min}/D_{max} = 1$. Such a decrease in the electric field has a positive effect on the functioning of the transistor due to a decrease in the number of hot carriers.

Various key transistor characteristics were analyzed depending on the magnitude of the D_{min}/D_{max} ratio at maximum control voltages. The obtained results show a strong dependence of the transistor characteristics on the D_{min}/D_{max} ratio. For the studied prototypes, the dependence of the threshold voltage (U_{th}) on D_{min}/D_{max} becomes less convex, different from the classical roll-off [19].

On Fig. 3 shows the results of calculations of the maximum current of the transistor at different values of D_{min}/D_{max} in the range 0.65–1.

Fig. 3 shows that a conical design in a certain range of the D_{min}/D_{max} ratio provides a larger drain current compared to a conventional cylindrical working area. This increase in

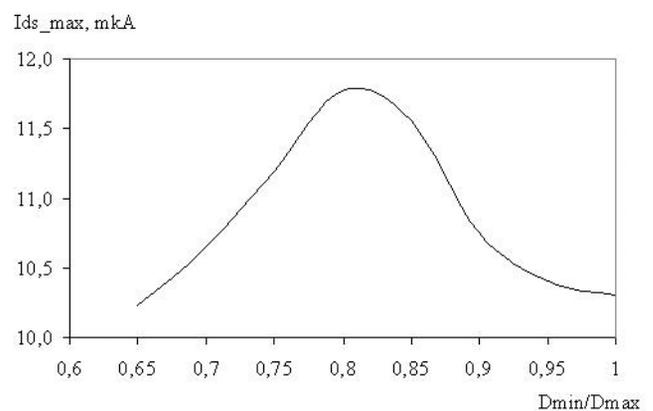


Fig. 3. Dependence of the maximum current (I_{ds_max}) of prototypes at $U_{ds} = U_{gs} = 0.6 \text{ V}$ on the ratio D_{min}/D_{max} .

the drain current is due to a more uniform distribution of carriers in the cross section of the working area, which occurs due to the narrowing of the channel at the drain. In this case, the most favorable conditions for the transport of carriers are created: the absence of hot carriers and a high accelerating potential, which determines the most effective transfer of carriers to the drain [1,7,20]. In experiments, the maximum current is $11.8 \mu\text{A}$ at $D_{\min}/D_{\max} = 0.81$. Compared to the cylindrical working area ($D_{\min}/D_{\max} = 1$), the maximum current is $1.5 \mu\text{A}$ lower, or 12% less. Reducing the D_{\min}/D_{\max} parameter reduces the silicon thickness, which negatively affects the mobility and conductivity of the prototype. At large values of D_{\min}/D_{\max} , the current decreases due to an increase in hot carriers due to deterioration of the drain shielding.

The I_{off} current plays an important role in the functioning of the nanotransistor. This current reduces the switching current and affects the performance of the transistor and the chip, ensures its noise immunity [4,5]. The leakage current must be kept very low to reduce static power dissipation. From Fig. 4, where the results of analytical modeling are presented, it can be concluded that conical prototypes provide a lower leakage current in the off state compared to a cylindrical design.

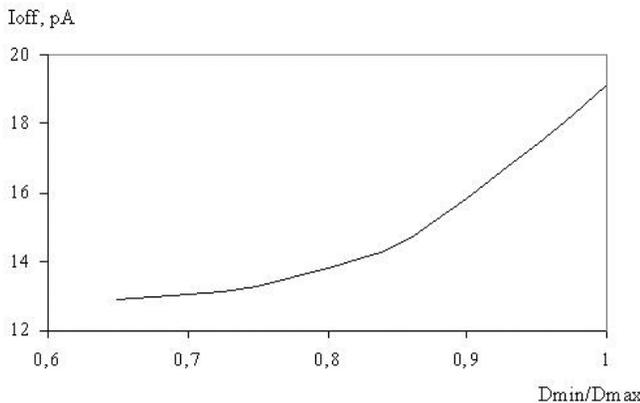


Fig. 4. Dependence of the I_{off} (D_{\min}/D_{\max}) at $U_{\text{ds}} = 0.6 \text{ V}$.

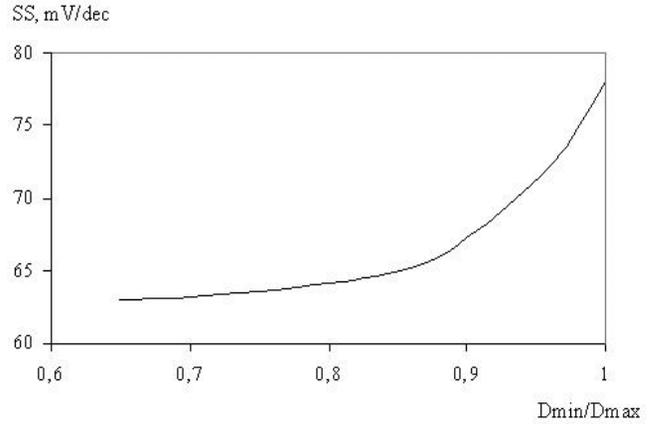


Fig. 5. Subthreshold slope (SS) from the D_{\min}/D_{\max} ratio.

The subthreshold slope (SS) is an important parameter for providing static and dynamic characteristics of short-channel nanotransistors. In particular, in order to obtain a high $I_{\text{on}}/I_{\text{off}}$ ratio, the subthreshold slope must be low, close to the theoretical aisle. On Fig. 5 shows the extracted values of the SS parameter.

From the results shown in Fig. 5, it can be seen that the SS characteristic improves as the D_{\min}/D_{\max} ratio decreases. At the same time, the steepness becomes more sensitive to changes in D_{\min}/D_{\max} at its small values due to the complex behavior of U_{th} with a decrease in D_{\min}/D_{\max} . Fig. 5 shows that the conical structure has a smaller subthreshold slope compared to the cylindrical one. For an optimized transistor structure, $SS = 63.0 \text{ mV/dec}$. With an increase in the D_{\min}/D_{\max} parameter, a steady increase in SS is observed. After the value $D_{\min}/D_{\max} = 0.86$, it becomes critical. This indicates an increase in the influence of SCE and limits the range of possible values of the D_{\min}/D_{\max} ratio.

For further discussion, we chose transistor prototypes with n - and p -conductivities, the technological parameters of which correspond to the case of $D_{\min}/D_{\max} = 0.81$. Among all the prototypes studied, they have the maximum I_{on} current value at control voltages of 0.6 V , the SS value, which is

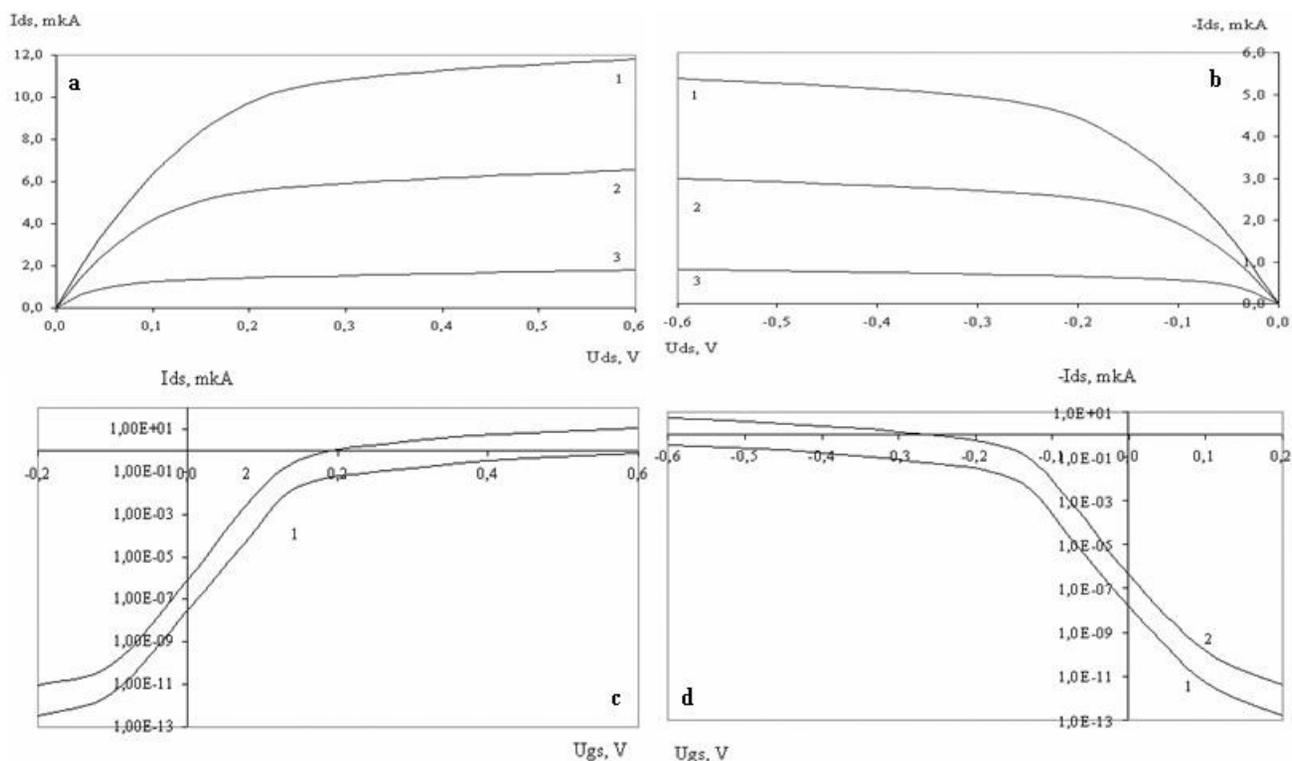


Fig. 6. *I-V data prototypes, where: a) n is the type of I-V data $I_{ds}(U_{ds})$ for different U_{gs} : 1-0.6 V, 2-0.4 V, 3-0.2; b) p-type of I-V data $I_{ds}(U_{ds})$ at different U_{gs} : 1- -0.6 V, 2- -0.4 V, 3- -0.2 V; c) n-type of I-V data $I_{ds}(U_{gs})$ at different U_{ds} : 1-0.05 V, 2-0.6 V; d) p-type of I-V data $I_{ds}(U_{gs})$ at different U_{ds} : 1- -0.05 V, 2- -0.6 V.*

comparable to other prototypes, and close to ideal. The results of the simulation of $I_{ds}(U_{ds})$ and $I_{ds}(U_{gs})$ I-V data prototypes in the control voltage range from 0 to 0.6 V are shown in **Fig. 6** and **Table 2**.

For the design and development of electronic devices, it is important to take into account the I_{on}/I_{off} current ratio [11,21]. For an optimized transistor structure, the current ratio is six orders of magnitude. It should be noted that the I_{on}/I_{off} ratio for transistors with different conductivities practically do not differ from each other and with a change in the supply voltage, they synchronously follow this change.

Table 2

Parameters of the transistor prototype (values for n- and p-types of conductivities are indicated via "/")

Parameter	Value
U_{th} , mV	108/-118
I_{on} , μA	11.8/-5.6
I_{off} , pA	13.9/-6.8

5. INVERTER SYNTHESIS

With the help of the TCAD program, using the models of the studied transistor structures of n- and p-types of conductivity developed in this work, the propagation of a high-frequency logic signal through an inverter on the prototypes selected above is simulated. A sketch of the placement of transistors included in the inverter is shown in **Fig. 7**.

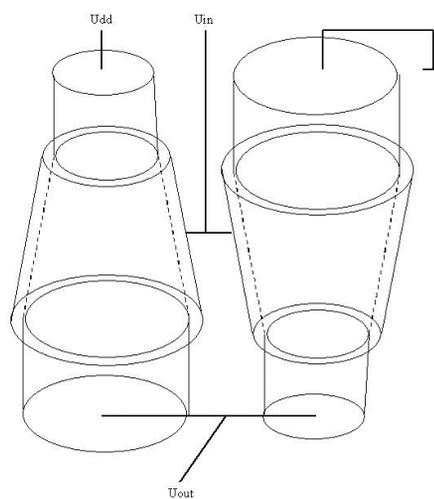


Fig. 7. *Sketch of the inverter.*

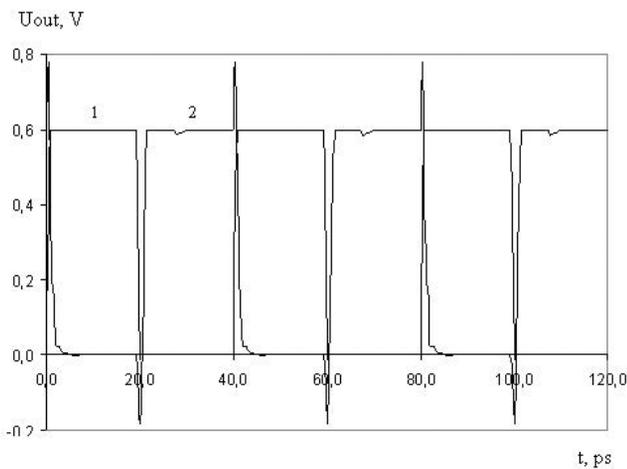


Fig. 8. Dynamic characteristic of the inverter, where 1 is the clock (input) signal, 2 is the response of the inverter.

Table 3

Characteristics of the inverter

Parameters		
f , psc	P , μW	P_{stat} , pW τ
1.5/0.72	0.21	4.4

The results of simulating the inverter switching at control voltages $U_{dd} = U_{in} = 0.6 V$ and a clock frequency of 25 GHz are shown in Fig. 8.

It should be noted that delays τ_{in} and τ_{out} , where τ_{in} and τ_{out} – the time delay of the rise and fall of the signal, respectively, differ from each other. The extracted values of delay and power (active and static) are given in Table 3.

It should be noted that due to the I-V data saturation effect, the $I_{ds}(U_{ds})$ dependencies are nonlinear. Therefore, the dependence of $U_{out}(U_{in})$ will also be nonlinear. The steepness of this characteristic is a measure of how well a given circuit can perform digital operations [22].

On Fig. 9 shows, the dependence of the total delay ($\tau_{in} + \tau_{out}$) at voltage U_{dd} obtained from the calculation results.

In this case, in the range from 0.6 to 0.52 V, the delay increases linearly (with a slope of 1.625 ps/V) with a decrease in U_{dd} . With a fixed supply voltage U_{dd} , the delay change can be represented as follows:

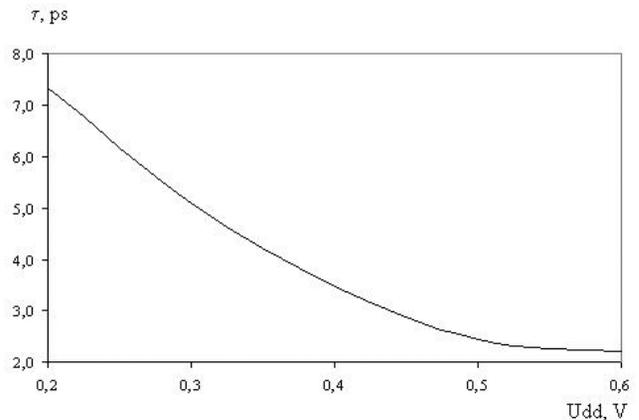


Fig. 9. Dependence $\tau(U_{dd})$.

$$\tau = \tau_0 \left(1 - \lambda \left(\frac{U_{in} - U_{th}}{SS} \ln 10 \right) \right), \quad (1)$$

where $\lambda = 2 - C_{ox} (1 / (1 + C_s / C_{ox}))$, it is associated with the capacitive interaction of adjacent working areas, the capacity of which is C_s , and the subgate oxide, the capacity of which is C_{ox} [23,24]. Expression (1) allows us to estimate the influence of topological and electrical parameters on the delay of the inverter. The growth leads to a decrease in the steepness of the linear section of the dependence $\tau(U_{dd})$. An increase in the value of SS leads to an increase in the steepness and a reduction in the linear section.

According to the established methodology, the assessment of the operability of logic circuits is formed based on the analysis of the propagation of a logic signal in a chain of a large odd number of identical inverters [5,15]. With the help of TCAD, the propagation of a logical signal through 11 similar inverters, the model of which is discussed above, is modeled. Fig. 10 shows the dependence $\tau(n)$ obtained from the calculation results, n is the number of the inverter.

Note the linear increase in the delay from the number of valves. The same nature of the dependence is preserved for control voltages of 0.52 V. The active power consumed by this circuit is 2.34 μW at $U = 0.6 V$ and 1.9 μW at 0.52 V. Static power are 48.5 and 42 pW,

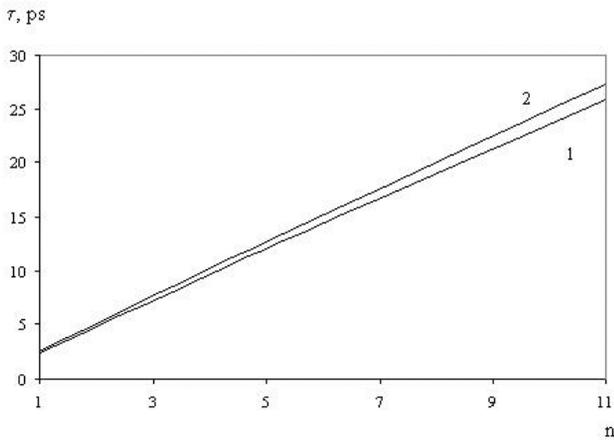


Fig. 10. Dependence $\tau(n)$ at different U_{dd} and clock frequency of 25 GHz, where 1 – $U_{dd} = 0.6$ V, 2 – $U_{dd} = 0.52$ V.

respectively. The combination of these factors determines the stable performance of deco nanoscale logic gates [25,26].

The totality of the results obtained allows us to consider the analyzed transistor architecture for creating low-power nanoscale electronic devices [4,5]. Thus, a conical architecture with an optimized diameter ratio can become a full member of the GAA (gate-all-around) family, which is a priority for the modern technological development of nanoelectronics.

6. CONCLUSION

The paper analyzes one of the possible approaches to the development of low-power electronics based on low-voltage logic gates made on SOI CMOS nanotransistors with cylindrical geometry, a surrounding gate and a variable diameter of the working area in the form of a truncated cone.

TCAD models of n- and p-types of conductivity of the nanotransistors under consideration have been developed. According to the results of mathematical modeling, it was revealed that conical prototypes demonstrate improved electrostatic characteristics compared to cylindrical ones. The electrophysical characteristics of the conical structure in

the range of control voltages from 0 to 0.6 V are characterized by a higher transistor current, a maximum I_{on}/I_{off} current ratio, a low leakage current and a slope of the subthreshold characteristic close to the theoretical chapel. Prototypes of transistors with optimal parameters with a working area length of 25 nm and a working area diameter ratio of 8.1/10 nm were selected for the synthesis of logic gates with a low supply voltage of 0.6 V.

The dynamic characteristics of the developed physical models of the inverter and the inverter chain are numerically investigated. At control voltages of 0.6 V and a frequency of 25 GHz, the inverter model predicts a maximum switching delay of 1.5 ps, the limit level of active power of 0.21 μ W, static 4.4 pW. The active power consumed by the circuit consisting of 11 identical inverters connected in series is 2.34 μ W at control voltages of 0.6 V.

REFERENCES

1. Kaushik BK (ed.). *Nanoelectronics: Devices, Circuits and Systems*. Elsevier Inc., 2018, 476 p. DOI: 10.1016/C2016-0-04936-7.
2. Sahay S, Kumar MJ. *Junctionless field-effect transistors: Design, Modeling, and Simulation*. Hoboken, New Jersey, Wiley-IEEE Press, 2019, 496 p.
3. Rahman F, Shakya B, Xu X, Forte D, Tehranipoor M. Security beyond CMOS: fundamentals, applications, and roadmap. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, 2017, 25(12):3420-3433.
4. International Technology Roadmap for Semiconductors (ITRS) Interconnect, 2020 Edition. [Online] Available: <https://irds.ieee.org/editions/2020>, access data 15.01.2021.
5. Tomar G, Barwari A. *Fundamental of electronic devices and circuits*. Singapore, Springer, 2019,

- 246 p.
6. Chiang T-K. A new quasi-3-D compact threshold voltage model for Pi-gate MOSFETs with the interface trapped charges. *IEEE Trans. on Nanotechnology*, 2015, 14(3):555-560.
 7. Gao H-W, Wang Y-H, Chiang T-K. A quasi-3-D scaling length model for trapezoidal FinFET and Its application to subthreshold behavior analysis. *IEEE Trans. on Nanotechnology*, 2017, 16(2):281-289.
 8. Masal'skii NV. Modeling the CMOS characteristics of a completely depleted surrounding-gate nanotransistor and an unevenly doped working region. *Rus. Microelectron.*, 2019, 48(6):394-398.
 9. Masal'skii NV. Simulation of the characteristics of low-voltage gates on combined cylindrical surrounding gate field-effect nanotransistors. *RENSIT: Radioelektronika. Nanosistemy. Informacionnyye Tehnologii*, 2021, 13(4):449-456. DOI: 10.17725/rensit.2021.13.449.
 10. Colinge J.-P. *FinFETs and Other Multi-Gate Transistor*. New York, Springer-Verlag, 2008, 340 p.
 11. Ferain I, Colinge CA, Colinge J. Multigate transistors as the future of classical metal-oxide-semiconductor field-effect transistors. *Nature*, 2011, 479:310-316.
 12. Lundstrom M, Guo J. *Nanoscale Transistors: Device Physics, Modeling and Simulation*. New York, Springer, 2006, 217 p.
 13. *TCAD Sentaurus Device User Guide* [Online]. Mountain View, CA 94043, Synopsys, Inc., 1296 p. Available: <https://www.synopsys.com/silicon/tcad/device-simulation/sentaurus-device.html>, access data 15.03.2021.
 14. Savio A, Monfray S, Charbuillet C, Skotnicki T. On the limitations of silicon for I-MOS integration. *IEEE Trans. Electron Devices*, 2009, 56(5):1110-1117.
 15. Neamen D. *Semiconductor physics & devices: basic principles*. New York, McGraw-Hill, 2011, 784 p.
 16. Samoju VR, Mahapatra K, Tiwari PK. Analytical modeling of subthreshold characteristics by considering quantum confinement effects in ultrathin dual-metal quadruple gate (DMQG) MOSFETs. *Superlattices Microstruct.*, 2017, 111:704-713.
 17. Shin YH, Bae M, Park C, Park J, Park H, Lee Y, Yun I. Universal core model for multiple-gate field-effect transistors with short channel and quantum mechanical effects. *Semiconductor Science and Technology*, 2018, 33(6):065010 DOI: 10.1088/1361-6641/AAC063.
 18. Yoon JS, Jeong EY, Baek CK, Kim YR, Hong JH, Lee JS, Baek RH, Jeong YH. Junction design strategy for Si bulk FinFETs for system-on-chip applications down to the 7-nm node. *IEEE Trans. Electron Dev. Lett.*, 2015, 36(10):994-996.
 19. Suh C. Two-dimensional analytical model for deriving the threshold voltage of a short channel fully depleted cylindrical/surrounding gate MOSFET. *J. Semiconductor Technology and Science*, 2011, 11(2):111-120.
 20. Karthigai Pandian M, Balamurugan NB. Analytical threshold voltage modeling of surrounding gate silicon nanowire transistors with different geometries. *J Electric Eng Technol.*, 2014, 9(6):742-751.
 21. Schwierz F, Wong H, Liou JJ. *Nanometer CMOS*. Singapore, Pan Stanford Publ., 2010, 338 p.
 22. Jaeger RC, Blalock TN. *Microelectronic circuit design*. New York, McGraw-Hill, 2011, 1190 p.

23. Lu W. Nanowire transistor performance limits and applications. *IEEE Trans. on Electron Dev.*, 2008, 55(11):2859-2876.
24. Sharma D, Vishvakarma SK. Precise analytical model for short channel cylindrical gate (CylG) gate-all-around (GAA) MOSFET. *Solid. State. Electron.*, 2013, 86(1):68-74.
25. Myunghwan R, Bien F, Kim Y. Optimal inverter logic gate using 10-nm doublegate-all-around (DGAA) transistor with asymmetric channel width. *AIP Advances*, 2016, 6:015311.
26. Majzoub S, Taouil M, Hamdioui S. System-level sub-20 nm planar and FinFET CMOS delay modelling for supply and threshold voltage scaling under process variation. *J. of Low Power Electron.*, 2019, 15(1):1-10.

DOI: 10.17725/rensit.2022.14.243

Heating dynamics of metal-dielectric structures with nanometer-thin conductive films under the influence of microwave fields

Vladimir V. Starostenko, Evgeniy V. Grigoriev, Alim S. Mazinov, Sergey P. Arsenichev, Evgeniy P. Taran, Aleksandr V. Starosek, Ibraim Sh. Fitaev

V.I. Vernadsky Crimean Federal University, <https://cfuv.ru/>

Simferopol 295007, Russian Federation

E-mail: starostenkovv@cfuv.ru, grigoriev@mail.ru, mazinovav@cfuv.ru, arsenichevsp@cfuv.ru, taran.ep@cfuv.ru, starosekav@cfuv.ru, fitaev.i@cfuv.ru

Received August 19, 2022, reviewed August 26, 2022, accepted August 31, 2022

Abstract: The paper presents the results of experimental studies of the heating dynamics of metal-dielectric structures (MDS) with aluminum conductive nanoscale films deposited by magnetron sputtering on low-cost substrates (glass, sitall, PET and PTFE) under exposure to monochromatic ultrahigh frequency (UHF) fields in a waveguide. At film thicknesses less than 2 nm, the experimental samples had mostly dielectric properties. Noticeable interaction of the films with microwave, causing thermal phenomena, was observed at thicknesses greater than 2 nm and reached a maximum at 5 nm. Thermal degradation processes of the metallized layer manifested in different ways in MDS with solid and polymer (flexible) substrates. Breakdown in conductive films generally occurred at temperatures below the melting temperature of the film material and occurred perpendicular to the electric field strength vector.

Keywords: metal-dielectric structure, nanometer conductive films, microwave, waveguide, thermal effect, breakdown

PACS 78.70.Gq

Acknowledgments: The study was supported by a grant from the Russian Science Foundation (Project no. 22-22-20126).

For citation: Vladimir V. Starostenko, Evgeniy V. Grigoriev, Alim S. Mazinov, Sergey P. Arsenichev, Evgeniy P. Taran, Aleksandr V. Starosek, Ibraim Sh. Fitaev. Heating dynamics of metal-dielectric structures with nanometer-thin conductive films under the influence of microwave fields. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2022, 14(3):243-248. DOI: 10.17725/rensit.2022.14.243.

CONTENTS

1. INTRODUCTION (243)
 2. EXPERIMENTAL METHODS (244)
 3. ELECTRODYNAMIC AND THERMAL EXPERIMENTAL STUDIES (245)
 4. CONCLUSION (247)
- REFERENCES (247)

1. INTRODUCTION

The intensive development of the ultrahigh frequency (UHF) range and the widespread adoption of nanotechnologies in microelectronics made the investigation of the resilience of electronic components to electromagnetic fields (EMF) particularly relevant. Numerous experimental studies conducted by different groups of scientists in

various countries [1,2] show that out of the range of such film microstructural components as active devices, dielectric and conductive elements that form the basis of integrated circuits, sensors, microcircuits, electronic modules, etc., the most vulnerable ones are conductive films, primarily the contact sites [1,2].

Studies of the properties of nanometer conducting films are mainly devoted to the research of their electrodynamic (diffraction) properties. In this paper we present the results of the dynamics of electrothermal processes in nanometer conductive films considered in previous works [3–9].

The study of physical phenomena in conductive structures of microelectronic products exposed to EMF is extremely difficult, so studies were conducted on macroobjects with metal-dielectric structures (MDS) of finite dimensions in the waveguide pathway. Conversion of microwave energy into heat energy during exposure of MDS is determined by the thickness and material of the conductive film, the generator power and the amount of absorbed energy (exposure time). All studies were conducted with a natural heat transfer, i.e., without forced dissipation of heat. The work uses the data of linear diffraction on MDS in the waveguide (unchanged values of the standing wave ratio VSWR and attenuation T), theoretical and experimental data on the dynamics of thermal processes in MDS. In experimental studies the effects of film thickness, generator power and exposure time (absorbed energy) on both the heating dynamics and the breakdown (burn-through) of MDS films were determined.

The current trend in electronics is to reduce the geometric dimensions of microstructural elements, which is related to the thicknesses of the films used in them. In conductive nanoscale films, depending on the thickness the

transition from dielectric (no film) to mirror (reflective) structure takes place. In this regard, this work focuses on physical phenomena in conductive films with a thickness of 1-10 nm exposed to high-power microwave radiation.

The aim of this paper is to experimentally investigate the processes of conversion of microwave field energy into thermal energy during exposure of metal-dielectric structures with nanometer-thick films.

2. EXPERIMENTAL METHODS

The solid dielectric substrates used were 18×18×0.15 mm cover glasses and sitalloy substrates. Before sputtering, the substrates were plasma polished and had a roughness of no more than 10 nm. In addition, as film substrates, we used polymeric PET (lavsan) and PTFE (fluoroplastic) sheets, the surfaces of which are quite different from those of glass and sitalbased substrates [10,11].

The diffraction characteristics were investigated in the waveguide on a panoramic VSWR and attenuation T meter (R2-56, $f=2.9-4.1$ GHz). In the microwave range, the ratios between the incident, reflected, transmitted, and absorbed waves do not depend on frequency, since the structural inhomogeneities are much smaller than the wavelength; they also do not change with the scale conservation of the cross-sectional dimension ratios of the MDS and waveguide [10,11].

Schematic diagram of the installation for direct exposure of MDS to monochromatic EMF is shown in **Fig. 1**.

The M105-1 magnetron operating at a frequency of 2450 MHz with an output power of up to 700 W was used as a microwave radiation generator. The waveguide path had dimensions of 34×72 mm. In all waveguide studies, the condition $S_{\text{MDS}} \ll S_{\text{wg}}$ was met, where S_{MDS} and S_{wg} are the cross sections of MDS and waveguide. It should be noted that

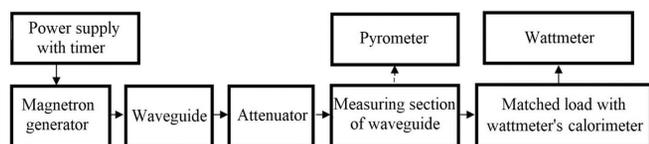


Fig. 1. The structural diagram of the installation for the study of electro-thermal processes in MDS.

the VSWR and attenuation T values are the same whether the MDS is located normal or parallel to the vector of the electric component of the H_{10} wave field. We chose an orientation of the MDS relative to the waveguide axis, that was convenient for conducting measurements.

During electrothermal exposure, the studied sample was placed in the measuring section of the waveguide path, and after dosed exposure to microwave radiation (multiple of 1 s), the temperature of the sample was measured with a MEGEON 16350 pyrometer through a special aperture in the waveguide. Then, the sample was allowed to cool down to ambient temperature and the process was repeated with a different generator power. Since significant errors are possible in temperature measurements, we conducted five cycles of experiments with subsequent averaging of the obtained values.

3. ELECTRODYNAMIC AND THERMAL EXPERIMENTAL STUDIES

In panoramic measurements of diffraction characteristics on objects in a waveguide, the VSWR characterizes the reflected wave by field, and the transmittance (attenuation) by power. Most often relative powers of reflected (R), transmitted (T) and absorbed (A) waves are used as diffraction characteristics – optical coefficients. Condition $R + T + A = 1$ follows from power balance, where R , T and A are normalized to input power diffraction characteristics; R is calculated through VSW, T is power transmittance (attenuation), absorption A is from power balance [10,11].

The most interesting is the behavior of R , T , and A in the thickness range from 1 to 10 nm. At such thicknesses the maximum possible conversion of EMF energy to thermal energy occurs. When using budget substrates and film deposition methods, the notion of thickness up to 7...10 nm is rather conventional. This is due to the roughness of the surface, on the one hand, and spatial and ohmic inhomogeneity of the film, on the other. At such thicknesses the film is inhomogeneous and has an island structure (a set of scattering centers, nanoparticles, etc. [12]). Usually, for such film thicknesses, the sputtering time correlation is used.

Temperature dependences of MDS on the thickness of deposited aluminum film at the same exposure time $t = 1$ s for microwave radiation power of 15, 30 and 60 W are shown in Fig. 2.

It follows from Fig. 2 that up to film thicknesses $d = 2$ nm MDS is more of a dielectric. At higher values of film thickness ohmic losses begin to appear. More detailed dynamics of EMF energy conversion to heat at $1.5 < d < 3$ nm are presented in the inset of Fig. 2. Up to thicknesses of $d = 2$ nm the film temperature is equal to ambient temperature, at $d > 2$ nm the film temperature increases due to ohmic losses, reaching maximum values at $d \approx 5$ nm. At $d > 5$ nm the film is practically formed. Reflection rather than absorption is predominant in this case. In general, at $d > 10$ nm a stationary relation between reflected

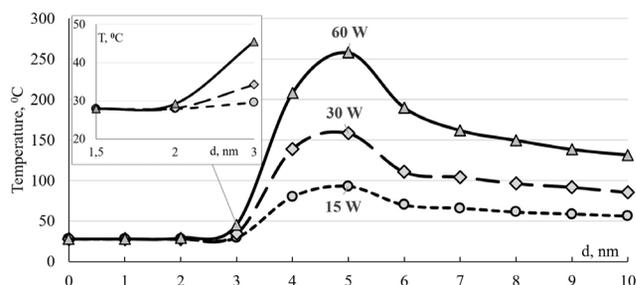


Fig. 2. Temperature dependences of MDS on thickness of Al film at $t = 1$ s at generator power of 15, 30 and 60 W.

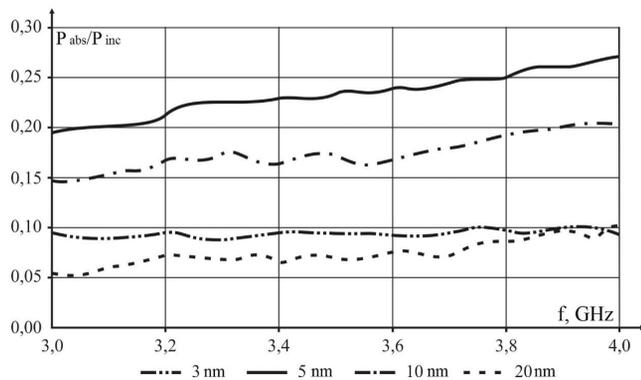


Fig. 3. Dynamics of the P_{abs}/P_{inc} frequency dependence in working range of R2-56 for MDS with aluminum films of different thicknesses.

and passed waves is established, the absorbed power is determined by the incident power.

Fig. 3 shows the dependences P_{abs}/P_{ins} for MDS with films of different thicknesses in the frequency range of the panoramic meter P2-56. These dependencies show that the maximum conversion of EMF energy to thermal energy occurs at film thicknesses near 5 nm. Similar results are obtained for MDS with films with specific conductivity more than 10^7 Sm/m, in particular, with films of gold, aluminum, etc. [5,12]. The insignificant dependence of P_{abs}/P_{ins} on the frequency (Fig. 3) is explained by the change of the electrical dimensions of MDS in the frequency range of the panoramic gauge. Comparison of the dependencies shown in Fig. 2 and 3 indicates that, at the values of generator power indicated in Fig. 2, only about 30% of the incident power is converted into heat (Fig. 3).

With an increase in generator power, it is possible to heat the film to the melting temperature (**Fig. 4**). As follows from Fig. 4, noticeable conversion of EMF energy into heat begins at film thicknesses of 3 nm or more. In an MDS with a film thickness of 3 nm, the temperature rose to 200°C in 3 s (Fig. 4). MDS with film thickness of 5 nm experienced breakdown in 1 s, similar situation takes place for MDS with film thickness of 7 and 10 nm.

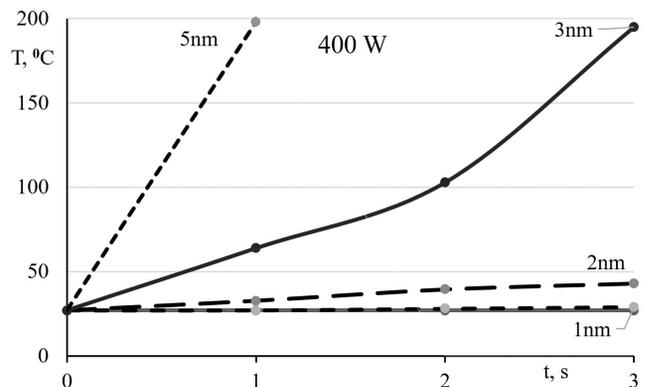


Fig. 4. Temperature dependences of MDS with films of 1, 2, 3 and 5 nm thickness exposed to 400 W EMF.

In MDS at catastrophic failure the film temperature may not reach the melting temperature – when exposed to EMF the film breakdown occurs earlier. This is due to the specifics of current and temperature distribution in the objects of finite size when exposed to EMF. Maximum values of the electric component of the field are produced on the edges parallel to the E -component of the EMF. The electrons' drift to the edges of the EMF is due to the Lorentz force, consequently, this leads to large current densities and thermal gradients on the edges of the EMF parallel to the E -component of the field. All this is the cause of the breakdown, which develops perpendicular to the E -component of the field.

Fig. 5 shows typical examples of breakdown in MDS on solid and polymer substrates.

Thermal changes in MDS with polymer substrates are observed as early as at film thicknesses of 2 nm: MDSs thus change their

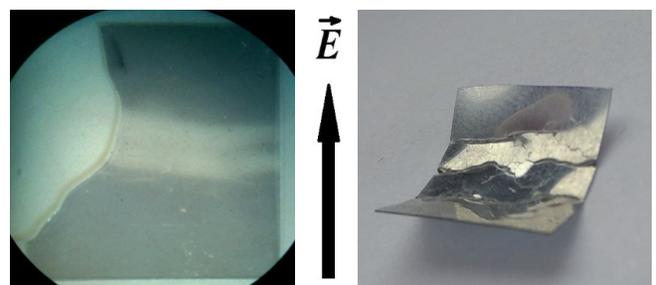


Fig. 5. Breakdowns and fractures in MDS on solid and polymer substrates, film thickness 10 nm, $P = 60$ W, $t = 2$ s.

geometry, the substrates begin to bend and twist.

4. CONCLUSION

The paper presents the results of studies of electrothermal processes in conducting films less than 10 nm thick when exposed to EMF in a waveguide. Noticeable temperature phenomena appear in conductive films at thicknesses greater than 2 nm, and thermal processes appear differently in MDS with solid and polymer (flexible) substrates. Breakdown in conductive films usually occurs at temperatures below the melting temperature of the film material and occurs perpendicular to the vector of the electric field component.

REFERENCES

1. Antinone RJ. *Electrical Overstress Protection for Electronic Devices*. New York, Noyes Publications, 1986, 492 p.
2. Chernyshev AA. *Osnovy nadezhnosti poluprovodnikovyykh priborov i integral'nykh mikroskhem* [Fundamentals of semiconductor and integrated circuit reliability]. Moscow, Radio i Svyaz Publ., 1988, 256 c.
3. Bosman H, Lau YY, Gilgenbach RM. Microwave absorption on a thin film. *Appl. Phys. Lett.*, 2003, 82(9):1353-1355.
4. Nimitz G, Panten U. Broad band electromagnetic wave absorbers designed with nano-metal films. *Ann. Phys.*, 2010, 19(1-2):53-59.
5. Li S, Anwar S, Lu W, Hang ZH, Hou B, Shen M, Wang C-H. Microwave absorptions of ultrathin conductive films and designs of frequency-independent ultrathin absorbers. *AIP Adv.*, 2014, 4(1):017130.
6. Pronin MS, Vdovin VA, Andreev VG. Issledovanie opticheskikh koeffitsientov nanometrovykh plenok medi i zolota v SVCh diapazone [A study of the optical coefficients of nanometric copper and gold films in the microwave range]. *Memoirs of the Faculty of Physics, Lomonosov Moscow State University*, 2016, 5:165411 (in Russ.).
7. Starostenko VV, Mazinov AS, Fitaev ISh, Taran EV, Orlenson VB. Dinamika formirovaniya poverkhnosti provodyashchikh plenok alyuminiya na amorfnyykh podlozhkakh [Dynamics of surface formation of conductive aluminium films on amorphous substrates]. *Prikladnaya fizika*, 2019, 4:60-65 (in Russ.).
8. Zuev SA, Starostenko VV, Taran EP, Shcherbakov SV, Arsenichev SP, Grigoriev EV, Fitaev ISh. Microwave Range Diffraction Properties of Structures with Nanometer Conductive Films on Amorphous Dielectric Substrates. *26th Telecommunications Forum (TELFOR)*, pp. 1-4, Belgrade, 2018. DOI:10.1109/TELFOR.2018.8611867.
9. Zuev SA, Zuev AS, Starostenko VV, Grigoriev EV, Mazinov AS, Taran EP, Fitaev ISh, Orlenson VB. Breakdown features in functional devices of telecommunication systems. *27th Telecommunications Forum (TELFOR)*, pp. 1-3, Belgrade, 2019. DOI:10.1109/TELFOR48224.2019.8971293.
10. Arsenichev SP, Grigor'ev EV, Zuev SA, Starostenko VV, Taran EP, Fitaev ISh. Difraktsiya elektromagnitnogo izlucheniya na tonkikh provodyashchikh plenkach metallodielektricheskikh struktur v pryamougol'nom volnovode [Diffraction of electromagnetic radiation on thin conductive films of metal-dielectric structures in a rectangular waveguide]. *Elektromagnitnye volny i elektronnyye sistemy*, 2017, 22(2) (in Russ.).

11. Mazinov AS Physical and electrodynamic properties of nanoscale conductive films on polymer substrates. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2020, 12(2):247-252. DOI: 10.17725/2020.12.247.
12. Maier SA. *Plasmonics: Fundamentals and Applications*. Moscow-Izshevsk, NIC "Regulyarnaya i khaoticheskaya dinamika" Publ., 2011, 296 c.

DOI: 10.17725/rensit.2022.14.249

Strange traces of a "strange" radiation

Anatoly I. Nikitin, Vadim A. Nikitin, Alexander M. Velichko, Tamara F. Nikitina

VL Talrose Institute for Energy Problems of Chemical Physics at NN Semenov Federal Research Center for Chemical Physics, Russian Academy of Sciences, <https://www.chph.ras.ru/>
Moscow 119334, Russian Federation

E-mail: anikitin@chph.ras.ru, vadim333@mail.ru, avelichko@chph.ras.ru, tnikitina1938@gmail.com

Received May 26, 2022, peer-reviewed June 3, 2022, accepted June 10, 2022

Abstract: A review of the main works on the study of "strange" radiation is presented – an intense glow that occurs above the explosion chamber during an electric discharge of metal foils in water in it. Photographs of the most diverse structures of tracks-traces of "strange" radiation are presented. Existing hypotheses of particles that leave such traces are outlined: Lochak's magnetic monopole hypothesis, the magneto-toro-electron radiation hypothesis, the tachyon hypothesis, the hypothesis of an intermediate quasi-molecular state of paired electrons, and the "dark" hydrogen hypothesis. Numerical estimates of all discussed hypotheses are given with an explanation of the mechanism of trace formation. The hypothesis of miniature ball lightnings – multiply charged clusters – spherical clusters with a radius of $2.14 \cdot 10^{-6}$ m, having a charge of $4.5 \cdot 10^{-12}$ C is presented. The electric field strength on the surface of such clusters can reach up to 10^{10} V/m, which is comparable to the electric field strength in an atom. When such a cluster is introduced into the crystal lattice, a strong polarization of the substance is possible, which can facilitate the conditions for the approach of protons and nuclei of elements. This forces us to consider the possibility of nuclear reactions proceeding through the process of overcoming the potential barrier, which lasts for a time that is many orders of magnitude longer than the nuclear collision time in a conventional nuclear process.

Keywords: type of traces of "strange" radiation, hypotheses of the structure of "strange" particles, model of a multiply charged cluster, explanation of the mechanism of traces formation

PACS: 36.40.Wa; 52.80.Mg

Acknowledgment: This work was carried out within the framework of Basic Scientific Research Program of the Russian Federation and it was funded from the budget. We thank Yuri Evdokimov, Vladislav Zhigalov, Keith Fredericks, and Leonid Urutskoev for permission to use their images of "strange" radiation traces.

For citation: Anatoly I. Nikitin, Vadim A. Nikitin, Alexander M. Velichko, Tamara F. Nikitina. Strange traces of a "strange" radiation. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2022, 14(3):249-268. DOI: 10.17725/rensit.2022.14.249.

CONTENTS

1. INTRODUCTION (250)	4.1. FORCES ACTING IN CHARGED CLUSTERS (258)
2. TYPE OF TRACES OF "STRANGE" RADIATION (250)	4.2. THE PROCESS OF TRACES FORMATION ON PHOTOGRAPHIC FILMS (260)
3. MODELS OF THE ARRANGEMENT OF "STRANGE" PARTICLES (252)	4.3. ACTION ON A CLUSTER OF A MAGNETIC FIELD (261)
3.1. MAGNETIC MONOPOLE (252)	4.4. MAGNETIC MOMENTUM OF A ROTATING CLUSTER (262)
3.2. MAGNETO-TORO-ELECTRIC RADIATION (254)	4.5. CURLS OF THE TRAJECTORY (263)
3.3. TACHYONS (255)	4.6. SYMMETRIC TRACKS (263)
3.4. COMPACT PAIRS OF ELECTRONS (255)	4.7. "STRANGE" PARTICLES AND NUCLEAR TRANSMUTATION (264)
3.5. DARK HYDROGEN (257)	5. CONCLUSION (265)
4. MULTIPLE CHARGED CLUSTERS (258)	REFERENCES (266)

1. INTRODUCTION

More than a quarter of a century ago, in experiments on the study of the electric explosion of metal foils in water, initiated by wide practical applications, an intense glow was noted that occurs above the explosion chamber, which received the name "strange" radiation accompanying the studied electric discharge in further studies. The registration of this radiation with various photographic materials revealed the presence on photographic materials of many micro- and macroeffects - tracks, traces in the form of lines, and imprints of round objects tens of microns in size [1,2]. Interest in such traces increased after the experiments of Urutskoev L.I. et al., who found that the appearance of traces on the films turned out to be somehow related to a change in the isotopic composition of materials that were in contact with the particles of this "strange" radiation [3,4]. It was found that these particles, flying parallel to the film at a speed of 20–40 m/s, are able to create traces in the photographic emulsion, the formation of which requires an energy of 10^{10} eV. They deflected in a magnetic field. Their footprints looked like solid or dotted lines with complex relief. The most interesting was the discovery of "twin tracks", when a large number of absolutely identical tracks were observed on an area of about 1 cm^2 . Sometimes these traces had chiral symmetry. In subsequent experiments, it was found that these traces appear not only in film emulsions, but also on the surfaces of any objects - on glass, mica, plastic [5,6].

This paper presents a review of the main studies of these tracks and a hypothesis about the object that leaves these strange tracks.

2. TYPE OF TRACES OF "STRANGE" PARTICLES

The result of the work of many researchers of this phenomenon was the determination of a fairly complete set of typical tracks of "strange" particles. The search for an explanation of the reasons for the formation of these tracks can become the basis for clarifying the nature of these objects.

It was found that the average speed of the "strange" particles was 20–40 m/s. However, the energy of the particles, determined from the blackening area of the photographic film

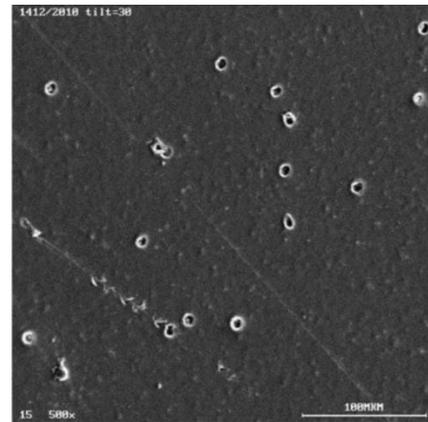


Fig. 1. Type of traces of "strange" particles [5,6].

under the assumption of Coulomb deceleration of charges, turned out to be greater than 700 MeV. **Fig. 1** shows the cavities left by such particles, and **Fig. 2** – section of such a cavity [5,6]. Let us assume that during the formation of the cavity, the detector material (polycarbonate) was heated to the melting point. The cavity volume is $V_c = 4.3 \cdot 10^{-20} \text{ m}^3$, and the mass of the melted material is $5 \cdot 10^{-17} \text{ kg}$. The energy required to heat the material, is $E = 10^{-11} \text{ J} = 75 \text{ MeV}$. The traces were located inside the photographic emulsion layer, which allowed us to conclude that the source of the blackening of the film flew parallel to the emulsion plane. Particles, moving along the surface of solid objects or photographic film, left continuous or intermittent traces on them (**Fig. 3**) [7]. Sometimes the solid line turned into a dotted line (or vice versa) (**Fig. 4**) [5,6]. These traces had a complex shape in the form

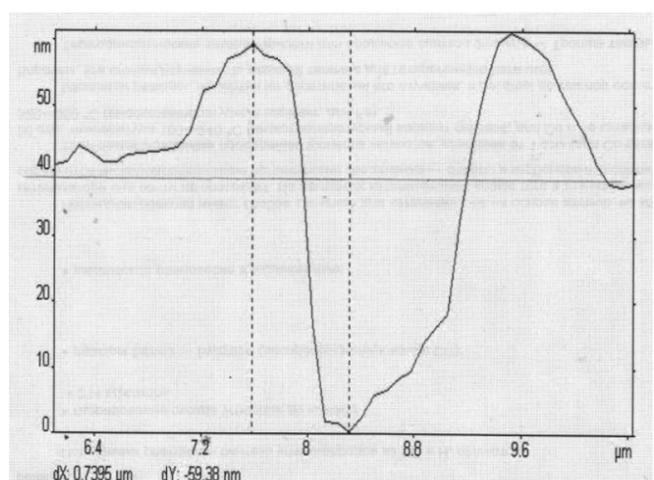


Fig. 2. Section of a cavity formed when a particle hits the surface of a polycarbonate disk [5,6].

