

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 **2 times a year** (that includes 2 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, 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).

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

- METHODS FOR CALCULATING RADIO HOLOGRAMS OF VOLUMETRIC OBJECTS
Valery A. Golunov, Vadim A. Korotkov 429
- OPTIMIZATION OF A POLARIZATION NEPHELOMETER
Viktor G. Oshlakov, Anatoly P. Shcherbakov, Yaroslav A. Ilyushin .. 437

SOLID STATE PHYSICS

- USING OF STOPPING AND RANGE OF IONS IN MATTER CODE TO STUDY OF RADIATION DAMAGE IN MATERIALS
Hiwa Mohammad Qadr, Ari Maghdid Hamad 451

NANOSYSTEMS

- NANOPARTICLES SURFACE AS A SPHEROIDAL DEFECT IN THE PARENT MATERIAL
Sergey P. Gubin, Yulia V. Ioni, Elene Yu. Buslaeva 457
- HIGHLY CONDUCTIVE POLYMER PEDOT: PSS - APPLICATION IN BIOMEDICAL AND BIOELECTROCHEMICAL SYSTEMS
Anatoly N. Reshetilov, Anna E. Kitova, Sergey E. Tarasov, Yulia V. Plekhanova, Alexander G. Bykov, Ashok K. Sundramoorthy, Iren E. Kuznetsova, Vladimir V. Kolesov, Pavel M. Gotovtsev 471

INFORMATION TECHNOLOGIES

- INTERFEROMETRIC PROCESSING OF ACOUSTIC INFORMATION BY USING EXTENDED ANTENNAS IN DISPERSING MEDIA
Venedikt M. Kuz'kin, Sergey A. Pereselkov, Sergey A. Tkachenko, Ilia V. Kaznacheev 483
- RESTORING MISSING FRAGMENTS OF A DISTORTED IMAGE DUE TO DEFOCUSING OR BLUR OF A KNOWN HARDWARE FUNCTION
Alexander V. Kokoshkin, Evgeny P. Novichikhin 495
- ANALYZE AND EVALUATE THE PERFORMANCE VELOCITY CONTROL IN DC MOTOR
Hazim M. Alkargole, Abbas S. Hassan, Raof T. Hussein 507
- INFORMATION PROCESSING ALGORITHMS IN AVIATION-BASED RADIOELECTRONIC SURVEILLANCE SYSTEMS
Ivan D. Biryukov, Pavel V. Buchuchan, Polina I. Timoshenko 517

SCIENCE METHODOLOGY

- TOWARDS A NEW PARADIGM
Gennady V. Mishinsky 529



RUSSIAN ACADEMY
OF NATURAL SCIENCES

DEPARTMENT OF
RADIOELECTRONICS,
NANOPHYSICS AND
INFORMATION TECHNOLOGIES
PROBLEMS

RENSIT:

**RADIOELECTRONICS.
NANOSYSTEMS.
INFORMATION
TECHNOLOGIES.**

2020, VOL. 12, № 4

FOUNDED IN 2009

4 ISSUES PER YEAR

MOSCOW

Editor-in-Chief

VLADIMIR I. GRACHEV

grachev@cplire.ru

Deputy Chief Editor

Alexander S. Ilyushin, 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. Cherepenin, DrSci, IRE

Alexander S. Dmitriev, DrSci, MSU

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

Vladimir A. Makarov, DrSci, MSU

Alexander V. Okotrub, DrSci, SB RAS

Aleksey P. Oreshko, DrSci, MSU

Igor B. Petrov, DrSci, CM RAS, MIPT

Alexander A. Potapov, DrSci, IRE

Vyacheslav S. Rusakov, DrSci, MSU

Alexander S. Sigov, DrSci, RAS, MIREA

Valentine M. Silonov, DrSci, MSU

Eugeny S. Soldatov, PhD, MSU

Arkady B. Tsepelev, DrSci, IMET

Lkhamsuren Enkhdorj, DrSci (Mongolia)

Yoshiyuki Kawazoe, DrSci (Japan)

Kayrat K. Kadyrzhanov, DrSci (Kazakhstan)

Peter Paul Mac Kenn, DrSci (USA)

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 2020

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.2020.12.429

Methods for Calculating Radio Holograms of Volumetric Objects

Valery A. Golunov, Vadim A. Korotkov

Kotelnikov Institute of Radioengineering and Electronics of RAS, Fryazino branch, <http://fire.relarn.ru/>