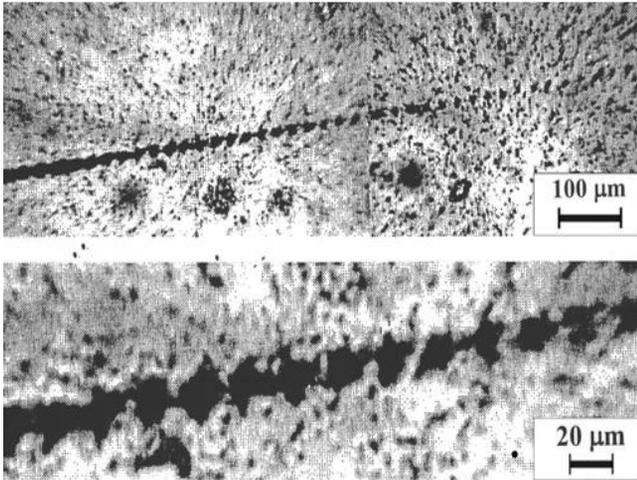


Fig. 3. Discontinuous trace of a "strange" particle on a photographic film [7].

of a periodically repeating pattern (**Fig. 5**) [8,9]. Both straight and curved trails were observed, and frequently the flying particle abruptly changed its direction of motion (**Fig. 6**) [8,9]. In a magnetic field directed perpendicular to the film surface, traces in the form of parabolas were observed on it (**Fig. 7**) [8,9]. Chaotic movement of particles that left traces was often observed (**Fig. 8**) [8,9]. Along with this, "twin tracks" were found when there were several identical copies of the track in an area of several square centimeters (**Fig. 9**) [5,6,8,9]. When their images were superimposed on each other, they almost completely coincided (**Fig. 10**) [8,9]. Moreover, the identity of the details of the traces was preserved when the image of



Fig. 4. Transition of a discontinuous trace to a continuous one [6].

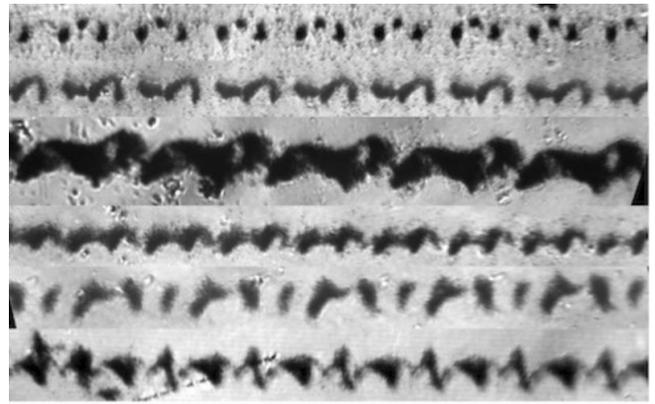


Fig. 5. Types of traces of "strange" particles [8,9].

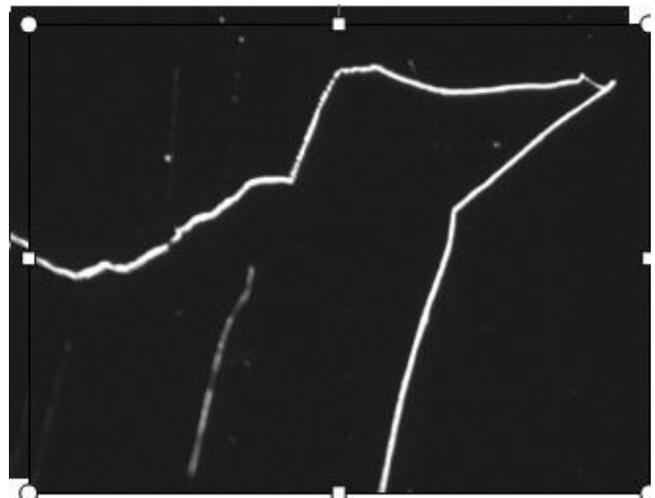


Fig. 6. Trajectories of "strange" particles [8,9].

the trace was magnified using a microscope (**Fig. 11**) [5,6]. "Chirally"-symmetric tracks were also found, traced by particles moving "in antiphase" (**Fig. 12**) [10,11].

Bogdanovich et al. studied results of action of current pulses of 15 kA with a voltage of 6 kV on a water jet. In these experiments,

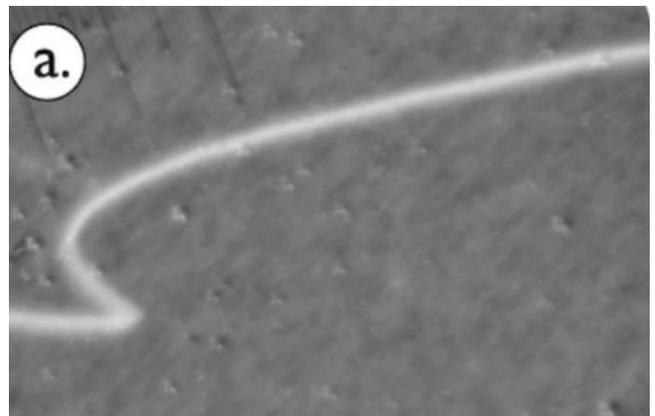


Fig. 7. Trail left by a particle moving in a magnetic field with an induction of 1 T [9].

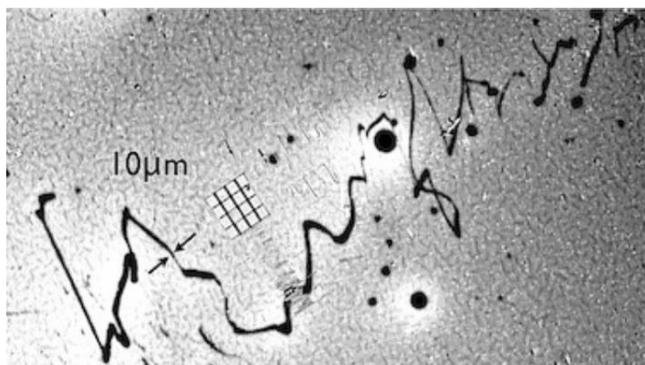


Fig. 8. Chaotic motion of particles [9].

particles were found that can penetrate into solid materials [12]. In this case, the titanium surface was hardened, and copper crumbled into a fine powder. Sections of copper plates were pierced with channels: sometimes straight, sometimes curved, sometimes helical. The copper plates after exposure to current retained the "activity" for a long time. Video filming in the dark showed that moving luminous spots appeared on the plate for tens of seconds [13].

Since the "strange" traces were discovered by nuclear scientists, most their models were based on the assumption that the "objects" are like elementary particles with high kinetic energy. The following were proposed: Lochak's magnetic monopole hypothesis [3,4,10,11,14-16], the Magneto-Toro-Electric radiation hypothesis [17,18], the tachyon hypothesis [8,9], the hypothesis of the Intermediate Quasimolecular State of paired



Fig. 9. Traces "twins" [8,9].

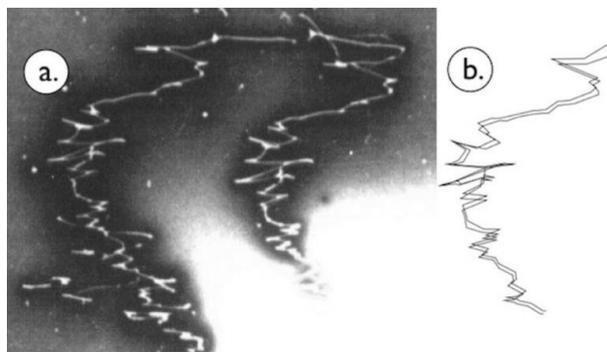


Fig. 10. Comparison of the type of traces of "twins": a) parallel tracks, b) the result of overlaying footprint images [9].

electrons [19], hypothesis of "dark" hydrogen [20]. We assumed that "strange" traces were left by miniature ball lightnings – "multiply charged clusters" [21-23]. In the following sections, we will present a description of the main ideas of the proposed hypotheses and discuss how they explain the appearance of the various traces of "strange" particles presented above.

3. MODELS OF THE ARRANGEMENT OF "STRANGE" PARTICLES

3.1. MAGNETIC MONOPOLE

Urutskoev L.I. et al. [3,4] suggested that these "objects" are particles with a magnetic charge – magnetic monopoles. Based on the fact that particle tracks in a magnetic field broadened and took the form of a "comet", they concluded that the particles do not have an electric charge, but they have a magnetic charge. Based on the

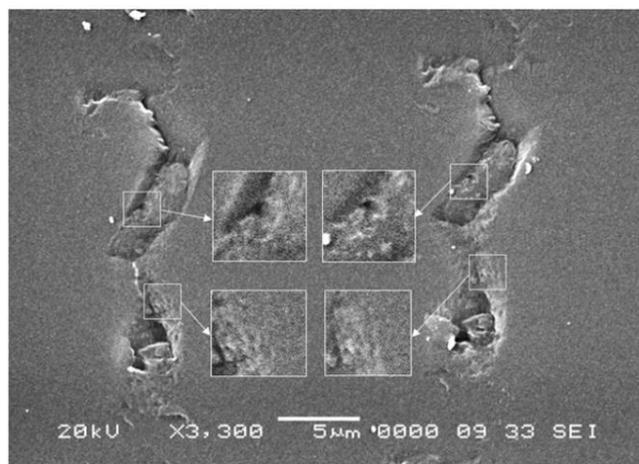


Fig. 11. Enlarged image of two parallel traces [5,6].

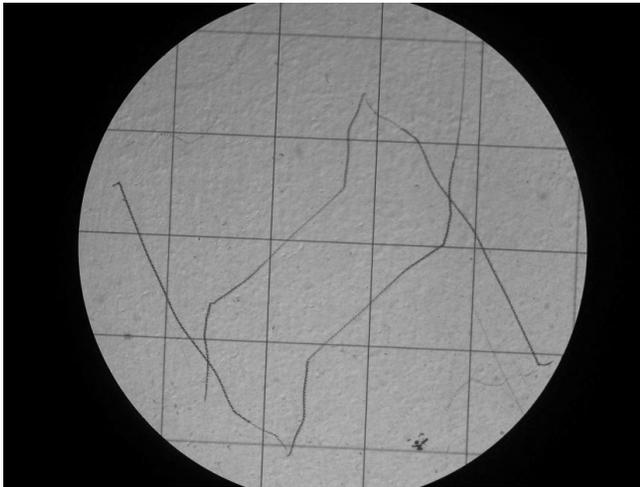


Fig. 12. "Chiral"-symmetric traces left by particles that simultaneously moved in the opposite direction. Square: $1\text{mm} \times 1\text{mm}$ [10,11].

results of observations, let's try to compile a list of properties of a magnetically charged particle that leaves traces in a photographic emulsion. This particle has no electric charge and interacts only with a magnetic field. Some of them move in the direction of the magnetic field vector, and some in the opposite direction. According to Ivoilov's observation [10,11], when a particle hits a photographic emulsion of a film, the direction of particle motion changes by 90° – it begins to move in the plane of the photographic emulsion under the action of a magnetic field whose vector is parallel to the plane of the photographic film. At this moment, the magnetic field perpendicular to the plane of the film ceases to act on the particle. Since the change in the direction of the particle's motion occurs almost instantly, it is assumed that the mass of the particle is zero or very small. The role of such particles is best suited to neutrino in an excited magnetic state, a particle theoretically predicted by Lochak [14]. Moving inside a photographic emulsion or other solid material, the particle periodically "emerges" and plunges into it again, melting the tunnel. The kinetic energy of the particle is clearly not enough to form such a tunnel. It does this by initiating exothermic nuclear reactions in the material. Fine details of track formation are not discussed in this case. The physical reason for the sharp change in the properties of

neutrinos under the action of a magnetic field is not discussed: the neutrino, which in the ground state freely passes through the thickness of the globe, after excitation begins to slow down in an emulsion layer several microns thick.

In Kyiv, in the Proton-21 laboratory, experimental evidence was obtained of the nuclear transformation of metal under the influence of coherent electron beams [15,16]. Tracks of magnetically charged particles in a multilayer MIS structure (metal-insulator-semiconductor) were studied. It was found that the particles behave like a needle in the shuttle of a sewing machine - they periodically pierce through the aluminum layer with a constant pitch ($60\ \mu\text{m}$), leaving a melted winding hollow tunnel about $1\ \mu\text{m}$ wide. An estimate of the energy release was made in [16]; it turned out to be about $10^6\ \text{GeV/cm}$. A hypothesis was proposed for the existence of a magnetically charged particle moving through a layer of a paramagnet. The authors drew attention to the fact that simple deceleration of particles is unable to generate such an amount of energy at a practically undecreasing particle velocity in the track. Therefore, they suggested that such particles are capable of initiating magnetic catalysis of energy-favorable nuclear reactions. This happens because magnetic monopoles, moving in matter, very strongly distort the electron shells of atoms which they come across on their way and thereby increase the probability of nuclear tunneling and their fusion. Getting into aluminum (paramagnet) as into a potential well, a magnetically charged particle stimulates nuclear reactions with the release of energy. Melting through the aluminum layer, the particle changes its magnetic properties (it becomes a diamagnet) and, as a result, it tends to "emerge" from this layer. After leaving the aluminum and passing some distance along the surface, the particle is again attracted by the potential well of the paramagnet, and the whole process is repeated. The authors suggest that an external magnetic field is essential for this

behavior of particles. An estimate of the speed of the monopoles showed that it is larger than 200 km/s: the particle must fly so fast that the magnetic field does not change during its flight. Indeed, the step on the track remains constant throughout the entire trace and, therefore, the entire track (2 mm long) must be created in a time significantly less than 30-50 ns (this is how long the current pulse lasts).

3.2. MAGNETO-TORO-ELECTRIC RADIATION

In 2010, Shishkin et al. [17,18] suggested that traces appear as a result of interaction with the film of MTER vortex solitons – Magneto-Toro-Electric radiation. Traces are obtained from the explosive unpacking of solitons, in each of which more than 10^{11} electrons and at least 10^5 ions are "packed". During the explosion of such a soliton, a significant part of the electrons acquires kinetic energy of 6-10 keV.

The authors of this hypothesis believe that "the physical vacuum is filled with the smallest material particles with a high penetrating power. This "hidden" matter consists of background "cold" neutrinos (BCN) (axions) coupled to each other through weak topological nontrivial bonds, which form a background neutrino (or axion) condensate. In the near field in the vicinity of the atoms, this neutrino condensate thickens, forming "field" shells of BCN, which, due to the interaction with the electrons of the atom, tending to "escape" from the nucleus, become extremely dense. The nucleus of an atom is a mini-neutron hole that absorbs very "compressed" vacuum background cold neutrinos (BCNs). Due to this mechanism of compaction, the shells made of BCN acquire a very high potential energy. As a result, a "concentrated" neutrino condensate is formed near the nucleus, which is in equilibrium with nuclear forces, and its residual potential, called the long-range Coulomb field of the nucleus, attracts electrons. But at a close distance from the nucleus, a repulsive "short-range" magnetic field begins to act on electrons, the power of which near the nucleus exceeds the Coulomb forces by an order of magnitude. Due

to repulsion, an electron with a large acceleration relative to the incident (electron) flies away from the nucleus. As a result of interaction with the magnetic field of the nucleus, the parameters of the electron change in such a way that the shell, consisting of BCN, becomes impenetrable for it. As a result, the electron begins to condense the shell, pushing the BCN from the nucleus to its periphery. At the Bohr radius from the nucleus, the electron loses energy, transferring it through the compacted shell of an atom to the BCN, external to the shell. After the transfer of part of its energy through the shell by the external BCN, the electron is again attracted by the Coulomb forces to the nucleus. The process is repeated. Thus, the nucleus receives energy, which, in turn, the electron transfers through the shell to the "vacuum" BCN. The "shell" of an atom, built by its nucleus, is a soliton, that is, a closed vortex structure of BCN dipoles. Based on this model, the shell has a "torus-like" structure similar to a "bun" with an extremely small central region. In this case, the nucleus of the atom hangs, as it were, on the "absorbing" and "radiating" vortices. When the shell is hit (mechanically, by a strong electric field, thermal or light radiation), with a certain probability, one of the "vortices" can break away from the core for a short time, as a result of which the core "falls out" of the shell. The shell also leaves the electrons that have lost their energy supply from the nucleus. An empty "field shell", being closed on itself, is a soliton. It has a high penetrating power and carries the characteristics of a "mother" nucleus. The shell flying away from the nucleus unpacks into a string-vortex soliton (SVS). The diameter of the head part of the string is 0.98 nm, the length of the SVS can reach 700 m, and a gimlet no longer than $0.1 \mu\text{m}$ and with a diameter of $d = 0.078420533 \cdot A$ "runs" along the SVS string, where A is the atomic weight of the "mother" nucleus. When interacting with a material object, the gimlet drills a cavity on the surface of the object. When colliding with the shells of atoms, the string-vortex soliton loses

energy and eventually collapses into the original torus-like structure. A large number of "empty" shells behave like "neutrino" – cluster radiation, called magneto-toro-electric radiation (MTER). A separate shell is called a magneto-toro-electric cluster (MTEC)."

The hypothesis was stated that "strange" traces "appear during the destruction of MTER clusters that release energy up to 10^{17} eV (according to the analysis of traces of strange radiation on X-ray detectors). Unlike the massless quantum of ordinary electromagnetic radiation (photon), the new formation (in fact, an empty "field" shell of an atom) - voluminous, but not heavy - carries a huge amount of energy. It occupies an intermediate place between corpuscular radiation and gamma radiation. The collapse of the cluster into separate vortices is possible – this is recorded on photodetectors in the form of black drops. On the basis of experiments, it was assumed that the diameter (d) of the microcrater is proportional to the atomic mass of the atom (A) that generated the corresponding MTER cluster, $d = k \cdot A$, where $k = 0.078$ (d is in μm). Starting from a certain atomic number, the MTER cluster is not able to "gnaw out" a large crater, but is able to ionize a zone corresponding to the cluster diameter. Hence the conclusion follows: the cluster, passing through the material, acquires the properties of this material."

3.3. TACHYONS

Fredericks [8,9] used sensitized photographic film to register "strange" marks that appear when fingers are applied to it. The appearance of these traces was the same as that of traces obtained by other methods. Solid lines, traces in the form of a dotted line of repeating figures, paired traces, traces with sharp breaks were observed. In the field of a permanent magnet applied to the film, traces in the form of parabolas were found (Fig. 7). Processing such traces, and assuming that they were left by particles with a magnetic charge, Fredericks found the average value of the particle moment $p = 20 \text{ eV}/c$ (c is the speed

of light). The kinetic energy of the particles was determined by taking into account the energy losses due to the formation of silver grains in the photographic emulsion. It was believed that for the formation of one grain it is necessary to expend energy of 750 eV. The kinetic energy of the particles turned out to be 10^8 - 10^9 eV. The resulting figures were plotted on a plot of kinetic energy versus momentum. For particles with a momentum of $20 \text{ eV}/c$, which speed is less than or equal to the speed of light, the kinetic energy cannot be greater than 1 eV. The found values of kinetic energy exceeded this figure by several orders of magnitude. Such a ratio of energy and momentum values can only belong to tachyons - particles moving faster than light. Tachyons have an imaginary mass. For them, unlike ordinary particles, the kinetic energy increases with decreasing momentum.

3.4. COMPACT PAIRS OF ELECTRONS

Kashchenko [19] believes that the repulsive force of the nuclei of chemical elements can be reduced by placing a negative charge in the space between them. "Such a charge can be created by two electrons located at a distance of $1 \text{ fm} = 10^{-15} \text{ m}$ from each other ((ee) -pair). Such electrons are bound by a force that exceeds the force of the Coulomb repulsion of charges. According to the uncertainty relation, the minimum kinetic energy of an electron pair should be about 10^3 MeV . This is three orders of magnitude higher than the rest energy of an electron pair. The presence of such energy determines the relativistic nature of the motion of electrons in a bound pair. The natural form of motion of a bound pair is rotation around a common center of mass. Let us choose the size of the pair d equal to the diameter of the circular orbit and take into account that on the scale $d \sim 1 \text{ fm}$ the stability of the pair is ensured by the contact interaction. An orbit with a small radius $r = d/2 = 0.5 \text{ fm}$ is possible at a speed close to the speed of light, $v \rightarrow c = 3 \cdot 10^8 \text{ m/s}$. In this case, the minimum value of rotational motion energy of the pair is $E_{\min}^* = 2mc^2$ (where m is the relativistic mass of the electron). The

value of m can be found based on the minimum value of the angular momentum of electrons $L = \hbar = h/2\pi = 1.054 \cdot 10^{-34} \text{ J}\cdot\text{s}$. Angular momentum $L = 2mvr = 2mcr = \hbar$. Hence $m = m_e [(1 - (v/c)^2)^{-1/2}] = \hbar/2cr$. Energy of relativistic rotation of a pair of electrons is $E_{\min}^* = \hbar c/r$. Substituting $L = \hbar$, $r = 0.5 \text{ fm}$ and $v = c$, we find $E_{\min}^* = 375 \text{ MeV}$." The author assumes the possibility of the existence of "IQS – an intermediate quasi-molecular state in the form of an annular electron current of radius R , flowing in the plane of the ring, orthogonal to a line connecting a pair of nuclei with equal charges. In the annular orbit, there are electrons in the form of bound pairs with opposite spins. The electrons are attracted to positively charged nuclei. The component of the force in the plane of the orbit is the centripetal force that holds the bound pairs in orbit. The nuclei are affected by the force of attraction in the direction of the electric field created by the electrons and the force of mutual repulsion. IQS model leads to the conclusion about the possibility of forming an ensemble of stable catalysts for low-temperature nuclear reactions under the conditions of electron flows."

"The synthesis of atoms with increasing charge and mass numbers corresponds to catalyzing ring orbits with an increasing number of (ee) -pairs. Ring orbits are capable of strongly polarizing the electron shells of atoms between which they are localized, reducing the shielding effect of atomic shells on the Coulomb field of nuclei. As a result, the radius of the catalyzed orbit rapidly decreases, which corresponds to an increase in the value of the effective negative charge at a point on the line connecting the centers of the nuclei. These (ee) -pairs remain stable if the energy of the quantum of action on them is small compared to the energy of the pair. This means that neither thermal vibrations of the lattice, nor photons, nor electron beams should destroy pairs."

According to the author of the IQS idea, the formation of traces on the surface of a detector (film, DVD, etc.) occurs as follows.

"There is an electrically neutral complex consisting of a catalytic ring of (ee) -pairs and an ionic component of a pair of approached nuclei (or already merged nuclei) with distorted and incompletely filled electron shells (called a KK -activator). Upon collision with the detector, the complex decomposes into a positively charged ionic component, which remains in the detector material, and a free KK -activator, which rapidly expands due to the repulsion of (ee) -pairs. The KK -activator contains rotating (ee) -pairs and has a mechanical moment L and a magnetic moment M . The propagation of KK -activators is responsible for the formation of tracks. 1) If the plane of the KK -activator is parallel to the plane of the detector, and the translational velocity of the center of inertia of the KK -activator is perpendicular to the plane of the detector, a trace appears in the form of a crater. 2) If the plane of the KK -activator is orthogonal to the plane of the detector, and the translational velocity of the center of inertia of the KK -activator is parallel to the plane of the detector, a wide straight line appears. A trace with discrete periodic point illumination of the detector surface is constructed on the basis of matching the translational and rotational motion of (ee) -pairs along helical trajectories. A track is a discrete set of intersection points of (ee) -pair trajectories with the detector surface. Smooth tracks are described by the movement of a ring with a plane sliding over the detector surface. The width of the track is equal to the diameter of the ring. The reason for the deviation of the shape of smooth lines from a straight line is due to ions in the detector material, which bend the trajectories of KK -activators."

The model proposed by the author for catalysis of the fusion of nuclei by rings of rotating pairs of electrons requires further analysis. It is unclear how the ring plane is localized exactly in the middle between the merging nuclei. In the absence of a mechanism that maintains the equality of distances between the ring and two nuclei, symmetry breaking and "attraction" of

the ring to one of the nuclei is inevitable. It is also unclear how a ring of (*ee*)-pairs which has collided with the detector and lost its "positively charged ionic component" (that is, the nuclei that the ring was "working" to bring together) can exist. But they were these nuclei that created the centripetal force which kept the (*ee*)-pairs in orbit. With the disappearance of this force, the electrons will fly away tangentially to the orbit.

3.5. DARK HYDROGEN

The idea of the existence of pairs of magnetically bounded electrons located at a distance of several femtometers (1 fm = 10^{-15} m) was used by Baranov and Zatelepin [20] in developing the concept of "dark hydrogen". Hydrogen is called "dark" because its spectrum lies in the X-ray region, and it cannot be detected by absorbing light in the optical range. According to this hypothesis, the existence of a compact atom with a nucleus of a pair of closely bounded electrons is possible. These electrons rotate in an orbit with a diameter of $2R = 62 \cdot 10^{-15}$ m. Around them a pair of protons rotates in an orbit with diameter of $2R_p = 65 \cdot 10^{-15}$ m. In "dark hydrogen", the electron attraction force is determined by several types of interaction. These are: *SS* – spin-spin interaction between intrinsic magnetic moments, *SO* – spin-orbit interaction between intrinsic magnetic moment and orbital moment of two electrons, *OO* – interaction of orbital magnetic moments of electrons, *K* – Coulomb repulsion of electrons. Interaction with protons can be neglected, since they are located in outer orbits, which dampens their electric fields, and the intrinsic magnetic moments of protons are three orders of magnitude smaller than those of electrons. The main contribution to the total *SS* + *SO* + *OO* + *K* – interaction comes from the *SO*-interaction. Accounting for *SS*-*OO*-*K* gives a correction of 50%. The speed of electrons is close to the speed of light *c*. The portable acceleration in the rotating system generates the centrifugal force $F_c = \gamma m_0 c^2 / R$. (Here $\gamma = (1 - v^2/c^2)^{-1/2}$, $m_0 = 9.1 \cdot 10^{-31}$ kg, *R* is radius of the orbital motion of electrons). Force $F_{SO} = 6\mu p^2 / 4\pi R^4$,

$\mu = 4\pi \cdot 10^{-7}$ H/m is the magnetic constant, $p = 0.927 \cdot 10^{-23}$ J/T is the Bohr magneton. Although the electron becomes γ times heavier, which reduces its magnetic moment, there is a γ times concentration of magnetic and electrical forces in the direction perpendicular to the electron's motion. These processes cancel each other out. We set $F_c = F_{SO}$. Electron orbit quantization rule: $\gamma m_0 c R = \hbar$. Here the unknowns are γ and *R*. Substituting constants, we find $R = 4.3 \cdot 10^{-14}$ m, $\gamma = 8.8$. Using these values, we get $F_c = F_{SO} = 16.8$ N and $\gamma m_0 = 8 \cdot 10^{-30}$ kg.

Dark hydrogen, when interacting with ordinary matter ("Bohr" atom), participates in three processes: 1) Oxidation of the "Bohr" atom (that is, the removal of an electron from it). 2) Creation of a molecule consisting of dark hydrogen and a "Bohr" atom. Creation of huge structures from millions of such joint molecules. 3) Transmutation of the nuclei of "Bohr" atoms. The first two processes proceed with the rearrangement of the electronic system of the "Bohr" atoms, that is, these are chemical processes. The third process is the change in the nucleus of the "Bohr" atom. This is not a chemical, but a nuclear process. Since there is no electric field outside the "atom" of dark hydrogen, a free electron or an ordinary atom can approach the relativistic electron pair at a distance of 10^{-13} m. Here the magnetic forces are greater than the Colombian, and the electron pair attracts a new electron. A new electron will "sit" in the orbit of an electron pair with a quantum number increased by one, which will increase the orbital magnetic moment of dark hydrogen. As a result, dark hydrogen will again be ready to accept the next electron, and its oxidizing properties increase, something like a chain reaction takes place. Track formation is a chemical process of oxidative destruction of the "Bohr" substance. After oxidation, the "Bohr" substance acquires a positive charge. Dark hydrogen and the ion of the "Bohr" substance create a molecule due to the Coulomb interaction. The process of formation of a molecule, like the oxidation

process, is exothermic with energy of several keV. The process of formation a molecule of the "Bohr" substance and dark hydrogen formation does not end there, since the formed molecule is able to continue oxidation due to magnetic forces. With each act of molecule enlargement, energy is released due to interaction with a new "Bohr" atom. During the oxidation reaction of dark hydrogen, not only the magnetic momentum increases, but also the mechanical momentum. Due to its small size, dark hydrogen has the ability to approach the nucleus of the "Bohr" atom at nuclear distances, this creates conditions for the tunneling of dark hydrogen protons into the "Bohr" nucleus (and vice versa, from the "Bohr" nucleus to the dark hydrogen proton), which changes the number of the nucleus. It is possible that dark hydrogen atoms can interact with each other. In this case a structure, resembling helium-4, is created.

To explain the appearance of mirror tracks, we can assume the possibility of the formation of associations of two dark hydrogen atoms located at a great distance from each other. In this case, something like a Cooper pair is obtained, but not from two electrons in the lattice, but from two dark hydrogen atoms on the surface of the detector.

4. MULTIPLE CHARGED CLUSTERS

We assumed that "strange" particles are multiple charged clusters, the structure of which is similar to the structure of ball lightning [21-23]. According to the electrodynamic model [24-33], ball lightning consists of a unipolar charged core and a dielectric shell (for example, water). In the electric field created by the charge of the core, the shell material is polarized and a force arises that make the shell to move towards the center of the sphere. The magnitude of this force is proportional to the first power of the charge of the nucleus $F_a \sim Q$. The force that stretches the shell due to the Coulomb repulsion of charges is $F_r \sim Q^2$. In the range of charge values from $Q = 0$ to Q_{\max} , the force F_a turns

out to be greater than F_r , and at $F_a = F_r$, the existence of a stable structure is possible. Due to the continuous charge leakage, the lifetime of such a structure is limited. When a multiple charged cluster approaches the conductor, the charge flows down towards the conductor. The mechanical impulse acquired by the carriers of the stacked charge when they move in the field of the main charge is transferred to the initial multiple charged cluster. This causes it to push off from the conductor. After some time, the multiple charged cluster will begin to approach the conductor, and the process of transferring part of the charge to it will be repeated. In the presence of an electric field whose vector is parallel to the plane of the conductor, a multiple charged cluster will move along the surface of the conductor (or photographic film) and leave a chain of discontinuous traces on it. If the velocity of the cluster along the surface is low, the points may merge into a solid line.

4.1. FORCES ACTING IN CHARGED CLUSTERS

Thus, it can be assumed that "strange" particles are multiple charged clusters, similar to microscopic ball lightnings, which are a certain amount of ions located inside a shell of water molecules. In the electric field of the charge located inside the shell, the dipole water molecules are oriented towards the center of the sphere, and a force arises that compresses the shell. At the same time, the molecules in the shell tend to push their "neighbors" out of it, which leads to a decrease in the compression force of the shell. The calculation showed that the force F_p of expulsion of water molecules from the shell of such a cluster is described by the formula [21,22]:

$$F_p = 2.69 \cdot 10^{-9} (R - R_0)^{-1.3} (H), \quad (1)$$

where R (inner radius of the sphere) is in angstroms (10^{-10} m), and $R_0 = -4.5$. Let us assume that there is one elementary charge inside the shell. Let us compare the force F_p with the force F_a of attraction of a water molecule (a dipole with moment $p_w = 6.327 \cdot 10^{-30}$ C·m) to the central

charge $q = 1.6 \cdot 10^{-19}$ C. $F_a = p_w \cdot \text{grad}E$, where E is the strength of the electric field created by the charge q :

$$F_a = p_w \text{grad}(q / 4\pi\epsilon_0 R^2) = -2p_w q / 4\pi\epsilon_0 R^3. \quad (2)$$

The force F_a is directed towards the center of the cluster and falls more steeply with increasing R than $F_p(R)$. At $R = 4 \cdot 10^{-10}$ m $F_a = 2.84 \cdot 10^{-10}$ N, which is 1.65 times greater than F_p , but already at $R = 6 \cdot 10^{-10}$ m $F_a = 0.843 \cdot 10^{-10}$ N, which is less than F_p ($1.27 \cdot 10^{-10}$ N).

Another reason that hinders the formation of an ordered cluster structure due to the action of the electric field of the ion is the thermal motion of molecules. According to the Langevin formula [34], the noticeable effect of the electric field on the dipole orientation with the momentum p_w stops at the strength $E_{\min} = 3k_B T / p_w$ (here $k_B = 1.38 \cdot 10^{-23}$ J/K is the Boltzmann constant, and T is the absolute temperature). At $p_w = 6.327 \cdot 10^{-30}$ C·m and $T = 300$ K, the value of $E_{\min} = 2 \cdot 10^9$ V/m. The electric field of the ion takes this value at a distance $R = 8 \cdot 10^{-10}$ m, that is, at the location of the second layer of water molecules in the cluster. Thus, we come to the conclusion that a cluster with $R = 4 \cdot 10^{-10}$ m can be formed in the electric field of a unit charge, however, further growth of the cluster, apparently, will occur without significant participation of the electric field of the central charge. This explains the conclusion obtained by calculation that when a water cluster grows in an electric field of an ion, the latter is always "pushed out" to the periphery of the cluster [35,36].

Let us discuss whether it is possible to keep a large number of ions inside the water cluster shell. Hope for the realization of this possibility is given by the fact that the force of attraction of a water molecule to the charge F_a can be increased due to an increase in the total charge of ions Q , and the magnitude of the force F_p does not depend on the charge (although the magnitude of the charge must be greater than a certain critical value determined by the requirement of complete polarization of molecules in the shell). To accommodate a large

number of ions, the size of the inner cavity of the cluster must be increased in comparison with the case of a singly charged cluster (the radius R must be increased). This should simultaneously reduce the magnitude of the force F_p .

Let us consider what a cluster with a total ion charge $Q_1 = 4.5 \cdot 10^{-12}$ C can look like. For example, $n = 2.8 \cdot 10^7$ hydroxyl (OH)⁻ ions have such a charge. Let the ion diameter be equal to $4 \cdot 10^{-10}$ m, the area occupied by one ion, $s = 16 \cdot 10^{-20}$ m², and the ions are placed on the surface of a sphere of radius r . The surface area of the sphere $S = 4\pi r^2 = s \cdot n$, hence $r = (s \cdot n / 4\pi)^{1/2} = 6 \cdot 10^{-7}$ m. The electric field created by the charge $Q = 4.5 \cdot 10^{-12}$ C at a distance $r = 6 \cdot 10^{-7}$ m, $E = Q / 4\pi\epsilon_0 r^2 = 10^{11}$ V/m, which is greater than $E_{\min} = 2 \cdot 10^9$ V/m determined by the Langevin criterion. This means that the water molecules in the shell are completely polarized. The ions stretch the shell with the force $F_Q = Q^2 / 8\pi\epsilon_0 r^2 = 2.53 \cdot 10^{-1}$ N. The water molecule is attracted to the charge placed in the center of the sphere with the force $F_a = p_w Q / 2\pi\epsilon_0 r^3 = 2.37 \cdot 10^{-12}$ N. At the same time, it is pushed out of the shell by neighboring molecules with the force $F_p = 2.69 \cdot 10^{-9} \cdot (6004.5)^{-1.3} = 3.3 \cdot 10^{-14}$ N. The resulting force acting on the molecule is $F_t = F_a - F_p = 2.337 \cdot 10^{-12}$ N. On the surface of the shell with radius $r = 6 \cdot 10^{-7}$ m can fit $n_w = 2.82 \cdot 10^7$ water molecules (we consider a molecule as a ball with a diameter of $4 \cdot 10^{-10}$ m). The shell compression force by one layer of water molecules $F_\Sigma = F_t \cdot n_w = 6.59 \cdot 10^{-5}$ N. This force is 3840 times less than the force F_Q . The force F_Q can be compensated if the number of layers of water molecules in the shell is 3840, and its thickness is $a = 4 \cdot 10^{-10} \times 3840 = 1.54 \cdot 10^{-6}$ m. As a result, we got a cluster with a cavity radius $r = 6 \cdot 10^{-7}$ m and outer radius $R = r + a = 2.14 \cdot 10^{-6}$ m. The mass of the cluster M_1 is equal to the mass of the shell. Its volume is $V_{sh} = 4\pi[(r + a)^3 - r^3] / 3 = 4 \cdot 10^{-17}$ m³, and its mass is $M_{sh} = \rho_w V_{sh} = 4 \cdot 10^{-14}$ kg (here $\rho_w = 10^3$ kg/m³ is the density of water). Carrying out similar calculations for a cluster with a cavity radius $r_2 = 6 \cdot 10^{-6}$ m, inside which there are ions with a total charge $Q_2 = 4.5 \cdot 10^{-11}$ C, we find its radius $R_2 = 3.75 \cdot 10^{-6}$ m and mass $M_2 = 8.62 \cdot 10^{-13}$ kg.

Fig. 2 shows a section of a "pit" that appeared as a result of the action of a "strange" particle on the surface of a polycarbonate disk [5,6]. Let us assume that this pit was formed due to the melting of the material and its ejection to the edges of the pit. The well depth is $h_p = 38$ nm, and the average radius is $r_p = 0.6$ μm . The volume of ejected material is $v_p = \pi r_p^2 h_p = 4.3 \cdot 10^{-20}$ m^3 , and its mass $m_{pc} = \rho_{pc} \cdot v_p = 5 \cdot 10^{-17}$ kg (here $\rho_{pc} = 1.2 \cdot 10^3$ kg/m^3 is density of polycarbonate). Heat capacity of polycarbonate $C_{pc} = 1.21 \cdot 10^3$ $\text{J}/\text{kg} \cdot \text{K}$. To heat $5 \cdot 10^{-17}$ kg of material to a softening temperature of 220°C , energy $E = 1.2 \cdot 10^{-11}$ $\text{J} = 75$ MeV is required. Let us assume that the same amount of energy is spent on the formation of a hole in the dotted track left by the "strange" particle on the surface of a polycarbonate disk. The number of these holes can reach up to 1000, therefore, the energy required for their formation can be 10^{-8} $\text{J} = 6 \cdot 10^{10}$ eV.

Let us assume that this energy is drawn from the energy of the electric field of a multiple charged cluster. Consider a cluster in the form of a sphere with a radius of $r = 6 \cdot 10^{-7}$ m, completely covered with ions with a total charge $Q = 4.5 \cdot 10^{-12}$ C. The expansion of ions is prevented by a spherical shell of water molecules. This cluster can be considered as a spherical capacitor with an inner electrode with a radius $r = 6 \cdot 10^{-7}$ m and an outer electrode with $R = \infty$. The electric capacitance of such a capacitor is $C = 4\pi\epsilon_0 r$, and the energy of the electric field is $W_e = Q^2/2C = 1.57 \cdot 10^{-7}$ $\text{J} = 10^{12}$ eV. This energy is quite sufficient for the formation of tracks on different materials.

4.2. THE PROCESS OF TRACES FORMATION ON PHOTOGRAPHIC FILM

Let us consider a possible scenario for the formation of traces on photographic films. Let us assume that a cluster with radius $R_1 = 2.14 \cdot 10^{-6}$ m, with charge $Q_1 = 4.5 \cdot 10^{-12}$ C and mass $M_1 = 4 \cdot 10^{-14}$ kg moves towards the film. When the edge of the cluster is at a distance $L = 15 \cdot 10^{-6}$ m from the film surface, a small cluster with charge $q_1 = 4.5 \cdot 10^{-14}$ C and mass $m_1 = 4 \cdot 10^{-16}$ kg separates from it and starts moving

towards the film. This cluster moves under the action of the force $F_{q_1} = Q_1 q_1 / 4\pi\epsilon_0 (R_1 + x)^2$, where x changes from 0 to L . Having passed the path L , the small cluster will acquire the energy $W_{q_1} = (Q_1 q_1 / 4\pi\epsilon_0) \cdot [1/R_1 - 1/(R_1 + L)] = 7.447 \cdot 10^{-10}$ $\text{J} = 4.65 \cdot 10^9$ eV. Its speed $v_1 = (2W_{q_1}/m_1)^{1/2} = 1.929 \cdot 10^3$ m/s, and moment of momentum is $m_1 \cdot v_1 = 7.7184 \cdot 10^{-13}$ $\text{kg} \cdot \text{m}/\text{s}$. A large cluster acquires the same momentum, it starts moving upwards from the film with a velocity $V_1 = (m_1 \cdot v_1)/M_1 = 19.296$ m/s and acquires an energy $W_{Q_1} = M_1 V_1^2 / 2 = 7.447 \cdot 10^{-12}$ J . Calculation for an ion with $R_2 = 7.5 \cdot 10^{-6}$ m, $Q_2 = 4.5 \cdot 10^{-11}$ C, $M_2 = 8.62 \cdot 10^{-13}$ kg, from which a small cluster with a charge $q_2 = 4.5 \cdot 10^{-13}$ C and mass $m_2 = 8.62 \cdot 10^{-15}$ kg separates at a height $L = 15 \cdot 10^{-6}$ m, leads to the following results: $W_{q_2} = 1.619 \cdot 10^{-8}$ $\text{J} = 1.012 \cdot 10^{11}$ eV, $v_2 = 1.938 \cdot 10^3$ m/s, $V_2 = 19.379$ m/s, $W_{Q_2} = 1.619 \cdot 10^{-10}$ J .

A cluster with a charge $Q_1 = 4.5 \cdot 10^{-12}$ C creates an electric field $E_1 = Q_1 / 4\pi\epsilon_0 L^2 = 1.8 \cdot 10^8$ V/m at a distance $L = 15 \cdot 10^{-6}$ m, and a cluster with a charge $Q_2 = 4.5 \cdot 10^{-11}$ C at the same distance creates a field $E_2 = 1.8 \cdot 10^9$ V/m. Such fields are comparable in magnitude with the intensity $E_{\text{min}} = 2 \cdot 10^9$ V/m, determined by the Langevin criterion. Therefore, we can assume that polarization of the film material will occur, and a force will act on the cluster, which will slow down its movement. When the work of this force becomes equal to the kinetic energy of the cluster, it will stop and begin to "fall" onto the film. The force acting on the cluster from the side of the polarized film is $F_d = P_1 Q_1 / 2\pi\epsilon_0 R^3$ (see formula 2), where P_1 is dipole moment of the polarized section of the film, and R is the distance between the charge center and the film. The work of this force on the section of the trajectory dR is equal to $dA = F_d dR$, and the total work until the cluster stops is:

$$A = \int_{R_0}^R F_d dR = -\frac{P_1 Q_1}{4\pi\epsilon_0} \cdot \frac{1}{R^2} \Big|_{R_0}^R = \frac{P_1 Q_1}{4\pi\epsilon_0} \left[\frac{1}{R_0^2} - \frac{1}{(R_0 + \Delta R)^2} \right]. \quad (3)$$

Here R_0 is the height of the cluster above the film surface at the moment it begins to move upward, and $R = (R_0 + \Delta R)$ is the height it will reach. In this problem, the unknown parameter is the value of the film dipole moment P_1 induced by the charge Q_1 . To estimate its value, we assume that the cluster, starting from a height $R_0 = L = 15 \cdot 10^{-6}$ m, could rise to a height $(R_0 + \Delta R) = 25 \cdot 10^{-6}$ m. Equating the kinetic energy of the cluster with the charge $Q_1 = 4.5 \cdot 10^{-12}$ C ($W_{Q_1} = 7.447 \cdot 10^{-12}$ J) in work A , from formula (3) we find $P_1 = 7.3 \cdot 10^{-21}$ C·m. The specific polarization of the medium P_V (C·m/m³) is related to the field strength E by the relation $P_V = \epsilon_0(\zeta - 1)E$ [34]. Assuming for the film material the value of dielectric constant $\zeta = 2.6$, at the field strength $E_{\min} = 2 \cdot 10^9$ V/m we obtain $P_V = 2.83 \cdot 10^{-2}$ C·m/m³. The dipole moment $P_1 = 7.3 \cdot 10^{-21}$ C·m can be created by a grain with a volume $V_f = P_1/P_V = 0.26 \cdot 10^{-18}$ m³ sized 0.64 μ m.

Fig. 13 shows the result of numerical calculation of the change in the cluster height above the level of separation of a small cluster from it $L = 15 \cdot 10^{-6}$ m. The parameters of the cluster are accepted as $Q_1 = 4.5 \cdot 10^{-12}$ C, $M_1 = 4 \cdot 10^{-14}$ kg and diameter $D_1 = 4.28 \cdot 10^{-6}$ m. It is believed that for a separated small cluster, the charge $q_1 = Q_1/100 = 4.5 \cdot 10^{-14}$ C and the mass $m_1 = M_1/100 = 4 \cdot 10^{-16}$ kg are one percent of the initial values of the charge and masses of the main cluster, and these quantities remain unchanged despite the fact that Q_1 and M_1 decrease with each "jump". For the dipole moment of the polarized film, the value $P_1(t)$

$= Q_1(t) \times 2 \cdot 10^{-8}$ C·m was taken (the decrease in the charge Q_1 with each act of separating a small cluster was taken into account). It is assumed that at $t = 0$ the lower edge of the cluster was at a distance $1.5 \cdot 10^{-6}$ m from the emission height of a small cluster and its velocity was zero. It can be seen that after five "jumps", the system switches to the mode of alternating short (0.5 μ s in duration) and long (1.4 μ s in duration) "jumps" with an average period of 1.86 μ s. The height of the "jump" and the duration of the "flight" grow with time (see Fig. 13). This can be explained by a gradual decrease in the charge of the large cluster and the magnitude of the induced dipole momentum of the film. The reason for the alternation of long (high) and short (low) "jumps" is as follows. When a cluster "falls" from a great height, it acquires a greater speed. To stop the cluster and change the direction of its velocity vector to the opposite, it is necessary to spend a significant part of the momentum acquired by it during the emission of a small cluster. As a result, the cluster begins to move upwards at a low speed and, accordingly, "falls" also at a low speed. The next pulse of interaction between the charges of a large and small cluster "tosses" it to a greater height. If a charged cluster has a velocity component directed along the plane of the film, it will leave a chain of double spots on it, separated by large intervals. These spots can merge due to charge spreading and form complex traces of the type shown in Fig. 5.

4.3. ACTION ON A CLUSTER OF A MAGNETIC FIELD Fredericks [8,9] observed the curvature of particle tracks on photographic film when he applied a permanent magnet to it (Fig. 7). The magnetic field was perpendicular to the film surface. Instead of the expected movement of magnetic monopoles along the magnetic field lines (that is, perpendicular to the plane of the film), they moved perpendicular to the lines of force parallel to the film surface, leaving traces in the form of parabolas on it. Let us discuss

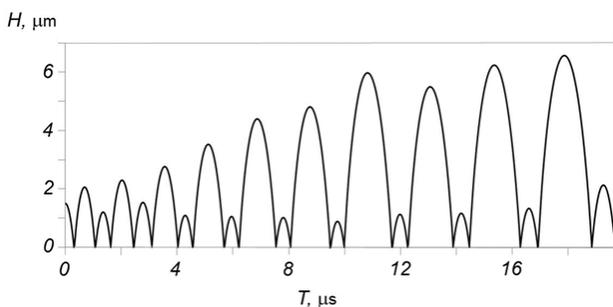


Fig. 13. Change in the rise height of the charged cluster above the level $L = 15 \mu$ m, at which a small charge was separated from it (calculation result) [21-23].

how this behavior of "strange" particles can be explained.

Let us assume that a multiple charged cluster with charge Q and mass m flies with velocity V into the region of action of a constant magnetic field with induction B . The value of this velocity is determined by the equality of the force of the external electric field E_x acting on the charge and the friction force of the cluster on air. Let the velocity V be perpendicular to the induction vector B . In this case, the force $F_m = QVB$ will act on the charge perpendicular to the velocity vectors V and induction B . Due to the low speed of the lateral displacement, let us assume that the friction force of the particle on the air is small and can be neglected. The particle's lateral displacement acceleration due to this force is $a = F_m/m$, and the displacement in time t is $S = at^2/2 = (QVB/2m) \cdot t^2$. Let's take $Q = 4.5 \cdot 10^{-12}$ C, $V = 5$ m/s, $m = 4 \cdot 10^{-14}$ kg. The time of passage of the track length $L = 10^{-2}$ m is equal to $t = 2 \cdot 10^{-3}$ s. During this time, the track deviated to a distance $S = 10^{-3}$ m. Substituting numerical values into the expression for S , we find $B = 2mS/QVt^2 = 0.89$ T. The line drawn during the passage of a particle of distance L , accompanied by a lateral displacement S , has the shape of a parabola.

4.4. MAGNETIC MOMENTUM OF A ROTATING CLUSTER

The momentum of inertia of a ball of radius R is $I_b = (2/5)mR^2$. The mass of the ball is $m = (4/3)\pi R^3 \rho$ (ρ is the density of the material of the ball), hence the momentum of inertia of the solid ball is $I = (8/15)\pi \rho R^5$. For a cluster in the form of a hollow sphere with outer radius $R = 2.12 \cdot 10^{-6}$ m and inner radius $R_{in} = 0.6 \cdot 10^{-6}$ m momentum of inertia $I = (8/15)\pi \rho (R^5 - R_{in}^5) = 71.58 \cdot 10^{-27}$ kg·m² (we took $\rho = 10^3$ kg/m³ equal to the density of water). Let's replace the cluster with a weight m_c rotating on a circle of radius R . For it $I = m_c R^2$ and $m_c = 15.927 \cdot 10^{-15}$ kg. The momentum of rotation of the load is equal to $I = m_c v R$ (v is the speed of the movement of the load). For a cluster $I = 3.376 \cdot 10^{-20} v$. If the load

rotates with a frequency of n_r revolutions per second, then its speed is $v = 2\pi R n_r$ m/s and $I = 4.49 \cdot 10^{-25} \cdot n_r$ kg·m²/s. Let us estimate the cluster rotation frequency. Let us take the kinetic energy of the load $m_c v^2/2$ equal to $k_b T = 1.38 \cdot 10^{-23}$ J/K $\times 300$ K = $4.14 \cdot 10^{-21}$ J. With $m_c = 15.927 \cdot 10^{-15}$ kg the velocity $v = 0.72 \cdot 10^{-3}$ m/s and $n_r = v/2\pi R = 54$ s⁻¹.

When the cluster rotates around an axis passing through its center, its inner sphere with a radius $R_{in} = 0.6 \cdot 10^{-6}$ m also rotates, on which there are ions with a total charge $Q = 2 \cdot 10^{-12}$ C. If the charges are uniformly distributed over the sphere, their surface density is $\sigma = Q/4\pi R_{in}^2 = 0.44$ C/m². The rotation of the charge is equivalent to the formation of a ring with current, which creates a magnetic moment, the vector of which is directed along the axis of rotation of the cluster. **Fig. 14** shows a cross section of such a cluster. The radius of the current ring is $r = R \cos \alpha$, where R is the radius of the sphere, and α is the angle between the plane of the equator and the line of intersection of the current ring with the sphere. Current ring area $s = 2\pi r R \cdot d\alpha = 2\pi R^2 \cos \alpha d\alpha$. Current ring charge $q = s \cdot \sigma = (Q/2) \cos \alpha d\alpha$. If the sphere makes n_r revolutions per second, then the current $i_r = q \cdot n_r = (Q/2) n_r \cos \alpha d\alpha$ flows through the ring of radius r . The magnetic momentum of this current is $dp_m = i_r \pi r^2 = (Q/2) n_r R^2 \cos^3 \alpha d\alpha$. Total magnetic flux of the rotating sphere is $p_m = Q n_r \pi R^2 \int_0^{\pi/2} \cos^3 \alpha d\alpha = (2/3) Q n_r \pi R^2$. At $Q = 4.5 \cdot 10^{-12}$ C and $R = 0.6 \cdot 10^{-6}$ m $p_m = 3.39 \cdot 10^{-24} n_r$ (C·m²/s). We found above that due to thermal perturbations, the cluster can rotate at a frequency of about 100 revolutions per second. Suppose that for some reason the rotational speed has become 1000 revolutions per second. Then $p_m = 3.39 \cdot 10^{-21}$ C·m²/s. In an inhomogeneous magnetic field, a force proportional to the magnetic induction gradient acts on the magnetic dipole, forcing the dipole to move in the direction of increasing magnetic field strength. This force is $F_m = p_m \cdot \text{grad} B$. Let

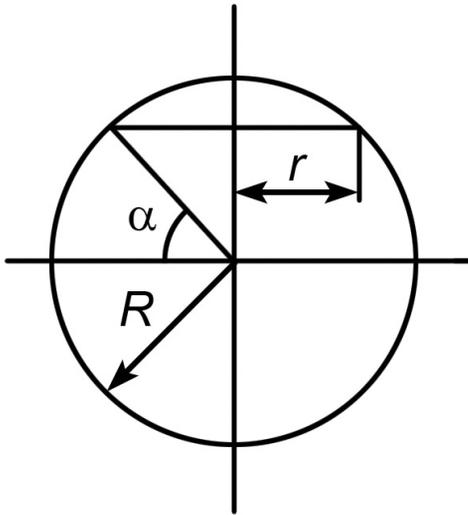


Fig. 14. Section of a rotating cluster.

$\text{grad}B = 10^2 \text{ T/m}$ near the pole of a strong magnet, then $F_m = 3.39 \cdot 10^{-19} \text{ N}$. This force is much less than the force of the weak electric field on the cluster charge (at $Q = 4.5 \cdot 10^{-12} \text{ C}$ and $E = 10 \text{ V/m}$ force $F_e = QE = 4.5 \cdot 10^{-11} \text{ N}$). Thus, the magnetic field can only act on a moving charged cluster (due to the Lorentz force). The "gyromagnetic" ratio for the cluster is $p_m/l = 33.9 \cdot 10^{-25} n_r / 4.49 \cdot 10^{-25} n_r = 7.55 \text{ C/kg}$ (for an electron $p_m/l = 1.76 \cdot 10^{11} \text{ C/kg}$).

4.5. CURLS OF THE TRAJECTORY

Characteristic feature of traces left by "strange" particles are sharp breaks in their trajectories (Fig. 6, Fig. 8). This gave a reason to some researchers [10,11] to believe that the traces were left by particles with a very small mass (magnetic monopoles or magnetically excited neutrinos) [3,4,10,11]. Let us discuss whether this behavior of particles can be explained if we consider them as charged spheres with a certain size and nonzero mass. Let us have a cluster with inner sphere radius $r = 6 \cdot 10^{-7} \text{ m}$, having charge $Q = 4.5 \cdot 10^{-12} \text{ C}$. The thickness of the water shell of such a cluster is $a = 1.54 \cdot 10^{-6} \text{ m}$, and its outer radius is $R = r + a = 21.4 \cdot 10^{-7} \text{ m}$. The volume of the shell is $V_{sh} = 4\pi[(r + a)^3 - r^3]/3 = 4 \cdot 10^{-17} \text{ m}^3$, and the mass of the shell, equal to the mass

of the cluster, $M_{sh} = 4 \cdot 10^{-14} \text{ kg}$. In an electric field E_x , the cluster is affected by the force $F_e = QE_x$, which makes it move with velocity v in the direction of the field strength vector. The air resistance force $F_D = (\pi C_D \rho_m R^2 v^2)/2$ acts opposite to the F_e force, where $C_D = 5.8$ is the medium resistance coefficient, and $\rho_m = 1.205 \text{ kg/m}^3$ is the air density. For $F_e = F_D$, the cluster moves at a constant speed. Let this speed be 10 m/s . In this case, the kinetic energy of the cluster $W_k = M_{sh} v^2/2 = 2 \cdot 10^{-12} \text{ J}$, and the force $F_D = 5 \cdot 10^{-9} \text{ N}$. Suppose that at some moment action of the force F_e is ceased and the cluster, slowing down under the action of the force F_D , will stop after passing a certain distance x . The magnitude of the force F_D will decrease with decreasing speed, but we will neglect this and assume that it is equal to its maximum value. The distance $x = W_k/F_D = 0.4 \text{ mm}$, that is, in order to stop the movement and change the trajectory, the cluster must describe an arc with a radius of about 0.25 mm . In fact, according to observations, the radius of this arc is 1000 times smaller. The reason for this discrepancy lies in the fact that we assumed that the cluster flies freely over the film surface without interacting with it. This is not true. According to Fig. 13 the cluster, moving along the film surface, oscillates in the direction perpendicular to the film plane. At some point, it is attracted to the film. This means that when the trajectory is bent, it makes a turn around a certain point lying on the surface of the film, and not a turn in free space. This situation is similar to what you might encounter when you are descending a steep mountain. You are moving fast, visibility is limited due to the fog. And suddenly you see that you are on the edge of the abyss and you no longer have time to turn sideways. You are saved by a tree that happened to be in your way. You grab at its trunk and turn sharply.

4.6. SYMMETRIC TRACKS

The movement of a spherical cluster occurs due to the action of two forces. The first is the action of the electric field E on the charge Q : $F_e = QE$.

Opposite to this force, a sphere of radius R moving at a speed v is affected by the air friction force [37]

$$F_D = (\pi C_D \rho_m R^2 v^2) / 2, \tag{4}$$

where $\rho_m = 1.205 \text{ kg/m}^3$ is the air density, and the drag coefficient of the medium at a speed $v = 10\text{-}20 \text{ m/s}$ can be set equal to $C_D = 5.8$. At $F_e = F_D$, the cluster velocity is:

$$v = \left(\frac{2QE}{\pi C_D \rho_m R^2} \right)^{1/2} = A \left(\frac{QE}{R^2} \right)^{1/2}. \tag{5}$$

In the images of traces of "strange" particles, symmetrical traces ("twin tracks") are sometimes observed (Fig. 9, Fig. 10, Fig. 11). It can be assumed that the particles that left them flew in the same direction, simultaneously making bends in the trajectory. According to our model, this was due to the fact that the particles had a charge and moved due to the action of electric fields on them, which changed direction at some points in time. In rare cases, "chiral" traces were observed when the particles that left them simultaneously moved in the opposite direction (Fig. 12). This can be explained by the fact that the pair of particles had charges of the opposite sign. An analysis of the lengths of the sections of the trajectory of two particles showed that the lengths of the straight segments are the same, that is, both of these particles moved at the same speed. In principle, this could happen if both particles had the same charge (modulo) and the same size. However, this is highly unlikely. It would be more realistic to assume that the particles had different charges and sizes, however, the size of particles with a smaller charge was smaller and therefore the force of their friction against air was also smaller. Ideally, one would expect particles that have the same ratio of charge Q to the square of radius R^2 to move in a constant electric field at the same speed.

Let us discuss whether there can be clusters for which the ratio $Q/R^2 = \text{const}$. In Section 4.1, we described the procedure for calculating the parameters of a cluster holding inside a shell with

a thickness $a = 1.54 \cdot 10^{-6} \text{ m}$ an electric charge $Q = 4.5 \cdot 10^{-12} \text{ C}$. The cluster radius turned out to be equal to $R = 2.12 \cdot 10^{-6} \text{ m}$. Calculating parameters of other clusters on described scheme we find that the cluster with a charge $Q = 4.5 \cdot 10^{-14} \text{ C}$ has a radius $R = 2.11 \cdot 10^{-7} \text{ m}$, the cluster with $Q = 4.5 \cdot 10^{-13} \text{ C}$ has a radius $R = 6.7 \cdot 10^{-7} \text{ m}$, and a cluster with $Q = 4.5 \cdot 10^{-11} \text{ C}$ has $R = 6.55 \cdot 10^{-6} \text{ m}$. The results of calculation of cluster parameters are presented in the **Table**.

As you can see, the value of the Q/R^2 parameter is equal to one. This value is close to the value $\sigma = 1 \text{ C/m}^2$, which is the surface charge density of dipoles of water molecules placed as a monolayer on the shell of ball lightning [24-33]. In twin tracks, there is synchronization of two orders. The first is external. On continuous tracks, it is seen that the particles change their trajectory simultaneously under the influence of a change in the external field. The second is internal order. On Fig. 11 it can be seen that the particles "draw" the strokes in a coordinated way, that is, they "reset" the charges at the same time. Perhaps the reason for this correlation is the exchange of pulses of electromagnetic radiation during each discharge. The appearance of multiple charged clusters with different polarity during electric discharges seems to be quite natural. In some cases, they are formed near the anode, and in others, near the cathode.

4.7. "STRANGE" PARTICLES AND NUCLEAR TRANSMUTATION

The idea was repeatedly expressed that the registered changes in the isotopic composition of elements during electric discharges can be somehow connected with the appearance of "strange" particles. Even the question of what is primary – the appearance of particles

Table
Cluster radius R as a function of its charge Q .

Q, C	R, m	R^2, m^2	$Q/R^2, \text{C/m}^2$
$4.5 \cdot 10^{-14}$	$2.11 \cdot 10^{-7}$	$4.45 \cdot 10^{-14}$	1.01
$4.5 \cdot 10^{-13}$	$6.70 \cdot 10^{-7}$	$44.89 \cdot 10^{-14}$	1.00
$4.5 \cdot 10^{-12}$	$2.14 \cdot 10^{-6}$	$4.58 \cdot 10^{-12}$	0.98
$4.5 \cdot 10^{-11}$	$6.55 \cdot 10^{-6}$	$42.90 \cdot 10^{-12}$	1.05

and the subsequent reaction of a change in the composition of nuclei, or vice versa [5,6], was even discussed. Let's assume that the first version is correct. Experiments have shown that "strange" particles can penetrate solids and then be emitting by them for a long time [12,13]. A multiple charged cluster embedded in a crystal lattice creates an electric field in the region around it, the strength of which is comparable to the strength of the electric field at the location of an electron in a hydrogen atom. If the cluster has a negative charge, this will lead to the repulsion of the electron cloud of the lattice atom, and the nucleus will be "bare", as it were. It is impossible to predict in advance what can happen in the region between the outer surface of the cluster shell and the "bare" nuclei during the time the cluster stays inside the crystal lattice (up to several days). For a thermonuclear reaction in deuterium-tritium plasma, there is the Lawson criterion, according to which the probability of a reaction is determined by the product of the plasma density and the time of its confinement. It is possible that the time parameter is also important in determining the probability of nuclear transformations in strong electric fields. Another aspect is important. In the search for the mechanism of nuclear transformations, it is always implicitly assumed that the reaction occurs in the form of a single act, the implementation of which requires energy of several mega-electronvolts. But processes also occur in nature when a "mega-transformation" occurs through a cascade of "small" steps. The most striking example of this is the absorption of two "red" quanta by plant chlorophyll. A polyatomic molecule (for example, SF₆) can be excited to the dissociation level ($E_d = 2$ eV) due to the absorption of CO₂ laser radiation quanta ($\lambda = 10.6$ μm , $E_\gamma = 0.11$ eV) [38]. A person cannot jump more than two meters, but can climb stairs to the 20th floor of a house. It is possible that such a process of step-by-step overcoming of the barrier may play some role in the reactions of "cold fusion" of elements.

5. CONCLUSION

Based on the above, the following conclusions can be drawn:

1. Particles of "strange" radiation are not elementary particles or atomic nuclei, for which the only reservoir of action on matter is their kinetic energy. It is also unlikely that these particles serve as a catalyst for exothermic nuclear reactions in ordinary matter. It is more realistic to consider that these are macroscopic particles with a large internal potential energy reserve.
2. The appearance of traces of "strange" particles is naturally explained under the assumption that they are left by particles with an electric charge exceeding the elementary charge by more than 1000 times.
3. The charge of these particles can be both positive and negative.
4. Charged particles move due to the action of random electric fields on them (the "fair weather" field strength near the earth's surface is about 100 V/m). This field changes chaotically in time.
5. The properties of "strange" particles are similar to the properties of ball lightning. They, like ball lightnings, are able to move along the surface of a solid body, make jumps, leave holes in the material when rebounding, and divide into parts. This makes it possible to identify them with miniature ball lightnings.
6. Like ball lightning, these particles are able to penetrate into a solid body. The estimate shows that a spherical cluster with a radius of $2.14 \cdot 10^{-6}$ m and a charge of $4.5 \cdot 10^{-12}$ C presses on the surface with a force of 10^{10} N/m².
7. The electric field strength on the surface of such a cluster can reach up to 10^{10} V/m. This is comparable to the strength of the electric field in an atom ($5 \cdot 10^{11}$ V/m). In the crystal lattice near the intercalated cluster, a strong polarization of the substance will occur, which can facilitate the conditions for the approach of protons and nuclei of elements.

8. This forces us to consider the possibility of nuclear reactions proceeding through the process of overcoming the potential barrier, which lasts for a time many orders of magnitude greater than the time of nuclear collision in a conventional nuclear process.

REFERENCES

1. Matsumoto T. Ball lightning during underwater spark discharges and the Matsumae earth-quarks. *Proc. 5th Intern. Symposium on Ball Lightning (ISBL97)*, p. 193-201. Japan, Tsugawa-Town, Niigata, 1997.
2. Matsumoto T. Micro ball lightning during underwater spark discharge. *Proc. 6th Intern. Symposium on Ball Lightning (ISBL-99)*, p. 249-254. Belgium, Antwerp, 1999.
3. Urutskoev LI, Liksonov VI, Tsinoev VG. Experimentalnoye obnaruzhenie "strannogo" izlucheniya i transformatsii khimicheskikh elementov [Experimental finding of "strange" radiation and transformation of chemical elements]. *Prikladnaya fizika*, 2000, (4):83-100 (in Russ.).
4. Urutskoev LI, Liksonov VI, Tsinoev VG. Observation of transmutation of chemical elements during electric discharge. *Annales de la Fondation Louis de Broglie*, 2002, 27(4):701-726.
5. Zhigalov VA. Treki na fotoplyonke ot strannogo izlucheniya: replikatsiya [Tracks on photo film from a strange radiation: replication]. *Zhurnal formiruyuschikhsya napravleniy nauki (ZhFNN)*, 2015, 3 (9):55-62 (in Russ.).
6. Zhigalov VA. Strannoye izlucheniye i LENR, kakaya svyaz'? [Strange radiation and LENR, what connection?]. *RENSIT: Radioelectronics. Nanosystems. Information technologies*, 2021, 13(3):329-348. DOI: 10.17725/rensit.2021.13.329.
7. Agapov AS, Kalensky VA, Kaitukov IB, Malyshev AV, Ryabova RV, Steblevsky AV, Urutskoev LI, Filippov DV. Obnaruzhenie "strannogo" izlucheniya i izotopnogo ischazheniya titana pri ispytaniyakh promyshlennogo elektrotekhnicheskogo oborudovaniya [Observation of "strange" emission and titan isotopes distortion at testing of the industrial electric equipment]. *Proc. 12th Russian Conference on Cold Nuclear Transmutation of Chemical Elements and Ball Lightning*, p. 24-40. Moscow, NIC FTP "Erzion", 2005, (in Russ.).
8. Fredericks KA. Possibility of tachyon monopoles detected in photographic emulsions? *Condensed Matter Nuclear Science*, 2015, 15:203-230.
9. Fredericks KA. Possible detection of tachyon monopoles in photographic emulsions. June 7, 2013, <https://www.researchgate.net/publication/289165518>.
10. Ivoilov NG. Nizkoenergetichnaya generatsiya "strannogo" izlucheniya [Low energy generation of the "strange" radiation]. *Georesursy*, 2005, 2(17):84-92 (in Russ.).
11. Ivoilov NG. Low energy generation of the "strange" radiation. *Annales de la Fondation Louis de Broglie*, 2006, 31(1):115-123.
12. Bogdanovich BY, Volkov NV, Kostochko YP, Len NA, Nesterovich AV, Starostin AI. Experimentalnoye issledovanie kvaziperiodicheskogo impulsnogo razryada, initsiirovannogo v potoke zhidkosti v prielectroдном prostranstve [Experimental investigation of quasi-periodical pulse discharge, initiated in the liquid flow in near-electrode space]. *Ingenernaya fizika*, 2000, 1:19-23 (in Russ.).
13. Nesterovich AV, Fetisov GP. Deformatsionnoye uprochnenie metalla pri vozdeystvii periodicheskogo razryada v potoke zhidkosti [Deformation strengthening of metal at action of periodical discharge in the liquid stream (PDLs)]. *Ingenernaya fizika*, 2007, 5:7-11 (in Russ.).
14. Lochak G. The symmetry between electricity and magnetism and the problem of the existence of magnetic monopole. In: *Advanced Electromagnetism*. TW Barnett and

- DM Grimes (eds), p. 105-147. Singapore, World Scientific Publishing Co., 1995.
15. Adamenko SV. Kontseptsiya iskusstvenno initsiiruemogo kollapsa veshchestva i osnovnye rezultaty pervogo etapa eyo experimentalnoi realizatsii [A conception of the artificial initiation of matter collapse and fundamental results of the first stage of its experimental realization]. *Preprint*, 36 p. 2004, Kiev, *Academperiodika*; <http://proton-21.com.ua/articles/.Preprint-ru.pdf> (in Russ.).
 16. Adamenko SV, Vysotsky VI. Experimentalnoye obnaruzhenie i modelirovanie orientatsionnogo dvizheniya gipoteticheskikh magnitozaryazhennykh chastits na mnogoslnoynoy poverkhnosti [Experimental finding and modeling of the oriented motion of hypothetic magneto-charged particles on the multi-layer surface]. *Poverkhnost*, 2006, 3:84-92 (in Russ.).
 17. Shishkin AL, Baranov VA, Vinogradova AV, Dubovik VM, Panyushkin VA, Tatur VYu. Issledovanie kharakteristik Magneto-Toro-Elektricheskikh izlucheni s pomoshchyu fotoplyonochnykh detektorov [Investigation of characteristics of Magneto-Toro-Electric radiation with a help of photographic films detectors]. *Academiya Trinitarizma*, Moscow, Publication 17244, 21.01.2012, URL <http://www.trinitas.ru/rus/doc/0231/004a/0231104.htm> (in Russ.).
 18. Shishkin AL, Dubovik VM, Kuroles VK, Tatur VYu. Issledovanie kharakteristik "neytrino" klaster'nogo izlucheniya [Investigation of characteristics of "neutrino" cluster radiation]. *Academiya Trinitarizma*, Moscow. Publication 23979, 20.11.2017, URL <http://www.trinitas.ru/rus/doc/0016/001f/00163510.htm> (in Russ.).
 19. Kashchenko MP, Kashchenko NM. *Nizkotemperaturnyi Yadernyi Sintez: Vvedenie v Problemy i eyo Konceptualnoye Reshenie* [Low Temperature Nuclear Synthesis: Introduction to the Problem and its Conceptual Solution]. Ekaterinburg, UGLTU Publ., 2022 (in Russ.).
 20. Baranov DS, Zatelepin VN. Mekhanizm formirovaniya i fiziko-khimicheskiye svoystva "tyomnogo vodoroda". Teoriya i experiment [Mechanism of formation and physicochemical properties of "dark hydrogen". Theory and experiment]. *Proc. 26th Russian Conference on Cold Nuclear Transmutation of Chemical Elements and Ball Lightning*, p. 64-86. Moscow, DeLibry Publ., 2020, (in Russ.).
 21. Nikitin AI, Nikitin VA, Velichko AM, Nikitina TF. "Strannye" chastitsy i microrazmernye sharovye molnii ["Strange" particles and micro-sized ball lightning]. *Proc. 26th Russian Conference on Cold Nuclear Transmutation of Chemical Elements and Ball Lightning*, p. 226-244. Moscow, DeLibry Publ., 2020, (in Russ.).
 22. Nikitin AI, Nikitin VA, Velichko AM, Nikitina TF. "Strange" particles and micro-sized ball lightning in some electric discharges. *Journ. Atm. Solar-Terr. Phys.*, 2021, 218:105525.
 23. Nikitin AI, Nikitin VA, Velichko AM, Nikitina TF. Sledy na plyonkakh i mnogozaryadnye klaster'y [Tracks on films and multi-charged clusters]. *RENSIT: Radioelectronics. Nanosystems. Information technologies*, 2021, 13(3):355-356. DOI: 10.17725/rensit.2021.13.355.
 24. Nikitin AI. Elektricheskii kondensator kak element energeticheskogo yadra sharovoy molnii [An electric capacitor as an element of energy core of ball lightning]. *Elektrichestvo*, 1998, 11:14-23 (in Russ.).
 25. Nikitin AI. An electrical capacitor as the element of the power core of ball lightning. *Electrical Technol. Russ.*, 1998, (4):70-85.
 26. Nikitin AI. The dynamic capacitor model of ball lightning. *Proc. of 6th Intern. Symposium on Ball Lightning (ISBL-99)*, p. 91-95. Belgium, Antwerp, 1999.
 27. Nikitin AI. Elektrodinamicheskaya model sharovoy molnii. [Electrodynamic model

- of ball lightning]. *Khimicheskaya fizika*, 2006, 25(3):38-62 (in Russ.).
28. Nikitin AI. Small-sized and composite ball lightning. *Intern. Journ. "Unconventional Electromagnetics and Plasmas (UEP)"*, 2012, 4(1-2):105-116.
29. Nikitin AI. Possible process of ball lightning training in nature. *Journ. Atm. Solar-Terr. Phys.*, 2019, 190:54-61.
30. Nikitin AI, Bychkov VL, Nikitina TF, Velichko AM. High-energy ball lightning observations. *IEEE Trans. Plasma Sci.*, 2014, 42(12):3906-3911.
31. Nikitin AI, Bychkov VL, Nikitina TF, Velichko AM, Abakumov VI. Sources and components of ball lightning theory. *J. Phys. Conf. Ser.*, 2018, 996:012011.
32. Bychkov VL, Nikitin AI, Dijkhuis GC. Ball lightning investigations. In: Bychkov VL, Golubkov GV, Nikitin AI (Eds). *The Atmosphere and Ionosphere. Dynamics, Processes, and Monitoring*. Dordrecht, Springer, 2010, p. 201-373.
33. Bychkov VL, Nikitin AI. Ball lightning. A new step in understanding. In: Bychkov VL, Golubkov GV, Nikitin AI (Eds). *The Atmosphere and Ionosphere. Elementary Processes, Monitoring, and Ball Lightning*. Cham, Springer, 2014, p. 201-367.
34. Blakemore JS. *Solid State Physics*. Cambridge, Cambridge University Press, 1985, 608 p.
35. Shevkunov SV. Vliyanie ionov khlorina na ustoychivost yader nukleatsii v kondensiruyushchikhsya parakh vody [Action of chlorine ions on stability of nucleation centers in the condensing water vapor]. *Jurnal Fizicheskoi Khimii*, 2011, 85(9):1702-1709 (in Russ.).
36. Shevkunov SV. Krizis ustoychivosti gidratnoi obolochki iona Na^+ v kondensiruyushchikhsya parakh vody [Stability crisis of the hydrate shell of ion Na^+ in the condensing water vapor]. *Kolloidnyi Jurnal*, 2011, 73(2):267-278 (in Russ.).
37. Reist PC. *Introduction to Aerosol Science*. New York/London, Macmillan Publishing Company, 1984.
38. Letokhov VS, Makarov AA. Mnogoatomnye molekuly v silnom infrakrasnom pole [Multi-atomic molecules in a strong infrared field]. *Uspekhi Fizicheskich Nauk*, 1981, 134(1):45-91 (in Russ.).

DOI: 10.17725/rensit.2022.14.269

Information Technologies Based on Noise-like Signals: III. Spectrum of Sequences Periods Forming by a Discrete Chaotic Algorithm

Nikita A. Ageykin, Vladimir I. Grachev, Nadezhda G. Petrova, Viktor I. Ryabenkov,
Vladimir V. Kolesov

Kotelnikov Institute of Radioengineering and Electronics of RAS, <http://www.cplire.ru/>
Moscow 125009, Russian Federation

E-mail: ageykin_niki@mail.ru, grachev@cplire.ru, petrova@cplire.ru, ryabenkov.vi@list.ru, kvv@cplire.ru

Alexander M. Popov

MIREA-Russian Technological University, <http://www.mirea.ru/>
Moscow 119454, Russian Federation

E-mail: popov@mirea.ru

Received September 05, 2022, peer-reviewed September 12, 2022, accepted September 16, 2022

Abstract: The patterns of formation of pseudo-random sequences by discrete chaotic algorithms with a delay of the Fibonacci type are studied. The algorithms are defined on a closed interval of integers; the generated numbers are returned to a given interval, which provides an effective mechanism for mixing in the phase space. For various values of the parameters, the spectra of periods of sequences are determined for an arbitrary set of initial values that uniquely determine the state of the system in its phase space. Relationships are obtained that make it possible to reduce the search time for the maximum period of a pseudo-random sequence formed by a recurrent chaotic Fibonacci-type algorithm.

Keywords: information technology, chaotic dynamics, pseudorandom sequences, redundant codes, noise-like signals

UDC 621.391

Acknowledgments: The work was carried out within the framework of the state task of the Kotelnikov IRE of RAS from the Ministry of Education and Science of the Russian Federation.

For citation: Nikita A. Ageykin, Vladimir I. Grachev, Nadezhda G. Petrova, Alexander M. Popov, Viktor I. Ryabenkov, Vladimir V. Kolesov. Information Technologies Based on Noise-like Signals: III. Spetrum of Sequences Periods Forming by a Discrete Chaotic Algorithms. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2022, 14(3):269-278. DOI: 10.17725/rensit.2022.14.269.

CONTENTS

1. INTRODUCTION (269)
 2. ALGORITHM FOR GENERATING A PSEUDO-RANDOM SEQUENCE (272)
 3. RESULTS OF CALCULATIONS AND REVEALED REGULARITIES (273)
 4. CONCLUSION (277)
- REFERENCES (278)

1. INTRODUCTION

The basic concept in the field of information technology in the transmission, reception and

processing of information is the concept of "information coding", usually interpreted as a synonym for the concept of "representation of information on a specific medium". Such information carriers can be: graphics (drawings), texts, numbers, musical notations, sequences of electromagnetic, optical, acoustic or other signals.

The word "code" in the most general sense refers to a specific way of presenting information. Often, the word "coding" is also understood not only as a way of presenting

information, but also as the process of converting from one representation of information to another. The process of inverse transformation in such a case is usually called decoding.

The applied application of information technologies involves the physical implementation of a specific coding process in the transmission, processing and storage of information in telecommunication systems and computer networks. Progress in this area is associated with an increase in speed and an increase in the noise immunity of information channels. This is primarily dictated by the need to develop effective channels for information exchange and control of distributed networks and automatic systems with remote control, where the cost of an error or partial loss of information can have catastrophic consequences, up to the loss of the entire system.

The trend observed in recent years of the global spread of various open telecommunication systems and a sharp increase in the number of subscribers lead to the need to protect information not only at the level of government agencies, special services or business circles, but also at the level of almost every individual user. In information networks, this problem is associated not so much with the confidentiality of information as with the loss of information due to the low noise immunity of various communication channels.

At present, the problem of increasing the noise immunity of communication channels is especially acute in radio relay lines. The saturation of the frequency range with traditional means of communication based on the principle of frequency division of channels has led to the development of communication means operating in a common frequency band and using coding

methods for the so-called code division of channels, in which streams of random (pseudo-random) numbers used as codes. When transmitting information, streaming coding provides the greatest noise immunity (and hence cryptographic strength) when using continuous random key streams with a uniform distribution function.

It should be specially noted that one of the main problems in the generation of pseudorandom sequences (PRS) is the need to form long implementations when using a short master key that determines the initial conditions. The development of new generating algorithms for generating pseudo-random sequences is an important modern task and requires an understanding of the patterns of formation of the PRS of numbers with certain specified statistical properties.

In the software implementation of algorithms for generating pseudorandom processes, the computer operates with discrete numbers in binary representation with a finite number of bits. Given this limitation on the finite bitness of numbers in a computer, the total volume of the phase space (PS), any point of which corresponds to an unambiguous state of the system, is limited. Accordingly, any algorithmic method of formation should sooner or later enter the periodic repetition of the same segments of the sequence being formed, that is, enter a cycle, although its period may be very large and even infinite from the point of view of a number of practical applications.

The requirements for the properties of sequences of pseudo-random numbers depend on specific applications and, as a rule, one algorithm is not able to satisfy all these requirements. In the general case, it is possible to formulate the main requirements for PRS [1]:

- *high quality*: according to statistical criteria, PRS should be close to a random process and have the longest possible period;
- *efficiency*: the algorithm should be fast and occupy as little memory as possible;
- *reproducibility*: with exact reproduction of the algorithm initial conditions, the same PRS should be formed on implementations of any duration, and minor changes in the initial procedure should lead to the generation of qualitatively different sequences;
- *simplicity*: the generator algorithm formula should be easy to implement and use.

In the available literature, there is practically no information about the methods for developing algorithms for generating pseudo-random numbers [2]. From the point of view of practical application in digital information technologies, algorithms defined on a closed interval of integers are of interest. Their advantage is related to the absence of the need to use any rounding in the process of calculating the terms of the sequence. Accordingly, the calculation results in this case will not depend on the data bus bitness in a particular computer and the number of significant digits in the representation of fixed-point numbers.

The successes achieved in recent years in understanding the mechanisms of the emergence of chaos in dynamic systems of a wide class have made it possible to take a fresh look at the ways of developing algorithms for the formation of pseudo-random numerical sequences [3]. In particular, the role of the mixing mechanism in ensuring stochastization has become clearer, which is especially important in the case of using algorithms defined on a limited numerical interval. The most well-known of the algorithms for forming the PRS described in the literature are the integer congruent algorithm proposed by the mathematician Lemaire [4] and the family

of algorithms of the Fibonacci type [5]. The congruent algorithm looks like:

$$X_n = (a \cdot X_{n-1} + C) \cdot (\text{mod}M),$$

where X_0, a, C are given integers (in this case, $X_0, a, C < M$, and some large integer is taken as M).

Fibonacci algorithms belong to the class of algorithms with delay. The generalized formula for their linear variant is:

$$X_n = \sum_{i=1}^{Nz} \left[a_i X_{n-i} \hat{A} \left(\sum_{j=1}^{Nz} b_j X_{n-j} \right) \right],$$

where a_i, b_j are equal to 0 or 1, Nz is the delay parameter, $\hat{A}()$ is some operator that takes into account the phase relations between the delay terms. The case $\hat{A}() \equiv 1$ corresponds to the generalized Fibonacci generator:

$$X_n = \sum_{i=1}^{Nz} a_{n-i} \cdot X_{n-i}.$$

The delay parameter Nz determines the number of given (or previously calculated) sequence members that must be stored in the memory of the computing device in order to be able to calculate a new sequence member at each next step of the algorithm, a_{n-i} are coefficients that are usually considered equal to 1 or 0. In some algorithms do not use sums, but the differences of these elements or their products. The classical Fibonacci algorithm takes into account only two terms of the series – the term calculated in the previous step ($n - 1$) of calculations and in step ($n - Nz$).

The dimension of the phase space (PS) of a dynamical system based on a computational algorithm is determined by the number of values that must be specified in order to uniquely determine its state and be able to determine the state of the system at the next computational step. Obviously, the dimension of the PS of a dynamical system described by a delay algorithm is determined by the delay parameter Nz .

Due to the PS limitation of the algorithm specified on the interval of integers, the set of which is significantly less than the entire set of numbers on this interval, the results of the calculation, sooner or later, must necessarily repeat the previously obtained values, that is, the system must go into a cycle. It can be expected that the magnitudes of these cycles depend on the initial conditions.

Thus, the purpose of the work was to identify, using the example of the simplest algorithms, the formation patterns of cycles ("trajectories") of the motion of the representing point of the state of the system in its phase space. We also studied the tendencies of changing these trajectories when changing the interval of numbers on which the algorithm is specified, and (or) changing the delay parameter, that is, the influence of changing the dimension of the phase space. As such an algorithm for the study, the classical Fibonacci-type algorithm was used.

2. ALGORITHM FOR A PSEUDO-RANDOM SEQUENCE GENERATING

To form a pseudo-random sequence, we used the classical Fibonacci-type algorithm based on integers $\{x_n\}$ on a given interval $[1, M]$ with a delay parameter Nz :

$$x_n = x_{n-1} + x_{n-Nz}, \quad (1)$$

which belongs to the class of algorithms whose general form is indicated, for example, in [6].

For this class of algorithms, numerical methods are used to analyze the influence of the main parameters of a generating algorithm with a delay on the statistical, correlation, structural, and fractal characteristics of non-periodic pseudo-random integer and binary sequences generated by the algorithm. It is shown that non-periodic pseudo-random sequences (PRSS) generated by the algorithm with delay, for all values of the main parameters, have good statistical, correlation, structural

and fractal characteristics, close to random sequences of independent trials [7].

Since, when performing algorithm calculation operations for individual pairs of added sequence members, the sum goes beyond the upper limit of the specified interval of the algorithm definition area, this algorithm must be supplemented with a return operator to the specified interval. The value of the sum x_n may be less than or equal to $2M$. The chosen return operator can in principle be different. Depending on this, different mixing modes are implemented, which leads, in particular, to different probability distributions for the sequence of calculated terms. For the algorithm under consideration, two variants of a simple operation of returning to a given interval can be presented. One can conditionally be called the "reflection from the wall" operator. If x_n exceeds the value of the upper limit of the interval M , the newly calculated member is assigned the value $x_n = M - (x_n - M)$, that is, the difference in the excess of the newly obtained number is subtracted from the maximum value of the interval M .

Another return option is "gluing" the ends of the interval. In this case, the excess over the value of M equal to the difference $(x_n - M)$ is added to the value of the lower boundary of the interval, i.e. $x_n = x_n - M$. In this work, this last version of the return operation was used, which provides the mixing mechanism.

To calculate the sequence generated by the algorithm under consideration, it is necessary to set the parameters M and Nz and the memory (delay) array $X_0(Nz) = (x_1, x_2, \dots, x_{Nz})$, that is, a series of initial values consisting of Nz integers from the interval $[1, M]$.

The phase space (PS) of such a system has the dimension M^{Nz} . It consists of a set of points with coordinates from the interval of integers $[1, M]$ that uniquely determine the state of the system. Obviously, due to

the boundedness of this interval $[1, M]$ and, consequently, the finite number of PS points in the process of successive calculation steps, the system will sooner or later necessarily get to the PS point where it has already been. Further, the movement will continue along the previously passed states in the PS. This means that the system is on a periodic "trajectory" with a repetition period equal to the corresponding number of counting steps (T_{M, N_z}). In principle, it can be assumed, and the results of calculations confirm this, that the dynamic system under study, depending on the values of the parameters M , N_z and the initial conditions in the phase transition, can have many cycles of different periods. The cycle with the largest period for the data M and N_z is called the maximum period cycle.

3. RESULTS OF CALCULATIONS AND REGULARITIES REVEALED

The study of the periods of sequences formed by algorithm (1) depending on the parameters M , N_z and initial values was carried out starting from small values of the parameters, although it is obvious that these cases are not important for practical applications. Nevertheless, this analysis can help in revealing patterns of sequence formation at arbitrary values of M and N_z .

Transitions from one state point to another will be depicted for the convenience of visualizing the sequence of system motion in the form of a phase trajectory (PT), although, strictly speaking, the system does not describe a trajectory in phase space, but jumps from one state to another. When the radius-vector $R_n \{x_{n-1}, \dots, x_{n-N_z}\}$ of the system state in the PS at the n -th step of the algorithm hits the state point in which it was already at the $(n-N)$ -th step, i.e. $R_n = R_{n-N}$, then at further steps of the calculation, due to the determinism of the algorithm, the vector will exactly repeat the previously passed movement. This means that

the system in the PS entered a cycle with a period $T = (N - 1)$ equal to the number of state points on the cycle. This periodic movement of the system in the PS will correspond to a periodic sequence of numbers generated by the algorithm.

Among all possible cycles in the PS of the system for given values of the parameters M and N_z , we single out the cycle with the largest period and call it the cycle of the maximum period $T_{M, N_z, \max}$. From the point of view of the formation of the PRS of the longest possible period, the configuration of the corresponding PT in the PS should have signs of a certain "chaotic" and not have spatial periodicity at the largest possible number of algorithm steps.

Computational experiments performed for different values of the parameters M , N_z and initial values have shown that the spatial configuration of the PT depends in a complicated way on the specified parameters. However, there are some general patterns:

1. The PS of the algorithm is a set of cycles of different periods. Each point of the system state belongs to only one cycle, which is a consequence of the rigid determinism and uniqueness of the algorithm. In other words, the cycles do not have common PS points.
2. For any M and N_z , one trajectory is singular and represents an isolated point, all of whose coordinates are equal to M , which follows from the formula of algorithm (1) and the return operation.
3. As a rule, for given M and N_z , there are several different cycles of the same period in the PS. They differ in the set of PS points. The number of cycles with the same period we denote by ν : $T(\nu)_{M, N_z}$. For example, $T(2)_{8, 2} = 6$ means that when $M = 8$, $N_z = 2$, there are 2 cycles in the PS with a period equal to 6.

4. The point with coordinates $(1, 1, \dots, 1)$ for any M and N_z always lies on the cycle with the maximum period. The points $(1, 1, \dots, 1)$ and (M, M, \dots, M) are singular throughout the PS.
5. The sum of state points in the PS of a dynamic system in all cycles is always equal to the total volume of the PS – $V = M^{N_z}$.

PT with PS dimension > 3 can be studied only by their projections onto the corresponding planes, and the complexity of their spatial configuration can be judged from a simpler two-dimensional case for $M = 8$, shown in Fig. 1.

For the convenience of visualization, the PS points depicting states sequentially passed by the system are connected by lines – "trajectories" of the system's motion. The trajectories shown in the figure correspond to cycles with a period: $T(4)_{8,2} = 12$ (B), $T(2)_{8,2} = 6$ (D), $T(1)_{8,2} = 3$ (F).

This Figure shows 4 out of 8 existing trajectories $T(1)_{8,2} = 1$, $T(1)_{8,2} = 3$, $T(2)_{8,2} = 6$ and $T(4)_{8,2} = 12$. Required note that the volume of the PS in this case is equal to 64 and

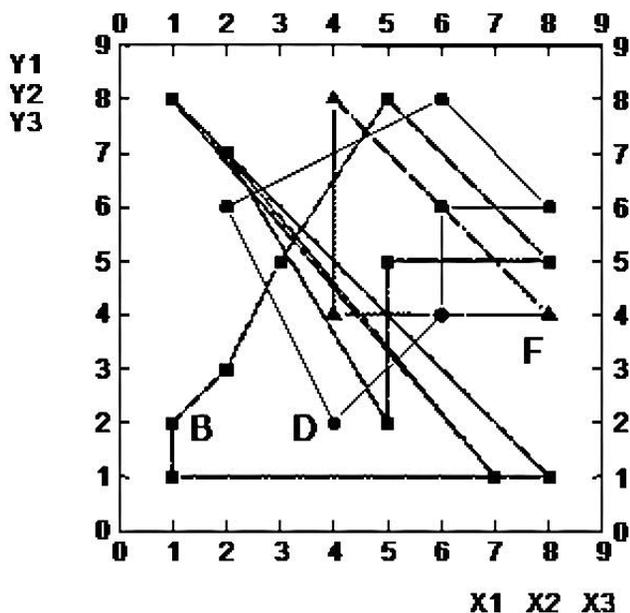


Fig. 1. Image on the plane of successive states traversed by a dynamical system in a 2-dimensional phase space for the algorithm parameters $M = 8$ and $N_z = 2$.

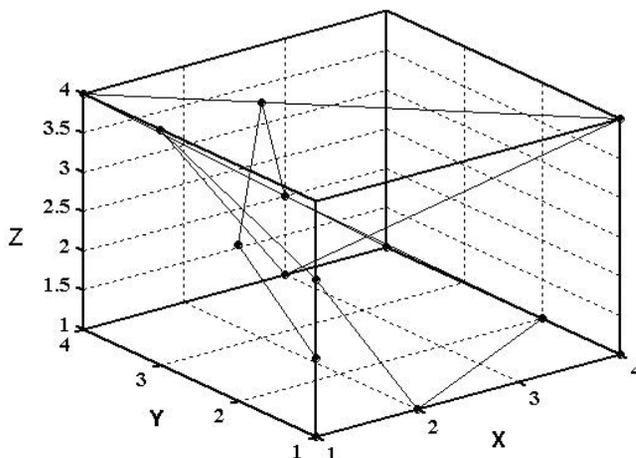


Fig. 2. Representation in 3D-space of the "trajectory" of successive states of the system for $M = 4$ and $N_z = 3$, corresponding to the cycle of the maximum period.

it coincides with the points sum over all PTs $(1 + 3 + 2 \cdot 6 + 4 \cdot 12 = 64)$. Fig. 2 shows the maximum length PT for $M = 4$ in 3D space, while there are 3 more PTs of the same length in the PT, one PT having 7 points, and one PT consisting of one point with coordinates $(4, 4, 4)$.

In all cases, when determining the period T_{M,N_z} , the following procedure was performed: the generation of sequence members continued until the vector of initial conditions (IC) was repeated. The dependence on the IC was studied by sort through of all possible IC vectors for given M and N_z . For clarity, the results of such studies for the case $N_z = 2$ were presented in the form of a matrix, the elements of which are the periods of the corresponding cycles. The numbers of rows and columns of this matrix for each of its elements determined the corresponding initial conditions. For the cases $M = 5, N_z = 2$ and $M = 6, N_z = 2$, these matrices have the form:

				24	24	24	24	24	24
20	20	4	20	20	24	8	24	8	24
4	20	20	20	20	24	24	3	24	24
20	20	20	4	20	24	8	24	8	24
20	4	20	20	20	24	24	24	24	24
20	20	20	20	1	24	8	3	8	24

Naturally, when passing to $N_z > 2$, the corresponding matrix becomes

multidimensional and its analysis and representation is a difficult task. Note that for evens M , the period matrix is symmetric with respect to the main diagonal (except for the cases when odd factors are present in the decomposition of M), and for any M , almost all periods lie on the main diagonal. In cases where M is not a prime number, but has factors, the period matrices (and hence the PT) "inherit" the features of the matrices of their factors, and the PT reproduce the PT factors on a larger scale and with the same period.

Similar calculations were performed up to the value of the PS dimensions $N_{\mathcal{Z}} = 6$, and the conclusions drawn for the two-dimensional case regarding the periods of the sequences formed by such algorithms turned out to be correct. The futility of using even values of M from the point of view of achieving the cycles of the largest period (this statement is the more rigid, the larger the dimension of the PS) is clearly demonstrated by the graph shown in **Fig. 3**, which shows the dependence of the maximum period $T_{M, N_{\mathcal{Z}}}$ from M (where M varies from 50 to 97) with a PS dimension of 6.

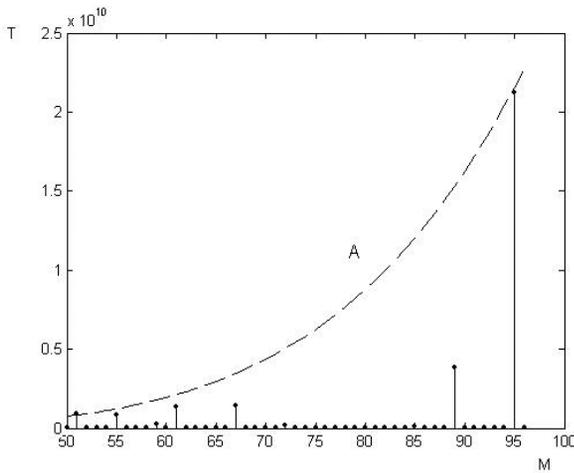


Fig. 3. Dependence of the maximum period $T_{M, N_{\mathcal{Z}}}$ on M for the sequence generated by the algorithm with the PS dimension $N_{\mathcal{Z}} = 6$ (where M varies from 50 to 80). The dotted curve corresponds to the power function $M^{(N_{\mathcal{Z}}-1)+\delta}$ approximating the maximum values of the periods, where in this case $\delta = 0.21$.

The complex nature of the dependence obtained indicates the nontriviality of the regularities in the formation of the PRS, despite the obvious simplicity of the algorithm (1). On Fig. 3 also shows the approximation of the maximum period value by the power function $M^{(N_{\mathcal{Z}}-1)+\delta}$ (in this case, $\delta = 0.21$, and for other parameters, as a rule, it is $\approx 0.1-0.25$), i.e. the maximum period increases with increasing M , but not monotonically. The graph also shows that not all prime numbers provide the achievement of the maximum period, therefore, this parameter is not a sufficient condition for this.

It should be noted that from the point of view of the formation of sequences with a maximum period, the presence of a large number of cycles in the PS is extremely undesirable, since this reduces the proportion of the volume of the PS that can be occupied by cycles with the maximum period. Having written in general form the first few members of the sequence of numbers formed using algorithm (1), for the simplest case $N_{\mathcal{Z}} = 2$ with initial conditions (X_1, X_2) , it is enough to simply make sure that at the n step of forming the sequence we will have, taking into account the return rule

$$X_n = f_n X_1 + f_{n+1} X_2 - \mu M, \tag{2}$$

where f_n and f_{n+1} are the well-known Fibonacci numbers, μ is an integer.

A similar procedure can be performed for $N_{\mathcal{Z}} > 2$, and for X_n we get an expression similar to (2)

$$X_n = f_n X_1 + f_{n-1} X_2 + \dots + f_{n+1} X_{N_{\mathcal{Z}}} - \mu M, \tag{3}$$

only in this case f_n are no longer classical Fibonacci numbers, as in the case of $N_{\mathcal{Z}} = 2$, but generalized Fibonacci numbers. The firsts $N_{\mathcal{Z}}$ of such generalized Fibonacci numbers are simply 1. Then the recurrence relation by which subsequent Fibonacci numbers are calculated is as follows

$$f_n = f_{n-1} + f_{n-N_z}, \quad (4)$$

and, for example, with $N_z = 4$, the first 10 generalized Fibonacci numbers are: 1, 1, 1, 1, 2, 3, 4, 5, 7, 10.

If the period of the generated sequence is equal to n , then starting from $n + 1$ steps, the IC will begin to repeat, which makes it possible to write down N_z equalities of the form

$$\begin{aligned} f_{n+1}X_1 + f_n X_2 + \dots + f_{n+2}X_{N_z} - \mu_1 M &= X_1, \\ f_{n+2}X_1 + f_{n+1}X_2 + \dots + f_{n+3}X_{N_z} - \mu_2 M &= X_2, \\ \dots \\ f_{n+N_z}X_1 + f_{n+(N_z-1)}X_2 + \dots + f_{n+(N_z+1)}X_{N_z} - \mu_{N_z} M &= X_{N_z}, \end{aligned} \quad (5)$$

where f_n are generalized Fibonacci numbers, the specific form of which is determined by N_z and (4). By introducing the matrix F_{N_z} , whose elements are obvious, equations (5) can be written in the matrix form

$$R_{N_z} F_{N_z} - \mu_{N_z} M = R_{N_z},$$

where R_{N_z} is the initial state vector determined by the initial conditions, and μ_{N_z} is a vector whose components are integers.