Fryazino 141190, Moscow Region, Russian Federation

E-mail: golsnow@mail.ru

Received July 07, 2020; peer reviewed July 27, 2020; accepted August 10, 2020

Abstract: A method for calculating holograms for volumetric objects based on the representation of objects in the form of ensembles of virtual point sources distributed on a set of parallel planes has been proposed. The proposed method is the development of the well-known method in which objects are represented as ensemble of real point scatterers. The possibilities of the proposed method are demonstrated by calculating a hologram of a fragment of a sphere, on which 1000 points are randomly selected, at which radiation emanating from the center of the sphere is scattered. The choice of a fragment of a sphere as an object under study is due to the fact that when calculating its hologram, phase errors inherent in approximate calculations are most pronounced. The calculations were performed for the frequency range of 2...100 GHz, the sphere radius of 0.5 m, a two-dimensional hologram size of 0.65×0.65 m, and a pixel count of 512×512. It is shown that, in comparison with the known method, the proposed method makes it possible to calculate the amplitude of a hologram with satisfactory accuracy if virtual sources are placed on parallel planes in an amount of more than 64 pieces. In the case of objects that require representation in the form of an ensemble of point scatterers in the amount of more than 1000 pieces, the calculation of their holograms by the proposed method turns out to be much more efficient than the known method. **Keywords:** radio holograms, virtual point sources, parallel planes, sphere.

Keywords: computer modeling, radio holograms, volumetric objects, point scatterers, virtual point sources, parallel planes, sphere

UDC 621.396

Acknowledgements: The work was carried out according to the state order of the Kotelnikov IRE RAS.

For citation: Valery A. Golunov, Vadim A. Korotkov. Methods for Calculating Radio Holograms of Volumetric Objects. *RENSIT*, 2020, 12(4):429-436. DOI: 10.17725/rensit.2020.12.429.

CONTENTS

- | | |
|--|--|
| <ol style="list-style-type: none"> 1. INTRODUCTION (430) 2. THE METHOD OF AN ENSEMBLE OF POINT SCATTERERS (THE METHOD 1) (430) 3. THE METHOD OF VIRTUAL SOURCES ON THE SAME PLAIN (THE METHOD 2) (430) 4. THE METHOD OF VIRTUAL SOURCES ON A SET OF PARALLEL PLANES (THE METHOD 3) (431) | <ol style="list-style-type: none"> 5. COMPARATIVE ANALYSIS OF THE RESULTS OF CALCULATION OF RADIO HOLOGRAMS BY METHODS 1, 2, 3 (432) 6. COMPARATIVE ANALYSIS OF THE COMPUTED TIME OF RADIO HOLOGRAMS BY VARIOUS METHODS (434) 7. CONCLUSION (435) <p>REFERENCES (435)</p> |
|--|--|

1. INTRODUCTION

Radio holography is one of the methods for solving such radar problems as detection, determination of the coordinates of the location and shape of objects [1]. Obtaining radio holograms by experimental methods involves measuring the fields scattered by real objects, which is usually quite difficult. In some cases, an alternative method may be the calculation of radio holograms based on computer simulation of the scattering of electromagnetic waves by objects with desired properties.

Until now, there is no analytical solution of the problem of wave scattering by an object of arbitrary shape. In the general case, the field, scattered by an object, is calculated using the Kirchhoff theorem [2-5]. Another method for modeling radio holograms is a calculation based on the representation of an object as a set of point scatterers [6]. In the case, if these scatterers lie on the same plane, one can use the property of Fourier transform to convolve image and use the fast Fourier transform (FFT) to perform the calculations [1]. However, if the objects are not flat, then using the FFT is difficult.

The aim of this work is to develop a method for calculating radio hologram of volumetric objects, which is based on objects representation in the form of a set of point scatterers.

2. THE METHOD OF AN ENSEMBLE OF POINT SCATTERERS (THE METHOD 1)

Let us represent the volumetric object under investigation in the form of an ensemble of point scatterers and place the radiation source at point I as shown in **Fig. 1**. Let

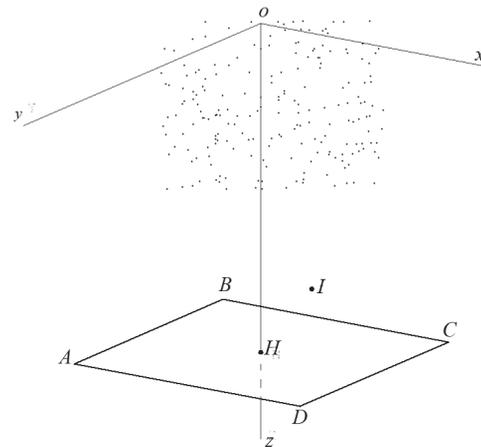


Fig. 1. Scheme for measuring a hologram of an object presented as an ensemble of point scatterers.

the scattered field be measured within the rectangle ABCD with center H on the OZ axis.