Let us rewrite (5) in the form

$$\begin{aligned} (f_{n+1} - 1)X_1 + f_n X_2 + \dots + f_{n+2}X_{N_z} &= \mu_1 M, \\ f_{n+2}X_1 + (f_{n+1} - 1)X_2 + \dots + f_{n+3}X_{N_z} &= \mu_2 M, \\ \dots \\ f_{n+N_z}X_1 + f_{n+(N_z-1)}X_2 + \dots + (f_{n+(N_z+1)} - 1)X_{N_z} &= \mu_{N_z} M. \end{aligned} \quad (6)$$

Considering that all μ_n are integers, system (6) can be written as equations

$$\begin{aligned} [(f_{n+1} - 1)X_1 + f_n X_2 + \dots + f_{n+2}X_{N_z}] \pmod{M} &= 0, \\ [f_{n+2}X_1 + (f_{n+1} - 1)X_2 + \dots + f_{n+3}X_{N_z}] \pmod{M} &= 0, \\ \dots \\ [f_{n+N_z}X_1 + f_{n+(N_z-1)}X_2 + \dots + (f_{n+(N_z+1)} - 1)X_{N_z}] \pmod{M} &= 0. \end{aligned} \quad (7)$$

In these equations, all parameters are known (M , N_z , and IC are given) and, therefore, it remains only to find the step n , which ends the period of the sequence, and from $n + 1$ steps, its repetition begins. Since we are interested in the maximum period, and it is achieved if the entire set of initial conditions is equal to 1, and then equations (7) are further simplified

$$\begin{aligned} [(f_{n+1} - 1) + f_n + \dots + f_{n+2}] \pmod{M} &= 0, \\ [f_{n+2} + (f_{n+1} - 1) + \dots + f_{n+3}] \pmod{M} &= 0, \\ \dots \\ [f_{n+N_z} + f_{n+(N_z-1)} + \dots + (f_{n+(N_z+1)} - 1)] \pmod{M} &= 0. \end{aligned} \quad (8)$$

Taking into account the recurrence relations (4), (8), we can write

$$\begin{aligned} (f_{n+(N_z+1)} - 1) \pmod{M} &= 0, \\ (f_{n+(N_z+2)} - 1) \pmod{M} &= 0, \\ \dots \\ (f_{n+(N_z+N_z)} - 1) \pmod{M} &= 0. \end{aligned} \quad (9)$$

From (9) it follows that in order to find the period of the formed sequence, it is necessary to determine N_z consecutive generalized Fibonacci numbers, which, after subtracting one from them, are divisible by M without a remainder. It should be noted that the conditions for entering the period are non trivial, which determines the rather complex dependence of T_{M,N_z} on M (see Fig. 3).

Conditions (9) can be reduced to one condition by addition the left and right parts, taking into account the recurrence relation (4):

$$(f_{n+(3N_z-1)} - N_z) \pmod{M} = 0. \quad (10)$$

Thus, to calculate the period of the sequence, it is necessary to determine the number of the generalized Fibonacci number, which, after subtracting N_z from it, is divisible by M without remainder.

The method for determining the period of sequences generated by the algorithm under consideration using conditions (9), gives some gain in computation time compared to the direct method of direct sequence generation and tracking the beginning of the repetition of the IC. It is this technique that was used to study the dependence of T_{M,N_z} on the PS dimension N_z . The results obtained are shown in Fig. 4. The calculations were performed for $M = 11, 15, 19$, the parameter N_z varied from 2 to 10, and IC was taken equal to 1. M was

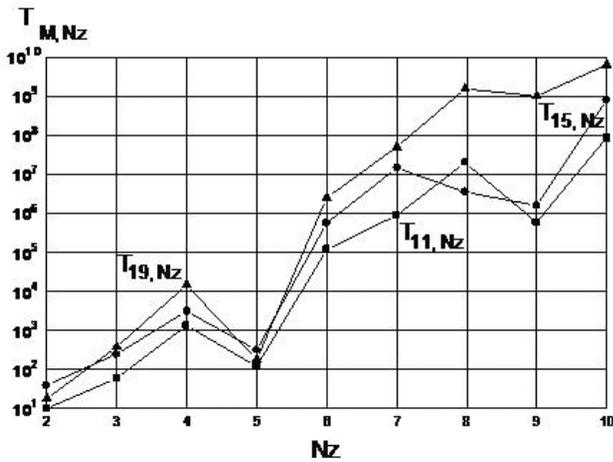


Fig. 4. Dependence of the cycle of the maximum period T_{M, N_z} on the delay parameters N_z for $M = 11, 15, 19$.

deliberately chosen to be odd, and two of them are prime numbers.

Curves in Fig. 4 show an increase in the period with an increase in N_z , but this increase is far from monotonic, and at $N_z = 5$ a clear dip is revealed. It is clear that the choice of N_z should be approached rather carefully, remembering that the dependence of the period on M is much more significant (see Fig. 3).

When forming the PRS, the question of the uniformity of the distribution of its numbers over a given interval is far from the last one $[1, M]$. It should be noted that from this point of view, algorithm (1) is not ideal. Nevertheless, it is possible to obtain a distribution of generated numbers close to uniform on the interval $[1, M]$. Obviously, a necessary condition for this is

$$T_{M, N_z} \pmod{M} = 0. \tag{11}$$

Such cases, in particular, are realized quite often, and if the required duration of the PRS is less than the corresponding period T_{M, N_z} , then they can be used in practice. It is this case that is realized when generating a PRS with the number of numbers 109 from the interval $[1, M]$, where $M = 257$ for $N_z = 16$. The corresponding distribution of the frequencies of the generated numbers in the specified

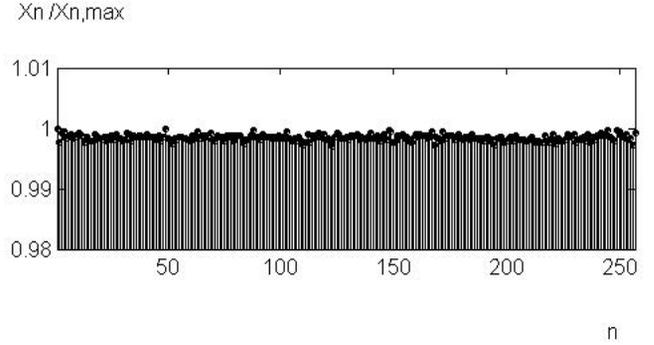


Fig. 5. Probability distribution of the frequency of occurrence of integers from the interval $[1, 257]$ in a pseudo-random sequence generated by the Fibonacci algorithm with a delay parameter $N_z = 16$. The frequencies of occurrence of numbers X_n are normalized to the maximum value in the distribution.

sequence X_n normalized to the maximum value in this distribution ($X_{n, \max}$) shown in Fig. 5. Here the maximum frequency $X_{n, \max} = 3896697$ corresponds to the value $n = 49$.

Pearson's χ^2 test allows us to evaluate the statistical significance of the differences between the two distributions. The χ^2 criterion is the most commonly used criterion for testing the hypothesis that the observed sample belongs to some theoretical distribution law. Performed using the χ^2 criterion, difference checking from the uniform distribution shown in Fig. 5, for a sequence of 109 numbers obtained by generating according to algorithm (1) for the interval $[1, 257]$ with $N_z = 16$, showed that for four degrees of freedom ($d = 4$) the probability $Pd(\chi^2 \geq \chi_0^2)$ of obtaining the value $\chi^2 \geq \chi_0^2$, for calculated by distribution value $\chi_0^2 = 0.239$, is $\sim 87\%$. This result indicates that the hypothesis of a uniform distribution of the sequence of numbers generated by the algorithm on the interval $[1, 257]$ can be accepted with a high degree of probability.

4. CONCLUSION

The paper studies the patterns of formation of pseudo-random sequences of numbers by a Fibonacci-type delay algorithm defined on a limited interval of integers of the natural series, using an operation that provides chaotic

mixing and returning to a given numerical interval.

The studies were carried out with a change in a wide range of values of the delay parameter Nz , which determines the dimension of the phase space of the algorithm, and the value of the numerical interval M of the domain of the algorithm.

For the delay parameter $Nz = 2$, period matrices of sequences generated by the algorithm are determined for the entire set of initial conditions.

A number of regularities are defined that characterize the sequences of numbers generated by the Fibonacci-type algorithm for small values of the delay parameter

Relationships are obtained that allow reducing the search time for the maximum period of a pseudo-random sequence generated by the Fibonacci algorithm.

These patterns can be used in developing methods for studying algorithms of this type at a large phase space dimension.

REFERENCES

1. "Numerical Methods and Software" David Kahaner, Cleve Moler and Stephen Nash. Prentice-Hall, Englewood Cliffs, NJ, 1989, 496 pp. DOI: 10.1137/1033033.
2. Knuth Donald E. *The Art of Computer Programming. Vol. 2. Seminumerical Algorithms*. USA, Addison-Wesley Publ., 704 pp.
3. Kuznetsov SP. *Dynamic chaos* (course of lectures). Moscow, Fizmatlit Publ., 2001, p. 295.
4. Lehmer DN. On the enumeration of magic cubes. *Bull. Amer. Math. Soc.*, 1934, 40:833-837. DOI: 10.1090/S0002-9904-1934-05976-7.
5. Hayes B. Computing Science: The Fibonacci Numbers. *American Scientist*, 1999, 87(4):296-301.
6. Nikita A. Ageykin, Vladimir I. Grachev, Viktor I. Ryabenkov, Vladimir V. Kolesov. Information Technologies Based on Noise-like Signals: I. Discrete Chaotic Algorithms. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2022, 14(1):47-64. DOI: 10.17725/rensit.2022.14.047.

7. Vadim V. Kashin, Vladimir I. Grachev, Viktor I. Ryabenkov, Vladimir V. Kolesov. Information Technologies Based on Noise-like Signals: II. Statistical and Fractal Properties of Chaotic Algorithms. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2022, 14(2):151-164. DOI: 10.17725/rensit.2022.14.151.

DOI: 10.17725/rensit.2022.14.279

Problems of Message Routing Algorithms in Streaming Data Processing Systems

George G. Bulychev, Alexey V. Chernykh

MIREA-Russian Technological University, <https://www.mirea.ru/>

Moscow 119454, Russian Federation

E-mail: geo-bulychev@mail.ru, meidm@yandex.ru

Received May 26, 2022, peer-reviewed June 6, 2022, accepted June 13, 2022

Abstract: Existing problems of routing algorithms, such as high overhead costs for scaling components and low fault tolerance of the system, are considered. An algorithm is proposed that eliminates the shortcomings of the existing ones due to the rejection of a centralized router and the strict separation of pipelines within the system. The algorithm is based on adding a special token to the message. For the proposed algorithm, the performance improvement is mathematically proved, and the practical load on the system is simulated to confirm the obtained theoretical results.

Keywords: streaming data processing, message routing, big data

UDC 004.62

For citation: George G. Bulychev, Alexey V. Chernykh. Problems of Message Routing Algorithms in Streaming Data Processing Systems. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2022, 14(3):279-290. DOI: 10.17725/rensit.2022.14.279.

CONTENTS

1. INTRODUCTION (279)
 2. ANALYSIS OF EXISTING ALGORITHMS (280)
 - 2.1. DIRECT JOIN ALGORITHM (280)
 - 2.2. ROUTER ALGORITHM (280)
 - 2.3. ALGORITHM COMPARISON (281)
 3. ALGORITHM FOR ROUTING MESSAGES USING TOKENS (284)
 4. SIMULATION ALGORITHMS (286)
 - 4.1. SIMULATION OF AN ALGORITHM WITH DIRECT CONNECTION (287)
 - 4.2. SIMULATION AN ALGORITHM WITH A ROUTER (287)
 - 4.3. SIMULATION AN ALGORITHM WITH MESSAGE TAGGING (288)
 - 4.4. SIMULATION FINDINGS (289)
 5. CONCLUSION (289)
- REFERENCES (289)

1. INTRODUCTION

The amount of data in the modern world is calculated in zettabytes [1] and every day their number is rapidly increasing. Such a volume of information opens up new and previously unseen opportunities for humanity to share knowledge and experience, statistical and marketing research, neural network training, early detection of threats, and much more. But in addition to opportunities, such volumes pose no less difficult tasks and challenges for researchers in implementing the mechanisms for collecting, processing and analyzing data.

The main approach currently used for big data analysis is the collection and aggregation of data into various storage systems with subsequent analysis. With this approach, the amount of resources consumed is proportional to the amount of data being processed. An alternative to this

approach is streaming data processing in real time, when a data processing pipeline is built and only the data that is directly needed to solve the task is stored. The most commonly used systems are ETL (Extract, Transform, Load) [2,3]. This class of systems consists of three categories of components: Extract – receiving data from external systems, Transform – data transformation, Load – transfer of processed data to third-party systems. Such systems are implemented in the form of pipelines, when data is transferred along the chain from one component to another.

For efficient data processing in real time, it is necessary to adapt data processing algorithms to work in the absence of storage facilities. It is necessary to develop universal algorithms [4] for solving a different range of problems, otherwise, for each specific study, it will be necessary to develop the required algorithms in the stream processing paradigm.

Separate requirements apply to the system itself. Each stage of processing can be performed on different equipment (including geo-distributed) to improve the efficiency of individual algorithms or data processing features related to the laws of various countries or the conditions for using information. The system should be centrally managed, with the ability to dynamically change the configuration of each individual component, and much more.

To solve all the problems described above, a streaming data processing system based on universal blocks is being developed. Such a system will allow implementing data processing scenarios in the shortest possible time by collecting chains of blocks through the user interface [5] and automatically launching them on any equipment.

One of the key aspects of the effectiveness of such systems is the message routing

algorithm within the system. If the message processing path is not optimal, there may be unnecessary overhead for message passing, scaling individual components, or system fault tolerance. At the moment, the routing algorithms used in streaming data processing systems have significant drawbacks in terms of data processing efficiency or system fault tolerance.

2. ANALYSIS OF EXISTING ALGORITHMS

In data processing systems, only two algorithms are mainly used: with direct connection of components [6] and through a router [7].

2.1. DIRECT JOIN ALGORITHM

The direct connection algorithm creates two independent pipelines, within each of them all components interact directly with each other (**Fig. 1**).

If scaling is necessary, the system is able to create additional instances of a particular component within the same pipeline (**Fig. 2**). At the same time, since the pipelines are isolated, the scaling of the components within each pipeline will occur independently.

2.2. ROUTER ALGORITHM

The router algorithm (**Fig. 3**), in turn, has one centralized component responsible for routing messages. Each component, except for the initiators, receives messages for processing

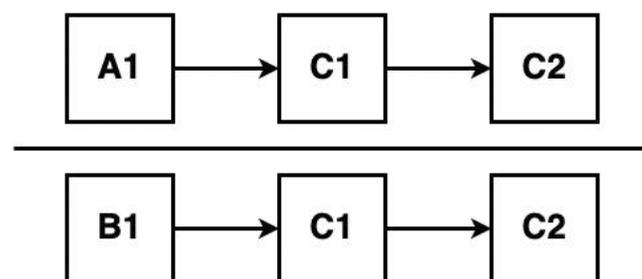


Fig. 1. System architecture with direct connection algorithm.

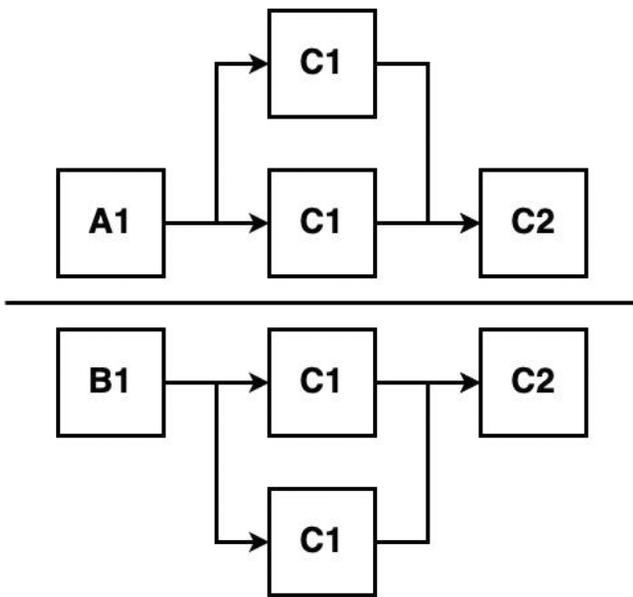


Fig. 2. Component scaling in a system with direct component connection algorithm.

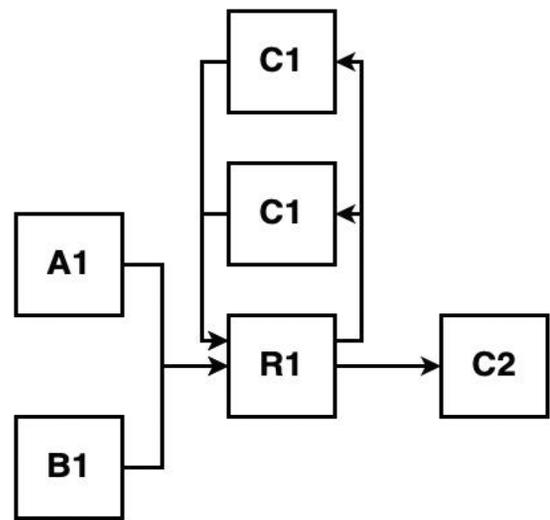


Fig. 4. Component scaling in a system with a message routing algorithm through a router.

from the central component, and after processing sends messages back to the router.

Since the essence of the pipeline becomes "virtual" in the algorithm, the components can be reused. In both pipelines, similar components perform the same task and have the same code base; components can process a message from one pipeline at one point in time, and from another at another point in time. The router, in turn, determines whether the message belongs to one of the pipelines and determines the further route of the

message. Due to this property, the scaling of such components will depend on the total load from all pipelines (Fig. 4).

2.3. ALGORITHM COMPARISON

Algorithms are compared according to six key parameters:

1. Number of message transfer operations between components.
2. The number of running component instances unique to each pipeline in the system.
3. The number of running instances of components reused in different pipelines in the system.
4. The number of operations for creating or deleting component instances during scaling.
5. Having a single point of failure.
6. Difficulty making changes to the pipeline.

Number of message transfer operations between components

In a direct connection algorithm, each message is passed directly from component to component, with no intermediate steps.

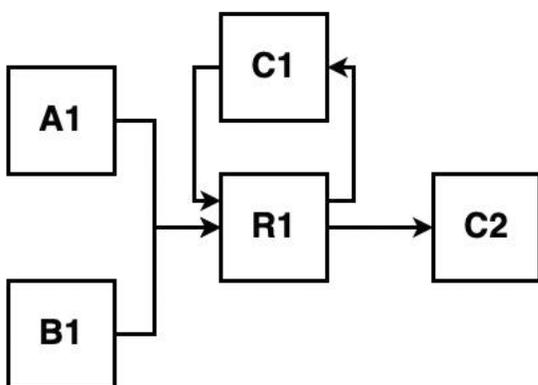


Fig. 3. System architecture with an algorithm for routing messages through a router.

Therefore, the number of message transfer operations between components will be as follows.

$$C = n - 1, \quad (1)$$

where C – number of message transfer operations; n – number of components inside the conveyor.

In the algorithm with a router, the efficiency of using transmission channels between components is reduced, since after processing by a component, the message is sent to the router, and only after the router – to the next component. Therefore, the number of message transfer operations between components will be as follows.

$$C = 2(n - 1), \quad (2)$$

where C – number of message transfer operations; n – number of components inside the conveyor.

In the algorithm with a direct connection, the number of message transfer operations between components is two times less than in the algorithm with a router. Due to this, the overhead costs for the transmission of messages within the system are reduced.

Number of running instances of components unique to each pipeline in the system

Some components are unique for each pipeline, and accordingly, these components will be scaled only considering the load on a specific pipeline, regardless of the selected algorithm. If the system has N unique components, then the total number of all instances is expressed by the formula (3):

$$C_1 = \sum_{i=1}^N \left[\frac{l_i}{p_i} \right], \quad (3)$$

where C_1 – number of unique component instances; N – number of unique components; l_i – load on the pipeline within which the component works; p_i – performance of a single component instance.

The number of running instances unique for each pipeline is the same in both algorithms.

Number of running instances of components reused in different pipelines in the system

In a direct connection algorithm, each pipeline is completely isolated, and the number of running instances of any of the components is determined by the load on a particular pipeline. Since reusable components are used in two pipelines, the total number of instances of a given component type is determined by the sum of the number of instances in each individual pipeline. If the system has Z pipelines and N components, then the total number of all instances is expressed by formula (4):

$$C_2 = \sum_{i=1}^N \sum_{j=1}^Z \left[\frac{l_{ij}}{p_i} \right], \quad (4)$$

where C_2 – number of instances of reusable components; N – number of reusable components; Z – number of pipelines in the system; l_{ij} – load on a component in a pipeline; p_i – performance of a single component instance.

In the algorithm with a router, the number of running instances of reused components depends only on the total load on the system. Then the number of copies is expressed by the formula (5):

$$\tilde{N}_3 = \sum_{i=1}^N \left[\sum_{j=1}^Z \frac{l_{ij}}{p_i} \right], \quad (5)$$

where C_3 – number of instances of reusable components; N – number of reusable components; Z – number of pipelines in the system; l_{ij} – load on a component in a pipeline; p_i – performance of a single component instance.

Since the sum of the ceilings is always greater than or equal to the ceiling of the sum, the following inequality is formed (6):

$$\sum_{i=1}^N \sum_{j=1}^Z \left[\frac{l_{ij}}{p_i} \right] \geq \sum_{i=1}^N \left[\sum_{j=1}^Z \frac{l_{ij}}{p_i} \right], \quad (6)$$

where N – number of reusable components; Z – number of pipelines in the system; l_{ij} – load on a component in a pipeline; p_i – performance of a single component instance.

The number of running instances of components reused in different pipelines in the router algorithm is less than or equal to that in the direct connection algorithm. Due to this, the efficiency of resource use in the algorithm with a router is higher. However, in the algorithm with a router, there is also an additional component - the router.

Number of operations to create or remove component instances when scaling them

To compare the two algorithms, consider the time interval on which the function of changing the load on conveyors monotonically increases or decreases. Since the number of unique component instances is the same in both algorithms, only reusable components will be considered.

The number of operations for creating or deleting component instances in a direct routed algorithm will be (7):

$$C_4 = \sum_{i=1}^N \sum_{j=1}^Z \left[\frac{|lend_{ij} - lstart_{ij}|}{p_i} \right], \quad (7)$$

where C_4 – number of instance creation or deletion operations; N – number of reusable components; Z – number of pipelines in the system; $lend_{ij}$ (load end_{ij}) – load on the component in the pipeline at the end of the time interval; $lstart_{ij}$ (load start_{ij}) – load on the component in the pipeline at the beginning of the time interval; p_i – performance of a single component instance.

The number of operations for creating or deleting component instances in the algorithm with a router will be (8):

$$C_5 = \sum_{i=1}^N \left[\sum_{j=1}^Z \frac{|lend_{ij} - lstart_{ij}|}{p_i} \right], \quad (8)$$

where C_5 – number of instance creation or deletion operations; N – number of reusable components; Z – number of pipelines in the system; $lend_{ij}$ (load end_{ij}) – load on the component in the pipeline at the end of the time interval; $lstart_{ij}$ (load start_{ij}) – load on the component in the pipeline at the beginning of the time interval; p_i – performance of a single component instance.

Since the sum of the ceilings is always greater than or equal to the ceiling of the sum, the following inequality is formed (9):

$$\sum_{i=1}^N \sum_{j=1}^Z \left[\frac{|lend_{ij} - lstart_{ij}|}{p_i} \right] \geq \sum_{i=1}^N \left[\sum_{j=1}^Z \frac{|lend_{ij} - lstart_{ij}|}{p_i} \right], \quad (9)$$

where N – number of reusable components; Z – number of pipelines in the system; $lend_{ij}$ (load end_{ij}) – load on the component in the pipeline at the end of the time interval; $lstart_{ij}$ (load start_{ij}) – load on the component in the pipeline at the beginning of the time interval; p_i – performance of a single component instance.

The number of operations for creating or deleting component instances during scaling in the routing algorithm is less than or equal to the number of operations in the direct connection algorithm. This results in lower component scaling overhead.

Having a single point of failure

In a direct connection algorithm, due to the independence of pipelines, there is no single point of failure for the entire system.

In the router algorithm, there is a central component responsible for routing packets throughout the system. If this component fails, the entire system will fail.

Difficulty making changes to the pipeline

In a direct connection algorithm, the pipeline is formed at the initialization stage. To make the change, you must completely reinitialize the pipeline, which will incur additional overhead.

In the router algorithm, the pipeline is formed based on the routing table, which is stored in the central component. To make changes to the pipeline, it is enough to change the routing table.

Conclusions from the comparison of algorithms

The direct connection algorithm has a high level of fault tolerance and resource efficiency for processing a single message, since the message is transmitted along the optimal route. However, the resource efficiency of the entire system is reduced by the inability to reuse components in multiple pipelines.

In turn, the algorithm with a router uses the resources of the entire system more efficiently, due to the reuse of components, however, fault tolerance is significantly reduced, and additional overhead costs for message transmission between components arise.

Algorithms have their pros and cons, however, to improve the efficiency of resource use, a new algorithm is needed that combines the key advantages of the presented algorithms and levels their shortcomings.

3. ALGORITHM FOR ROUTING MESSAGES USING TOKENS

It's necessary to achieve the transfer of messages along the optimal route, while maintaining the possibility of reusing components in different pipelines.

To solve this problem, it is proposed to abandon the use of a central router in favor of additional modules integrated into components. The message marking module is integrated into the first component. The marker indicates that the message belongs to a particular pipeline. A routing module is integrated into the processing components, which is responsible for determining the

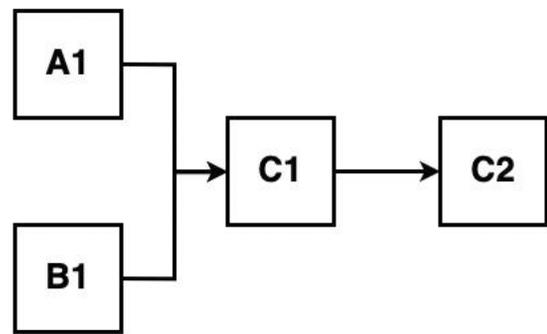


Fig. 5. System architecture with message routing algorithm using tokens.

further route of the message based on the token (Fig. 5).

When the load on the entire system increases, individual components are scaled up (Fig. 6). When the load drops – the number of instances of individual components will decrease.

Thanks to this architecture, the efficiency of using system resources will be approximately the same as in the case of the router algorithm, since the number of component instances will depend on the load on the entire system due to the reuse of components. However, due to the additional module, the performance of the components will be reduced. The number of component instances will be expressed by the formula (10):

$$C_6 = \sum_{i=1}^N \left[\sum_{j=1}^Z \frac{l_{ij}}{p_i - pr} \right], \quad (10)$$

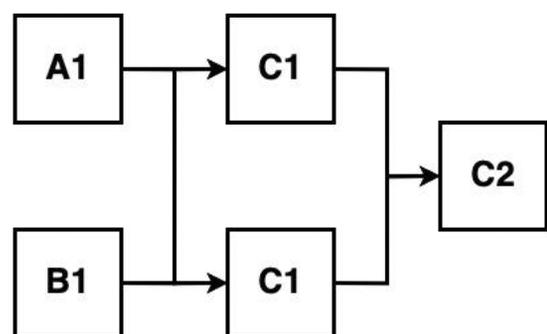


Fig. 6. Scaling of "C₁" in a system with message routing algorithm using markers.

where C_3 – number of instances of reusable components; N – number of reusable components; Z – number of pipelines in the system; L_{ij} – load on a component in a pipeline; p_i – performance of a single component instance; p_r – performance of the routing module in the component.

The resulting difference in the number of components will directly depend on the actual implementation of the routing module.

The efficiency of using message passing channels between components will be similar to the direct connection algorithm, since messages are transferred directly from component to component within the pipeline route. Since the routing module is integrated directly into each component, the cost of sending a message to the module will be significantly lower than sending it over a network link to the central routing component. Therefore, the number of message transfer operations between components will be as follows.

$$C = n - 1, \tag{11}$$

where C – number of message transfer operations; n – number of components inside the conveyor.

Due to the absence of a central component, the fault tolerance of the system will be similar to the algorithm with a direct connection of components.

However, a way to route the message based on the token needs to be developed.

Similar algorithms are used in network packet routing systems through the use of a routing table [8]. When applying such an algorithm to a streaming data processing system, it is necessary to keep the routing table in the router of each component up to date. Then the router can use the message token to find information in the routing table about which pipeline the given message

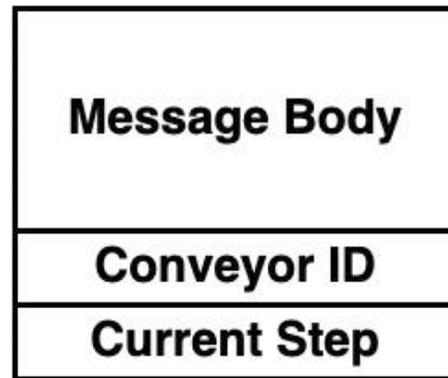


Fig. 7. *Message structure with marker.*

belongs to and the current position of the message relative to the route of the pipeline. Depending on the route, the message is sent to the next component (**Fig. 7**).

The main disadvantage of this approach is the need to keep the routing tables in each component up to date.

An alternative is to consider special markers, inside which the entire message route will be located (**Fig. 8**). When the router inside the component receives messages, it retrieves the route and the current position of the message relative to the route from the marker. Depending on the route, the message is passed to the next component.

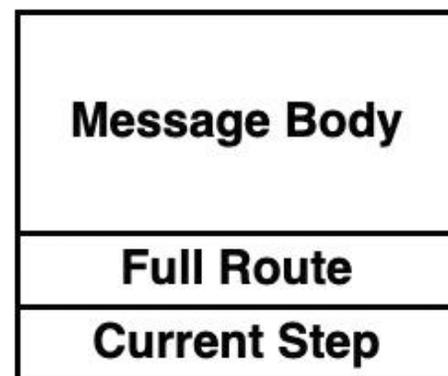


Fig. 8. *Message structure with route.*

Such message routing algorithms within a streaming data processing system require the introduction of an additional module in the first component of each pipeline, which is responsible for marking the message with a marker. The main disadvantage of the algorithm with a marker will be leveled, since it is necessary to synchronize routes with only one component inside each pipeline, and not with all.

4. SIMULATION ALGORITHMS

Consider a system consisting of two conveyors. The Alpha pipeline (Fig. 9) consists of three components: A1 – receiving messages from the “AlphaNet” social network, C1 – evaluating the emotional coloring of the message, C2 – saving messages to the database.

The Beta pipeline (Fig. 10) consists of three components: B1 – receiving messages from the “BetaNet” social network, C1 – evaluating the emotional coloring of the message, and C2 – saving messages to the database.

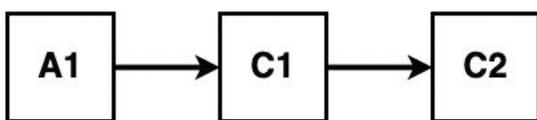


Fig. 9. Alpha pipeline.

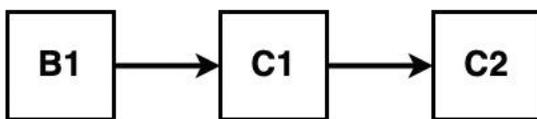


Fig. 10. Beta pipeline.

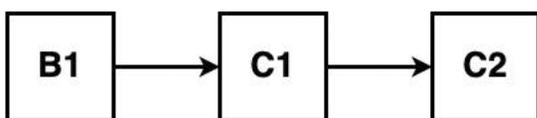
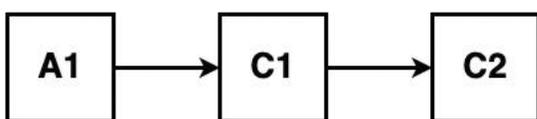


Fig. 11. Simulation system architecture with direct connection algorithm.

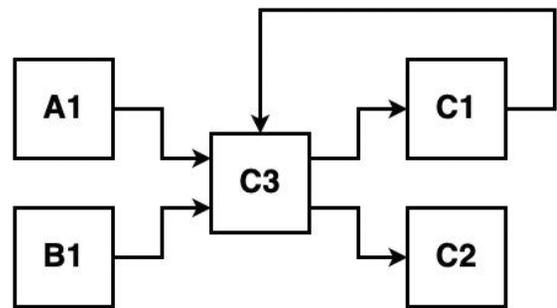


Fig. 12. Simulation system architecture with an algorithm for routing messages through a router.

Components A1 and B1 are unique for each conveyor. Components C1 and C2 are the same in the pipelines.

If you use the algorithm with direct connection of components, you will get the following system (Fig. 11).

When using the algorithm with the router, the following system is formed (Fig. 12). Also a new component C3 – message router – is added to the system.

If you use the routing algorithm with message labeling, you get the following system (Fig. 13).

The performance of each component is shown in Table 1.

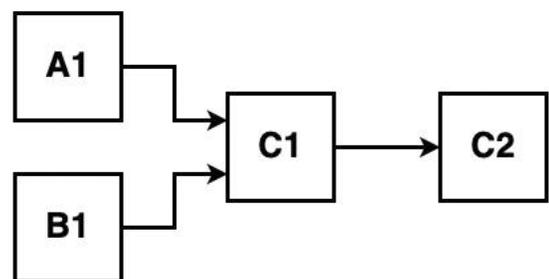


Fig. 13. Simulation system architecture with message routing algorithm using tokens.

Table 1

Component performance.	
Component	Performance (msg./sec.)
A1	30000
B1	40000
C1	9000
C2	55000
C3	90000

Table 2
Load on conveyors.

Time	Alpha load (msg./sec.)	Betta Load (msg./sec.)
1	100000	0
2	180000	30000
3	180000	100000
4	120000	140000
5	80000	180000
6	40000	230000
7	20000	280000
8	40000	200000
9	100000	180000
10	150000	120000
11	200000	80000

Since in a routing algorithm with message tagging the performance of components may decrease, due to the need to tag messages or determine the further route of messages, let us introduce a performance degradation factor.

$$K = 0.95, \tag{12}$$

where, K – degradation factor

Each conveyor has a unique load applied to it. The load changes every hour. An increase in the number of instances of a component occurs when the maximum load that an instance can handle is reached. A decrease in the number of instances of a component occurs when the load on an instance decreases to zero.

The load on each conveyor for 11 hours is shown in **Table 2**.

4.1. SIMULATION OF AN ALGORITHM WITH DIRECT CONNECTION

In a system with a direct component linkage algorithm, the number of instances of each component depends only on the load on the component in a particular pipeline, the final number of component instances C1 and C2 is calculated as the sum of the component instances in each pipeline.

The number of components of each type is shown in **Table 3**.

There were 464 instances of a component in the system during the whole time of operation.

Table 3
The number of instances of components in a system with direct connection.

Time	Alpha load (msg./sec.)	Betta load (msg./sec.)	Number of A1	Number of B1	Number of C1	Number of C2
1	100000	0	3	0	12	2
2	180000	30000	5	1	24	5
3	180000	100000	5	4	32	6
4	120000	140000	3	5	30	6
5	80000	180000	2	6	29	6
6	40000	230000	1	8	31	6
7	20000	280000	1	10	35	7
8	40000	200000	1	7	28	5
9	100000	180000	3	6	32	6
10	150000	120000	4	4	31	6
11	200000	80000	5	3	32	6

Table 4 shows the number of component instances created or deleted.

There have been 162 operations to create or delete component instances over the lifetime of the system.

4.2. SIMULATING AN ALGORITHM WITH A ROUTER

In a system with a router algorithm, the number of instances of each component depends only on the total load on the component. There is also a C3 router in the system.

Table 4
The number of operations in a system with direct connection of components.

Time	Alpha load (msg./sec.)	Betta load (msg./sec.)	Operations A1	Operations B1	Operations C1	Operations C2
1	100000	0	3	0	12	2
2	180000	30000	2	1	12	3
3	180000	100000	0	3	8	1
4	120000	140000	2	1	10	2
5	80000	180000	1	1	9	2
6	40000	230000	1	2	10	2
7	20000	280000	0	2	8	1
8	40000	200000	0	3	11	2
9	100000	180000	2	1	10	1
10	150000	120000	1	2	11	2
11	200000	80000	1	1	11	2

Table 5

The number of component instances in a system with a router.

Time	Alpha load (msg./sec.)	Betta load (msg./sec.)	Number of A1	Number of B1	Number of C1	Number of C2	Number of C3
1	100000	0	3	0	12	2	2
2	180000	30000	5	1	24	4	3
3	180000	100000	5	4	32	6	4
4	120000	140000	3	5	29	5	3
5	80000	180000	2	6	29	5	3
6	40000	230000	1	8	30	5	3
7	20000	280000	1	10	34	6	4
8	40000	200000	1	7	27	5	3
9	100000	180000	3	6	32	6	4
10	150000	120000	4	4	30	5	3
11	200000	80000	5	3	32	6	4

The number of components of each type is shown in **Table 5**.

There have been 489 instances of a component in the system over the lifetime of the system.

Table 6 shows the number of component instances created or deleted.

There have been 108 component instance creation or deletion operations in the system over the lifetime of the system.

Table 6

The number of operations in a system with a router.

Time	Alpha load (msg./sec.)	Betta load (msg./sec.)	Operations A1	Operations B1	Operations C1	Operations C2	Operations C3
1	100000	0	3	0	12	2	2
2	180000	30000	2	1	12	2	1
3	180000	100000	0	3	8	2	1
4	120000	140000	2	1	3	1	1
5	80000	180000	1	1	0	0	0
6	40000	230000	1	2	1	0	0
7	20000	280000	0	2	4	1	1
8	40000	200000	0	3	7	1	1
9	100000	180000	2	1	5	1	1
10	150000	120000	1	2	2	1	1
11	200000	80000	1	1	2	1	1

Table 7

The number of components instances in a system with message tagging.

Time	Alpha load (msg./sec.)	Betta load (msg./sec.)	Number of A1	Number of B1	Number of C1	Number of C2
1	100000	0	3	0	12	2
2	180000	30000	5	2	25	5
3	180000	100000	5	4	33	6
4	120000	140000	4	5	31	5
5	80000	180000	3	7	31	5
6	40000	230000	2	9	32	6
7	20000	280000	1	10	36	6
8	40000	200000	2	8	29	5
9	100000	180000	3	7	33	6
10	150000	120000	4	5	32	6
11	200000	80000	6	3	33	6

4.3. SIMULATION AN ALGORITHM WITH MESSAGE TAGGING

In a system with a message-labeled algorithm, the performance of each component is reduced. The number of instances of each component depends only on the total load on the component.

The number of components of each type is shown in **Table 7**.

There have been 483 instances of a component in the system over the lifetime of the system.

Table 8 shows the number of component instances created or deleted.

Table 8

The number of operations in a system with message tagging.

Time	Alpha load (msg./sec.)	Betta load (msg./sec.)	Operations A1	Operations B1	Operations C1	Operations C2
1	100000	0	3	0	12	2
2	180000	30000	2	2	13	3
3	180000	100000	0	2	8	1
4	120000	140000	1	1	2	1
5	80000	180000	1	2	0	0
6	40000	230000	1	2	1	1
7	20000	280000	1	1	4	0
8	40000	200000	1	2	7	1
9	100000	180000	1	1	4	1
10	150000	120000	1	2	1	0
11	200000	80000	2	2	1	0

There have been 94 operations to create or delete component instances over the lifetime of the system.

4.4. SIMULATION FINDINGS

The lowest total number of component instances in the system over all time is generated in the algorithm with direct connection - 464, against 489 in the algorithm with router and 483 in the algorithm with message tagging. The difference is due to the need to scale an additional component in the router algorithm and the reduced component performance in the message tagging algorithm. However, the difference with the message labeling algorithm is 4.09%.

The message tagging algorithm has the lowest number of operations to create or remove instances of components - 94 operations, compared to 162 operations for the direct join algorithm and 108 for the router algorithm. The difference is 41.98% with the direct join algorithm.

An algorithm with message tagging shows better results in terms of reducing the number of component creation or deletion operations than an algorithm with a router or a direct connection. Depending on the implementation, the performance reduction factor may vary, and the identified difference in the number of instances may be evened out.

5. CONCLUSION

The proposed message routing algorithm using a token eliminates the main disadvantages of previously used algorithms, such as: reduced efficiency of using system resources in the direct routing algorithm, low fault tolerance and efficient use of message transmission channels in the algorithm with a router.

However, the message marking algorithm will require additional resources for routing and an increased message size, in contrast to the direct connection algorithm, and the final

efficiency of use will depend on the optimal implementation of the routing module, without restrictions from the algorithms.

Thanks to the developed algorithm, the possibility of analyzing large amounts of data increases, the fault tolerance of the system increases, and the cost of the resources used is reduced.

However, in this article, only the classic scaling algorithm was considered - increasing or decreasing the number of component instances depending on the current load. However, this algorithm may be inefficient if the function is not monotone on the time interval. Optimization of the scaling algorithm will improve system performance by reducing the overhead costs for creating and deleting component instances.

REFERENCES

1. David Reinsel, John Rydning, John F. Gantz. IDC's Global DataSphere Forecast Shows Continued Steady Growth in the Creation and Consumption of Data. <https://www.idc.com/getdoc.jsp?containerId=prUS46286020>.
2. David Loshin. *ETL (Extract, Transform, Load). Business Intelligence: the Savvy Manager's Guide*, 2nd Edition. Elsevier, 2012, 400 p.
3. Vassiliadis P, Karagiannis A, Tziouva V, Simitsis A. Towards a Benchmark for ETL Workflows. *5th International Workshop on Quality in Databases*, 2007, 49-60 pp.
4. Simitsis A, Vassiliadis PA. A method for the mapping of conceptual designs to logical blueprints for ETL processes. *Decision Support Systems*, 2008, 45(1):22-40. DOI:10.1016/j.dss.2006.12.002.
5. Skoutas D, Simitsis A. Designing ETL processes using semantic web technologies. *Proceedings ACM 9th International Workshop on Data Warehousing and OLAP*, 2006, 67-74 pp.

6. Alkis Simitsis, Panos Vassiliadis, Timos Sellis. Optimizing ETL processes in data warehouses. *Proceedings of the International Conference on DataEngineering (ICDE)*, 2005.
7. Camunda Platform 7: Architecture Overview. URL: <https://docs.camunda.org/manual/7.16/introduction/architecture/> (Accessed 04/12/2022).
8. Baker F. *RFC 1812 Requirements for IP Version 4 Routers*. URL: <https://datatracker.ietf.org/doc/html/rfc1812> (Accessed 04/12/2022).

DOI: 10.17725/rensit.2022.14.291

QR-decomposition based receivers for downlink MIMO-SCMA

Islam Kanatbekuli, Dmitriy A. Pokamestov, Yakov V. Kryukov, Eugeni V.

Rogozhnikov, Renat R. Abenov

Tomsk State University of Control Systems and Radioelectronics, <https://tusur.ru/>

Tomsk 634050, Russian Federation

E-mail: islam-mabad@bk.ru, dmaltomsk@mail.ru, kryukov.tusur@gmail.com, udzhon@mail.ru, abenov.rr@gmail.com

Received 24.06.2022, peer-reviewed 30.06.2022, accepted 07.07.2022

Abstract: For the signal receiver of multiantenna communication systems, various multiple access schemes based on sparse SCMA (Sparse Code Multiple Access) codes are considered, applied in conjunction with MIMO (Multiple Input-Multiple Output) multiantenna systems, which make it possible to provide a high level of noise immunity and spectral efficiency of communication and network connections, being the basis of a possible concept for organizing the physical layer of 6G networks. Various MIMO-SCMA schemes are considered and the main stages of signal generation and processing are given. One of the problems in the implementation of MIMO-SCMA is the high computational complexity of signal processing algorithms in the receiver. To solve this problem, it is proposed to use QR decomposition for MIMO decoding in the MIMO-SCMA scenario. It is shown that this method allows not only to significantly reduce the computational complexity due to received signal transformations, but also to provide higher noise immunity compared to classical approaches. The results are obtained using the message passing algorithm for various MIMO coding schemes and antenna system configurations.

Keywords: NOMA access methods, SCMA codes, MIMO systems, 5G and 6G standards, QR decomposition

UDC 621.396.4

Acknowledgments: The work was supported by a grant from the President of the Russian Federation for state support of scientific research of young Russian scientists. Grant MK-4952.2021.4.

For citation: Islam Kanatbekuli, Dmitriy A. Pokamestov, Yakov V. Kryukov, Eugeni V. Rogozhnikov, Renat R. Abenov. QR-decomposition based receivers for downlink MIMO-SCMA. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2022, 14(3):291-300. DOI: 10.17725/rensit.2022.14.291.

CONTENTS

1. INTRODUCTION (291)
 2. SYSTEM MODEL (292)
 3. MIMO-SCMA DECODING ALGORITHMS (290)
 4. ANALYSIS OF THE COMPUTATIONAL COMPLEXITY OF DECODING ALGORITHMS (292)
 5. SIMULATION RESULTS AND DISCUSSION (293)
 6. CONCLUSION (294)
- REFERENCES (295)

1. INTRODUCTION

The development of science and technology leads to the emergence of new applications, such as virtual reality, the internet of things, and swarms of unmanned aerial vehicles. There is work underway on the formation of requirements for 6G networks [1]. To ensure mass connectivity at the physical level, it is proposed to use non-orthogonal multiple access methods (NOMA) [2], among

which multiple access based on sparse codes (SCMA) is one of the most promising [3]. In the SCMA, user bits are directly converted into sparse multidimensional code words. By introducing controlled interference, it is possible to use SCMA for networks with the number of channels exceeding the number of available resources. To eliminate the influence of inter-user interference, the message passing algorithm (MPA) has been proposed [4].

Combining the SCMA technology with multiple-input multiple-output (MIMO), called MIMO-SCMA, makes it possible to increase the spectral and energy efficiency of SCMA systems. There are many recorded works on the subject of MIMO-SCMA. Thus, spatial modulated SCMA (SM-SCMA) systems are considered in [5-6]. Works [6-8] consider spatial multiplexed SCMA (MUX-SCMA). Another variant of MIMO techniques is space-time block-coded SCMA (STBC-SCMA) [6-7,9-11]. The single-input multiple-output SCMA (SIMO-SCMA) systems are considered in [7,12] as a special case of MIMO-SCMA.

All MIMO-SCMA techniques in the uplink scenario use joint MPA (JMPA) to decode transmitted signals. The computational complexity of the JMPA of MIMO-SCMA systems with N_t transmitting and N_r receiving antennas is proportional to $O(KN_r M^{N_r d_k})$. To reduce the computational complexity, works [7,13] use the expectation propagation algorithm (EPA) instead of MPA, which reduces the complexity to $O(KN_r M^{N_r d_k})$.

In the downlink scenario of MUX-SCMA, JMPA is a sub-optimal detector with the same performance as an ML detector but with less complexity. However, in both JMPA and Joint EPA detectors, the computational

complexity depends on N_t and N_r , which can be impractical when N_t and N_r are large. In the downlink scenario, due to the fact that subscriber signals are added to the transmitter, it is possible to separate MIMO and SCMA decoding to reduce the overall computational complexity. In this paper, we propose to use the QR -decomposition for MIMO decoding of MIMO-SCMA. To the authors' knowledge, the QR -decomposition has not been used to decode MIMO-SCMA. In [12-13], the QR -decomposition is used to reduce the number of resource nodes for multiple receiving antennas. Separate decoding of MIMO and SCMA makes it possible not only to reduce computational complexity by converting the received signal but also is a universal scheme where MPA, EPA, or any other detector can be used as a SCMA detector.

The following notations are used in the work: to denote scalars, vectors, and matrices, x , \mathbf{x} , and \mathbf{X} notations are used, respectively; \mathbf{I}_L is a unit matrix of the $L \times L$ size; $(\cdot)^H$ denotes Hermitian conjugate; $(\cdot)^*$ stands for complex conjugation, and $(\cdot)^T$ denotes transposition.

2. SYSTEM MODEL

We consider an MIMO-SCMA system with J users, K orthogonal subcarriers, N_t transmitting and N_r receiving antennas. A generalized model of the MIMO-SCMA system is shown in **Fig. 1**.

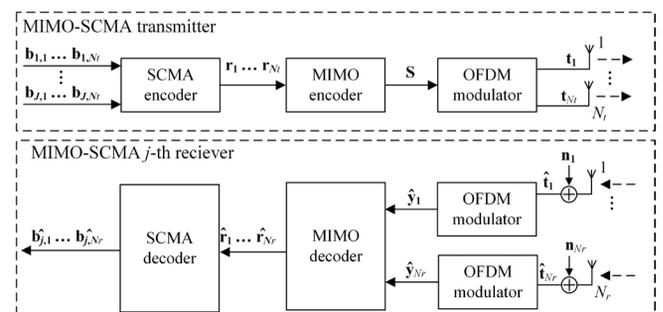


Fig. 1. Generalized diagram of a MIMO-SCMA communication system model.

2.1. MUX-SCMA

In the downlink channel, the codewords of the subscribers $\mathbf{x}_j, j = 1, \dots, J$ are added in the transmitter:

$$\mathbf{r} = \sum_{j=1}^J \mathbf{x}_j.$$

In downlink MUX-SCMA (in general, its corresponding scheme is shown in Fig. 1), j -user's bit streams \mathbf{b}_j^n from the j -th subscriber ($j = 1, \dots, J$), intended for transmission through the i -th antenna ($i = 1, \dots, N_t$) come to the block SCMA encoder. In this block, the selection of subscribers' codewords \mathbf{x}_j from the codebooks and their addition. Formed in this way the vectors go to the block MIMO encoder, which, depending on the algorithm chosen matrix is formed by the transmitted symbols \mathbf{S} . In block OFDM (Orthogonal Frequency Division Multiplexing) modulation symbols are divided into N_t streams, each of which are formed OFDM symbols $\mathbf{t}_i (i = 1, \dots, N_t)$: perform an inverse fast Fourier transform (IFFT) and added a cyclic prefix. With spatial multiplexing, the symbols $\mathbf{r}_1, \dots, \mathbf{r}_{N_t}$ are transmitted in parallel, which allows to increase the spectral efficiency by N_t times compared to conventional SCMA systems. The received signal in the m -th receiving antenna ($m = 1, \dots, N_r$) in the frequency domain on the k -th subcarrier ($k = 1, \dots, K$) is:

$$y_{m,k} = \sum_{i=1}^{N_t} h_{i,m,k} r_{i,k} + n_k^m, \tag{1}$$

where $h_{i,m,k}$ is the channel coefficient from the i -th transmitting to the m -th receiving antenna on the k -th subcarrier, $h_{i,m,k} \sim CN(0,1)$; and n_k^m is additive white Gaussian noise in the m -th receiving antenna ($m = 1, \dots, N_r$) on the k -th subcarrier, and $n_k^m \sim CN(0, N_0)$.

As a vector notation, expression (1) can be written as:

$$\mathbf{y}_m = \sum_{i=1}^{N_t} \text{diag}(\mathbf{h}_{i,m}) \mathbf{r}_i + \mathbf{n}_m,$$

where $\text{diag}(\mathbf{x})$ is a diagonal matrix with \mathbf{x} diagonal elements. As a matrix notation, the received signal for N_r receiving antennas is written as:

$$\begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{N_r} \end{pmatrix} = \begin{pmatrix} \text{diag}(\mathbf{h}_{1,1}) & \cdots & \text{diag}(\mathbf{h}_{1,N_t}) \\ \vdots & \ddots & \vdots \\ \text{diag}(\mathbf{h}_{N_r,1}) & \cdots & \text{diag}(\mathbf{h}_{N_r,N_t}) \end{pmatrix} \cdot \begin{pmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_{N_t} \end{pmatrix} + \begin{pmatrix} \mathbf{n}_1 \\ \vdots \\ \mathbf{n}_{N_r} \end{pmatrix}, \tag{2}$$

$$\mathbf{y} = \mathbf{H} \cdot \mathbf{r} + \mathbf{n}.$$

2.2. ALAMOUTI SCHEME BASED ON MIMO-SCMA

To increase noise immunity in channels with fading, STBC is used in systems with multiple transmitting antennas. Alamouti scheme is orthogonal STBC (OSTBC) with two transmitting antennas [11,14]. The STBC-SCMA coding matrix for Alamouti scheme is

$$\mathbf{C}_2 = \begin{pmatrix} \text{diag}(\mathbf{r}_1) & \text{diag}(\mathbf{r}_2) \\ -\text{diag}(\mathbf{r}_2^*) & \text{diag}(\mathbf{r}_1^*) \end{pmatrix}.$$

The received signal can be recorded using a channel matrix as follows:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2^* \end{bmatrix} = \begin{bmatrix} \text{diag}(\mathbf{h}_1) & \text{diag}(\mathbf{h}_2) \\ \text{diag}(\mathbf{h}_2^*) & -\text{diag}(\mathbf{h}_1^*) \end{bmatrix} \cdot \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{pmatrix} + \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2^* \end{bmatrix}.$$

For N_r receiving antennas, the received signal can be represented as:

$$\begin{pmatrix} \mathbf{y}_{1,1} \\ \mathbf{y}_{2,1}^* \\ \vdots \\ \mathbf{y}_{1,N_r} \\ \mathbf{y}_{2,N_r}^* \end{pmatrix} = \begin{pmatrix} \begin{bmatrix} \text{diag}(\mathbf{h}_{1,1}) & \text{diag}(\mathbf{h}_{2,1}) \\ \text{diag}(\mathbf{h}_{2,1}^*) & -\text{diag}(\mathbf{h}_{1,1}^*) \end{bmatrix} \\ \vdots \\ \begin{bmatrix} \text{diag}(\mathbf{h}_{1,N_r}) & \text{diag}(\mathbf{h}_{2,N_r}) \\ \text{diag}(\mathbf{h}_{2,N_r}^*) & -\text{diag}(\mathbf{h}_{1,N_r}^*) \end{bmatrix} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{pmatrix} + \begin{pmatrix} \mathbf{n}_{1,1} \\ \mathbf{n}_{2,1}^* \\ \vdots \\ \mathbf{n}_{1,N_r} \\ \mathbf{n}_{2,N_r}^* \end{pmatrix}, \tag{3}$$

$$\bar{\mathbf{y}} = \bar{\mathbf{H}} \cdot \mathbf{r} + \bar{\mathbf{n}}.$$

The Alamouti scheme has a code rate of $R = 1$. However, there is no OSTBC with $R = 1$ for $N_t > 2$. To achieve the same MIMO code rate, quasi-orthogonal STBC (QOSTBC) is used. [15] has been chosen as QOSTBC for four transmitting antennas. Distinct from QOSTBC in [11], the proposed code has less

coding and decoding complexity. The coding and channel matrices for one receiving antenna are presented below:

$$C_4 = \begin{pmatrix} \text{diag}(\mathbf{r}_1 + \mathbf{r}_3) & \text{diag}(\mathbf{r}_2 + \mathbf{r}_4) & \text{diag}(\mathbf{r}_3 - \mathbf{r}_1) & \text{diag}(\mathbf{r}_4 - \mathbf{r}_2) \\ \text{diag}(-\mathbf{r}_4^* - \mathbf{r}_2^*) & \text{diag}(\mathbf{r}_1^* + \mathbf{r}_3^*) & \text{diag}(-\mathbf{r}_4^* + \mathbf{r}_2^*) & \text{diag}(\mathbf{r}_3^* - \mathbf{r}_1^*) \\ \text{diag}(\mathbf{r}_3 - \mathbf{r}_1) & \text{diag}(\mathbf{r}_4 - \mathbf{r}_2) & \text{diag}(\mathbf{r}_1 + \mathbf{r}_3) & \text{diag}(\mathbf{r}_2 + \mathbf{r}_4) \\ \text{diag}(-\mathbf{r}_4^* + \mathbf{r}_2^*) & \text{diag}(\mathbf{r}_3^* - \mathbf{r}_1^*) & \text{diag}(-\mathbf{r}_4^* - \mathbf{r}_2^*) & \text{diag}(\mathbf{r}_1^* + \mathbf{r}_3^*) \end{pmatrix},$$

$$H = \begin{pmatrix} \text{diag}(\mathbf{h}_1 - \mathbf{h}_3) & \text{diag}(\mathbf{h}_2 - \mathbf{h}_4) & \text{diag}(\mathbf{h}_3 + \mathbf{h}_1) & \text{diag}(\mathbf{h}_4 + \mathbf{h}_2) \\ \text{diag}(\mathbf{h}_2^* - \mathbf{h}_4^*) & \text{diag}(\mathbf{h}_3^* - \mathbf{h}_1^*) & \text{diag}(\mathbf{h}_4^* + \mathbf{h}_2^*) & \text{diag}(-\mathbf{h}_1^* - \mathbf{h}_3^*) \\ \text{diag}(\mathbf{h}_3 - \mathbf{h}_1) & \text{diag}(\mathbf{h}_4 - \mathbf{h}_2) & \text{diag}(\mathbf{h}_3 + \mathbf{h}_1) & \text{diag}(\mathbf{h}_4 + \mathbf{h}_2) \\ \text{diag}(\mathbf{h}_4^* - \mathbf{h}_2^*) & \text{diag}(\mathbf{h}_3^* - \mathbf{h}_1^*) & \text{diag}(\mathbf{h}_2^* - \mathbf{h}_4^*) & \text{diag}(-\mathbf{h}_1^* - \mathbf{h}_3^*) \end{pmatrix}.$$

The signal received on the m -th receiving antenna is written as follows in the vector notation:

$$\mathbf{y}'_m = \begin{pmatrix} \mathbf{y}_{1,m} \\ \mathbf{y}_{2,m}^* \\ \mathbf{y}_{3,m} \\ \mathbf{y}_{4,m}^* \end{pmatrix} = \mathbf{H}_m \cdot \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \\ \mathbf{r}_4 \end{pmatrix} + \begin{pmatrix} \mathbf{n}_{1,m} \\ \mathbf{n}_{2,m}^* \\ \mathbf{n}_{3,m} \\ \mathbf{n}_{4,m}^* \end{pmatrix}.$$

For N_r receiving antennas:

$$\begin{pmatrix} \mathbf{y}'_1 \\ \vdots \\ \mathbf{y}'_{N_r} \end{pmatrix} = \begin{pmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_{N_r} \end{pmatrix} \cdot \mathbf{r} + \begin{pmatrix} \mathbf{n}'_1 \\ \vdots \\ \mathbf{n}'_{N_r} \end{pmatrix}, \tag{4}$$

$$\hat{\mathbf{y}} = \hat{\mathbf{H}} \cdot \mathbf{r} + \hat{\mathbf{n}}.$$

3. ALGORITHMS FOR MUX-SCMA

3.1. LINEAR MIMO DECODERS

In MUX-SCMA, we can use well-known linear decoding algorithms for the MIMO part. To decode (2) using the Zero Forcing (ZF) decoder, the following transformation can be written

$$\hat{\mathbf{r}}_{ZF} = (\mathbf{H}^H \cdot \mathbf{H})^{-1} \mathbf{H}^H \cdot \mathbf{y},$$

and for minimum mean square error (MMSE) of the decoder:

$$\hat{\mathbf{r}}_{MMSE} = (\mathbf{H}^H \cdot \mathbf{H} + \sigma_n^2)^{-1} \mathbf{H}^H \cdot \mathbf{y},$$

where \mathbf{I} is a unit matrix.

After the MIMO decoding, symbols $\hat{\mathbf{r}}_i$ can be decoded using MPA/EPA simultaneously. Due to the effect of noise amplification of the ZF and MMSE decoder, the noise power

for each symbol on each subcarrier will be different and uncertain, which affects the noise immunity of the SCMA detector. Further, for these two decoders, noise power

$$\hat{\sigma}_i^2 = \sigma_n^2 / N_{tx} \text{ for } i = 1, \dots, N_{tx}.$$

3.2. JMPA DECODER

In [6,8], a JMPA detector is used for MIMO-SCMA joint decoding. JMPA is implemented in several steps:

1. *Plotting an extended factor graph:*

$$\bar{\mathbf{F}} = N_{rx} \overbrace{\left\{ \begin{pmatrix} \mathbf{F} & \dots & \mathbf{F} \\ \vdots & \ddots & \vdots \\ \mathbf{F} & \dots & \mathbf{F} \end{pmatrix} \right\}}^{N_{tx}}.$$

After the transformation of the graph, we obtain $\bar{J} = J \cdot N_t$ virtual users and $\bar{K} = K \cdot N_r$ virtual subcarriers. As a consequence, the received signal from all receiving antennas is converted as:

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{N_r} \end{pmatrix}.$$

2. *Developing virtual codebooks.* A $\bar{j} = J(n-1) + j$ virtual codebook of the $\bar{K} \times M$:size is developed for each virtual user:

$$\mathbf{x}_{\bar{j}} = \text{diag} \left[(\mathbf{h}_{1,n}^j)^T, \dots, (\mathbf{h}_{N_r,n}^j)^T \right] \begin{pmatrix} \mathbf{x}_j \\ \vdots \\ \mathbf{x}_j \end{pmatrix} \Bigg\} N_r,$$

$$j = 1, \dots, J, \quad n = 1, \dots, N_t.$$

3. *MPA.* The MPA detector inputs \mathbf{y} with the corresponding $\mathbf{x}_{\bar{j}}$. Channel coefficients are equal to 1.

Analyzing the second point of the algorithm and the MPA detector algorithm, we can conclude that it is not necessary to build virtual codebooks. It is enough to ensure correct indexing of channel coefficients in MPA algorithm. For virtual user $\bar{j} = J(n-1) + j$ on virtual subcarrier $\bar{K} = K(r-1) + k$:

$$h_{\bar{k},\bar{j}} = h_{k,j}^{n,r}, \quad r = 1, \dots, N_r.$$

3.3. QR-OSIC-MPA DECODER

In this subsection, we propose the QR-decomposition-based algorithm, named QR-OSIC-MPA, for decoding MUX-SCMA. Using the QR-decomposition, we can represent matrix $\mathbf{H} = \mathbf{Q} \cdot \mathbf{R}$, then (2) can be represented as

$$\mathbf{y} = \mathbf{Q} \cdot \mathbf{R} \cdot \mathbf{r} + \mathbf{n},$$

where \mathbf{Q} is unitary and \mathbf{R} is an upper triangular matrix.

The decoding process using QR-OSIC-MPA is performed in several stages:

1. QR-decomposition of matrix \mathbf{H} and subsequent signal preprocessing using matrix \mathbf{Q} :

$$\tilde{\mathbf{y}} = \begin{pmatrix} \tilde{y}_1 \\ \vdots \\ \tilde{y}_{N_r} \end{pmatrix} = \mathbf{Q}^H \cdot \mathbf{y} = \mathbf{R} \cdot \mathbf{r} + \tilde{\mathbf{n}}.$$

The matrix \mathbf{R} has $N_r K \times N_r K$ dimension and looks like:

$$\mathbf{R} = \begin{pmatrix} \text{diag}(\mathbf{R}_{1,1}) & \cdots & \text{diag}(\mathbf{R}_{1,N_r}) \\ 0 & \ddots & \vdots \\ \vdots & \vdots & \text{diag}(\mathbf{R}_{N_r,N_r}) \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{pmatrix},$$

where \mathbf{R}_{ij} is a vector of $K \times 1$ size, $\mathbf{0}$ is a zero matrix of $K \times K$ size.

Due to the unitarity property of matrix \mathbf{Q} , the noise component of the processed signal $\tilde{\mathbf{n}} = \mathbf{Q}^H \cdot \mathbf{n}$ does not change its statistical characteristics, that is, the noise power does not change. Since matrix \mathbf{R} is upper-triangular, then, after multiplication by \mathbf{Q}^H , transmitted signals $\mathbf{r}_1, \dots, \mathbf{r}_{N_t}$ contain only $\tilde{y}_1, \dots, \tilde{y}_{N_r}$.

2. OSIC-MPA. For the i -th iteration, $i = N_r, \dots, 1$, the following holds:

Channel coefficients are $\hat{\mathbf{h}}_i = \mathbf{R}_{i,i}$. If $i = N_r$, then $\hat{\mathbf{y}}_i = \tilde{y}_i$, in other way

$$\hat{\mathbf{y}}_i = \tilde{y}_i - \sum_{p=i+1}^{N_r} \text{diag}(\mathbf{R}_{i,s}) \cdot \hat{\mathbf{r}}_p.$$

The i -th character is restored. $\hat{\mathbf{y}}_i$ and $\hat{\mathbf{h}}_i$ are applied to the input of the MPA detector. After decoding, the symbol $\hat{\mathbf{r}}_i$ is restored. In the case of systems with encoding, at this stage, there should be the SCMA decoding and channel decoding performed. After restoring the transmitted subscribers' bits, it is necessary to re-encode the bit vectors and restore the $\hat{\mathbf{r}}_i$.

3.4. H-MPA FOR ALAMOUTI SCHEME

The classical processing of equation (3) is performed using matrix \mathbf{H} [11]:

$$\tilde{\mathbf{y}} = \begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{pmatrix} = \bar{\mathbf{H}}^H \cdot \bar{\mathbf{y}} = \begin{pmatrix} \text{diag}(\tilde{\mathbf{h}}) & 0 \\ 0 & \text{diag}(\tilde{\mathbf{h}}) \end{pmatrix} \cdot \mathbf{r} + \bar{\mathbf{H}}^H \cdot \bar{\mathbf{n}}, \tag{5}$$

where

$$\tilde{h}_k = \sum_{i=1}^2 \left(\sum_{r=1}^{N_{rx}} |h_{i,r,k}|^2 \right).$$

Expression (5) can be represented as two received signals of SISO-SCMA:

$$\tilde{y}_1 = \text{diag}(\tilde{\mathbf{h}}) \cdot \mathbf{r}_1 + \tilde{\mathbf{n}}_1,$$

$$\tilde{y}_2 = \text{diag}(\tilde{\mathbf{h}}) \cdot \mathbf{r}_2 + \tilde{\mathbf{n}}_2.$$

Further decoding of the signal is performed using the MPA/EPA detector. Symbols $\tilde{\mathbf{r}}_1$ and $\tilde{\mathbf{r}}_2$ can be decoded simultaneously, since they are independent from each other. The power of the converted noise is $\sigma_{\tilde{n}}^2 = 2N_{rx} \sigma_n^2$.

3.5. QR-MPA FOR ALAMOUTI SCHEME

We can use $\bar{\mathbf{Q}}^H$ instead of $\bar{\mathbf{H}}^H$ in expression (5), then the MIMO transformation will be written as

$$\tilde{\mathbf{y}} = \begin{pmatrix} \tilde{\mathbf{y}}_{1,1} \\ \tilde{\mathbf{y}}_{2,1} \\ \vdots \\ \tilde{\mathbf{y}}_{1,N_{rx}} \\ \tilde{\mathbf{y}}_{2,N_{rx}} \end{pmatrix} = \mathbf{Q}^H \cdot \bar{\mathbf{y}} = \begin{pmatrix} \text{diag}(\mathbf{a}) & 0 \\ 0 & -\text{diag}(\mathbf{a}) \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix} \cdot \mathbf{s} + \tilde{\mathbf{n}},$$

where $\mathbf{a} = \mathbf{R}_{1,1}$.

After the MIMO processing, transmitted signals \mathbf{r}_1 and \mathbf{r}_2 are contained only in $\tilde{\mathbf{y}}_{1,1}$ and $\tilde{\mathbf{y}}_{2,1}$. The processed signal can also be represented as two SISO-SCMA signals:

$$\tilde{\mathbf{y}}_{1,1} = \text{diag}(\tilde{\mathbf{h}}_1) \cdot \mathbf{r}_1 + \tilde{\mathbf{n}}_1,$$

$$\tilde{\mathbf{y}}_{2,1} = \text{diag}(\tilde{\mathbf{h}}_2) \cdot \mathbf{r}_2 + \tilde{\mathbf{n}}_2,$$

where $\tilde{\mathbf{h}}_1 = \mathbf{a}$ and $\tilde{\mathbf{h}}_2 = -\mathbf{a}$. The power of the converted noise is $\sigma_{\tilde{\mathbf{n}}}^2 = \sigma_n^2$.

3.6. H-MPA FOR QOSTBC

The decoding process (4) applying matrix \mathbf{H} is carried out in several steps:

1. Signal preprocessing: $\bar{\mathbf{y}} = \begin{pmatrix} \bar{\mathbf{y}}_1 \\ \bar{\mathbf{y}}_2 \\ \bar{\mathbf{y}}_3 \\ \bar{\mathbf{y}}_4 \end{pmatrix} = \hat{\mathbf{H}}^H \cdot \hat{\mathbf{y}};$

2. Deinterleaving symbols: $\begin{pmatrix} \tilde{\mathbf{y}}_1 \\ \tilde{\mathbf{y}}_2 \\ \tilde{\mathbf{y}}_3 \\ \tilde{\mathbf{y}}_4 \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{y}}_1' + i\bar{\mathbf{y}}_3^o \\ \bar{\mathbf{y}}_2' + i\bar{\mathbf{y}}_4^o \\ \bar{\mathbf{y}}_3' + i\bar{\mathbf{y}}_1^o \\ \bar{\mathbf{y}}_4' + i\bar{\mathbf{y}}_2^o \end{pmatrix};$

3. Calculating coefficients:

$$\mathbf{a} = \sum_{r=1}^{N_{rx}} \left(2 \cdot \left(\sum_{i=1}^4 |\mathbf{h}_{i,r}|^2 - 2 \cdot (\mathbf{h}_{1,r} \odot \mathbf{h}_{3,r}^* + \mathbf{h}_{2,r} \odot \mathbf{h}_{4,r}^*)^l \right) \right);$$

$$\mathbf{b} = \sum_{r=1}^{N_{rx}} \left(2 \cdot \left(\sum_{i=1}^4 |\mathbf{h}_{i,r}|^2 + 2 \cdot (\mathbf{h}_{1,r} \odot \mathbf{h}_{3,r}^* + \mathbf{h}_{2,r} \odot \mathbf{h}_{4,r}^*)^l \right) \right);$$

where \odot is the operator of element-wise matrix or vector multiplication.

4. Signal converting:

$$\begin{cases} \tilde{\mathbf{y}}_l = \sqrt{\mathbf{b}} \odot \tilde{\mathbf{y}}_l^l + i\sqrt{\mathbf{a}} \odot \tilde{\mathbf{y}}_l^o, l=1,2 \\ \tilde{\mathbf{y}}_l = \sqrt{\mathbf{a}} \odot \tilde{\mathbf{y}}_l^l + i\sqrt{\mathbf{b}} \odot \tilde{\mathbf{y}}_l^o, l=3,4 \end{cases}$$

5. Subscribers' codebook converting:

$$\begin{cases} \hat{\mathbf{x}}_j^o = \sqrt{\mathbf{b}} \odot \mathbf{a} \odot \mathbf{x}_j^l + i\sqrt{\mathbf{a}} \odot \mathbf{b} \odot \mathbf{x}_j^o, \text{ for } \tilde{\mathbf{y}}_1 \text{ and } \tilde{\mathbf{y}}_2 \\ \hat{\mathbf{x}}_j^o = \sqrt{\mathbf{a}} \odot \mathbf{b} \odot \mathbf{x}_j^l + i\sqrt{\mathbf{b}} \odot \mathbf{a} \odot \mathbf{x}_j^o, \text{ for } \tilde{\mathbf{y}}_3 \text{ and } \tilde{\mathbf{y}}_4 \end{cases}$$

6. Decoding MPA/EPA according to the equation:

$$\tilde{\mathbf{y}}_i = \tilde{\mathbf{r}}_i + \tilde{\mathbf{n}}_i, \quad i=1, \dots, 4$$

where $\tilde{\mathbf{r}}_i$ is the superimposed SCMA symbol derived from $\hat{\mathbf{x}}_j^o$.

Symbols can be processed simultaneously. Noise power for the k -th subcarrier is $\sigma_{n,k}^2 = 8N_{rx}\sigma_n^2(a_k + b_k)$.

3.7. QR-MPA FOR QOSTBC

A similar decoding method can be performed using the QR -decomposition:

1. Signal preprocessing: $\bar{\mathbf{y}} = \begin{pmatrix} \bar{\mathbf{y}}_1 \\ \bar{\mathbf{y}}_2 \\ \bar{\mathbf{y}}_3 \\ \bar{\mathbf{y}}_4 \end{pmatrix} = \mathbf{Q}^H \hat{\mathbf{y}}.$

2. Deinterleaving symbols:

$$\begin{pmatrix} \tilde{\mathbf{y}}_{1,1} \\ \tilde{\mathbf{y}}_{2,1} \\ \tilde{\mathbf{y}}_{3,1} \\ \tilde{\mathbf{y}}_{4,1} \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{y}}_{1,1}' + i\bar{\mathbf{y}}_{3,1}^o \\ -(\bar{\mathbf{y}}_{2,1}' + i\bar{\mathbf{y}}_{4,1}^o) \\ \bar{\mathbf{y}}_{3,1}' + i\bar{\mathbf{y}}_{1,1}^o \\ -(\bar{\mathbf{y}}_{4,1}' + i\bar{\mathbf{y}}_{2,1}^o) \end{pmatrix}.$$

3. Calculating coefficients:

$$\mathbf{a} = \mathbf{R}_{1,1},$$

$$\mathbf{b} = \mathbf{R}_{3,3},$$

4. Subscribers' codebook converting:

$$\begin{cases} \hat{\mathbf{x}}_j^o = \mathbf{a} \odot \mathbf{x}_j^l + i\mathbf{b} \odot \mathbf{x}_j^o, \text{ for } \tilde{\mathbf{y}}_{1,1} \text{ and } \tilde{\mathbf{y}}_{2,1} \\ \hat{\mathbf{x}}_j^o = \mathbf{b} \odot \mathbf{x}_j^l + i\mathbf{a} \odot \mathbf{x}_j^o, \text{ for } \tilde{\mathbf{y}}_{3,1} \text{ and } \tilde{\mathbf{y}}_{4,1} \end{cases}$$

5. Decoding MPA/EPA according to the equation:

$$\tilde{\mathbf{y}}_i = \tilde{\mathbf{r}}_i + \tilde{\mathbf{n}}_i, \quad i=1, \dots, 4.$$

Noise power is $\sigma_{\tilde{\mathbf{n}}}^2 = \sigma_n^2$.

4. DECODING COMPLEXITY ANALYSIS

Table 1 presents the computational complexity for $JMPA$ and QR -OSIC-MPA.

Table 1

Algorithms Complexity for MUX-SCMA

Detector	Exhibitors	Addition	Multiplication
JMPA	$I_{iter} M^{N_{tdk}} N_r K N_t d_k$	$I_{iter} [M^{N_{tdk}} (N_t d_k + 1) - 1] N_r K N_t d_k$	$I_{iter} N_r K N_t d_k [M^{N_{tdk}} (2d_k + 2) + M(N-1)]$
QR-OSIC-MPA	-	$2K^3 N_{rx} N_{tx}^2 + KN_{rx} - KN_{tx} \cdot (1 + 1.5KN_{tx} - 0.5N_{tx}) + N_{tx} \cdot N_{op}(MPA)$	$2K^3 N_{rx} N_{tx}^2 + KN_{rx} \cdot (2 + KN_{rx}) + KN_{tx} (1 + 0.5KN_{tx} + 0.5N_{tx}) + 0.5KN_{tx} \cdot (N_{tx} - 1) + N_{tx} N_{op}(MPA)$

Complexity is calculated as the number of operations on complex numbers. The computational complexity of the JMPA algorithm is equal to the complexity of the MPA detector with \bar{J} , number of users, \bar{K} number of subcarriers, $\bar{d}_k = d_k \cdot N_{tx}$, and $\bar{N} = N \cdot N_{rx}$. The Gram-Schmidt process is used to calculate the QR-decomposition, and the computational complexity is calculated for the decomposition of a non-sparse matrix.

Table 2 shows a numerical representation of the computational complexity. The JMPA complexity is chosen as a benchmark.

5. SIMULATION RESULTS AND DISCUSSION

This section presents the results of the MUX-SCMA and STBC-SCMA simulations. A codebook [16] with parameters $J = 6, K = 4, M = 4, d_k = 3$ has been used for modeling. MPA with five iterations has been used as a detector.

On Fig. 2 results of modeling STBC-SCMA (bit error rate (BER) performance) for two transmitting antennas are presented.

Table 2

Numerical comparison of computational complexity, %

Model MIMO	Detector					
	QR-OSIC-MPA			JMPA		
	exp	+, -	*, /	exp	+, -	*, /
MIMO-SCMA 2x2	0.3906	0.2565	0.2459	100	100	100
MIMO-SCMA 2x3	0.1736	0.1223	0.1142	100	100	100
MIMO-SCMA 3x3	0.0027	0.0015	0.0013	100	100	100

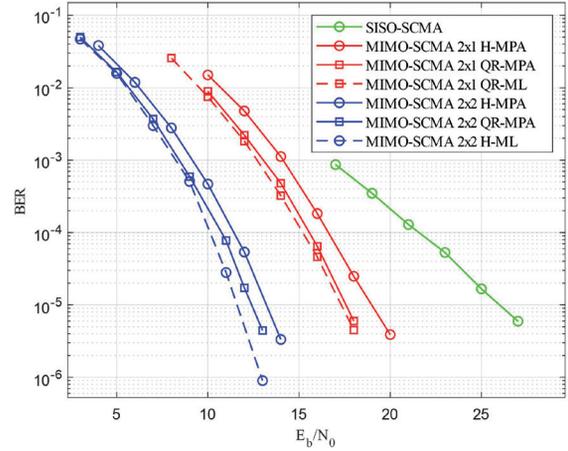


Fig. 2. STBC-SCMA simulation results 2xNr

As can be seen from Fig. 2, using the MIMO encoding made it possible to increase the noise immunity compared to a single-antenna circuit by 7 dB and 12.8 dB for the H-MPA processing at BER = 10⁻⁵ for one and two receiving antennas, respectively. For QR-MPA, similar gains were 8.4 and 13.6 dB.

Compared to the optimal ML detector, the performance degradation was about 1.6 dB for H-MPA. For QR-MPA, the loss was 0.3 and 0.8 dB for one and two receiving antennas at BER = 10⁻⁵. This deterioration in noise immunity can be reduced by increasing the number of the MPA iterations or by using more advanced MPA options providing fast convergence.

Fig. 3 shows the results of the STBC-SCMA simulation for four transmitting antennas.

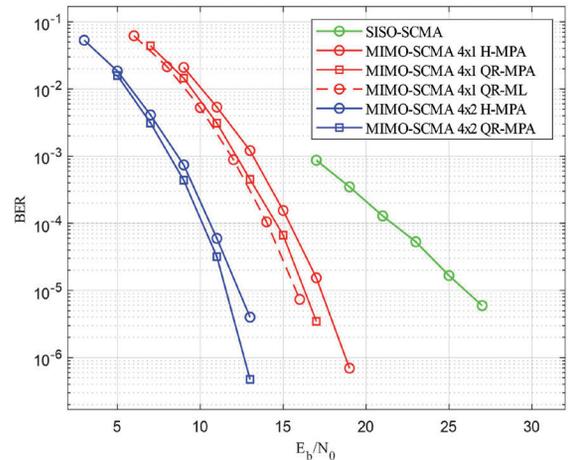


Fig. 3. STBC-SCMA simulation results 4xNr

As can be seen from Fig. 3, STBC-SCMA with four transmitting and one receiving antennas made it possible to increase the noise immunity compared to a single-antenna circuit by 8.7 dB and 9.7 dB at BER = 10⁻⁵ for H-MPA and QR-MPA, respectively. For two receiving antennas, similar gains were 13.7 dB and 14.45 dB.

Comparatively to the optimal ML detector, the performance degradation at BER = 10⁻⁵ was about 1.5 and 0.5 dB for H-MPA and QR-MPA, respectively.

Analysis of Fig. 2 and Fig. 3 shows that QR processing allows for greater noise immunity compared to classical processing using a channel matrix. Fig. 4 shows a comparison of STBC-SCMA schemes.

As can be seen from Fig. 4, STBC-SCMA 4×N_r ensures full dispersing during the transmission. At the same time, the gain compared to STBC-SCMA 2×N_r at BER = 10⁻⁵ is about 1.3 dB and 0.85 dB for one and two receiving antennas, respectively.

Fig. 5 shows the results of the MUX-SCMA simulation for two transmitting antennas.

As can be seen from Fig. 5, for two receiving antennas at BER = 10⁻², the MMSE detector provides gain of 1.7 dB compared

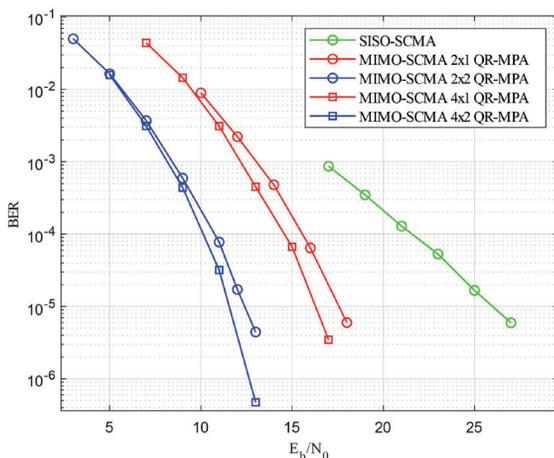


Fig. 4. Comparison of STBC-SCMA schemes.

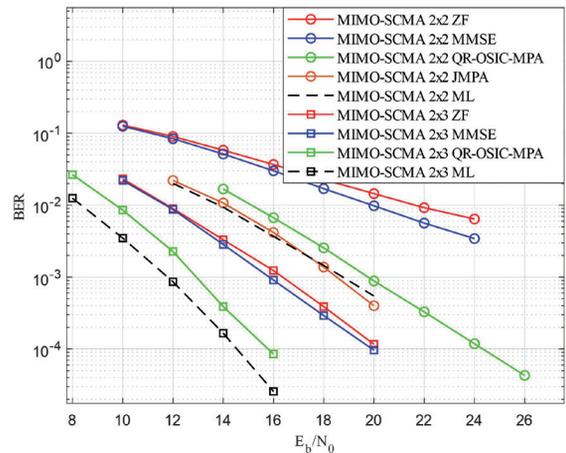


Fig. 5. MUX-SCMA simulation results 2×N_r.

to the ZF detector. QR-OSIC-MPA with the same BER made it possible to increase noise immunity by 4.8 dB against MMSE. The loss was about 1 dB compared to JMPA and ML. For three receiving antennas, QR-OSIC-MPA exceeds MMSE at BER = 10⁻³ by 2.9 dB. The loss was 1.15 dB compared to ML.

When analyzing the noise immunity and computational complexity of the MUX-SCMA decoding methods, QR-OSIC-MPA is concluded to be the best solution that provides high noise immunity with relatively low complexity. The final gains depend on the codebook used.

6. CONCLUSIONS

The article discusses the receiver scheme of communication systems with SCMA-MIMO based on the QR-decomposition and MPA for various MIMO coding schemes and configurations of antenna systems. The proposed method allows for significant reduction of the computational complexity of processing algorithms and shows lower BER performance compared to classical approaches. Thus, compared to JMPA, the computational complexity decreases from about 3 times (for N_{tx} = N_{rx} = 2) up to almost 70 for (for N_{tx} = N_{rx} = 3). Simulation

modeling of the proposed methods showed that, compared to SISO-SCMA, MIMO-SCMA has a difference of up to 15 dB in E_b/N_0 . The gain from using QR processing compared to classical SCMA-MPA (JMPA) processing methods ranges from 0.8 to 1.3 dB and is inferior to optimal ML processing from 0.3 to 0.8 dB. The use of MIMO with the SCMA technology can be applied in communication systems of future generations (6G) to increase noise immunity and the number of network connections. The use of QR-MPA processing will solve one of the key problems of SCMA-MIMO, that is, increasing computational complexity, providing a high level of BER performance.

REFERENCES

1. Katz M, Pirinen P, Posti H. Towards 6G: Getting ready for the next decade. *Proc. 16th international symposium on wireless communication systems (ISWCS)*, 2019:714-718.
2. Dai L, Wang B, Yuan Y, Han S, C.-L. I, Wang Z. Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends. *IEEE Commun. Mag.*, 2015, 53(9):74-81.
3. Nikopour H, Baligh H. Sparse code multiple access. *IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE, 2013:332-336.
4. Pokamestov DA, Demidov AYA, Kryukov YAV, Rogozhnikov EV, Abenov RR. Formirovanie i obrabotka signalov mnozhestvennogo dostupa s razrezhennym kodom [Sparse code multiple access generation and processing]. Moscow, Elektrosvyaz' Publ., 2016, 10:56-61.
5. Pan Z, Luo J, Le, J, Wen L, Tang, C. Uplink spatial modulation SCMA system. *IEEE Communications Letters*, 2018, 23(1):184-187.
6. Jibreel N, Elkawafi S, Younis A, Mesleh R. Performance analysis of sparse code multiple access with variant MIMO techniques. *Physical Communication*, 2020, 39:1-12.
7. Mheich Z, Hemadeh IA, Liu Z, Xiao P. Low-complexity expectation propagation detection for uplink MIMO-SCMA systems. *International Conference on UK-China Emerging Technologies (UCET)*, IEEE, 2020:1-4.
8. Tang S, Hao L, Ma Z. Low complexity joint MPA detection for downlink MIMO-SCMA. *IEEE Global Communications Conference (GLOBECOM)*, 2016:1-4.
9. Wu Z, Zhang C, Shen X, Jiao H. Low complexity uplink SFBC-based MIMO-SCMA joint decoding algorithm. *3rd IEEE International Conference on Computer and Communications (ICCC)*, IEEE, 2017:968-972.
10. Li HY, Liu ZJ, Yu QY, Xiang W, Adachi F. Uplink SCMA with STBC in Fading Channels. *IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, IEEE, 2020:1-5.
11. Pan Z, Liu W, Lei J, Luo J, Wen L, Tang C. Multi-dimensional space-time block coding aided downlink MIMO-SCMA. *IEEE Transactions on vehicular technology*, 2019, 68(7):6657-6669.
12. Kang GM, Kim HM, Shin Y, Shin OS. Message passing algorithm based on QR decomposition for an SCMA system with multiple antennas. *International Conference on Information and Communication Technology Convergence (ICTC)*, IEEE, 2017:941-944.
13. Wang P, Liu L, Zhou S, Peng G, Yin S, We, S. Near-optimal MIMO-SCMA uplink detection with low-complexity expectation propagation. *IEEE Transactions on Wireless Communications*, 2019, 19(2):1025-1037.

14. Alamouti SM. A simple transmit diversity technique for wireless communications. *IEEE Journal on selected areas in communications*, 1998 16(8):1451-1458.
15. Pham V B, Qi B Y, Sheng W X, Wang M. An improved full rate full diversity QOSTBC with linear decoding in MIMO systems. *Wireless personal communications*, 2013, 69(1):121-131.
16. Cai D, Fan P, Lei X, Liu Y, Chen D. Multi-Dimensional SCMA Codebook Design Based on Constellation Rotation and Interleaving. *IEEE 83rd Vehicular Technology Conference (VTC Spring)*, IEEE, 2016:1-5.

DOI: 10.17725/rensit.2022.14(3):301

Extraction of Noise Source Modes in a Shallow Sea Based on Holographic Interferometry in the Presence of Intense Internal Waves

Venedikt M. Kuz'kin

Prokhorov Institute of General Physics of RAS, <http://www.gpi.ru/>
Moscow 119991, Russian Federation

E-mail: kumiov@yandex.ru

Sergey A. Pereselkov, Elena S. Kaznacheeva, Sergey A. Tkachenko, Pavel V. Rybyanets

Voronezh State University, <http://www.vsu.ru/>

Voronezh 394006, Russian Federation

E-mail: pereselkov@yandex.ru, elena.kaznacheeva@yandex.ru, sega-th@mail.ru, rybyanets.edu@yandex.ru

Vladimir I. Grachev

Kotelnikov Institute of Radioengineering and Electronics of RAS, <http://www.cplire.ru/>

Moscow 125009, Russian Federation

E-mail: grachev@cplire.ru

Received 24.06.2022, peer-reviewed 01.07.2022, accepted 08.07.2022

Abstract: The results of a numerical experiment on approbation of the holographic method of mode selection in a shallow water area against the background of intense internal waves, which cause interaction of waves of a broadband sound source. An analysis of the resolution of modes and the restoration of their parameters has been carried out. It is shown that the effects of mode interaction do not lead to the loss of the identity of the mode characteristics of the signal.

Keywords: shallow sea, holographic interferometry, mode extraction, intense internal waves, mode interaction, moving sound source

UDC 004.052.34

Acknowledgments: This work was supported by the Russian Foundation for Basic Research (project no. 19-29-06075). The work of E.S. Kaznacheeva was supported by a grant from the President of the Russian Federation MK-6144.2021.4. The work of S.A. Tkachenko was supported by a grant from the President of the Russian Federation MK-4846.2022.4.

For citation: Venedikt M. Kuzkin, Sergey A. Pereselkov, Elena. S. Kaznacheeva, Vladimir I. Grachev, Sergey A. Tkachenko, Pavel V. Rybyanets. Mode extraction in a shallow sea based on holographic interferometry in the presence of intense internal waves. *RENSIT: Radioelectronics. Nanosystems. Information technologies*, 2022, 14(3):301-308. DOI: 10.17725/rensit.2022.14.301.

CONTENTS

1. INTRODUCTION (301)
 2. CHARACTERISTICS OF INTENSE INTERNAL WAVES (302)
 3. PHYSICAL FOUNDATIONS OF THE HOLOGRAPHIC METHOD OF MODE SELECTION (303)
 4. IDENTIFICATION OF MODES AGAINST THE BACKGROUND OF INTENSE INTERNAL WAVES (305)
 5. CONCLUSION (307)
- REFERENCES (307)

1. INTRODUCTION

The problem of separating individual propagating modes and determining their characteristics is important for many branches of ocean acoustics. For example, when localizing low-noise sources [1,2], when constructing low-mode acoustic tomography [3], when restoring the parameters of the bottom [4-7], when modeling the propagation of acoustic waves in methods of processing hydroacoustic signals consistent with the

propagation medium [8,9], when restoring the depth of the sound source [10,11], etc.

Determination of the mode composition of the acoustic field can be carried out both with the use of receiving antennas consisting of a large number of elements, and with the use of single receivers. Mode resolution methods using antennas are very complex and expensive to implement in practice. For this reason, more and more attention has recently been paid to single-receiver mode selection methods based on waveguide signal dispersion.

At present, for mode selection using single receivers, the Wigner transform [6,7] and the time-warping method [4,5] are mainly used, which allow restoring the group velocities of individual modes. The main advantage of the Wigner transform is the compensation of the intramode dispersion, which makes it possible to more accurately estimate the arrival times of the selected modes. The time-warping method is based on the use of the so-called sweeping transformation or deformation operator, which makes it possible to straighten dispersion curves and greatly simplify mode separation. Strictly, the time-warping method is applicable to an ideal waveguide, but under certain conditions it also approximately works in real water areas of a shallow sea. However, within the framework of the application of these two methods, the question of restoring the remaining mode parameters remains open.

To overcome this difficulty, a holographic method for selecting modes and estimating their parameters is proposed, based on the two-dimensional time-frequency Fourier transform of the field of a broadband moving source [12]. The reconstructed mode parameters are: amplitude, real (propagation constant) and imaginary (modal attenuation coefficient) components of the horizontal wave number, group velocity. This makes it possible to reconstruct the transfer function

of the waveguide in a parametric form and use this information to model the propagation of acoustic waves. The holographic method of mode selection [12] is an alternative to adaptive methods for localizing low-noise underwater sources [13,14], which do not require information about the characteristics of the oceanic medium.

In real shallow water waveguides, there are always space-time inhomogeneities, which can lead to the loss of identity of the mode parameters of individual modes. Therefore, for the practical application of mode selection methods, it is undoubtedly important to consider their performance in randomly inhomogeneous media. The issues of mode resolution under such conditions were not considered. An exception is the work [15], where the time-warping operator is generalized to the case of a waveguide with small smooth variations in the water area depth. In this paper, on the basis of numerical simulation, we present the results of approbation of the holographic method of mode selection against the background of intense internal waves, which determine the interaction of the waves of a broadband source.

2. CHARACTERISTICS OF INTENSE INTERNAL WAVES

In shallow water areas, the intense internal waves (IIW) are trains of intense short-term oscillations of the water surface of constant density, caused by internal tides, interpreted as trains of solitons that propagate towards the coastline. According to the experimental data, soliton trains are characterized by the following parameters: velocity $u \sim 0.5-1$ m/s, periods of calm $\delta L \sim 5-10$ km, length $L \sim 2-4$ km, period $D \sim 200-400$ m (distance between crests of adjacent solitons), half-width $\eta \sim 50-150$ m, amplitude $B \sim 10-30$ m [16,17]. Soliton trains are characterized by: a) anisotropy in the horizontal plane, radius of

curvature of the front $r = 15\text{-}25$ km; *b*) quasi-sinusoidality in the direction of propagation, i.e. spatial narrow spectrum; *c*) synchronism of vertical displacements in depth, which indicates the dominance of the first gravity mode. These factors determine the horizontal refraction and interaction of sound waves of the source if the path is located at small and large angles to the wave front of the soliton train, respectively.

3. PHYSICAL FOUNDATIONS OF THE HOLOGRAPHIC MODE SELECTION METHOD

Let a point source move in a horizontally homogeneous waveguide at a fixed depth z_s with a constant radial velocity w . We write the real component of the sound pressure of a broadband source at a horizontal distance r from a single receiver as the product of the signal spectrum $S(\omega) = |S(\omega)| \cos[i\varphi_s(\omega)]$ and the channel transfer function, presented as the sum of propagating modes

$$p(r, \omega) = S(\omega) \sum_m A_m(\omega, r) \exp[-\gamma_m(\omega)(r_0 + r)] \times \cos[ih_m(\omega)(r_0 + r)] = \sum_m p_m(r, \omega). \quad (1)$$

Here $\omega_1 \leq \omega \leq \omega_2$, $\omega_{1,2} = \omega_0 \pm (\Delta\omega/2)$, $\Delta\omega$ and ω_0 are the width and average frequency of the spectrum, $\omega = 2\pi f$ is the circular frequency; r_0 is the initial horizontal distance at time $t_0 = 0$; $A_m(\omega, r)$, $h_m(\omega)$, and $\gamma_m(\omega)$ are the amplitude, propagation constant, and modal damping coefficient of the m -th mode. The cylindrical discrepancy of the field, the source depth z_s and the receiver depth z_q are taken into account in the mode amplitude.

In relation (1), we pass from the distance variable r to the time variable t , $r = wt$, and apply the two-dimensional Fourier transform to it:

$$F(\tau, \tilde{\nu}) = \int_0^{\Delta t} \int_{\omega_1}^{\omega_2} p(t, \omega) \exp[i(\tilde{\nu}t - \omega\tau)] dt d\omega = \sum_m F_m(\tau, \tilde{\nu}), \quad (2)$$

where Δt is the observation time; τ and $\tilde{\nu} = 2\pi\nu$ are the time and circular frequency of the hologram; $F_m(\tau, \tilde{\nu})$ is the partial hologram of the m -th mode. The hologram (2) records the source field in all intermediate states that it successively passes during the observation time. In the linear approximation of the waveguide dispersion and phase spectrum of the signal, assuming $r_0 \gg w\Delta t$, the partial hologram is [12]

$$F_m(\tau, \tilde{\nu}) = |S(\omega_0)| A_m(\omega_0, r_0) \exp[-\gamma_m(\omega_0)r_0] \Delta\omega \Delta t \times \frac{\sin\left\{\left[r_0 \frac{dh_m(\omega_0)}{d\omega} + t_s - \tau\right] \frac{\Delta\omega}{2}\right\}}{\left[r_0 \frac{dh_m(\omega_0)}{d\omega} + t_s - \tau\right] \frac{\Delta\omega}{2}} \times \frac{\sin\left\{\left[wh_m(\omega_0) + \tilde{\nu}\right] \frac{\Delta t}{2}\right\}}{\left[wh_m(\omega_0) + \tilde{\nu}\right] \frac{\Delta t}{2}} \cos[i\Phi_m(\tau, \tilde{\nu})], \quad (3)$$

where $t_s = d\varphi_s(\omega_0)/d\omega$, phase

$$\Phi_m(\tau, \tilde{\nu}) = \left(\tilde{\nu} \frac{\Delta t}{2} - \tau\omega_0 + \varphi_s(\omega_0)\right) + h_m(\omega_0) \left(r_0 + w \frac{\Delta t}{2}\right). \quad (4)$$

The spectral density of the hologram (2) is localized in a small region in the form of focal spots corresponding to individual modes. The peak amplitude of the m -th focal spot is $\Delta\omega\Delta t$ times greater than the field amplitude of the m -th mode. Information about the propagation constants $h_m(\omega_0)$ and group velocities $u_{gm}(\omega_0) = d\omega/dh_m(\omega_0)$ of the modes is determined by the coordinates of the focal spot peak (3)

$$\tilde{\nu}_m = -wh_m(\omega_0), \quad \tau_m = \frac{r_0}{u_{gm}(\omega_0)} + t_s. \quad (5)$$

The estimates of the mode parameters, in contrast to their true values, are indicated by a dot above. The value of τ_m is interpreted as the time of arrival of the signal of the m -th mode.

According to the Rayleigh criterion, the resolvability of the maxima of neighboring focal spots is reduced to the simultaneous satisfaction of the inequalities

$$\Delta\omega r_0 \left| \frac{dh_m(\omega_0)}{d\omega} - \frac{dh_{m+1}(\omega_0)}{d\omega} \right| \geq 2\pi, \quad (6)$$

$$\Delta t |w| \| \mathbf{h}_m(\omega_0) - \mathbf{h}_{m+1}(\omega_0) \| \geq 2\pi. \quad (7)$$

To separate the spectral regions of neighboring focal spots, it suffices to satisfy one of these inequalities. The inequalities impose certain restrictions on the bandwidth, distance, observation time, and radial velocity. The fulfillment of inequalities (6), (7) becomes easier with an increase in the mode number. According to condition (6), the difference in the arrival times of adjacent mode signals must exceed the time $2\pi/\Delta\omega$. Condition (7) reduces to the requirement that the phase difference between adjacent modes during the observation time exceed 2π .

A hologram is an unambiguous representation of the source field. The inversion, carried out by the reverse application of the two-dimensional Fourier transform to the spectral densities of the selected focal spots, restores the fields of individual modes. Let us assume that the focal spots are resolved, i.e. one of the inequalities (6), (7) holds. We filter the spectral regions of focal spots with a two-dimensional filter

$$\theta_m(\tau, \tilde{\nu}) = \mu_m(\tau)\eta_m(\tilde{\nu}), \quad (8)$$

where

$$\mu_m(\tau) = \begin{cases} 1, & -(\Delta\tau_m/2) + \tau_m \leq \tau \leq \tau_m + (\Delta\tau_m/2), \\ 0, & \text{in other cases,} \end{cases} \quad (9)$$

$$\eta_m(\tilde{\nu}) = \begin{cases} 1, & -(\Delta\tilde{\nu}_m/2) + \tilde{\nu}_m \leq \tilde{\nu} \leq \tilde{\nu}_m + (\Delta\tilde{\nu}_m/2), \\ 0, & \text{in other cases.} \end{cases} \quad (10)$$

Here $\Delta\tau_m, \Delta\tilde{\nu}_m$ are the linear dimensions of the focal spots along the $\tau, \tilde{\nu}$ axes. Next, we perform the inverse two-dimensional Fourier transform. As a result, we obtain the reconstructed spectral densities $\dot{p}_m(t, \omega)$ of individual modes. If the initial distance

r_0 and the radial velocity w are known, then, according to (5), the group velocities and mode propagation constants are reconstructed from the measurements of the coordinates of the focal spot peaks. Otherwise, the initial distance r_0 and the radial velocity w are reconstructed by adaptive algorithms [13,14].