The field at the point (x_j, y_j, z_j) can be estimated using the Huygens-Fresnel approximation [1]:

$$E(\vec{r}_h) = E_0 \sum_{n=0}^{M-1} \beta(\vec{r}_n) \frac{\exp(ik|\vec{r}_l - \vec{r}_n|)}{|\vec{r}_l - \vec{r}_n|} \frac{\exp(ik|\vec{r}_h - \vec{r}_n|)}{|\vec{r}_h - \vec{r}_n|}, \quad (1)$$

where M is the number of point scatterers; $\vec{r}_l, \vec{r}_n, \vec{r}_h$ – vectors of the coordinates of the position of the source, the n -th scatterer and receiver, accordingly; E_0 – the amplitude of the source; $\beta(\vec{r}_n)$ – coefficient of reflection (scattering) of the n -th scatterer [1].

3. THE METHOD OF VIRTUAL SOURCE ON THE SAME PLAIN (THE METHOD 2)

Let each point scatter of an object be correspond to a virtual source of a spherical wave in the $PQRS$ plane as shown in **Fig. 2**. The x, y coordinates of the point scatter and the virtual source coincide, and the phase of the wave of the virtual source is adjusted taking into account its change in the distance from the point scatter to the virtual source.

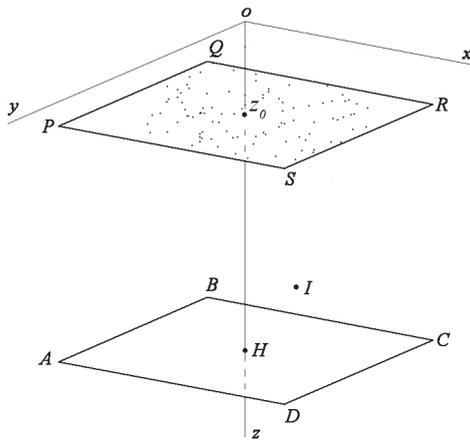


Fig. 2. Scheme for measuring a hologram of an object in the form of an ensemble of virtual sources on the PQRS plane when they are irradiated from point I.

The field at the point (x_h, y_h, z_h) on the ABCD plane is determined taking into account the phase correction of the virtual sources with respect to the corresponding point scatterers using the following relation similar to relation (1):

$$E_1(\vec{r}_h) = E_0 \sum_{n=0}^{M-1} \beta(\vec{r}_n) \frac{\exp(ik|\vec{r}_I - \vec{r}_n|)}{|\vec{r}_I - \vec{r}_n|} \frac{\exp(ik|\vec{r}_h - \vec{r}_n|)}{|\vec{r}_h - \vec{r}_n|} \exp(i|\vec{r}'_n - \vec{r}_n|), \quad (2)$$

where \vec{r}_n is the coordinate vector of the n -th virtual source, $|\vec{r}'_n - \vec{r}_n| = |z_0 - z_n|$ is the distance from the plane ABCD to the n -th point scatterer, $|\vec{r}_h - \vec{r}'_n| = \sqrt{(x_h - x_n)^2 + (y_h - y_n)^2 + (z_h - z_0)^2}$ is the distance from the receiver to the n -th virtual source.

Expression (2) can be calculated using the FFT.

4. THE METHOD OF VIRTUAL SOURCES ON A SET OF PARALLEL PLANES (THE METHOD 3)

When a size of the object along the OZ axis is significant, the errors in the calculation of the hologram, as follows from (2), will increase in comparison with the method 1.

To reduce the errors, we have designed point scatters that make up the object on a series of parallel planes, for example, on three planes $P_1Q_1R_1S_1, P_2Q_2R_2S_2, P_3Q_3R_3S_3$, as shown in Fig. 3, so that on each plane there are virtual sources corresponding to the nearby point scatterers.