The reconstructed amplitude of the m -th mode at distance r and frequency ω_0 is equal to

$$\dot{A}_m(r, \omega_0) = \frac{|F_m(\tau_m, \tilde{\nu}_m, r)|}{|S(\omega_0)| \Delta\omega \Delta t}, \quad (11)$$

where $|F_m(\tau_m, \tilde{\nu}_m, r)|$ is the peak value of the m -th focal spot. The expression for the reconstructed modal damping coefficient, taking into account the cylindrical divergence of the mode amplitude, has the form

$$\dot{\gamma}_m(\omega_0) = -\frac{1}{\Delta r} \ln \left[\sqrt{r_2/r_1} g_m(r_1, r_2) \right], \quad r_2 > r_1. \quad (12)$$

Here

$$g_m(r_1, r_2) = |F_m(\tau_m, \tilde{\nu}_m, r_2)| / |F_m(\tau_m, \tilde{\nu}_m, r_1)|, \\ \Delta r = r_2 - r_1.$$

The mode resolution method is implemented as follows. During the observation time Δt , in the spectral band $\Delta\omega$, J independent realizations are accumulated with duration δt_1 and with time interval δt_2

$$J = \frac{\Delta t}{\delta t_1 + \delta t_2}. \quad (13)$$

A two-dimensional source field $p(t, \omega)$ is formed in time-frequency variables, and a two-dimensional Fourier transform is applied to it. Reconstruction of the modal attenuation coefficient is based on the recording of holograms for two different observation intervals. At the initial time $t_0 = 0$, processing (2) is implemented in the time interval Δt . At the time $t_1 = t_0 + \delta T$, the next processing is carried out during the time Δt . Thus, holograms are formed for two different observation intervals. In this case, the coordinates of the focal spot peaks shift along the time axis, while remaining constant along the frequency axis.

4. SELECTION OF MODES AGAINST THE BACKGROUND OF INTENSE INTERNAL WAVES

In the framework of numerical simulation, we consider the performance of the described method of mode resolution in the case of a waveguide with IIW causing mode interaction.

The depth distribution of sound speed in an undisturbed waveguide is shown in Fig. 1. Parameters of the liquid absorbing bottom: density ratio of soil and water $\rho = 1.8$, complex refractive index $n = 0.849(1 + i0.03)$. The frequency range $f = 100-120$ Hz, the sound field was formed by three modes. The propagation constants h_m and group velocities u_{gm} of modes for an unperturbed waveguide at the middle frequency of the spectrum $f_0 = 110$ Hz are given in Table 1. The scheme of the numerical experiment is shown in Fig. 2.

Let us consider the cases of a noise source stationary and moving at a speed $w = 1$ m/s towards a single source. The signal spectrum is assumed to be uniform, and the phase is

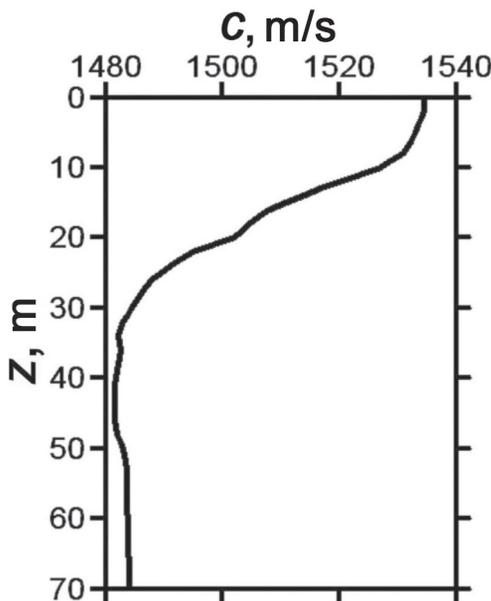


Fig. 1. Sound velocity profile.

Table 1

Mode parameters of the unperturbed waveguide at the frequency $f_0 = 110$ Hz.

Mode numbers, m	1	2	3
h_m, M^{-1}	0.4635	0.4557	0.4450
$u_{gm}, M/c$	1478.76	1468.75	1448.98

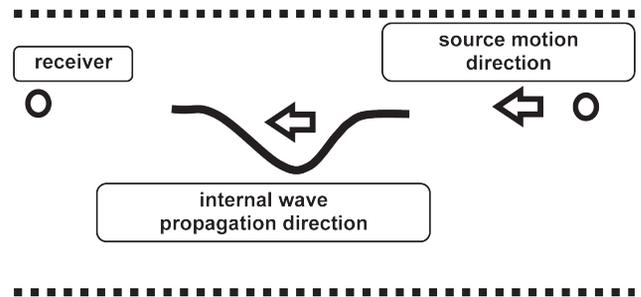


Fig. 2. Geometry of the numerical experiment.

constant. The source is located at a depth of $z_s = 20$ m, and the receiver is at a depth of $z_q = 45$ m. A single soliton moving in the direction from the receiver towards the source was used as the IIW model. Soliton parameters: amplitude $B = 15$ m, half-width $\eta = 150$ m, velocity $u = 0.7$ m/s. At the initial time $t_0 = 0$: the distance between the source and the receiver is $r_0 = 10$ km, the distance between the soliton and the receiver is $r_* = 2.5$ km.

At the initial time $t_0 = 0$, time-frequency processing (2) of mode selection was implemented in the time interval $\Delta t = 20$ s, the duration of the noise signal realization $\delta t_1 = 0.75$ s (frequency step $\delta f = 1.4$ Hz), the time interval between realizations $\delta t_2 = 0.25$ s, the number of time intervals $J = 20$. The processing was repeated for subsequent time intervals of the same duration, but shifted relative to the previous one by a fixed time interval $\delta T = 1$ min. Processing time $T = 20$ min 20 s. During the experiment T with a frequency of $\delta T = 1$ min for time points $t_i = 0, 1, \dots, 19$ min, $N = 20$ holograms were formed. During the processing time, the soliton traveled the distance $\Delta r = 854$ m. The sound field at the receiving point was calculated using the modal approach of mode interaction.

In the case of a stationary source, the processing results for three time points $t_i = 0, 10, 19$ min are shown in Fig. 3-5. The signal envelope was reconstructed by applying the inverse two-dimensional Fourier transform to the hologram and making a horizontal section of the reconstructed field at a frequency $f_0 = 110$ Hz. On the holograms, the coordinates

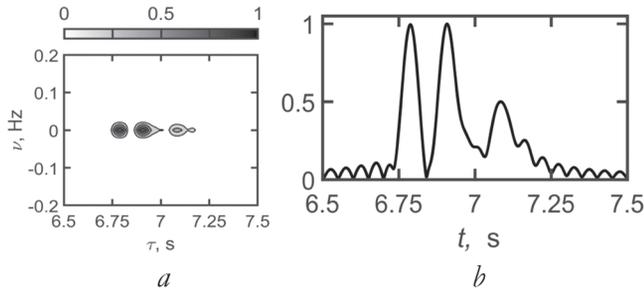


Fig. 3. Normalized hologram (a) and signal envelope (b). Processing start: time $t_0 = 0$. The source is stationary.

of the focal spot maxima are located on the time axis. The mode group velocities u_{gm} were determined from the coordinates of the maxima τ_m of the focal spots (5) or from the positions of the maxima t_m of the signal envelope: $\dot{u}_{gm}(\omega_0) = r_0 / \tau_m$ or $\dot{u}_{gm}(\omega_0) = r_0 / t_m$. When the source is stationary, the mode propagation constants are not restored.

According to the results of processing for twenty time points $t_i = 0, 1, \dots, 19$ min:

- a) relative variations of the group velocities of the modes with respect to the model values – $0.9955 \leq \dot{u}_{g1} / u_{g1} \leq 0.9997$, $0.9848 \leq \dot{u}_{g2} / u_{g2} \leq 0.9860$, $0.9739 \leq \dot{u}_{g3} / u_{g3} \leq 0.9751$;

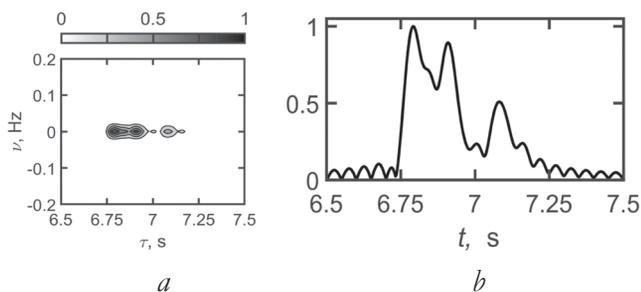


Fig. 4. Normalized hologram (a) and signal envelope (b). Time $t_{10} = 10$ min. The source is stationary.

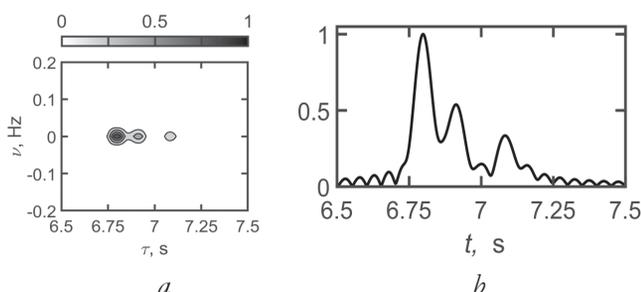


Fig. 5. Normalized hologram (a) and signal envelope (b). End of processing: time $t_{19} = 19$ min. The source is stationary.

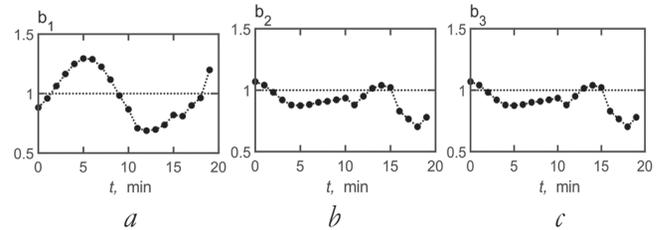


Fig. 6. Dependences of the relative variations of the reconstructed mode amplitudes on the processing time: (a) the first mode, (b) the second mode, (c) the third mode. The source is stationary.

- b) maximum relative errors of the reconstructed group velocities of the modes – the first mode, $m = 1$, $\max \delta_1 = 0.45\%$; second mode, $m = 2$, $\max \delta_2 = 1.52\%$; third mode, $m = 3$, $\max \delta_3 = 2.61\%$.

The relative variations of the mode amplitudes $b_m(t_i) = \dot{A}_m(t_i) / A_m$ are shown in **Fig. 6**, where $\dot{A}_m(t_i)$ and A_m are the amplitude of the m -th mode, reconstructed at time t_i , and its model value for the unperturbed waveguide, respectively.

In the case of a moving source, the processing results for three time points $t_i = 0, 10, 19$ min are shown in **Fig. 7-9**. The mode group velocities u_{gm} were determined from the coordinates of the maxima τ_m of the focal spots (5) or from the positions of the maxima t_m of the signal envelope: $\dot{u}_{gm}(t_i, \omega_0) = r(t_i) / \tau_m(t_i)$ or $\dot{u}_{gm}(t_i, \omega_0) = r(t_i) / t_m(t_i)$. Here $r(t_i)$ is the distance between the source and the receiver at time t_i . The mode propagation constants were determined from the coordinates of the maxima $\nu_m(t_i)$ of focal spots (5).

According to the results of processing for twenty time points $t_i = 0, 1, \dots, 19$ min, it was established:

- a) relative variations of the group velocities of the modes with respect to the model values – $0.9977 \leq \dot{u}_{g1} / u_{g1} \leq 1.0049$, $0.9904 \leq \dot{u}_{g2} / u_{g2} \leq 0.9945$, $0.9750 \leq \dot{u}_{g3} / u_{g3} \leq 0.9835$;

- b) maximum relative errors of the reconstructed group velocities of the modes – the first mode, $m = 1$, $\max \delta_1 = 0.49\%$; second

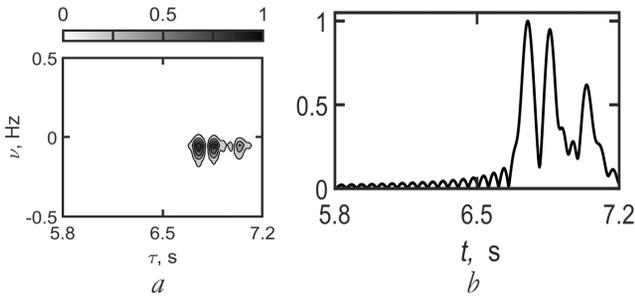


Fig. 7. Normalized hologram (a) and signal envelope (b). Processing start: time $t_0 = 0$. The source is moving.

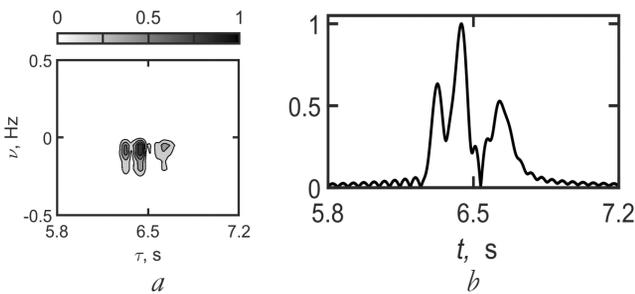


Fig. 8. Normalized hologram (a) and signal envelope (b). Time $t_{10} = 10$. The source is moving.

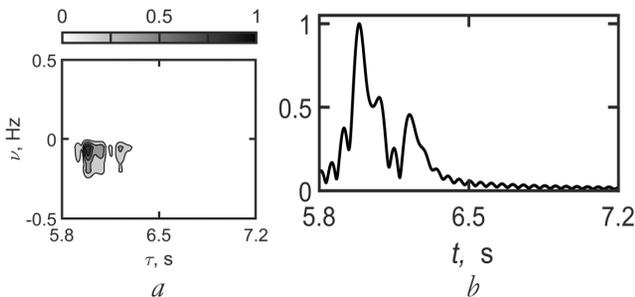


Fig. 9. Normalized hologram (a) and signal envelope (b). End of processing: time $t_{19} = 19$. The source is moving.

mode, $m = 2$, $\max \delta_2 = 0.96\%$; third mode, $m = 3$, $\max \delta_3 = 2.50\%$;

c) relative variations of the mode propagation constants with respect to the model values – $0.8923 \leq \dot{h}_1 / h_1 \leq 1.0705$, $0.7382 \leq \dot{h}_2 / h_2 \leq 0.9914$, $0.7229 \leq \dot{h}_3 / h_3 \leq 1.0081$.

The relative variations of the mode amplitudes $b_m(t_i) = \dot{A}_m(t_i) / A_m$ are shown in

Fig. 10.

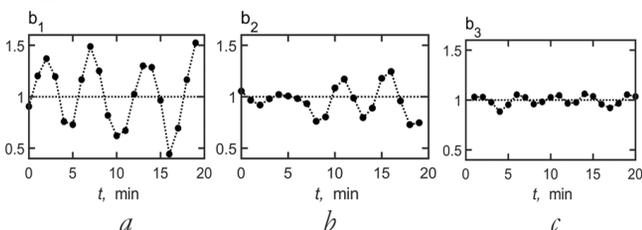


Fig. 10. Dependences of the relative variations of the reconstructed mode amplitudes on the processing time: (a) first mode, (b) second mode, (c) third mode. The source is moving.

5. CONCLUSION

The data of numerical simulation of the selection of modes and the reconstruction of their parameters against the background of IIW, which cause the interaction of the modes of the source field, are presented. The cases of stationary and moving sources are considered. It is shown that in both cases there is no loss of identity of the mode parameters: the modes are resolved and their parameters are restored. Variations in the amplitudes, propagation constants, and mode group velocities were determined from the results of a numerical experiment. It has been established that in the low-frequency range the relative error in reconstructing the group velocities of the modes with respect to the unperturbed waveguide does not exceed one percent.

REFERENCES

1. Kuznetsov GN, Kuzkin VM, Pereselkov SA. Spectrogram and localization of a sound source in a shallow sea. *Acoustical Physics*, 2017, 63(4):449-461. DOI: 10.1134/S1063771017040078.
2. Kaznacheev I.V., Kuznetsov G.N., Kuzkin V.M., Pereselkov S.A. Interferometric method for detecting a moving sound source with a vector-scalar receiver. *Acoustical Physics*, 2018, 64(1):37-48. DOI: 10.1134/S1063771018010104.
3. Kuzkin VM, Pereselkov SA. *Interferometric diagnostics of hydrodynamic disturbances in a shallow sea*. Moscow, Lenand Publ., 2019, 200 p.
4. Bonnel J, Chapman NR. Geoacoustic inversion in a dispersive waveguide using warping operators. *J. Acoust. Soc. Am.*, 2011, 130:EL101-EL107.
5. Zeng J, Bonnel J, Chapman NR. Inversion of seabed attenuation using time-warping of close range data. *J. Acoust. Soc. Am.*, 2013, 134:EL394-EL399.

6. Belov AI, Kuznetsov GN. Estimation of the acoustic parameters of a bottom model in a shallow sea using a priori geological and geophysical information and the Wigner transformation. *Acoustical Physics*, 2014, 60(2):190-195.
7. Belov AI, Kuznetsov GN. Evaluation of the acoustic characteristics of the surface layers of the seabed using four-component vector-scalar receivers. *Acoustical Physics*, 2016, 62(2):194-202.
8. Baggeroer AB, Kuperman WA, Mikhalevsky PN. An overview of matched field methods in ocean acoustics. *IEEE J. Oceanic Eng.*, 1993, 18(4):401-423.
9. Sazontov AG, Malekhanov AI. Consistent spatial signal processing in underwater sound channels (review). *Acoustical Physics*, 2015, 61(2):233-253.
10. Besedina TN, Kuznetsov GN, Kuz'kin VM, Pereselkov SA, Prosovetskiy DYU. Estimation of the depth of an immobile sound source in shallow water. *Phys. wave phenom.*, 2015, 23(4):292-303.
11. Kuznetsov GN, Kuz'kin VM, Pereselkov SA, Prosovetskiy DYU. Wave method for estimating the sound source depth in an oceanic waveguide. *Phys. wave phenom.*, 2016, 24(4):310-316.
12. Kuz'kin VM, Matvienko YuV, Pereselkov SA, Prosovetskiy DYU, Kaznacheeva ES. Mode selection in oceanic waveguides. *Phys. wave phenom.*, 2022, 30(2):111-118.
13. Kaznacheeva ES, Kuznetsov GN, Kuz'kin VM, Lyakhov GA, Pereselkov SA. Measurement capability of the interferometric method of sound source localization in the absence of data on the waveguide transfer function. *Phys. wave phenom.*, 2019, 27(1):73-78.
14. Kaznacheeva ES, Kuz'kin VM, Lyakhov GA, Pereselkov SA, Tkachenko SA. Adaptive algorithms for interferometric processing. *Phys. wave phenom.*, 2020, 28(3):267-273.
15. Fershalov MYu, Petrov PS, Manulchev DS, Zakharenko AD. Generalization of the method of geoacoustic inversion by recording a pulsed signal with a single hydrophone, taking into account bathymetry inhomogeneities. *Underwater research and robotics*, 2021, 1(35):51-59.
16. Zhou J, Zhang XZ, Rogers PH. Resonant interaction of sound wave with internal solitons in the coastal zone. *Acoust. soc. Am.*, 1991, 90(4):2042-2054.
17. Hsu MK, Liu AK, Liu C. An study of internal waves in the China seas and yellow sea using SAR. *Continental Shelf Res.*, 2000, 20(4-5):389-410.

DOI: 10.17725/rensit.2022.14.309

Light Diffraction by Flat Geometric Bifractals

Galina V. Arzamastseva, Mikhail G. Evtikhov, Fedor V. Lisovsky, Ekaterina G. Mansvetova

Kotelnikov Institute of Radioengineering and Electronics of RAS, Fryazinsky Branch, <http://fire.relarn.ru/>
Fryazino 141190, Moscow Region, Russian FederationE-mail: arzamastseva@mail.ru, emg20022002@mail.ru, lisf@df.ru, mansvetova_eg@mail.ru

Received May 26, 2022, peer-reviewed June 6, 2022, accepted June 13, 2022

Abstract: In the Fraunhofer zone, an experimental study of the diffraction of a collimated beam of light with a wavelength of 0.63 microns after passage through computer-generated images of flat geometric bifractals, which are a combination of two geometric fractals of different dimensions, was performed. The studies used bifractals based on the classical Vicsek fractal and two less well-known fractals of L-systems. The selected images were transferred to a transparent film using a phototypesetter with a resolution of 1333 dots per centimeter (3386 dpi) and a point size of 7.5 microns. Diffraction patterns were visually observed on the screen and recorded using a digital camera and then transmitted to a computer for processing. The diffraction patterns observed in optical experiments were compared with "digital" diffractograms, that is, with Fourier images of bifractal images approximated by a grid function on a uniform square grid at different values of the parameter used in calculations that determines the ratio between the overall size of the smallest element of the prefractal and the grid period.

Keywords: bifractal, diffraction, Fraunhofer zone, monofractal, prefractal, grid function, Vicsek fractal, L-system fractal, Fourier image

UDC 51.74; 535.4

Acknowledgments: The work was carried out within the framework of the state task.

For citation: Galina V. Arzamastseva, Mikhail G. Evtikhov, Fedor V. Lisovsky, Ekaterina G. Mansvetova. Light diffraction by flat geometric bifractals. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2022, 14(3):309-316. DOI: 10.17725/rensit.2022.14.309.

CONTENTS

1. INTRODUCTION (309)
 2. METHODS OF CONDUCTING EXPERIMENTS (310)
 3. METHODS OF CONSTRUCTING BIFRACTALS AND DESCRIPTION OF EXPERIMENTAL RESULTS (311)
 4. CONCLUSION (315)
- REFERENCES (316)

1. INTRODUCTION

Previously, we proposed a method for constructing geometric bifractals based on the use of the "fractal-carrier – carryable prefractal" algorithm [1]. A number of bifractals were constructed, which are a

pairwise combination of four geometric monofractals (the Sierpinsky carpet or the Vicsek fractal and two L-system monofractals with an axiom in the form of a square), their digital megapixel pictures were obtained, and their digital Fourier images were calculated, the comparison of which with the corresponding diffraction patterns observed in the Fraunhofer region, in [1], however, was not performed due to the lack of experimental data. In this paper, which was carried out in order to eliminate this gap, the results of an experimental study of light diffraction for two of the bifractals mentioned in [1] are presented, in the first of which the carrier is

the Vicsek fractal, and the carryable is the LS1 fractal of the L-system (hereinafter referred to as bifractal No. 1), and in the second the carrier is the fractal L-system LS2, and the carryable is the LS1 fractal (hereinafter referred to as bifractal No. 2). General information and description of the fractals mentioned above can be found, for example, in [2-5].

To obtain Fourier spectra, we used the computer modeling method previously described by us in [6], which was first applied to real thermodynamically stable fractal-like domain structures [7] and was finally worked out on test objects, as which ordered domain structures realized in thin magnetic films were selected, allowing diffraction "to the lumen" [8]. The method was also applied in the study of abstract fractal objects: Koch's snowflake, Sierpinsky carpet, Vicsek fractal, and others [9-13].

2. METHODS OF CONDUCTING EXPERIMENTS

For experimental research of diffraction by fractals, black-and-white bitmap pictures of selected objects were created on a computer using specially developed programs. The resulting pictures were printed by imagesetter on a transparent film with a resolution of 1333 dots per centimeter (3386 dpi) and a point size of 7.5 microns. The smallest fractal element was formed from 4-10 points, that is, its linear size ranged from 30 to 75 microns. The images obtained by the described method had an extremely high contrast, almost unattainable for traditional photography.

A standard technique was used to observe the diffraction pattern formed after the passage of a light beam through a transparent film with an picture of the prefractal diffraction pattern in the Fraunhofer zone. A narrow beam of light with a wavelength of 0.63 microns from a helium-neon laser was expanded and collimated

to a diameter of 5 to 8 cm using a system of confocal lenses, after that it was directed to a film depicting a fractal object whose overall dimensions did not exceed the diameter of the collimated light beam. The picture of the diffraction pattern was formed by a lens in the diffraction plane where the screen was located. Instead of a screen, one can use a digital camera paired with a computer and subject the observed diffraction patterns to the necessary processing.

The computer-generated black-and-white bitmap pictures of the prefractals were approximated by a grid function on a uniform square grid with a number of nodes $n_1 \times n_2$, where the values n_1 and n_2 were chosen sufficiently large (up to 8192) to adequately approximate the details of the smallest size of the prefractal (in computer representation) and to enable the study of prefractals with high generation numbers. For the image digitized in this way, using the fast Fourier transform, the values of the square of the modules of the Fourier components were determined, that is, the spectral distribution of the intensity of diffracted radiation I in the Fraunhofer zone. To display the intensity of diffraction maxima on the plane, circles with a radius proportional to the intensity (or logarithm of intensity) were used, where the proportionality coefficient was chosen for reasons of obtaining optimal clarity of images; for the same purposes, Gaussian blurring of circles was additionally used. The degree of correspondence between the calculated Fourier images and the experimentally observed diffraction patterns depends on the value of the parameter p , equal to the ratio of the smallest element of the fractal to the mesh cell size. The larger the p , the better the match. In our calculations, the value of p was from 4 to 8.

3. METHODS OF CONSTRUCTING BIFRACTALS AND DESCRIPTION OF EXPERIMENTAL RESULTS

As it was indicated in the introduction, that three geometric monofractals were used for the construction of bifractals: a fractal (snowflake) of a Vicsek [2] and two monofractals of the L-system with an axiom ("seed") in the form of a square [4,5] (hereinafter referred to as the LS1 and LS2 fractals). The first fractal has a scaling factor $m = 3$, and theoretical value of dimension $D_f = \ln 5 / \ln 3 = 1.465$ the last two have $m = 2$ and $D_f = \ln 3 / \ln 2 = 1.585$. For each of the listed monofractals, megapixel black-and-white bitmap images were created, according to which Fourier images were located.

Fig. 1a shows an image of the 3rd generation of the Vicsek prefractal, and **Fig. 1b** (close-up, fractal part) and **Fig. 1c** (general view) – digital diffractograms of this 5th generation prefractal for $p = 8$. Experimental diffractogram for the 5th generation prefractal, presented at two different scales in **Fig. 1d** (close-up, fractal part) and **Fig. 1e** (general view), agree well with the calculated for $p = 8$.

Algorithms for constructing two monofractals LS1 and LS2 with an axiom in the form of a quadrangle can be described on the basis of a standard approach for L-systems using a set of rules "turtle graphics", however, in our case, the representation is simpler using the recurrence relations given below, displaying sequential transformations of a single seed a square on the complex plane. Note that this uses transformations using only integer translations of objects along the real and (or) imaginary axes, as well as rotations only by multiples of angles $\pi/2$.

If we choose the orientation of the coordinate system on the complex plane in such a way that the imaginary axis is directed horizontally to the right, and the real axis is directed vertical down, and place the unit seed

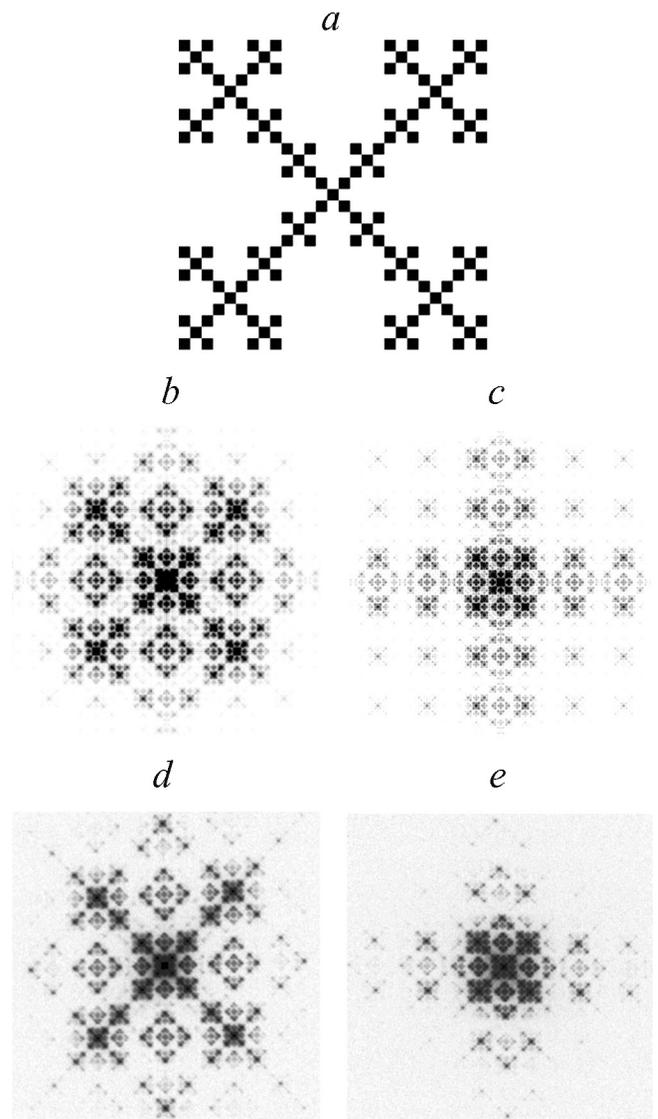


Fig. 1. Image of the 3rd generation Vicsek prefractal (a), digital diffractograms for the same type prefractal of the 5th generation with $p = 8$ (b – close-up, c – general view) and experimental diffractograms (d – close-up, e – general view).

square (set $Z^{(0)}$) in the first quadrant (coordinates of vertices $(0,0)$, $(0,i)$, $(1,i)$ and $(1,0)$), then LS1 prefractals of any order (generation) $Z^{(n)}$ are obtained by the operation of combining sets using recurrent relations

$$Z^{(n+1)} = Z^{(n)} \cup (iZ^{(n)} + (1+i)2^n) \cup (-iZ^{(n)} + (1+i)2^n) \quad (1)$$

As an example in **Fig. 2a** shows an image of the LS1 fractal of the 4th order resulting from the application of the described procedure, corresponding to the 5th iteration according to formula (1). The calculated for $p = 8$ Fourier

images for a similar fractal of the 10th order are shown in Fig. 2*b* (fractal part, close-up) and Fig. 2*c* (general plan), and experimentally obtained - in Fig. 2*d* (fractal part, close-up) and Fig. 2*e* (general plan). Digital diffractograms for the LS1 fractal correspond well to experimental ones.

For the LS2 fractal, the chain of successive transformations of the set in the form of a single seed square is given by the expressions $z^{(n+1)} = z^{(n)} \cup (-z^{(n)} + (2+i)2^n) \cup (-z^{(n)} + (1+2i)2^n)$. (2)

For example at $n = 5$, we obtain a prefractal of the 4th order, shown in Fig. 3*a*; diffractograms of the same type of fractal of the 10th order (11th iteration in (2)) for $p = 8$ are shown in Fig. 3*b* (fractal part, close-up) and Fig. 3*c* (general plan). They correspond well to the experimentally obtained Fourier images shown in Fig. 3*d* (fractal part, close-up) and Fig. 3*e* (general plan).

The diffraction patterns, both calculated and experimentally obtained, show that the

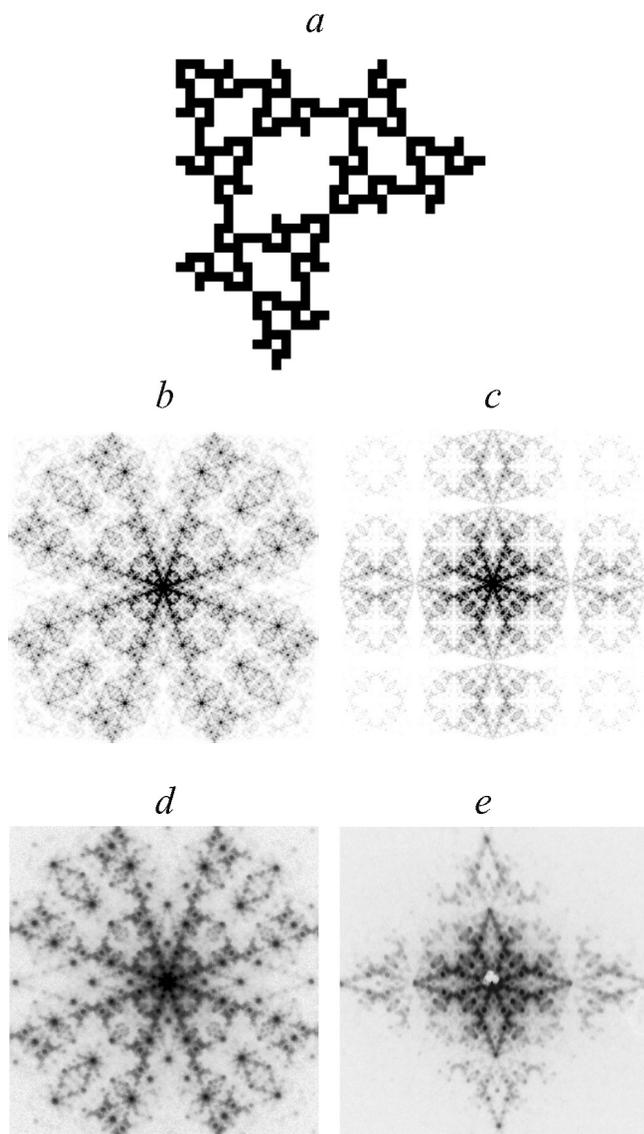


Fig. 2. Image of the LS1 prefractal of the 4th generation (a), digital diffractograms for the same type of fractal of the 10th generation with $p = 8$ (b – close-up, c – general view) and experimentally obtained diffractograms (d – close-up, e – general view).

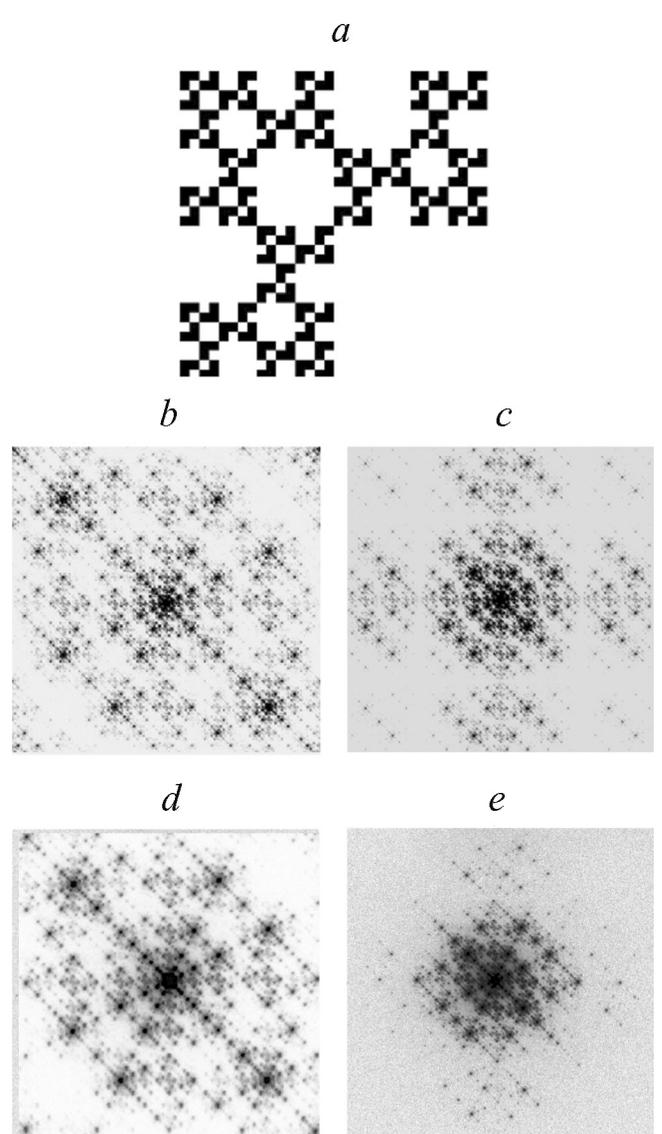


Fig. 3. Image of the LS2 prefractal of the 4th generation (a), digital diffractograms for the same type of fractal of the 10th generation with $p = 8$ (b – close-up, c – general view) and experimentally obtained diffractograms (d – close-up, e – general view).

fractal part has self-similarity with the scaling factor $m = 3$ for the Vicsek fractal (Fig. 1*b*, Fig. 1*d*) and $m = 2$ for LS1 fractals (Fig. 2*b*, Fig. 2*d*) and LS2 (Fig. 3*b*, Fig. 3*d*).

The bifractals LS1 and LS2 would be formed from two simple geometric monofractals, in one of which (the carrier fractal) the other fractal (the carryable fractal) stood as a seed object. The order of the bifractal obtained in this case is characterized by two indices $k_1 \times k_2$, where k_1 is the order of the carrier fractal, and k_2 is the order of the carryable prefractal. This procedure is fundamentally different from the long-used procedure for creating bifractals (both geometric and "signal") by simply combining two sets.

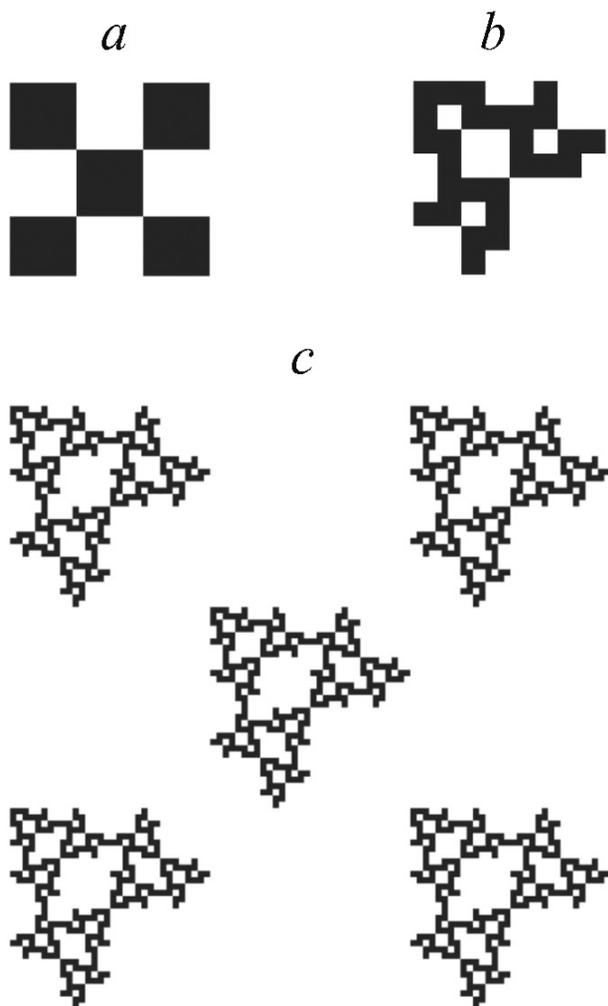


Fig. 4. The procedure for constructing bifractal No. 1; *a* – the first generation of Vicsek fractal, *b* – the second generation of LS1 fractal, *c* – bifractal No. 1 of 1×4 order.

The procedure for constructing bifractal No. 1 is shown in **Fig. 4**; the first orders of the Vicsek and LS1 fractals are shown in Fig. 4*a* and Fig. 4*b*, respectively. If the black squares in Fig. 4*a* is replaced by a prefractal LS1 of the 4th order, then you get a bifractal No. 1 of the 1×4 row, the image of which is shown in Fig. 4*c*. In **Fig. 5** shows the diffractograms calculated for different scales for : general plan (Fig. 5*a*), medium (Fig. 5*b*) and large (Fig. 5*c*). Experimental diffractograms are shown in Fig. 5*d* (general plan), Fig. 5*e* (medium plan) and Fig. 5*f* (close-up); they correspond well to the numerical one.

It can be seen that the fine plan of the Fourier image of this bifractal (when considering the "afar"), medium plan on the calculated (Fig. 5*a* and Fig. 5*b*) and experimentally obtained (Fig. 5*d* and Fig. 5*e*) diffractograms generally corresponds to the Fourier image of the LS1 fractal (see Fig. 2*b*-Fig. 2*e*) with scaling factor $m = 2$. But at the same time on the diffractograms of the medium plan (Fig. 5*b* and Fig. 5*e*) the elements characteristic of the diffraction pattern for the Vicsek fractal are already visible (see Fig. 1*b* and Fig. 1*d*). With a further increase in the center of the diffraction pattern (close-up, Fig. 5*c* and Fig. 5*f*) it becomes similar to the Fourier image of the Vicsek

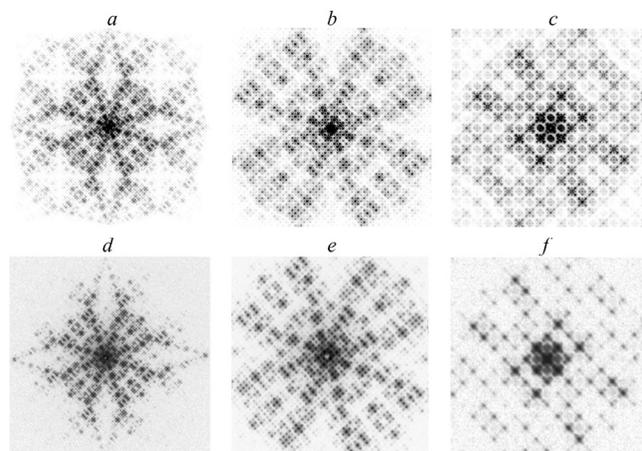


Fig. 5. Digital diffractograms for bifractal No. 1 for $p = 4$ (*a* – general plan, *b* – medium, *c* – large) and experimental diffractograms (*d* – general plan, *e* – medium, *f* – large).

fractal (see Fig. 1*b* and Fig. 1*d*) with a scaling factor $m = 3$. At an intermediate scale (between the scales in Fig. 5*b* (Fig. 5*e*) and Fig. 5*c* (Fig. 5*f*)) self-similarity on diffractograms, calculated and experimental, is lost.

In addition, the entire background of the diffraction pattern is "modulated" in accordance with the pattern for the Vicsek fractal. This phenomenon becomes noticeable at a sufficiently large magnification for any (not only the central) part of the diffraction pattern. Thus, the periphery and rough features of the Fourier image of bifractal No. 1 are called the LS1 fractal, and the central part and "subtle" features are called the Vicsek fractal. With the bifractal itself, the situation is reversed: when viewed "from afar" (small plan), the observer will see the Vicsek fractal, and close up – the LS1 fractal. Thus, both the bifractal itself and its Fourier image are scale-dependent. These features of objects of the "fractal-carrier – carryable fractal" type are typical and were observed in all the bifractals considered in this paper.

The properties of bifractals with the same values of fractal dimension D and scaling coefficients m for the carrier and the carryable object were also studied. In particular, this was done for bifractal No. 2, using as such objects the fractals LS2 and LS1 having the same values $D = 1.585$ and $m = 2$. In Fig. 6*a* and Fig. 6*b* shows the second generations used in the construction of fractals, in Fig. 6*c* is the resulting bifractal with size 1×4 . Fig. 7 shows the calculated for $p = 4$ diffractograms at different scales: general plan (Fig. 7*a*), medium (Fig. 7*b*) and large (Fig. 7*c*). The experimentally obtained diffractograms are shown in Fig. 7*d* (general plan), Fig. 7*e* (medium plan) and Fig. 7*f* (close-up), which agree well with the calculated ones. Diffractograms in general and medium-scale scales, both calculated and

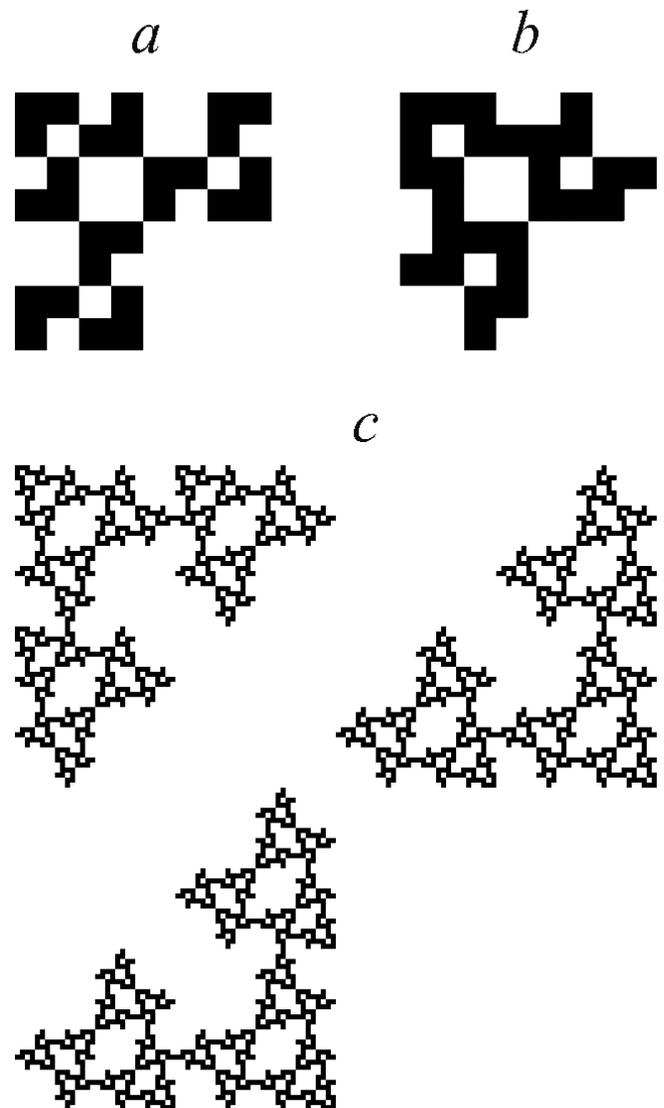


Fig. 6. The procedure for constructing bifractal No. 2; *a* and *b* are the second orders of fractals LS2 and LS1, respectively, *c* is bifractal No. 2 of 1×4 order.

experimentally obtained, correspond well to those for the LS1 fractal (Fig. 2*b* - Fig. 2*e*) and have self-similarity with a scaling factor $m = 2$. The central part of the diffractograms (close-up, Fig. 7*c* and Fig. 7*f*) is similar to the Fourier image of the LS2 fractal (Fig. 3*b* and Fig. 3*d*) and also has self-similarity with a scaling factor $m = 2$. Although both the wearable fractal and the carrier fractal have the same scaling coefficient, at some scale intermediate between the scales for medium and small plans, the diffraction pattern loses its self-similarity due to a change in the algorithm for constructing the bifractal.

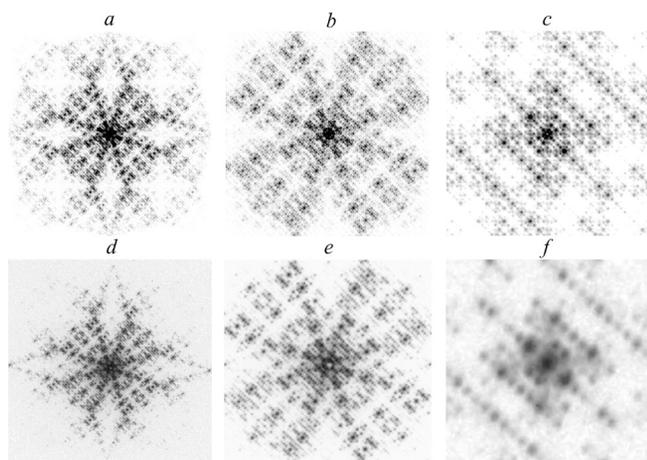


Fig. 7. Digital diffractograms for bifractal No. 2 for $p = 4$ (a – general plan, b – medium, c – large) and experimental diffractograms (d – general plan, e – medium, f – large).

For all the studied geometric fractals, the difference between the central and peripheral parts is observed on the digital diffractograms and diffraction patterns experimentally obtained by the optical method. The part localized near the center has the self-similarity inherent in fractal objects (see Fig. 1b and Fig. 1d, Fig. 2b and Fig. 2d, Fig. 3b and Fig. 3d), while the peripheral part, which does not possess self-similarity, is characterized by an equidistant arrangement of diffraction maxima along a certain number of radial directions (rays) (see, for example, Fig. 1c and Fig. 1e). For the first time, this fact was noticed by the authors of the work [14] devoted to the study of Fraunhofer optical diffraction on the classical Koch snowflake, where they used the terms "fractal part" and "lattice part" to denote the central and peripheral parts and associated the appearance of the latter with the fact that in a two-dimensional set of elements forming a fractal, one-dimensional diffraction gratings consisting of identically oriented elements (e.g. segments) can be distinguished.

For bifractals, this effect also takes place. The rough pattern of the lattice part of

the bifractal diffractogram is determined by the carryable fractal, but, in addition, it (like the entire diffraction pattern) is modulated by a carrier fractal. The rough drawing of the periphery of the fractal part is also determined by the carryable fractal, and only the very center of the diffraction pattern corresponds to the diffractogram of the carrier fractal; the size of this part depends on the number of orders in the carrier fractal and the ratio of orders in both fractals.

4. CONCLUSION

A careful comparison of the diffraction patterns observed in experiments for bifractals based on the classical Vicsek fractal and two fractals of the L-system with Fourier images of bifractals approximated by a grid function on a uniform square grid allowed us to formulate certain requirements for the approximation process. It has been shown that almost complete correspondence of experimental and digital diffractograms and good reproduction of the fine structure of the former is achieved with a sufficiently high ratio of the size of the smallest element of the bifractal to the period of the grid function.