Then the field at the point (x_h, y_h, z_h) can be estimated taking into account the above phase correction as follows:

$$E_1(\vec{r}_h) = E_0 \sum_{p=0}^{N_p-1} \sum_{n=0}^{M_{np}-1} \beta(\vec{r}_n) \frac{\exp(ik|\vec{r}_I - \vec{r}_n|)}{|\vec{r}_I - \vec{r}_n|} \frac{\exp(ik|\vec{r}_h - \vec{r}'_n|)}{|\vec{r}_h - \vec{r}'_n|} \exp(i|\vec{r}'_{np} - \vec{r}_n|), \quad (3)$$

where $|\vec{r}'_{np} - \vec{r}_n| = |z_{op} - z_n|$ is the distance from the $P_pQ_pR_pS_p$ plane to the n -th point scatterer, $|\vec{r}_h - \vec{r}'_n| = \sqrt{(x_h - x_{np})^2 + (y_h - y_{np})^2 + (z_h - z_{op})^2}$ is the distance from the receiver to the np -th virtual source, which is the projection of the n -th point source on the $P_pQ_pR_pS_p$, N_p и M_{np} are the number of planes and the number of virtual sources on each plane, respectively.

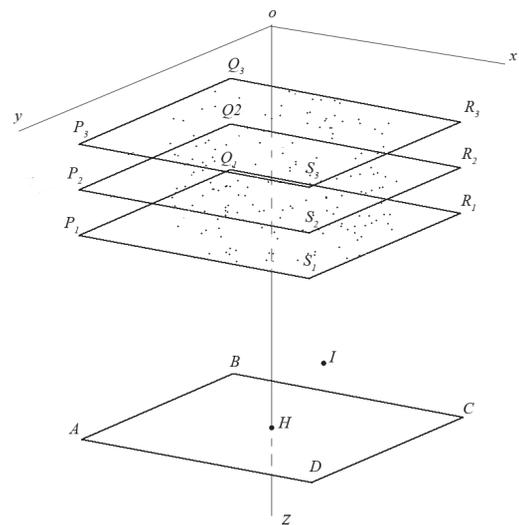


Fig. 3. Scheme for measuring a hologram of an object in the form of an ensemble of virtual sources on the planes $P_1Q_1R_1S_1, P_2Q_2R_2S_2, P_3Q_3R_3S_3$ which are irradiated from point I.

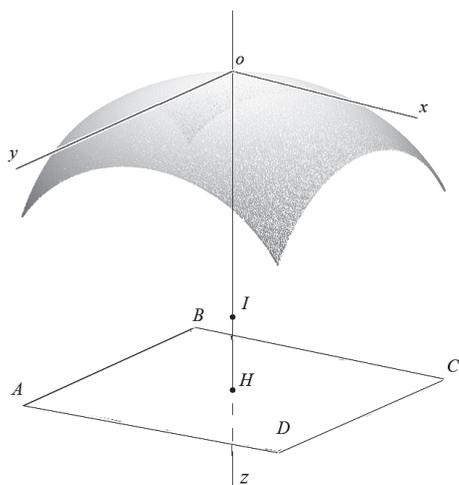


Fig. 4. Scheme for measuring a hologram on the ABCD plane when a fragment of a sphere is irradiated from its center I.

5. COMPARATIVE ANALYSIS OF THE RESULTS OF CALCULATION OF RADIOGOLOGRAMS BY METHODS 1, 2, 3

To compare the efficiency of calculations of radio holograms by the methods considered above, we choose a part of the metal sphere as a test object which, as shown in **Fig. 4**, is irradiated by a spherical wave from a point lying in the center of the sphere.

The choice of such a test object is due to the property of the sphere to focus the waves emitted from its center in its center. Therefore, as a criterion for the effectiveness of the application of the compared methods, one can take their ability, as a result of calculations, to provide high-quality focusing of the scattered radiation.

Next, we perform calculations for a radiation wavelength of $\lambda = 0.0025$ m, the sphere radius $R = 0.5$ m, a two-dimensional hologram size of 0.65×0.65 m with a number of pixels 512×512 . We choose the scattering fragment of the sphere in such a way that its projection onto the measurement plane of the hologram coincides with it in shape and size. Let 1000 point scatterers be randomly selected on the surface of a sphere fragment.

In the calculations, we used the transformation of the back projection algorithm to the form of a two-dimensional FFT developed in [4].