All the obtained Fourier images of bifractals, both calculated and experimentally observed, were scale-dependent. The diffractograms of the small plan ("far") of the lattice and fractal parts corresponded to the wearable fractal, the close-up ("near") – to the carrierable fractal. At some intermediate scale, self-similarity was violated. All parts of the diffraction pattern, both central and peripheral, had a fine structure modulated in accordance with the Fourier image of the wearable fractal, which could be seen with sufficient magnification.

REFERENCES

1. Arzamastseva GV, Evtikhov MG, Lisovsky FV, Mansvetova EG. Svoystva ploskih geometric bifractals. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2012, 4(2):93-107 (in Russ.).
2. Vicsek T. Fractal model for diffusion controlled aggregation. *J. Phys. A*, 1983, 16(17):1647-1652.
3. Loskutov AYu, Mihaylov AS. *Osnovy teorii slozhnykh sistem*. Moscow-Izhevsk, Institut comp'yuternykh issledovaniy Publ., 2007, 620 s (in Russ.).
4. Linder Mayer A. Mathematical model for cellular interaction in development. I. Filaments with one-sided inputs. *J. Theor. Biol.*, 1968, 18(3):280-299.
5. Linder Mayer A. Mathematical model for cellular interaction in development. II. Simple and branching filaments with two-sided inputs. *J. Theor. Biol.*, 1968, 18(3):300-315.
6. Arzamastseva GV, Evtikhov MG, Lisovsky FV, Lukashenko LI. Comp'yuternoe modelirovanie difraktsii na fractal'nyh domennykh strukturah. *Tr. XIX Merzdunarod. Shkoly-seminara "Novye magnitnye materialy microelektroniki"*, Moscow, 2004, c. 632-634 (in Russ.).
7. Lisovsky FV, Lukashenko LI, Mansvetova EG. Termodinamicheski ustoychivye fractalopodobnye domennye struktury v magnitnykh plenках. *Pis'ma v ZHETF*, 2004, 79(7):432-435 (in Russ.).
8. Arzamastseva GV, Evtikhov MG, Lisovsky FV, Mansvetova EG, Temiryazeva MP. Amorfizatsiya biperiodicheskikh domennykh struktur v kvaziodnoosnykh magnitnykh plenках kriticheskoy tolshchiny. *ZHETF*, 2008, 134(2):282-290 (in Russ.).
9. Arzamastseva GV, Evtikhov MG, Lisovsky FV, Mansvetova EG. Fur'e obrazy fraktal'nyh ob'ektov. *Izv. RAN, Ser. fiz.*, 2010, 74(10):1430-1432 (in Russ.).
10. Arzamastseva GV, Evtikhov MG, Lisovsky FV, Mansvetova EG. Comp'yuternoe modelirovanie difraktsii Fraungofera na H-fraktalakh i krivykh Peano. *Electromagnitnye volny i elektronnyye sistemy*, 2012, 17(7):48-58 (in Russ.).
11. Arzamastseva GV, Evtikhov MG, Lisovsky FV, Mansvetova EG. Family of Generalized Triadic Koch Fractals: Dimensions and Fourier Images. *RENSIT: Radioelectronics. Nanosystems. Information technologies*, 2016, 8(1):81-90. DOI: 10.17725/rensit.2016.08.081.
12. Arzamastseva GV, Evtikhov MG, Lisovsky FV, Mansvetova EG. Fourier Images Quasisymmetry of Generalized Triadic Koch Prefractals. *RENSIT: Radioelectronics. Nanosystems. Information technologies*, 2016, 8(2):207-214 DOI: 10.17725/rensit.2016.08.207.
13. Arzamastseva GV, Evtikhov MG, Lisovsky FV, Mansvetova EG. Light Diffraction by Fractals: Comparison of Experimental Data With the Obtained by Numerical Methods Fourier Images of the Object Pictures. *RENSIT: Radioelectronics. Nanosystems. Information technologies*, 2017, 9(2):221-229. DOI: 10.17725/rensit.2017.09.221.
14. Uozumi J, Kimura H, Asakura T. Fraunhofer diffraction by Koch Fractals. *J. Mod. Optics Phys.*, 1990 37(6):1011-1031.

DOI: 10.17726/rensit.2022.14.317

Method of wave train electrical activity analysis – the theoretical basis and application

Olga S. Sushkova, Alexei A. Morozov, Nadezhda G. Petrova, Margarita N. Khokhlova

Kotelnikov Institute of Radioengineering and Electronics of RAS, <http://www.cplire.ru/>
Moscow 125009, Russian Federation

E-mail: o.sushkova@mail.ru, morozov@cplire.ru, petrova@cplire.ru, margokhokhlova@gmail.com

Alexandra V. Gabova, Karine Yu. Sarkisova

Institute of Higher Nervous Activity and Neurophysiology of RAS, <http://ihna.ru/>
Moscow 117485, Russian Federation

E-mail: agabova@yandex.ru, karine.online@yandex.ru

Alexei V. Karabanov, Larisa A. Chigaleychik

Research Center of Neurology, <https://www.neurology.ru/>
Moscow 125367, Russian Federation

E-mail: doctor.karabanov@mail.ru, chigalei4ick.lar@yandex.ru

Received August 18, 2022, peer-reviewed August 25, 2022, accepted August 31, 2022

Abstract: Classical methods for signal analysis are limited to describe either the global features or the local features. This paper proposes a new mathematically founded concept called wave train electrical activity analysis to investigate both local and global features in biomedical signals simultaneously. First, mathematical means for the investigation of the properties of the wave trains observed in the biomedical signals, histograms of wave train parameters and AUC diagrams, are discussed. Second, several examples of the practical application of the method of the wave train electrical activity analysis are considered. Specifically, its application is demonstrated in the investigation of epileptic seizures as well as the differential diagnosis of neurodegenerative diseases, Parkinson's disease and essential tremor.

Keywords: wave train electrical activity, biomedical signals, wave train, wavelet spectrogram, AUC diagrams, ROC analysis, epileptic seizure, differential diagnosis, Parkinson's disease, essential tremor

UDC 519.67, 612.8, 53.083, 519.24, 004.93

Acknowledgments: The authors are grateful to Dr. Yu.V. Obukhov, who proposed the idea of analyzing local maxima on wavelet spectrograms, as well as Academician of the Russian Academy of Sciences S.N. Illarioshkin for help with the research. The research was supported by the Russian Science Foundation grant No. 22-75-10079, <https://rscf.ru/en/project/22-75-10079/>.

For citation: Olga S. Sushkova, Alexei A. Morozov, Nadezhda G. Petrova, Margarita N. Khokhlova, Alexandra V. Gabova, Alexei V. Karabanov, Larisa A. Chigaleychik, Karine Yu. Sarkisova. Method of wave train electrical activity analysis – the theoretical basis and application. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2022, 14(3):317-330. DOI: 10.17725/rensit.2022.14.317.

CONTENTS

- | | |
|---|--|
| 1. INTRODUCTION (318) | 4. AUC DIAGRAMS OF THE WAVE TRAIN |
| 2. THE DEFINITION OF THE WAVE TRAIN (319) | PARAMETERS (325) |
| 3. HISTOGRAMS OF THE WAVE TRAIN PARAMETERS (323) | 5. CONCLUSIONS (327) |
| | REFERENCES (328) |

1. INTRODUCTION

The development of highly-sensitive methods for detecting early features of brain diseases is a vital problem [1–7]. State-of-the-art methods for analyzing biomedical signals used to solve this problem are mainly based on the Fourier analysis, wavelet analysis, autoregressive models, recurrent neural networks, etc. [8–17]. In our opinion, a disadvantage of the standard methods is that they are aimed either at identifying generalized spectral properties over long-time intervals (that is typical for the Fourier analysis) or at identifying local time-frequency features of the signal (that is typical for wavelet analysis). Thus, the standard methods miss a significant amount of useful information including generalized properties of the local time-frequency features.

Let us consider an example of an electroencephalographic signal (EEG) in a rat brain that contains epileptic activity (see Fig. 1). Fig. 2 demonstrates the Fourier spectrum of the signal computed using the Welch method (Hann window is used with window width equal 3 s and overlap 7/8). One can see pronounced peaks in the spectrum at frequencies 7.48, 14.79, and 22.11 Hz. The Fourier spectrum of the signal indicates the generalized properties of the signal but loses the information on the local features of the signal which can be found in the wavelet-spectrogram of

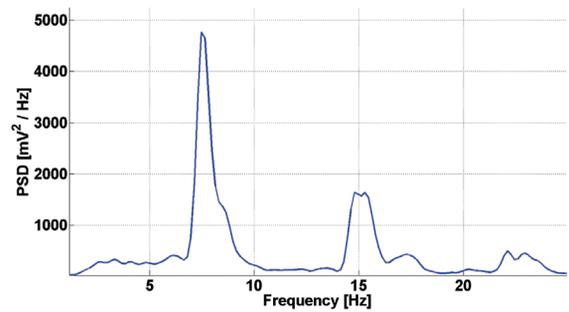


Fig. 2. The Fourier spectrum of the EEG signal in the rat brain.

the signal (see Fig. 3, the complex Morlet wavelet is used, $F_b = 1$, $F_c = 1$).

The wavelet spectrogram indicates that the signal is non-stationary and contains parts with different spectral characteristics. In particular, there is an epileptic activity in the middle of the signal from 16 to 23 seconds. The disadvantage of the Fourier analysis is that it loses information about the local features while the disadvantage of the wavelet analysis is that it does not indicate generalized time-frequency properties of the signal. For instance, the frequency characteristics of the background EEG are not visible in the Fourier spectrum because the power spectral density of the epileptic activity is bigger. On the other hand, the investigation of the wavelet spectrogram of a complete EEG record would take a lot of time and it still does not allow us to get a general idea of the time-frequency properties of the signal because the amount of the information is huge.

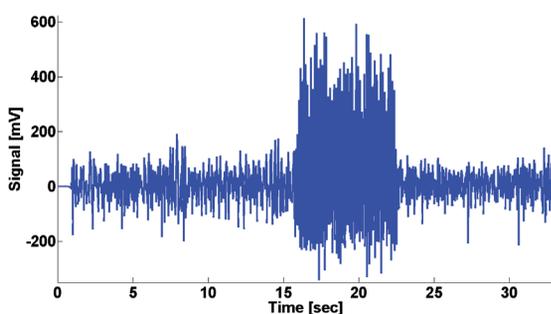


Fig. 1. An example of EEG signal in a rat brain.

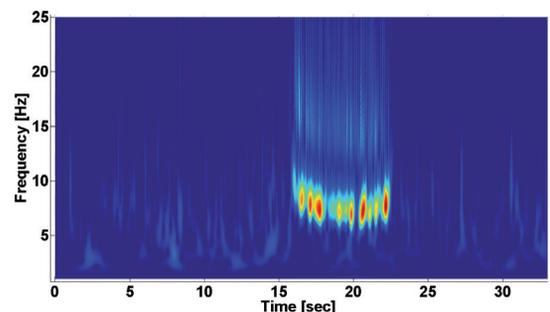


Fig. 3. The wavelet spectrogram of the EEG signal in the rat brain.

The idea of the wave train analysis is that we instruct the computer to investigate a long wavelet spectrogram, extract local features in this spectrogram, and describe the general time-frequency properties of the signal in the terms of the properties of the detected local features.

The idea can be implemented differently, depending on what kind of local features of the signal are investigated and what parameters of these local features are important for us. Our approach is to consider the simplest local features of the signal that are manifested in the form of so-called wave trains in the wavelet spectrograms.

The wave train is an increase of the signal power spectral density localized in time and frequency. The wave trains are displayed on the wavelet spectrograms in the form of local maxima – bright spots of different shapes. For example, in Fig. 3, there are notable wave trains with a frequency of about 7 Hz and weak wave trains with a frequency of about 15 Hz. Computational processing of the wavelet spectrogram allows to identify local maxima of small amplitude that are invisible against the background of the epileptiform activity. We assume that the investigation of the local maxima in the wavelet spectrogram offers a complete idea of the time–frequency properties of the signal; the local features of a complex form are considered as a set of several local maxima distributed in time and frequency.

In this paper, we analyze the following parameters of the wave trains: the center frequency, the maximum power spectral density (PSD), the duration, the bandwidth, and the instantaneous phase. Note that all these parameters can be defined in different ways. For instance, the duration of the wave train can be measured on the half-height of the wave train or $1/\sqrt{2}$ of the wave train height. We assume that any rational

definition of the wave train parameters can be used to investigate the wave trains. The main condition is that the formal definition must be appropriate for the most observed local features of the signals and it cannot be changed during the investigation of the signals. A good illustration of this analysis approach is the comparison of EEG signals in Parkinson’s disease (PD) and essential tremor (ET) patients. The analysis of the signals in the groups of patients revealed a statistically significant difference in the number of wave trains that have certain parameters. The absolute numbers of the wave trains detected in EEG depend on how we measure the parameters of the wave trains. Nonetheless, we will observe the differences between the groups of patients in all cases.

In the following, Section 2 gives a formal definition of wave train parameters. In Section 3, an investigation of the epileptic activity in the rat brain is considered using special histograms. In Section 4, the problem of differential diagnosis is discussed on the base of an example of PD and ET neurodegenerative diseases. The AUC diagrams are introduced as mathematical means for solving this problem.

2. THE DEFINITION OF THE WAVE TRAIN

We use the complex Morlet wavelet for the computation of wavelet spectrograms. This wavelet is selected because the mother function of the Morlet wavelet is a sinusoid inside a Gaussian window:

$$\Psi(x) = \frac{1}{\sqrt{\pi F_b}} \exp(2\pi i F_c x) \exp\left(-\frac{x^2}{F_b}\right),$$

where $F_b = 1, F_c = 1$.

Naturally, any other wavelets can be used for the analysis of the wave train activity. Moreover, the windowed Fourier analysis

can be used instead of the wavelets in many cases. An advantage of the Morlet wavelet is that it is easy to interpret. The Morlet wavelet spectrogram is well-understandable and convenient to use because it is like the windowed Fourier spectrogram.

The time and frequency resolutions of the wavelet spectrogram are automatically changed depending on the considered frequency domain of the signal; that is the advantage of the wavelets in comparison with the windowed Fourier analysis. As the frequency increases, the time resolution of the wavelet spectrogram increases and the frequency resolution decreases. This property of the wavelet spectrogram makes it possible to observe simultaneously fast and slower processes in the investigated object.

From the mathematical point of view, any local maxima in the wavelet spectrogram can be considered as wave trains to be analyzed. The situation is more complicated on the level of the software implementation. The problem is that plenty of "jags" appear on the spectrogram due to the discrete nature of the algorithms used for the computation of the wavelet spectrogram (see an example in Fig. 4). Thus, a smoothing of the wavelet spectrogram is necessary prior to the detection of the wave trains in the spectrogram.

We smooth the wavelet spectrograms using an adaptive 2D Gaussian window filter. The

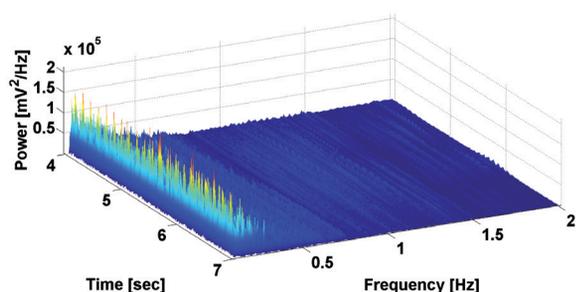


Fig. 4. An example of the wavelet spectrogram before the smoothing operation.

window is called adaptive because the width of the window in time and frequency is changed depending on the considered frequency:

$$G(i, j) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{x(i)^2}{2\sigma_x^2} - \frac{y(j)^2}{2\sigma_y^2}\right),$$

where σ_x is the standard deviation of the normal distribution along the x -axis; σ_y is the standard deviation of the normal distribution along the y -axis.

In our experience, it is sufficient to apply the Gaussian window that has the width in time and frequency equal to half of the width of the mother function of the Morlet wavelet. In addition, we consider only the wave trains in the smoothed wavelet spectrogram that satisfy the following additional condition: the full width of the considered peak on $1/\sqrt{2}$ height of the peak (see Fig. 5, on the left) must be at least $1/10$ period of the signal on the center frequency of the considered peak. This additional condition is necessary to exclude outliers that occurred during the signal recording.

The set of wave train parameters to be analyzed is based on our practical experience in the investigation of EEG and electromyograms (EMG). The following wave train parameters are currently investigated:

1. The center frequency in Hz.
2. The maximum power spectral density in $\mu\text{V}^2/\text{Hz}$.
3. The duration on $1/\sqrt{2}$ height of the peak measured in seconds and the number of periods of the center frequency (see Fig. 5, on the left).

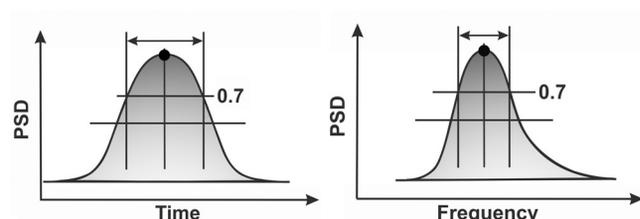


Fig. 5. The time and frequency slices of the local maximum in the wavelet spectrogram.

4. The bandwidth on $1/\sqrt{2}$ height of the peak in Hz (see Fig. 5, on the right).
5. The instantaneous phase in the local maximum in radians; the instantaneous phase is calculated as the four-quadrant arctangent on the imaginary and real parts of the complex wavelet value.

Let us consider an example of the wave train in EEG in a rat (see Fig. 6). This rat belongs to the WAG/Rij genetic line, which has the genetic absence epilepsy. The wave train is indicated by the red ellipse. The signal contains several periods of the so-called spike-wave seizure that is a characteristic of absence epilepsy.

The wavelet spectrogram of the EEG signal is given in Fig. 7. The wave train is indicated by a black ellipse. One can see an increase in the power spectral density localized in time and frequency in the spectrogram. This local maximum in the wavelet spectrogram corresponds to the given formal definition of the wave train and has the following attributes:

1. The center frequency is 6.9 Hz which is a typical frequency of epileptic seizures in rats.
2. The maximum power spectral density is $245260 \mu\text{V}^2/\text{Hz}$.
3. The duration is 0.75 s which corresponds to 5.21 periods of the signal at 6.9 Hz.
4. The bandwidth is 2 Hz.
5. The instantaneous phase is -1 radian.

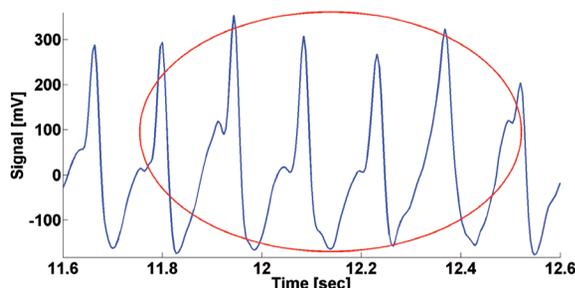


Fig. 6. An example of the wave train in EEG in a rat that belongs to the WAG/Rij genetic line.

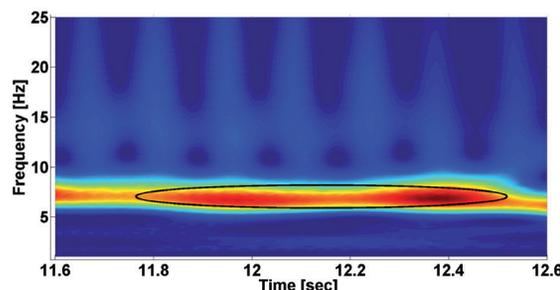


Fig. 7. This is an example of the wave train in the wavelet spectrogram of the EEG signal which is given in Fig. 6.

The physical meaning of the center frequency and maximum power spectral density of the wave train corresponds to the spectral characteristics of the signal indicated by the Fourier spectrum (see Fig. 8). Other parameters of the wave train are explained below.

The duration of the wave train characterizes the duration of the observed oscillatory process. Note that the mathematical definition of the wave train allows the wave train with small durations including one period on the center frequency of the wave train and even less. This is not a mistake. The physical meaning of such short wave trains is that we observe short increases in the power spectral density. Such short wave trains, of course, cannot be named oscillatory processes. Another possible reason for the occurrence of the short wave trains is that long oscillatory processes can be divided into short parts corresponding to local

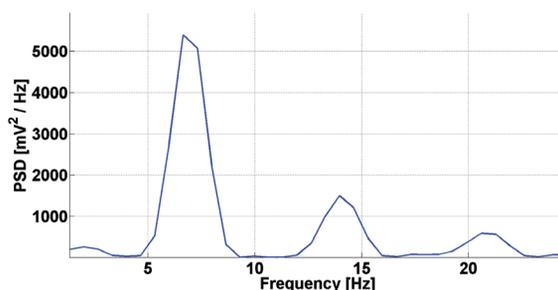


Fig. 8. This is the Fourier spectrum of the EEG signal which is given in Fig. 6. Hann window is applied; the width of the window is 0.75 s.

increases of the instantaneous amplitude of the oscillations.

The bandwidth of the wave trains characterizes the shape of the observed signal. In particular, a narrow bandwidth corresponds to signals that are close to harmonic signals. A wide bandwidth corresponds to signals of complex shape including the signals that contain sharp peaks and steps. Note that a signal of a complex shape that is a superposition of harmonic signals can be represented as several wave trains with different center frequencies; the bandwidth of the separate wave trains can be small.

The instantaneous phase of the wave train characterizes the asymmetry of the local maximum in the wavelet spectrogram. Note that this parameter is very sensitive and allows to reveal properties of the signal invisible to the eye. For example, the instantaneous phase of the wave train given in Fig. 6 is -1 radian; this is evidence of the asymmetry of the wave train.

Note that the wave trains must never be considered as a special kind of signals observed in EEG, EMG, or other biomedical signals. From the mathematical point of view, any signal is a superposition of the wave trains with different attributes. We believe that such a representation of biomedical signals is much closer to reality than the Fourier analysis which considers any signal as a superposition of sinusoids. Meanwhile, it is necessary to account that long signals with complex shapes such as

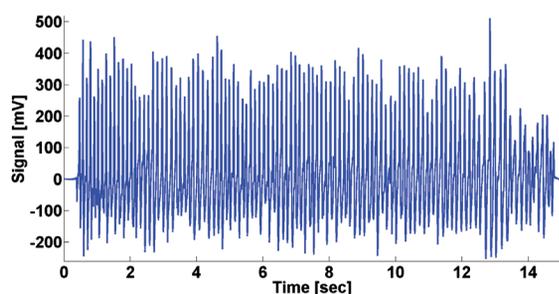


Fig. 9. An example of epileptic seizure in the frontal cortex of a WAG/Rij rat.

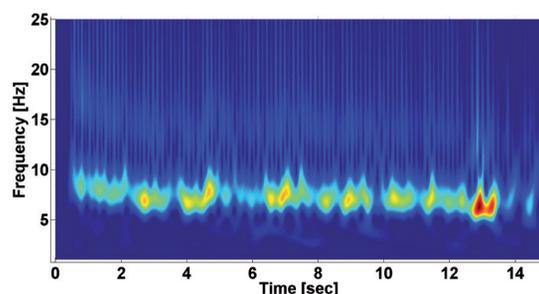


Fig. 10. This is an example of a wavelet spectrogram of the epileptic seizure which is presented in Fig. 9.

epileptic seizures can be represented using a big number of wave trains.

Let us consider an example of an epileptic seizure in the cerebral cortex of a rat that belongs to the WAG/Rij genetic line (see Fig. 9). One can see the spike-wave process on 7 Hz in the wavelet spectrogram in Fig. 10. Besides, separate bright spots are presented in the wavelet spectrogram on 15 Hz that corresponds to the second harmonic in the spike-wave seizure.

The set of wave trains detected in the wavelet spectrogram of the epileptic seizure is presented in Fig. 11. Two chains of wave trains corresponding to 7 Hz and 15 Hz can be observed. Thus, the oscillatory process at 7 Hz does not interfere with the observation of wave trains at 15 Hz. The duration, bandwidth, and instantaneous phase of the wave trains indicate additional information on the epileptic seizure that cannot be obtained using the standard spectral analysis means.

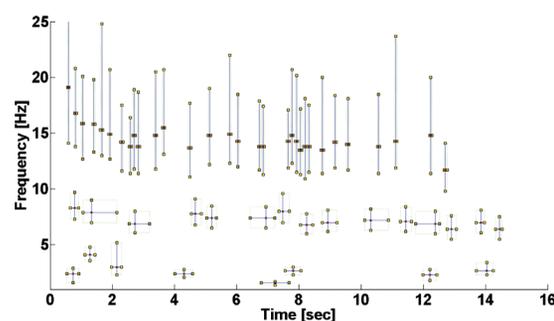


Fig. 11. This is the set of wave trains detected in the wavelet spectrogram of the epileptic seizure which is given in Fig. 10.

Next, we will consider a method of investigation of the wave train electrical activity using histograms of parameters of the wave trains. The method will be explained using an example of epileptic activity in the rat brain.

3. HISTOGRAMS OF THE WAVE TRAIN PARAMETERS

Let us consider the parameters of wave trains detected in EEG signals in absence epileptic WAG/Rij rat. The wave trains were detected in fragments of EEG that contain epileptic seizures (29 fragments) and background EEG (29 fragments).

The data were collected in IHNA&NPH RAS. In total, we have investigated 16 rats. EEG was recorded using implanted electrodes. The experimental setting is described in [18,19]. The sampling rate was 250 Hz. We have analyzed EEG signals in the frontal cortex (the F1 electrode). EEG signals were pre-filtered using 50 Hz and 100 Hz notch filters. After that, the Butterworth bandpath filter was applied (the frequency band was from 0.1 to 120 Hz). Wavelet spectrograms of the fragments of EEG signals were computed. Wave trains were detected in the wavelet spectrograms. The values of parameters of the wave trains were computed.

Histograms of the center frequency of the wave trains detected in EEG are given in Fig. 12. The left histogram corresponds to the wave trains detected in the epileptic seizures. The right histogram corresponds to the background EEG. These histograms are an analog of the Fourier spectrum. By analogy with the spectral analysis, one can conclude that

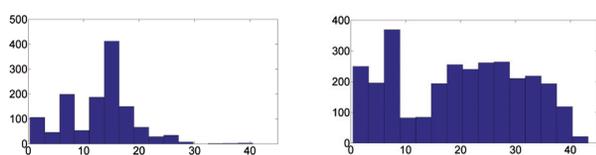


Fig. 12. The histograms of the center frequency of the wave trains. On the left, the epileptic seizures are analyzed. On the right, the background EEG is analyzed.

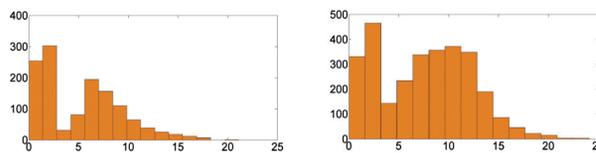


Fig. 13. The histograms of the bandwidth of the wave trains (on the left, the epileptic seizures; on the right, the background EEG).

there are two main rhythms in both datasets. We observe the frequency peak at 7 and 15 Hz in the epileptic seizures and 7 and 24 Hz in the background EEG.

Histograms of the bandwidth of the wave trains (see Fig. 13) demonstrate information that cannot be obtained using the Fourier spectrum. Two pronounced peaks are observed in both groups of datasets. It means that at least two clusters are presented in the data. Suppose that these two clusters correspond to the peaks observed in the histograms of the center frequency.

Histograms of the instantaneous phase of the wave trains (see Fig. 14) indicate a sufficient difference between the epileptic seizures and the background EEG. The histograms show that the epileptic seizures are characterized by inhomogeneity of the instantaneous phase values. This is evidence that wave trains of a certain shape are presented inside the epileptic seizures, but not in the background EEG. The question is what cluster of wave trains observed in Fig. 12 (left) contains the wave trains with a certain shape?

To answer this question, let us investigate the parameters of the wave trains detected in the epileptic seizures in more detail. Let us conduct the following experiment. Let us

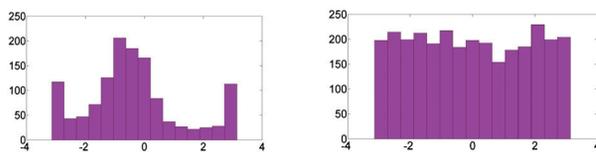


Fig. 14. The histograms of the instantaneous phase of the wave trains (on the left, the epileptic seizures; on the right, the background EEG).

extract the set of the wave trains that have the instantaneous phase in the interval from $-\pi$ to -2.5 radians and in the interval from 2.5 to π radians. Given intervals correspond to the increase of the histogram on the left and right borders (see Fig. 14, left). Let us create the histogram of the center frequency of the extracted wave trains (see Fig. 15, left). Fig. 15 (right) is created by analog. The set of wave trains that have the instantaneous phase in the interval from -1.5 to 0 radians was extracted. This interval corresponds to the peak in the central of the histogram given in Fig. 14 (left). The histogram of the center frequency of the extracted wave trains is demonstrated in Fig. 15 (right).

The histograms in Fig. 15 indicate that the histogram of the center frequency (Fig. 12, left) can be separated into complementary histograms that contain peaks at 7 and 15 Hz. This proves the fact that two clusters of the wave trains observed in Fig. 12 (left) correspond to the peaks observed in the histogram of the instantaneous phase (Fig. 14, left). Thus, we can deduce that each kind of wave trains observed in epileptic seizures is characterized by a certain shape.

An advantage of the method of the wave train electrical activity analysis is that one can determine the occurrence time of each wave train. Thus, one can find and display examples of wave trains that have given

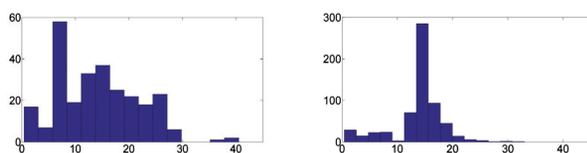


Fig. 15. The histograms of the center frequency of the wave trains in the epileptic seizures. On the left, the instantaneous phase belongs to the intervals from $-\pi$ to -2.5 radians and from 2.5 to π radians. On the right, the instantaneous phase belongs to the interval from -1.5 to 0 radians.

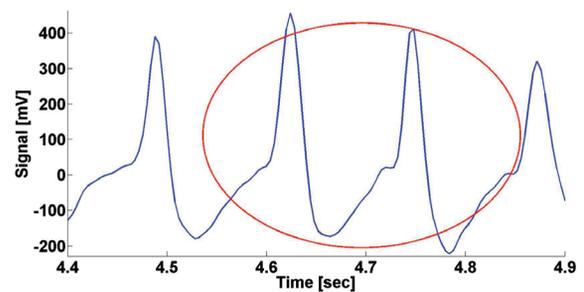


Fig. 16. An example of the wave train detected inside the epileptic seizure. The center frequency of the wave train is 7.8 Hz, the maximal PSD is $300000 \mu V^2 / Hz$, the duration is 0.32 s (2.5 periods), the bandwidth is 2.3 Hz, and the instantaneous phase is -2.7 radians.

values of parameters in the EEG record. For instance, Fig. 16-19 demonstrate examples of the wave trains detected inside the epileptic seizures and background EEG that have the parameters predicted by the histograms considered above.

Thus, the analysis of the wave train electrical activity in the rat brain indicated that:

1. One can observe at least two clusters of wave trains in epileptic seizures. These clusters differ in the center frequency and bandwidth of the wave trains. The clusters of the wave trains correspond to certain intervals of the instantaneous phases; that is evidence of the fact that these clusters differ by the shape of the signals.

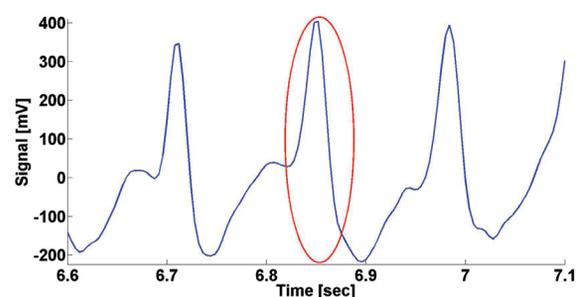


Fig. 17. An example of the wave train detected inside the epileptic seizure. The center frequency of the wave train is 13.8 Hz, the maximal PSD is $80000 \mu V^2 / Hz$, the duration is 0.06 s (0.93 periods), the bandwidth is 6.1 Hz, and the instantaneous phase is -0.62 radians.

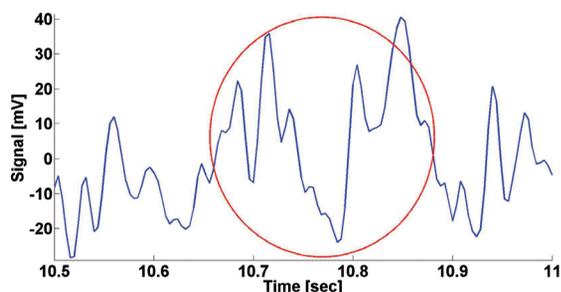


Fig. 18. An example of the wave train detected in the background EEG. The center frequency of the wave train is 7 Hz , the maximal PSD is $2800 \mu\text{V}^2/\text{Hz}$, the duration is 0.22 s (1.54 periods), the bandwidth is 2.3 Hz , and the instantaneous phase is 2 radians.

2. The background EEG also contains two clusters of wave trains that differ in the center frequency and bandwidth. These clusters have no characteristic instantaneous phase. It means that the wave trains observed in the background EEG have no characteristic shape.

Note that the considered examples contain the spike-wave seizures which are characteristic for the typical absence epilepsy. The spike-wave seizures of the genetic absence epilepsy differ from other types of epileptic seizures including the atypical absence epilepsy and convulsive forms of epilepsy. A promising direction of further research is an investigation of differences between various forms of epilepsy using the wave train analysis.

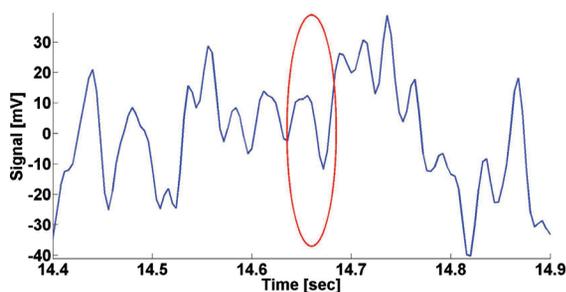


Fig. 19. An example of the wave train detected in the background EEG. The center frequency of the wave train is 25.6 Hz , the maximal PSD is $241 \mu\text{V}^2/\text{Hz}$, the duration is 0.04 s (1.22 periods), the bandwidth is 10 Hz , and the instantaneous phase is -1.64 radians.

4. AUC DIAGRAMS OF THE WAVE TRAIN PARAMETERS

A promising application area of the wave train analysis is the differential diagnosis of neurodegenerative diseases as well as solving other problems related to the comparison of multiple experimental datasets. In the framework of the wave train analysis, the problem of differential diagnosis can be considered as a search of certain intervals of the wave train parameters such that the wave trains with appropriate values of parameters are typical for one data category but are rarely observed in another data category. Thus, the problem of the differential diagnosis is reduced to the mathematical problem of the search for appropriate intervals of the wave train parameters.

We have developed a visualization method that facilitates the comparison of datasets and the search for the intervals of the wave train parameters [20-23]. The visualization method includes the following steps:

1. All possible intervals of the given wave train parameter are to be considered. The intervals are tested using all wave trains detected in the wavelet spectrograms of given datasets.
2. ROC curve is created for each considered interval. The ROC curve indicates the result of the comparison of given datasets, for instance, the set of patients and the set of healthy volunteers. The comparison is conducted using the given parameter of the wave trains.
3. The area under the ROC curve (AUC) is computed.
4. 2D and 3D diagrams that indicate how AUC depends on the intervals of the given wave train parameter are created.

Let us consider an example of the AUC diagram that compares the number of wave trains per second detected in the wavelet

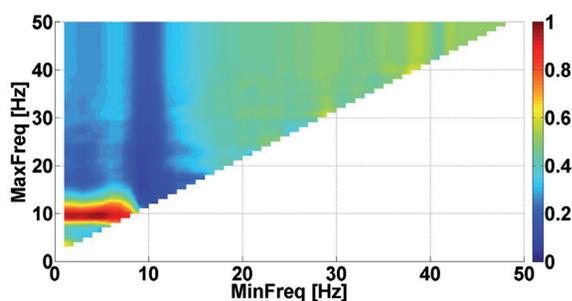


Fig. 20. An example of an AUC diagram that compares the number of wave trains per second detected in EEG signals in PD and ET patients.

spectrograms of EEG signals in patients with PD and ET (see **Fig. 20**). The data are recorded in the Research Center of Neurology. The dataset contains data from 11 patients with PD and 15 patients with ET. Patients with the first stage of PD that demonstrate hyperkinetic movements in the right arm were investigated. The patients were not taking antiparkinsonian drugs. During the data acquisition, the patients sat in an armchair with arms outstretched forward. The experimental setting is described in more detail in [20,21,24,25]. In the example under consideration, the occipital region of the cerebral cortex is investigated (the O2 EEG channel). The sampling rate of recorded signals was 500 Hz. The 50, 100, 150, and 200 Hz notch filters were applied to suppress the network interference. The 0.1-240 Hz Butterworth bandpass filter was used. The wavelet spectrograms of the EEG signals were computed. The wave trains were detected in the wavelet spectrograms. The values of the parameters of the wave trains were computed.

In **Fig. 20**, the abscissa indicates the lower bound of the wave train center frequency interval; the ordinate indicates the upper bound of the interval. AUC values are indicated using a blue-red color scale. The blue color corresponds to 0. The

red color corresponds to 1. The green color corresponds to 0.5.

The diagram contains a prolonged red area with coordinates from 0 to 9 Hz along the abscissa and from 8 to 12 Hz along the ordinate. In addition, there is a bright blue region with coordinates from 8 to 12 Hz along the abscissa and from 11 to 50 Hz along the ordinate. These areas correspond to the intervals of the wave train center frequencies where strong differences between the groups of patients with PD and ET are observed. Note that the diagram contains the areas of both red and blue colors. This is evidence that we obtained two independent and fundamentally different features that differentiate the patients with PD and ET. The red area indicates the center frequency intervals of the wave trains that are typical for patients with PD but are rare in patients with ET. The blue area, on the contrary, indicates the center frequency intervals of the wave trains that are typical for patients with ET but are rare in patients with PD. The values of AUC in the red area reach 0.97 (see the ROC curve in **Fig. 21**). The values of AUC in the blue area reach 0.06 (see the ROC curve in **Fig. 22**). These AUC values indicate that founded features of PD and ET can be used for the differential diagnosis of the patients.

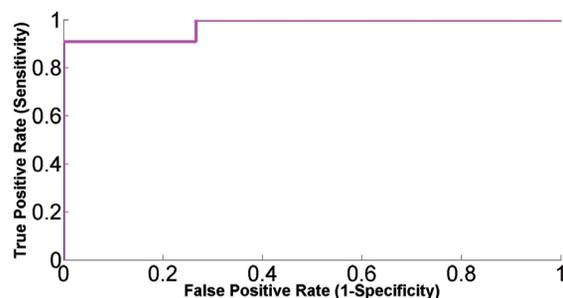


Fig. 21. An example of a ROC curve comparing the number of wave trains per second detected in EEG signals in the patients with PD and ET. The wave train center frequency interval is from 5 to 10 Hz.

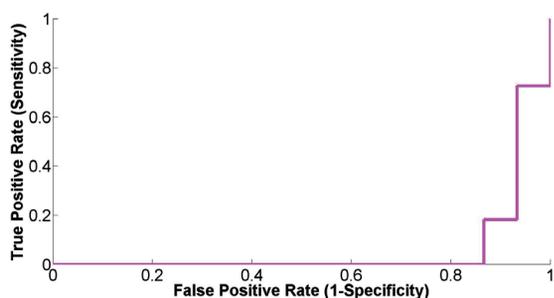


Fig. 22. An example of a ROC curve comparing the number of wave trains per second detected in EEG signals in patients with PD and ET. The wave train center frequency interval is from 9 to 18 Hz.

Examples of wave trains that are typical for PD and ET are demonstrated in **Fig. 23** and **Fig. 24**.

Other types of AUC diagrams are considered in [23]. The experience of data analysis demonstrated that the wave train electrical activity analysis method can be successfully applied to different kinds of biomedical signals, including EEG [20,21,24-26], EMG [22,23,27,28], and accelerometer signals [22,29]. Further development of the method may include the usage of additional parameters of the wave trains as well as a statistical analysis of relationships between the wave trains with different attributes. For instance, the occurrence times of the wave trains belonging to the mu (7.5-13.5 Hz) and beta (18-30 Hz) frequency bands were compared [26]; it was demonstrated that the beta wave

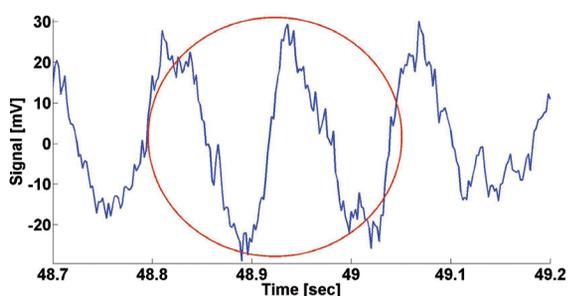


Fig. 23. An example of a wave train detected in the EEG in a patient with PD. The center frequency of the wave train is 8 Hz.

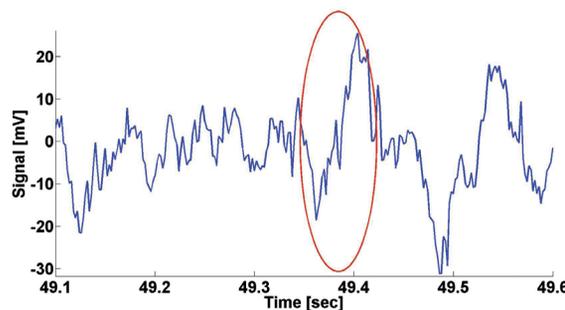


Fig. 24. An example of a wave train detected in the EEG in a patient with ET. The center frequency of the wave train is 15.1 Hz.

trains were not a harmonic of the mu wave trains. Another promising direction of the research is the investigation and comparison of the effects of pharmacological drugs using AUC diagrams.

5. CONCLUSIONS

The main idea of the proposed method of analyzing the wave train electrical activity is the statistical analysis and search for regularities in the parameters of local maxima (wave trains) detected in wavelet spectrograms of signals. The developed method allows revealing the regularities in the biomedical signals that cannot be found using wavelet analysis and standard methods of spectral analysis. The method is universal and can be applied for the investigation of signals of various kinds including EEG, EMG, and accelerometer signals. Previously, the method was successfully applied for the early and differential diagnosis of PD and ET using EMG and accelerometers. A patent based on these results was registered [30]. In this paper, it was demonstrated that the method can be applied for the investigation of epileptiform activity and clinical diagnosis of neurodegenerative diseases using EEG. The results of the investigation of the spike-wave seizures in the rats that belong to the WAG/Rij genetic line indicate that the method is promising for the investigation of other forms of epilepsy.

REFERENCES

1. Illarioshkin SN. Modern view on etiology of Parkinson's disease. *Neurological Journal*, 2015, 20(4):4-13.
2. Ivanova EO, Fedin PA, Brutyan AG, Ivanova-Smolenskaya IA, Illarioshkin SN. Analysis of tremor activity of antagonist muscles in essential tremor and Parkinson's diseases. *Neurological Journal*, 2014, 19(4):11-18.
3. Wolters E, Baumann C. *Parkinson Disease and Other Movement Disorders*. VU University Press, 2014:154.
4. Brockmann K, Berg D. *Diagnosis of Parkinson's Disease. Parkinson Disease and Other Movement Disorders* (eds. E. Wolters, C. Baumann). Int. Ass. of Parkinsonism and Related Disorders, VU University Press, 2014:229-242.
5. Parekh A, Selesnick I, Rapoport D, Ayappa I. Sleep spindle detection using time-frequency sparsity. *IEEE Signal Processing in Medicine and Biology Symposium*, Philadelphia, PA, IEEE, 2014:16.
6. O'Reilly C, Nielsen T. Automatic sleep spindle detection: benchmarking with fine temporal resolution using open science tools. *Frontiers in Human Neuroscience*, 2015, 9:353.
7. Deb R, Bhat G, An S, Ogras U, Shill H. Trends in Technology Usage for Parkinson's Disease Assessment: A Systematic Review. *medRxiv*, 2021, Jan 1.
8. Oktay AB, Kocer A. Differential diagnosis of Parkinson and essential tremor with convolutional LSTM networks. *Biomedical Signal Processing and Control*, 2020, 56:101683.
9. Andreeva Y, Khutorskaya O. EMGs spectral analysis method for the objective diagnosis of different clinical forms of Parkinson's disease. *Electromyography and Clinical Neurophysiology*, 1996, 36(3):187-192.
10. Rissanen SM, Kankaanpaa M, Meigal A, Tarvainen MP, Nuutinen J, Tarkka IM, Airaksinen O, Karjalainen PA. Surface EMG and acceleration signals in Parkinson's disease: feature extraction and cluster analysis. *Medical & biological engineering & computing*, 2008, 46(9):849-858.
11. O'Sullivan S, Murray N, Rodrigues TB. A telehealth and sensor-based system for user-centered physical therapy in Parkinson's disease: research proposal. In *Proceedings of the 13th ACM Multimedia Systems Conference*, 2022, June, 383-387.
12. Avci D, Dogantekin A. An expert diagnosis system for Parkinson disease based on genetic algorithm-wavelet kernel-extreme learning machine. *Parkinson's disease*, 2016, ID 5264743, doi: 10.1155/2016/5264743.
13. Loconsole C, Trotta GF, Brunetti A, Trotta J, Schiavone A, Tatò SI, Losavio G, Bevilacqua V. Computer vision and EMG-based handwriting analysis for classification in Parkinson's disease. *International Conference on Intelligent Computing*, Springer, Cham, 2017, August: 493-503.
14. Lacy SE, Smith SL, Lones MA. Using echo state networks for classification: A case study in Parkinson's disease diagnosis. *Artificial intelligence in medicine*, 2018, 86:53-59.
15. Fattorini L, Felici F, Filligoi G, Trallesi M, Farina D. Influence of high motor unit synchronization levels on non-linear and spectral variables of the surface EMG. *Journal of neuroscience methods*, 2005, 143(2):133-139.
16. Robichaud JA, Pfann KD, Vaillancourt DE, Comella CL, Corcos DM. Force control and disease severity in Parkinson's disease. *Movement Disorders*, 2005, 20(4):441-450.
17. Flament D, Vaillancourt D, Kempf T, Shannon K, Corcos D. EMG remains

- fractionated in Parkinson's disease, despite practice-related improvements in performance. *Clinical neurophysiology*, 2003, 114(12):2385-2396.
18. Sarkisova KY, Gabova AV. Maternal care exerts disease-modifying effects on genetic absence epilepsy and comorbid depression. *Genes, Brain, and Behavior*, 2018, Sep., 17(7): e12477.
19. Gabova AV, Sarkisova KY, Fedosova EA, Shatskova AB, Morozov AA. Developmental Changes in Peak-Wave Discharges in WAG/Rij Rats with Genetic Absence Epilepsy. *Neuroscience and Behavioral Physiology*, 2020, February, 50(2):245-252.
20. Sushkova OS, Morozov AA, Gabova AV, Karabanov AV. Application of brain electrical activity burst analysis method for detection of EEG characteristics in the early stage of Parkinson's disease. *Zhurnal Nevrologii i Psichiatrii imeni S.S. Korsakova*, 2018, 118(7):45-48.
21. Sushkova OS, Morozov AA, Gabova AV. Data mining in EEG wave trains in early stages of Parkinson's disease. *Advances in Soft Computing: MICAI 2016/Obdulia Pichardo-Lagunas, Sabino Miranda-Jimenez (Eds.)*. Springer, 2017, V. 10062 LNAI:403-412.
22. Sushkova OS, Morozov AA, Gabova AV, Karabanov AV. Investigation of surface EMG and acceleration signals of limbs' tremor in Parkinson's disease patients using the method of electrical activity analysis based on wave trains. *Advances in Artificial Intelligence: IBERAML4 2018 /G.R. Simari, F. Eduardo, F. Gutiérrez Segura, J.A. Rodríguez Melquiades (Eds.)*. Springer, 2018, V. 11238 LNAI:253-264.
23. Sushkova OS, Morozov AA, Gabova AV, Karabanov AV, Illarioshkin SN. A statistical method for exploratory data analysis based on 2D and 3D area under curve diagrams: Parkinson's disease investigation. *Sensors*, Ernest N. Kamavuako (Eds.). MDPI, 2021, 21(14):4700.
24. Sushkova OS, Morozov AA, Gabova AV. Investigation of Specificity of Parkinson's Disease Features Obtained Using the Method of Cerebral Cortex Electrical Activity Analysis Based on Wave Trains. *13th International Conference on Signal-Image Technology and Internet-Based Systems*. December 4-7, 2017, Jaipur, India. – Jaipur: MNIT, 2017: 168-172.
25. Sushkova OS, Morozov AA, Gabova AV. A method of analysis of EEG wave trains in early stages of Parkinson's disease. *Proc. Intern. Conf. on Bioinformatics and Systems Biology*. March 4-6, 2016, Allahabad, India. – Allahabad: IIIT, 2016:1-4.
26. Sushkova OS, Morozov AA, Gabova AV. EEG Beta Wave Trains are not the Second Harmonic of Mu Wave Trains in Parkinson's Disease Patients. *Proceedings of the International conference on Information Technology and Nanotechnology. Session Image Processing, Geoinformation Technology and Information Security (ITNT 2017, Samara, Russia, April 24-27, 2017)*. Samara: CEUR, 2017, V. 1901: 226-234.
27. Sushkova OS, Morozov AA, Gabova AV, Karabanov AV. Investigation of the multiple comparisons problem in the analysis of the wave train electrical activity of muscles in Parkinson's disease patients. *Journal of Physics: Conference Series*. IOP Publishing Ltd, 2019, 1368(5):052004.
28. Sushkova OS, Morozov AA, Gabova AV, Karabanov AV, Chigaleychik LA. Investigation of the 0.5-4 Hz low-frequency range in the wave train electrical activity of muscles in patients with Parkinson's disease and essential tremor. *RENSIT:*

- Radioelectronics. Nanosystems. Information Technologies*, 2019, 11(2):225-236. DOI: 10.17725/rensit.2019.11.225.
29. Sushkova OS, Morozov AA, Gabova AV, Karabanov AV, Chigaleychik LA. An Investigation of Accelerometer Signals in the 0.5-4 Hz Range in Parkinson's Disease and Essential Tremor Patients. *Advances in Intelligent Systems and Computing. Proceedings of International Conference on Frontiers in Computing and Systems (COMSYS-2020)*, Jalpaiguri Government Engineering College, West Bengal, India, January 13-15, 2020, Proceedings / D. Bhattacharjee, D.K. Kole, N. Dey, S. Basu, D. Plewczynski (Eds.). Springer, Singapore, 2021. V. 1255:455-462.
30. Sushkova OS, Morozov AA, Gabova AV, Karabanov AV. Method for Differential Diagnosing of Essential Tremor and Early and First Stages of Parkinson's Disease Using Wave Train Activity Analysis of Muscles. *Russian Federation Patent No. 2741233*, 22 January 2021.

DOI: 10.17725/rensit.2022.14.331

Computational experiment – nondimensionalization of equations, computational stability and program testing

Mikhail G. Evtikhov

Kotelnikov Institute of Radioengineering and Electronics of RAS, Fryazinsky branch, <http://fire.relarn.ru/>
Fryazino 141190, Moscow region, Russian Federation

E-mail: emg20022002@mail.ru

Vladimir G. Evtikhov

Moscow Polytechnic University, <https://mospolytech.ru/>
Moscow 107023, Russian Federation

E-mail: evg2002@yandex.ru

Received 17 June 2022, peer-reviewed 24 Juni 2022, accepted 01 July 2022

Abstract: From the stages of the computational experiment, the stage of non-dimensionalization of the initial equation (system of equations) of the problem is considered - the replacement of its variables by the product of the corresponding dimensionless quantities by their units of measurement with subsequent transformations. Such a transition from a physical model to a mathematical (dimensionless) one makes it possible to obtain software implementations for research. A critical evaluation of its complexity is carried out and possible errors in the results are evaluated. At the same time, new versions of software are formed. Object-oriented programming tools and version control systems (for example, git) allow you to create versions of software tools adapted to different conditions of their use and for different types of users. Parallelization of work on versions is carried out. At the same time, for further software implementation, the set-theoretic language of formulas with partially recursive functions is effective. To implement versions with large amounts of calculations and data, high-performance computing systems based on software and hardware acceleration, parallel information processing and cloud architectures are used. As a rule, a difference model of the problem and iterative methods for solving it are constructed for a program version. Computational stability conditions are usually stipulated in modern instructions for standard program libraries. For new algorithms, it is necessary to analyze the stability of difference schemes based on the refinement of their spectral properties and the use of functional analysis methods. For storage and subsequent application of the results of computational experiments, it is advisable to use modern databases. As a kind of computational experiment, testing of alpha and beta versions of programs and their releases is also considered.

Keywords: computational experiment, problem statement, nondimensionalization, partial recursive function, numerical stability, computer literacy, tester, software development, software versioning

UDC 167.2, 004.004, 004.007

For citation: Mikhail G. Evtikhov, Vladimir G. Evtikhov. Computational experiment – nondimensionalization of equations, computational stability and program testing. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2022, 14(3):331-340. DOI: 10.17725/rensit.2022.14.331.

CONTENTS

- | | |
|---|---|
| <ol style="list-style-type: none"> 1. INTRODUCTION (332) 2. STAGES OF A COMPUTATIONAL EXPERIMENT (333) <ol style="list-style-type: none"> 2.1. PROBLEM STATEMENT (333) 2.2. NONDIMENSIONALIZATION OF EQUATIONS (334) 2.3. MODEL PROGRAMMING (334) 2.4. PROBLEM SOLVING (335) | <ol style="list-style-type: none"> 2.5. EVALUATION OF THE PROBLEM SOLUTION. CYCLING BACK TO THE LEVEL OF THOUGHT EXPERIMENTS (335) 2.6. RETURN TO MODEL PROGRAMMING (335) 2.7. RETURN TO THE LEVEL OF NATURAL EXPERIMENTS (335) 2.8. MODEL LEVELS (335) |
|---|---|

- 3. COMPUTATIONAL INSTABILITY (336)
- 4. VERSIONS OF SOFTWARE PRODUCTS (337)
- 5. CONCLUSION (338)
- REFERENCES (339)

1. INTRODUCTION

A computational experiment is an experiment conducted not on the original real object, but on a mathematical (information, simulation) model of the object using computational and logical procedures implemented by appropriate software tools on computer systems. Closures to the computational experiment concepts are numerical, mathematical and simulation experiments on a computer [1].

In [2], various approaches to the formation of mathematical models based on scientific laws and principles are considered. These approaches can be considered as ways of setting up computational experiments.

A special view of the computational experiment, closely related to the theory of probability, is being developed, for example, within the framework of the "planning of experiments" direction [3]. When designing new types of computers in the 70s, the possibilities of replacing some physical parts of a computer system with its software models were analyzed [4]. This efficient approach has now found its way into programming methods in the form of separating interfaces from their implementations (eg *b*-files and *cpp*-files in C++). The development of computational experiment methods and its central part, mathematical modeling, was accompanied by the clarification of a number of fundamental features of computational methods. In [5], a profound idea was expressed that programs for computers from a mathematical point of view are described by partially recursive functions. With strict definitions of algorithms, such a statement is considered as a hypothesis and is called "Church's thesis". From the point

of view of an applied programmer, we mean mathematical constructions that are more widely understood than formulas based on radicals that are familiar to many researchers. For example, numerically solving an equation of degree 10 is a standard problem, but there is no general solution in radicals for an equation of degree 5. This does not mean that formulas can be dispensed with. In the formulas of the mathematics of partially recursive functions, the language of set theory and recursion, substitution, and enumeration of elements of sets are more widely used. For example, in [6], to description of fractals the L-system is used, i.e. recursive expression for sets: if $Z^{(0)}$ is a unit black square in the complex plane, then

$$Z^{(n+1)} = Z^{(n)} \cup (iZ^{(n)} + (1+i)2^n) \cup (-iZ^{(n)} + (1+i)2^n),$$

where \cup is the operation of union of sets.

Generalizations of recursive formulas of set theory adequate to the subject of research, when necessary accurately and concisely record the results that already received and researched. However, it is difficult to find errors in these mathematical constructions, make changes and check them. Accordingly, debugging takes more than 90% of the development time of a new program and continues throughout its entire life cycle.

When writing already obtained theoretical solutions to problems, the language of formulas with recursive functions is effective for their further software implementation. This approach not only simplifies programming, but also allows professional mathematicians and designers to take part in solving complex problems of algorithm optimization and information storage.

A significant difficulty of the computational experiment is the computational instability [7]. The instructions for the standard libraries of programs (classes) usually stipulate the conditions for computational stability and other

conditions investigated for the correct use of software tools. The use of libraries for general sparse matrices [8] expands the possibilities for developing new numerical methods.

There are developed software tools and software libraries for solving various problems of modeling and computational experiment. These tools are deeply thought out and researched. It would seem that there is opportunity to further develop physics, other sciences and technologies. At the same time, the stock of achievements in the field of functional analysis is far from exhausted. A solid groundwork in the theory of probability [9,10] is the basis for many successful applications of the computational experiment. Obviously, programming cannot develop without experimental research also in the field of programming itself.

In recent years, outsourcing (attraction of external specialists) in the field of software product testing has become inevitable. The work of a tester can be interpreted as an experimental study of a software product; this activity determines the quality and competitiveness of software products. A few years ago, no special skills or education were required from a tester. Currently, special training is already required. Simulation and computational experiment become relevant at all levels of modern programming. Related to testing interesting and practically valuable ideas are discussed in many Internet publications. A general view, a "pyramid of tests", is given, for example, in [11].

The purpose of this article is to reveal the concept of a computational experiment, using the example of a simple problem, taking into account the point of view of applied programmers: the role of nondimensionalization of equations, the computational stability of numerical

methods, and the development of programs by versions.

2. STAGES OF THE COMPUTATIONAL EXPERIMENT

It seems appropriate to state some topical problems of a computational experiment using the example of solving a relatively simple problem.

2.1. PROBLEM STATEMENT

For example: on a steep seashore, a cannon fires at a given angle to the horizon. How far will the projectile fly if you can ignore air resistance and assume the Earth is flat.

The flight of the projectile must satisfy Newton's second law:

$$m \frac{d^2 y}{dt^2} = -mg; \quad m \frac{d^2 x}{dt^2} = 0,$$

m is the mass of the projectile, x is the horizontal coordinate of the projectile in flight, y is the vertical coordinate, t is the time, g is the acceleration of gravity near the surface of the Earth.

Note that the mass can be reduced. Therefore, the solution will be independent of the mass of the projectile. If there were also additional terms describing air resistance in the equations, then they would be divided on mass. The larger the mass, the less their influence would be. We temporarily neglect air resistance, so the equations are reasonable and adapted to the massive and small bodies.

V_0 is the speed of the projectile at the moment of its departure from the cannon, α is the angle at which the cannon fires, b is the height at which the cannon stands.

2.2. NONDIMENSIONALIZATION OF EQUATIONS

The non-dimensionalization of equations is a subtle, not obvious to everyone, and, perhaps, the most important stage of the study. This is sometimes underestimated and leads to serious and even catastrophic consequences. This stage

requires consultations with specialists involved in full-scale experiments.

The laws of the natural sciences are usually written in the SI system of units. The advantage of the SI system of units is its linkage to the scale of ordinary everyday experience.

We write the physical variables as the product of a dimensionless quantity (let's mark it with a wave) and its units of measurement (in square brackets):

$$x = \tilde{x}[\text{m}]; y = \tilde{y}[\text{m}]; h = \tilde{h}[\text{m}]; t = \tilde{t}[\text{s}]$$

$$g = \tilde{g}[m \cdot s^{-2}], \tilde{g} = 9,8; V_0 = \tilde{V}_0[m \cdot s^{-1}]$$

Here all lengths are measured in meters, time is in seconds, g is the acceleration of gravity near the Earth's surface, speed is measured in meters per second.

The standard library functions assume that the angles are given in radians, i.e. α is already a dimensionless quantity. The angle given in degrees is nondimensionalized using the relation $\alpha = \frac{\pi \tilde{\alpha}[\text{deg}]}{180[\text{deg}]}$, π is the ratio of the circumference to the diameter, $\tilde{\alpha}$ is the angle in degrees.

The procedure for dimensionless equations is that the variables, together with their dimensions, are substituted into the original equations.

It turns out that the dimensions can be reduced. Moreover, all dimensions must be canceled out. After non-dimensionalization, only dimensionless variables should remain in the equations:

$$\tilde{y} = -\frac{\tilde{g}\tilde{t}^2}{2} + \tilde{t}\tilde{V}_0 \sin \alpha + \tilde{h},$$

$$\tilde{x} = \tilde{t}\tilde{V}_0 \cos \alpha.$$

The form of the dimensionless equation must not have to exactly repeat the form of the original equation. If it turns out to be necessary to set, for example, the speed not in meters per second, but in kilometers per hour, then additional factors will appear in the non-dimensional equations. Sometimes

artificial units of measurement are introduced specifically for this task.

When analyzing some models, the physical dimensions of quantities not can be reduced. This case is result of serious misunderstanding of problem, either result of equations errors. Often with the help of non-dimensionalization, we can identify untenable theories claims.

2.3. MODEL PROGRAMMING

The nondimensionalization procedure is a transition from a physical model to a mathematical one. There are no grams, meters, or rubles in the computer's memory. Standard programs help to convert the dimensionless values of the model into computer bits and bytes. Only with the help of dimensionless mathematical model, programmers get the opportunity to write programs. In the case of a cannon problem, you should write functions that calculate \tilde{x} and \tilde{y} for a given \tilde{t} . A good manner would be to make the model as a class and put the acceleration \tilde{g} , height and angle in class fields, and \tilde{x} and \tilde{y} put calculation as class methods.

The mathematical model is useful precisely for the task for which it was developed. Its extended interpretations are possible, but doubtful. A simplified model can be a tool for solving more complex problems and a source of test cases for other more complex models. You should start with an extremely simplified model, and then develop versions of programs. Newton's idea of extremely simple experiments is especially relevant in a numerical experiment and is based on the possibility of using inherited classes.

2.4. PROBLEM SOLVING

The projectile flight time is the solution of the equation $\tilde{y} = 0$. The class method for obtaining this time can be implemented

completely by yourself, but it is better to use standard programs. It's not about whether it's easy or hard to solve a quadratic equation. Standard numerical methods and programs can numerically solve even such equations for which there are no solutions in the form of familiar formulas with radicals. After calling the methods of the library classes, in our case, it is necessary to provide for the selection of a solution in which $\tilde{t} > 0$. Having obtained the flight time of the projectile, we obtain a solution to the problem using the class method to calculate \tilde{x} . It remains to formalize the entire solution of the problem as another class method.

2.5. EVALUATION OF THE PROBLEM SOLUTION. CYCLING BACK TO THE LEVEL OF THOUGHT EXPERIMENTS

After solving the problem, the turn comes to the most unpleasant surprise question: how to evaluate the error of this solution? Without an estimate of the error, physical quantities are meaningless. Estimates of errors in the initial parameters of the model correspond to the stages of a thought experiment and non-dimensionalization. The numerical model makes it possible to vary its parameters and estimate the result error. Sometimes this is easier than looking for derivatives of solution formulas.

At the stage of critical evaluation of the solution, ideas for the next version of the program arise. Work on the current version is still very far from completion, but it is already becoming possible and of fundamental importance to work on a new version. The technology of parallelization of work on versions appeared, apparently, in the aircraft industry, but it is possible even earlier. All modern programs are developed by versions. Existing version control systems allow you to do this quite conveniently.

2.6. RETURN TO MODEL PROGRAMMING

How to convert the solution for its storage and further use? Collecting, storing and processing big data is becoming a difficult problem. In [12], various types of modern databases and high-performance computing based on hardware-software acceleration, parallel information processing, and cloud architectures are discussed.

2.7. RETURN TO THE LEVEL OF NATURAL EXPERIMENTS

The data array obtained as a result of solving modeling problems and experiments must be stored in the database, and also brought to the consumer in a form understandable to him. Such tasks of developing database interfaces are becoming more and more relevant. Ability to develop web applications (including gadgets) has already become very desirable competence of a modern programmer [13]. It intersects with tasks computational experiment and has numerous directions. For example, wiki projects. They allow users to use the well-known Wikipedia interface. Some firms create their own specialized wikis independent of the main Wikipedia. They contain related materials for technically complex products, including software. It is possible to organize sophisticated automated searches on the main Wikipedia and other wikis. In [14], an example of a complex search, comparison and visualization of historical data is given. All these possibilities are promising for application in research and experimental work.

2.8. MODEL LEVELS

As a result, the computational experiment can be represented as interconnected experiments with 4 layers of models:

- Thought experiment based on literature data, understanding how to evaluate model errors.

- Full-scale experiment with the development of interfaces.
- Mathematical model and experiments with databases.
- Solving mathematical problems, analysis of computational stability of algorithms using functional analysis tools.

3. COMPUTATIONAL INSTABILITY

Now in the previous problem we will have to take into account air resistance. In the new version we will have to build difference model, as well as an iterative method for its solutions. Iterative algorithms are efficient and simple, it is almost impossible to refuse them. Usually computational instability appears in iterative algorithms.

The essence of computational instability is the accumulation of computational errors. Computational instability is the main problem of numerical methods and experiments. The first attempts to use computational models of physical processes often turned out to be computationally unstable. They give results, that far from true solutions (for example, Richardson's scheme [7]).

The simplest model of oscillations with frequency ω can be described by a complex equation:

$$\frac{dX}{dt} = i\omega X; X(t=0) = A.$$

(for example, the Schrödinger equation at zero momentum in a homogeneous potential).

If A is a real number, then the solution to this equation will be oscillations with amplitude A , frequency ω and period $T = 2\pi/\omega$:

$$X = A \exp(i\omega t) = A \cos(\omega t) + iA \sin(\omega t),$$

$$\operatorname{Re}(X) = A \cos(\omega t); \operatorname{Im}(X) = A \sin(\omega t).$$

On the plane $(\operatorname{Re}(X), \operatorname{Im}(X))$ you should get a circle with radius A .

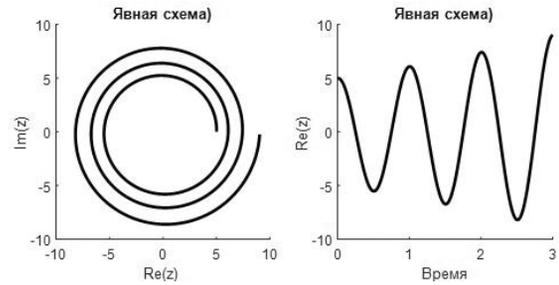


Fig. 1. Explicit difference scheme.

We choose a time step $\tau \ll T$ and approximate the derivative as

$$\frac{dX}{dt} \approx \frac{X_{i+1} - X_i}{\tau}.$$

we get

$$\frac{X_{i+1} - X_i}{\tau} = i\omega X_i \Rightarrow X_{i+1} = X_i(1 + i\omega\tau); X_1 = A.$$

Let, for simplicity, $\omega = 2\pi$, then the oscillation period should be 1. Let $\tau = 0.01$. Such an approximation of the equation seems natural, but this appearance is deceptive (Fig. 1).

The reason for this parasitic increase in amplitude is that small errors accumulate at each iteration step, and as a result, the solution deviates more and more from the true one. This phenomenon is called computational instability. As the step τ decreases, the parasitic increase in the amplitude decreases, but does not disappear.

A difference scheme, when the source value is taken at the current time layer, is called an explicit difference scheme. An implicit difference scheme is possible when the source function is taken not on the current, but on a new time layer (Fig. 2):

$$\frac{X_{i+1} - X_i}{\tau} = i\omega X_{i+1} \Rightarrow X_{i+1} = X_i / (1 - i\omega\tau); X_1 = A.$$

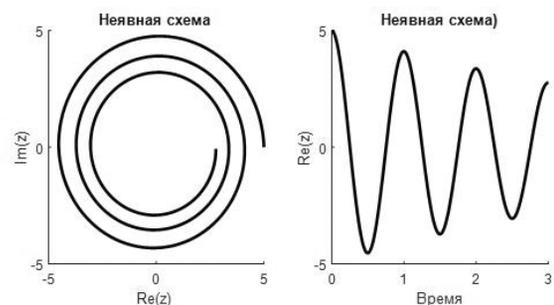


Fig. 2. Implicit difference scheme.

This difference scheme is considered to be more successful. But now there is parasitic attenuation of the amplitude. The amplitude is also not preserved, which is also far from always acceptable. The computational stability of the difference scheme does not guarantee the absence of such computational attenuation. In [15], one can get acquainted with the methods of spectral analysis of difference schemes.

The examples in Fig. 1 and Fig. 2 reveal the opposite of the concepts of "computational stability" and "robustness" (resistance to outliers). The preferred difference scheme for a computational experiment should be the least robust of the stable ones. The question arises: "Is it possible, after all, to calculate the oscillations so that the amplitude is preserved?". It is enough, for example, to use the Crank-Nicolson difference scheme [16] (**Fig. 3**):

$$\frac{X_{i+1} - X_{i-1}}{2\tau} = iwX_i \Rightarrow X_{i+1} = X_{i-1} + 2iw\tau X_i; X_1 = A; X_2 = (1 + iw\tau)A.$$

The effects of computational instability manifest themselves not only in oscillation models, but also when oscillations are not visible, for example, in Richardson's three-layer explicit difference scheme for the heat conduction problem [7].

Many problems of the computational experiment have long been solved, and the

solutions have been repeatedly tested. In order to evaluate other people's non-trivial solutions, in the new version of the program it would be worthwhile to give more attention to the recommendations of the libraries of standard programs. Theorems on the properties of difference schemes [7] are quite general. In [17], for example, for the equations of mathematical physics, the possibility of combining difference schemes with a fast Fourier transform and embedding solutions in the Sobolev space was realized. Using this approach, it was possible to obtain by numerical methods not only the self-constructed of dissipative structures and magnetic domains, but also solitons.

In the case of nonlinear problems, it is difficult to prove computational stability, but additional checks can be included in the computational algorithms. The emergence of programs for working with sparse matrices [8] expands the possibilities for implementing numerical methods and makes the application of the basic definitions and approaches of the theory of difference schemes even more relevant [7]. The mathematical approach to the analysis and justification of the synthesis of difference schemes implies familiarity with functional analysis and the ability to immerse difference schemes in completed spaces [18].

4. VERSIONS OF SOFTWARE PRODUCTS

In the 1990s, a certain classification developed for software versions. Alpha version – the first versions of the developed program. They are workable, but only the developers themselves can really use them. Beta version is a more user-friendly version of the program. It is assumed that the user

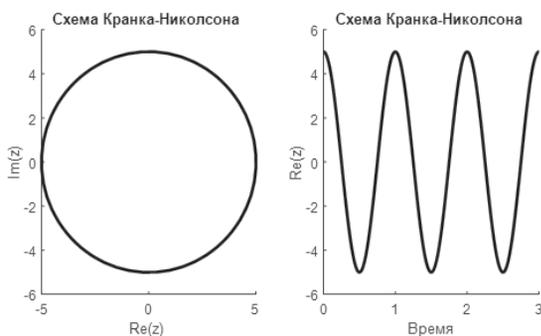


Fig. 3. *Almost constant oscillation amplitude.*

uses it at his own peril and risk, that he himself will deal with his questions.

Initially, beta versions were intended for a limited circle of consumers qualified enough to use the program for free and without claims, despite possible errors and failures. Now a wider audience is participating in beta testing. The formation of this circle of people ensures further promotion of the product on the market.

GitHub cloud service allows programmers to share their developments and software. It supports materials quite conveniently. Releases – further improvement of the software product. The release numbering includes the version number and the revision number. For example, Windows 3.1 was the first widespread version of Windows. Sometimes versions are referred to by the year they were released, or they are given special names.

When developing and using software, most of the work time is spent on finding and fixing bugs, i.e. errors, inconsistencies with program specifications, inconsistencies with user expectations. These inconsistencies can be eliminated only in the simplest cases by formulating and testing hypotheses. Bug finding and testing problems are advisable to consider as a kind of computational experiment with version of program.

Modern browsers have modes that allow you to track document structures and related data, display formats, messages sent between form elements. Known software tools can record and reproduce the actions of tester during testing. Test results should be accumulated in a database with the ability to display in a form convenient for further use. The volume of required manual testing can and should be reduced by automating the synthesis of tests and the implementation of the testing process [11]. To perform many

operations in parallel in the development and maintenance of software products, some firms use a outsourcing policy. There are firms that specializes in software product testing and new specializations of programmers.

If the elimination of a bug affects the work of other parts of the program, then would reasonable its postponing eliminating to a new version. Another simplified way to get rid of complex bugs is to create a specialized version with more narrow capabilities. Sometimes a demo version is useful. These bugs handling methods are to the levels of a mathematical model and a full-scale experiment, programming. But a level is deepest, at which programs are not explored, but are used as a means of solving problems.

5. CONCLUSION

The paper considers three aspects of a computational experiment: non-dimensionalization of the original model, computational instability of solutions, and testing as a kind of computational experiment. It is shown that at the stage of non-dimensionalization, the physical model is transformed into a mathematical one. This transformation makes it possible to reasonably proceed to the next stage of the computational experiment – programming. Parallelization of the development of algorithms and program versions for different users and purposes is based on an assessment of data complexity and errors. Object-oriented programming tools and version control systems are used.

For complex calculations and big data, it is advisable to develop software versions based on high-performance computing tools: on hardware-software acceleration, parallel information processing and cloud architectures.

As a rule, the new version of the program and a difference model are built to solve the problem by iterative methods. Conditions for computational stability usually specified in modern instructions for standard software libraries. For new algorithms, when analyzing stability of difference schemes their spectral properties are refined and methods of functional analysis are used.

For storage and subsequent use, the obtained solution data are converted for their accumulation in modern databases. If necessary, you can use distributed databases and cloud storage.

As a kind of computational experiment, testing and use of alpha and beta versions and releases are also considered.

REFERENCES

- Mikhailov AP, Chetverushkin BN. COMPUTATIONAL EXPERIMENT. *Great Russian Encyclopedia*, vol. 6, p. 161. Moscow, 2006. Electronic version (2016), <https://bigenc.ru/mathematics/text/2379725>.
- Samarskii AA, Mikhailov AP. *Principles of Mathematical Modeling. Ideas, Methods, Examples*. London and New York, Taylor and Francis Publ., 2002, 349 pp.
- Krasovskii GI and Filaretov GF. *Planirovanie eksperimenta* (Design of Experiment), Minsk: Belarus State University Publ., 1982, 302 p.
- Baytser B. *Architecture of computer systems*, v. 1. Moscow, Mir Publ., 1974, 498 p.
- Kolmogorov AN(author), Shiryaev AN (editor). *Selected Works III: Information Theory and the Theory of Algorithms*. Astoria, NY 11106, Springer, 2012, 304 p.
- Arzamastseva GV, Evtikhov MG, Lisovsky FV, Mansvetova EG. Light diffraction by fractals: comparison of experimental data with the obtained by numerical methods fourier images of the object pictures. *RENSIT: Radioelectronics. Nanosystems. Information technologies*, 2017, 9(2):221-228. DOI: 10.17725/rensit.2017.9.221.
- Samarskii AA. *The Theory of Difference Schemes*. Boca Raton, CRC Press, 2001, 786 p.
- Tewarson RP. *Sparse Matrices*. NY, Acad. Press, 1973.
- Cramer H. *Mathematical Methods of Statistics*. Princeton University Publ., 1999, 575 p.
- Kolmogorov AN. *Selected Works II: Probability Theory and Mathematical Statistics*. Astoria, NY 11106, Springer, 1992, 615 p.
- Ham Vocke. *The Practical Test Pyramid*. [Electronic resource] // martinowler.com. URL: <https://martinowler.com/articles/practical-test-pyramid.html> (date of the application – 12.04.2022).
- Evtikhov VG, Evtikhova NV, Evtikhov MG, Evtikhov MV. *Vysokoproizvoditelnye vychisleniya* [High-performance computing]. Kazan, Buk Publ., 2020, 150 p.
- Evtikhov VG, Evtikhova NV. *WEB-razrabotka v obrazovatelnykh i vikipediynykh proektakh* [WEB-development in educational and Wikipedia projects]. Kazan, Buk Publ., 2018, 112 p.
- Evtikhov MV, Evtikhov VG. Vizualizatsiya istoricheskikh dannykh [Visualization of historical data]. *Molodoy uchenyi*, 2019, 1(239):1-5 (in Russ.).
- Godunov SK, Ryabenky VS. *Raznostnye skhemy* [Difference schemes]. Moscow, Nauka Publ., 1973, 440 p.
- Marchuk GI. *Methods of Numerical Mathematics*. Berlin-Heidelberg-New York, Springer-Verlag, 1975, 316 p.

17. Evtikhov MG. On the numerical solution of the equations of mathematical physics with nonlinear sources. *Journal of Communications Technology and Electronics*, 2007, 52(8):1-4.
18. Anuchina NN, Babenko KI, Godunov SK, Dmitriev AN, Dmitrieva LV, D'yachenko VF, Zabrodin AV, Lokucievskij OV, Malinovskaya EV, Podlivaev IF, Prokopov GP, Sofronov ID, Fedorenko RP. *Teoreticheskie osnovy i postroenie chislennykh algoritmov zadach matematicheskoy fiziki* [Theoretical foundations and design of numerical algorithms for problems of mathematical physics]. Moscow, Nauka Publ., 1979, 295 p.

DOI: 10.17725/rensit.2022.14.341

Triad matter - substance - material

Sergey P. Gubin

Kurnakov Institute of General and Inorganic Chemistry of RAS, <http://www.igic.ras.ru/>
Moscow 119991, Russian Federation

Email: gubin@igic.ras.ru

Received April 01, 2022, peer-reviewed April 08, 2022, accepted April 15, 2022

Abstract: An essay on the philosophical aspects of materials science in the absence of a modern philosophical category "material" is presented. The features of the categories "substance" and "material" are considered, which express the distinctive properties of the corresponding objects of nature - the formlessness of the substance and the proper form of the material, the size of the particles of the substance and the structural form of the material. Substance-material mutual transformations, their energy aspects, a certain conditionality of the substance-material boundary, and adequate filling of the corresponding categories are analyzed. Also noted is the virtual reality of the elements of the structure of matter and material, its dependence on the distorting influence of experimental tools. The prospects of translating the regularities of matter-material transformation to extraterrestrial objects are noted.

Keywords: materials science, material shape, substance particle size, substance-material transformation

UDC 1:001; 001.8; 02.31.31; 620.22: 81.09.01

For citation: Sergey P. Gubin. Triad matter–substance–material. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2022, 14(3):341-348. DOI: 10.17725/rensit.2022.14.341.

CONTENTS

1. INTRODUCTION (341)
 2. MATTER (342)
 3. SUBSTANCE (342)
 4. MATERIAL (343)
 5. SUBSTANCE-MATERIAL INTERCONVERSIONS (345)
 6. CONCLUSION (348)
- REFERENCES (348)

1. INTRODUCTION

A person, having once been on the planet Earth, was forced, in order to ensure his existence, first of all to establish interaction with the surrounding material world. This interaction, the development of the surrounding world proceeded empirically, by the method of "trial and error". And it turned out that the surrounding world, perceived by Man, with all its diversity, consists of objects of two types – having their own FORM and FORMLESS media – gaseous, liquid, dispersed (loose), etc.

Being himself an integral part of the material world, from the first steps of his formation, Man began not only to "master" the surrounding material world, but also to adapt (remake) it to his

needs. Since then, this process has not stopped even for a minute.

However, since ancient times, it has turned out that in the philosophical understanding of the surrounding world, the categories of *spirit* and *matter* (*primary substance, substance*) turned out to be not equivalent from the point of view of Man: "mundane" matter somehow attracted his attention less than everything connected with the spiritual side his life. Somehow, there was an idea behind the scenes that the material world is something "simple" that does not require special attention. Philosophy of nature, ancient and medieval – natural philosophy from the pre-Socratics and Aristotle [1] to Thomas Aquinas [2] was richly kneaded in the "expression of will" of spiritual forces, mystical within the existing level of scientific knowledge.

REMARK

It should be noted that and at the present time both *matter* as a whole and its categories such as *substance* and, in particular, *material*, do not often appear as a basis (subject) for discussion on the pages of philosophical publications [3,4]. There

is a section of the science – materials science, but there is no definition of what a material is. There is no material classification. At the same time, any Person always deals with materials, is immersed in the material world.

There is an urgent need for a modern philosophical interpretation of such forms of reason as the categories of *substance* and *material*, which interpret the corresponding states of matter, their mutual transformations and place among other types of matter. While Man in everyday life and activity deals with objects of precisely these two categories of both mineral and biological origin, their modern philosophical interpretation leaves much to be desired. Therefore, the concentration on these issues of the attention of specialists from different areas within the framework of anthropocentric perception of the material world does not seem superfluous. In this paper, an attempt is made to place accents in the interpretation of the categories matter – substance – material.

2. MATTER

The world around us in the Human consciousness (perception) is divided into material – given to us in direct sensations by our senses, and "non-material" – created by us in our consciousness on the basis of the field attributes of matter, including information, and their analysis and synthesis. Although there is often no sharp boundary between the two worlds in the mind, such a habitual dichotomy allows a Human to navigate in the environment, trying to use the objects of the material world not only to ensure his own existence, but also to cognize the "non-material" world with their help.

Usually, both specialists working in the field of natural science and technology and ordinary people deal with matter in three types – either with compact objects in the form of subjects, products, samples, pieces of rock, etc. (we will call them materials), or with shapeless objects in the form of powders of varying degrees of dispersion, liquids or gases (let's classify them as

substances), or, finally, with field attributes of the surrounding world given to us in sensations. Both that, and another, and the third are parts of a more general philosophical category of *matter*, which thus represents to Man three types of existing objects – objectively rigid, or completely formless, as well as field.

Of all the types (kinds, varieties) of matter, we will here consider only substances and materials with which Man constantly deals and voluntarily or involuntarily studies (feels, observes, classifies, etc.) directly (real, not speculative), with using his senses: sight, touch, hearing. Staying at the same time in the fields of the surrounding world (gravitational, electromagnetic) and using them as far as possible and necessary.

REMARK

Having set foot on another Celestial body, the first thing a Man will not be interested in is whether he has many electrons, bosons, etc. under his feet. First of all, he will be interested in what kind (state) the matter is under his feet and whether it is possible from it and how to create a material world around him that is close to the Earthly one. That it is a free-flowing shapeless powder in which a heavy object sinks without finding support, or is it a solid compact object that you can lean on. At the same time, the composition, structure, electronic structure, etc. will be of interest to him later. On Earth, Man answered these questions a long time ago (hundreds of thousands of years ago) using the "trial and error" method. On other Celestial bodies, he will never have such a reserve of time. It is necessary to know more deeply the properties of matter in various states, the possible variants of its morphological transformations.

3. SUBSTANCE

At the anthropocentric approach, substance is one of the types of matter perceived by Man as such and at the same time as what the entire material world perceived by him on Earth consists of.

To the category of *substance*, we will refer objects of the surrounding world that do not have their own FORM under normal conditions – gases, liquids, loose powders and other dispersions. Being left to themselves, these objects, under the influence of the Earth's force field, spread (scatter) over the surface in a thin layer. The constituent parts (powders, flakes, drops, etc.) of a substance are too loosely interconnected to resist the action of the Earth's force field, which, when freely poured out (pouring out), usually leads to uniform distribution (spreading) over the surface so that individual parts experienced the same force of Earth's gravity. Such objects (substances) usually take the form of the that volume (vessel, can, box, etc.) in which they are located.

The key parameter of a substance is the SIZE of its constituent macro- and microparticles (crystallites, molecules, clusters, atoms, elementary particles) and their interaction. And although the category – *substance* appeared in philosophy quite a long time ago, its current definition is completely (without any participation of philosophy) formulated in terms of modern physics.

3.1. NANO-CONDITION OF SUBSTANCE

The dispersion (and dispersibility) of a substance (including spontaneous) is determined, depends on the ratio of the energies of intra- and interparticle interaction. At the same time, large material objects (large crystals, large ingots, etc.) do not differ in properties from smaller samples. In contrast, when materials (and substances) are crushed to sizes nano-, there comes a moment when the properties change, moreover sharply. Such a nano-condition of substance can be preserved not only in dispersed media, but also in compact materials. Similar discreteness is due to energy factors.

Nanosized particles of substance – nanoparticles – one of the forms of existence of matter. A person cannot fix the nano-condition of matter without instruments

and equipment, but special properties of nanoobjects are detected in some cases.

REMARK

A nanoparticle (NP) is a part of the whole, having dimensions of several nanometers; "whole" is a substance in the form of a gas, liquid or solid body (material). Having the same composition, NPs differ in properties (often significantly, sometimes radically) from the "parent" material (gas, liquid, solid). In other words, without changing the composition of the sample, it is possible to transform it by grinding into a substance (material) with fundamentally different properties. This feature requires the separation of NPs into a separate class of objects that differ from atoms, ions, molecules, or clusters, on the one hand, and micropowders, granules, crystallites, etc. – with another.

Thus, from the point of view of Man, the category – *substance*, denoting the corresponding condition of matter, has an ubiquitous character on Earth, but is perceived by Man rather abstractly. The category *substance* is broader than what is usually called material.

4. MATERIAL

Category *material* has long been considered as part of a more general philosophical category of *matter*, while materials (of which there are a great many in the environment of Man) remained out of discussion, as well as their mutual transformations.

In classical philosophy, this category is absent.

A material is a kind of matter that has its own FORM under the Earth's p, T conditions – the key parameter of the material. In other words, a material is a substance in which the individual components interact with each other strongly enough to resist gravity and not crumble.

When the FORM is lost (during crushing, melting, etc.), the material turns into a substance, as a rule, of the same

composition. This process is reversible, there are a significant number of techniques and methods for giving the substance its own form, i.e. transformation it into material. Material objects can take various forms, the role of which is to distinguish this material object from the environment. If the material is in the form of one or another object, then the Person identifies it primarily by its shape, structural features and weight (mass).

Unlike a substance, a material, being in the form of pieces, products, etc. objects, usually has one or another form of its own, which is able to resist the action of the Earth's force field and keeps the substance contained in its composition in the form of various three-dimensional objects – from kilometer-long mountains to thin films. Individual components in the material are linked into a single whole.

A person has long learned to give a free-flowing formless substance one or another form he needs, turning it into a material object. The substance turned into a material begins to show new qualities (properties) that are not characteristic of a shapeless mass. For example, quartz single crystals (*material*) are used in spectral experiments, but are unsuitable for polish, while fine quartz powder (*substance*) is excellent for polish, but it is unsuitable for using in spectral devices. In other words, the shape of a material can change his properties when transformed from a substance and vice versa; when a material is created from a substance, the latter often acquires new properties. That is Man in his practical activity has long learned to turn substance into material and back material into substance.

As a result, a Man at any given moment can have matter at hand in the form that is convenient for his practical tasks (goals). The ability to transform the surrounding matter into the necessary state is the greatest achievement of Mankind.

REMARK

Herewith we are talking ONLY about inanimate nature, biotransformations of matter are beyond the scope of this consideration.

4.1. FORM

The substance has a composition and an structure, but does not have its own form, while any material reality cannot remain formless for a long time. The form limits the place in space that a given material object occupies. The form allocates (isolates, separates) the object in the general flow of matter.

What is new in a substance when it takes on a form? Being in a dispersed state, all parts of matter interact equally (equally) both with the force field of the Earth and with the components of the environment – radiation, gases, liquids. Giving shape to a formless substance is a "fight against entropy", the material has better organization and less chaos than in the original substance. As soon as the transformation of substance into material has taken place, the differentiation of substance into internal and external immediately appears. These two parts of the same substance begin to differ not only in morphology and structure, but often also in composition. The transformation of a completely homogeneous substance into a material is always accompanied by the appearance of a certain amount of heterogeneity (differentiation) in it. The original homogeneous substance is transformed into a more complex, non-single-component object. At the same time, one of the essential components of the material is a special structure – the shell.

4.2. SHELL

Any material object exists in the environment. Between the material and the environment, by definition, there is a border divide where the objects in contact interact with each other. So, any surface to some extent sorbs the components of the environment. In addition, parts of the material that are on its surface are in an asymmetric environment

due to energy differences inside the material and outside it. At the same time, individual structural elements do not stick out on the surface as fragments of broken bonds, but find opportunities to interact with each other, which leads to the emergence of a new substance – SHELL, covering the surface of a material object.

POSTULATE

The shell that separates the material from the environment is an integral his component, which has its own structure and properties.

Since the shell is an self-dependent component of the system, it is counterproductive to neglect its existence. The shell of the material always exists, although it may be atomically thin and hardly noticeable, or it may be quite perceptible.

The material is in the shell as in a cocoon, often it supports it, protects it, determines the connection with the outside world. The composition and structure of the shell (interface) always differs to a greater or lesser extent both from what is inside and, of course, from the environment in which the object is located. The shell is always the result (product) of the interaction of internal and external. And the contribution of each of these two parts of the material can vary from almost 0 to almost 100%, but both of these components are always present in any sample.

The shell, being atomically thin, is such that it is usually considered that it does not exist, but it certainly exists and will manifest itself under certain conditions. The shell-core ratio of the material can be arbitrarily small, but there will always be a situation in which the influence of the shell will be noticeable.

Friction, corrosion, absorption or reflection of energy, and many other processes begin (and often are realized) precisely in the surface layer (shell) of material. Thus, the shell is an integral part of the material.

5. SUBSTANCE-MATERIAL INTERCONVERSIONS

Within the framework of Kant's empirical realism taken as a basis in this work, matter in any form on Earth is a reality given to us in direct sensations. This is the first, "SURFACE" level of Man interaction with the material world. There is no place for electrons, atoms, ions, clusters, nanoparticles, etc. – that is, elements with the use of which they usually distinguish, classify, build the structure of substances and materials. This second level is no already based on the results of Man direct observations.

The reality of the above elements of the structure of substance and material is usually based on the data of devices and the results of calculations. A Man cannot directly receive information about these structural elements. For a Man, this is another, virtual reality, the fruit of our consciousness. Their perception is due to the presence of the necessary instrumental base and computer technology in the Man. Without them, the presence of these structural elements in substances and materials remains speculative.

A person was forced to place between matter and himself a certain device (let's call it a device in a broad sense), with the help of which it became possible to look inside a substance or material. Over time, such "devices" appeared a great many.

But any devices to some extent distort information about the object. A Man is forced to constantly recheck the information received in this way in order to confidently navigate in the surrounding material world.

From the frequent use of these virtual concepts (nothing more), they are often operated on as if they had become quite real, tangible; they are drawn, molded, models are built, etc. But it's still a typical virtual reality. The boundary between the virtual and real worlds in the Man consciousness is gradually blurred.

The internal structure of substances and materials is outside the scope of our anthropocentric research – since the Human senses cannot directly and reliably identify objects smaller than 10 microns, and even more so – their internal structure.

5.1. MUTUAL TRANSFORMATIONS SUBSTANCE - MATERIAL

Nature knows how to turn materials into substances and substances into materials. A person has mastered (created new ones) many techniques (technologies) that allow such transformations to be carried out.

The conditionality of the substance–material boundary is illustrated by the following example: a finely crystalline powder of table salt is a typical substance, and a cubic single crystal grown from it is a material. Plates are cut out of it and used in spectral studies. At the same time, the composition and structure remain unchanged. This is a typical transformation of substance into material.

On the other hand, you can take a microcrystal of salt from a salt shaker and print an electronic circuit of nanoscale on it. In this case, the microcrystal, remaining a substance, turns into a substrate for the circuit – a typical material, while the composition and internal structure remain unchanged. This example illustrates the conventionality of the boundaries between the categories *substance* and *material*. Any material is a substance, but not all substance is a material.

With such reasoning, it is necessary to introduce one or another coordinate system. Such transformations are constantly taking place in the material world surrounding Man. They are divided into natural and artificial. The first occur without the participation of Man. Materials once created by Nature in the form of massive ridges, mountains, plates and other similar material objects over time (billions of years) spontaneously transform from shaped materials into shapeless fine dispersions of various types – sands, water dispersions, solutions, etc.

But Nature is able not only to transform materials into substances, but also to transform substances into materials. Most often these are thermal processes under pressure, etc.

From the moment of his formation, man has existed in the environment of substances and materials created by Nature, but gradually their "range" ceases to satisfy him. Man began to master techniques that allowed him to give the substance the shape and functionality he needed, and to mold various material objects from loose substances. A person has mastered (created) many techniques (technologies) that allow such transformations to be carried out.

Such transformations occur only under certain conditions: for example, a liquid substance – water (a typical formless substance) turns into a material – ice below 0°C, CO₂ gas (substance) below minus 80°C turns into a fairly strong material, the so-called dry ice. These typical materials with increasing temperature (changing p, T conditions) pass from materials to matter. These are phase transitions of the first kind. Many other types of such transformations are known.

Thus, the matter surrounding Man on Earth appears to him in the form of two categories of substance and/or material, which are in constant mutual transformations.

5.2. ENERGY

Matter in the form of a dispersed substance has a large reserve of internal energy compared to that which has a compact material of the same mass and composition. The transformation of a compact material into a dispersed substance (chaotization of matter) is accompanied by an increase in the internal energy of the system, while when an object is ordered (for example, when a substance is transformed into a crystalline monolith), excess energy is released in one form or another.

When solving the problems that arise before it, Nature makes extensive use of this regularity in the processes of conservation and reconversion of energy not only on a local,

but also on a planetary scale. A person is still (thank God!) too weak technically to effectively participate in these processes.

Note, however, that as the Earth ages, the processes of dispersion of materials on its surface gradually begin to prevail over compaction: in other words, there is more and more matter in the dispersed state on the surface of the Earth; perhaps this is a temporary phenomenon.

Man, together with the matter surrounding him, being on the surface of the Earth, is exposed to the environment of the force and magnetic field of the Earth, thermal radiation of the Earth and the Sun, cosmic rays, atmosphere, pressure, aggressiveness of the environment, etc. The types of matter surrounding Man – substances and materials – react differently to the action of these environmental factors. In a compact material, the main "impact" is taken by the surface layer (shell). Under the influence of the factors listed above, the shell can change not only the structure, but also the composition, continuity, etc. The change in the phase state of matter (melting, dissolution, evaporation), which is in the form of a substance or material, occurs differently, the boundaries of phase stability depend on the degree of dispersion of the object. Since matter in the form of a substance has a much larger surface area per unit mass than that of a compact material, the results of external influences are manifested more intensely, distinctly, brightly.

5.3. ABOUT SIZES AND DISPERSION

Human interaction with the material world is limited by the capabilities of his senses. Most of the micro- and the entire nano-world remain outside the limits of direct perception by Man.

The properties of the material do not change when the dimensions of the sample are changed up to certain limits. A giant (weighing tens of kilograms) sample of quartz in the form of a single crystal and a small quartz plate have the same set of properties. On the other hand, the material in the form of one or another

three-dimensional object loses its shape during grinding and turns into a powder (substance), as a rule, without changing its composition (in this case, into quartz powder). But in the usual sense, the object ceases to be a material and passes from a real quartz to a virtual state, when a Person is no longer able to directly (without any auxiliary devices) identify it as quartz, and not table salt. In other words, it is difficult for a person to distinguish fine quartz powder from other equally dispersed powders. The sample loses its individuality.

Changing the size of an object first of all changes the ratio of the surface layer (shell) to the total mass of the object. For the majority of real material objects surrounding Man, these changes are insignificant and, as a rule, they are neglected. The size of objects from 100 to 10 microns is the limit where the possibilities of direct (only using touch and vision) interaction between a Human and the outside world end. Deeper – only with the help of auxiliary devices (devices, etc.).

AXIOM

All auxiliary research devices to some extent distort information about the object under study. Therefore, when creating a picture of an object less than 10 microns close to reality, a combination of instruments and methods and a cross-analysis of the results are used.

Based on the above, the following definition can be given.

Materials science is a branch of science about a substance that is in the form of compact materials of various shapes, and about the relationship between their structure, physical and chemical properties and performance characteristics.

REMARK

The term "materials science" is widely represented in educational, scientific and technical literature. As usual, each author gives his own definition; their diversity reflects the versatility of the topic and the relative youth of this branch of science.

6. CONCLUSION

Being in the Earth conditions, a Man from all types of matter primarily interacts (communicates) with substances and materials that ensure his existence not only as a person (subject), but also allow him to claim the role of a global player in the UNIVERSE, while still not far from the Earth, but in the future this role is quite definite. In this regard, it can be assumed that having ever been on another celestial body, where other chemical (composition of the surface and gaseous shell) and physical (p, T , gravitational and electromagnetic fields) conditions on the surface, Man, as well as on Earth, first of all, will meet matter in the form of a formless mass of substance and various compact material objects of various shapes and hardness. It will be necessary to know in what form the matter is on the surface of a given celestial body (substance-material ratio) and how these types of matter will interact with Man. Later, it will be necessary to know how the conditions of substance-material interconversions on a given celestial body differ from those on Earth. Without this, it will hardly be possible to gain a foothold in the new territory. Transporting matter (substances and materials) from the Earth is counterproductive. Matter is everywhere – Man needs to learn how to turn it into substances, materials and vital objects that he needs at a given point in the Universe.

If we figure out how such transformations take place on Earth, the translation of these regularities to extraterrestrial objects will become more reliable.

It can be assumed that the two types of matter discussed above will, just as on Earth, constitute the main environment of Man. And he will need not only to understand, but also to manage the processes of their mutual transformations in conditions different from those on Earth. It is quite possible that matter exists on other celestial bodies in other forms than on Earth,

but the essence of their primary anthropocentric perception is unlikely to be different.

The quite natural desire of Man to understand the internal structure of material objects diverted attention from the "lying on the surface" properties of matter. It can be seen that even before "looking inside" it is possible to investigate the interaction between Man and matter.

REMARK

Biological materials and, in general, issues of materials science of biological systems are not considered here. The main attention here is paid to the issues of compatibility of "ordinary" materials with biological (including living) systems. Solid parts of living organisms such as bones, shells, shells, corals, etc. are not usually the subject of materials science research, although their role in the formation of the earth's crust can hardly be overestimated. When considering these phenomena, one would inevitably have to touch upon the philosophical aspects of the "living-non-living" problem, which is beyond the scope of this essay.

REFERENCES

1. Aristotle. Works, vol. 3. Physics. Moscow, Mysl' Publ., 1981, 616 p.
2. Aquinas Thomas. Works. Moscow, Lenand Publ., 2021, 264 p.
3. Gaidenko PP. The history of modern European philosophy in its connection with science. Moscow, Librokom Publ., 2009, 376 p.
4. Lange FA. A history of materialism and a critique of its current meaning: a history of materialism after Kant. Moscow, URSS Publ., 2010, 746 p.