Let us compare the root mean square deviations (RMSD) between the amplitudes of the holograms calculated by methods 1 and 3, depending on the number of planes with virtual sources in method 3.

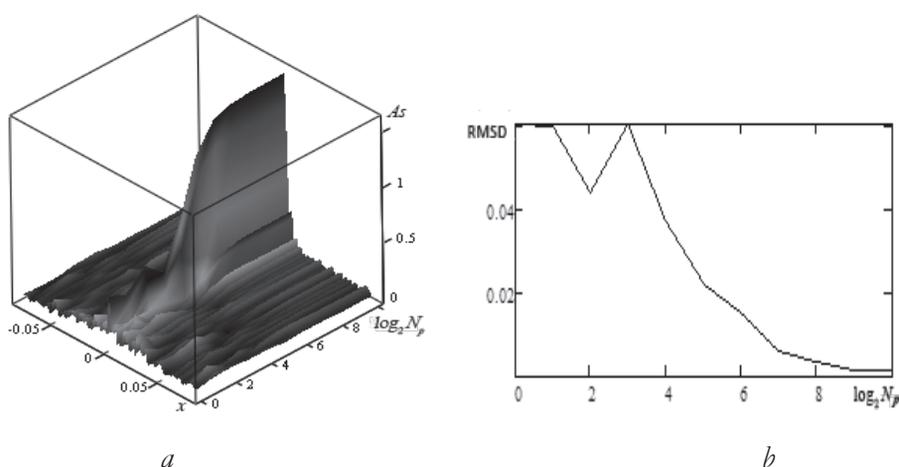


Fig. 5. Sections of the surface of the amplitude of the hologram along the OX axis at $y = 0$ (a) and the graph of the dependence of the RMSD in the center of the hologram (b) depending on the number of planes with virtual sources.

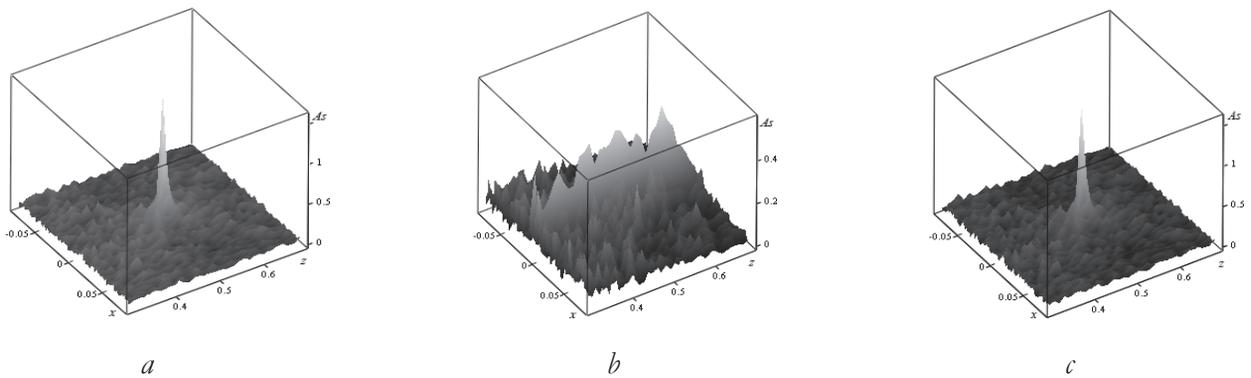


Fig. 6. Sections of the hologram amplitude surface along the OX axis at $y = 0$, depending on the distance z from its plane to the surface of the fragment of the sphere, calculated by methods 1 (a), 2 (b), and 3 (c).

Fig. 5a shows sections of the hologram amplitude surface along the OX axis at $y = 0$, and **Fig. 5b** shows the RMSD diagram at $x = y = 0$ depending on the N_p number of planes with virtual sources. Note that the center of the hologram and the center of the sphere have the same coordinates $x = y = 0$ for different values of z .

When $\log_2 N_p > 6$ (i.e., $N_p > 64$), the amplitude is stabilized as shown in Fig. 5a, and the standard deviation becomes less than 0.01 (i.e. it is negligible) as shown in Fig. 5b. In this regard, in the future, applying method 3, we will use 64 planes with virtual sources.

Next, we consider the dependences of the amplitude calculated by methods 1, 2, 3 at the center of the hologram (i.e., at $x = y = 0$) on the distance to the object along the OZ axis. **Fig. 6** shows sections of the surface of the hologram amplitude along the OX axis at $y = 0$, depending on the distance z from the hologram plane to the surface of the sphere.

It can be seen that the holograms calculated by both methods 1 and 3 have pronounced peaks in its center that are not observed in the hologram calculated by the method 2. More accurate quantitative data of

these dependences are presented in **Fig. 7**, which shows graphs of the dependences of the hologram amplitudes at the central point $x = y = 0$ on the distance z .

From a comparison of these graphs, it follows that, firstly, the results calculated by both the methods 1 and 3 practically coincide, and, secondly, the focus of the hologram calculated by the method 2 is absent.

Next, we consider the dependence of the hologram amplitude on the radiation frequency.

For this purpose, dependences of the hologram amplitude at $y = 0$ were computed on both x and the frequency in the range 2-100 GHz using methods 1, 2,

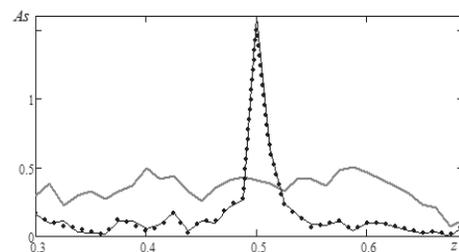


Fig. 7. Amplitudes of holograms at $x = y = 0$, depending on the distance to the surface of the sphere fragment, calculated by methods 1 (solid line), 2 (dashed line), and 3 (dotted line).

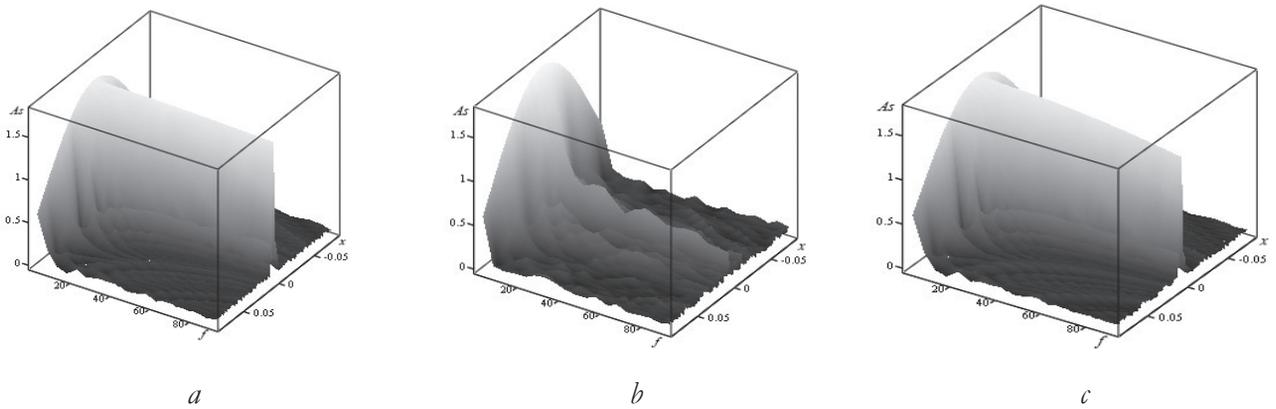


Fig. 8. Sections of the amplitude surface of the holograms along the OX axis at $y = 0$ as a function of x and frequency f computed by methods 1 (a), 2 (b) and 3 (c).

3. The computed data are shown in **Fig. 8a, 8b, 8c.**

Comparing the data in Fig. 8a, 8b, 8c, it follows that the results computed by method 2 at high frequencies differ significantly from those computed by methods 1 and 3. To compare method 1 with method 3, graphs of the frequency dependence of the amplitudes at the center point of the hologram coinciding with the center of the sphere are shown in **Fig. 9.** It is seen that method 3, with a total number of planes equal to 64, makes it possible to calculate the amplitudes of holograms of volumetric objects with an acceptable error in comparison with the

calculation results obtained by the known method 1.

6. COMPARATIVE ANALYSIS OF THE COMPUTED TIME OF HOLOGRAMS BY VARIOUS METHODS

From (1) it follows that the time of direct calculation t_1 of a hologram of size $N \times N$ and an object consisting of M scatterers is estimated as

$$t_1 \sim M \times N \times N \tag{4}$$

that is, it is proportional to the number of scatterers. Since the total number of scatterers in the object under study is 1000, and the number of pixels in the hologram $512 \times 512 = 262\,144$, then, for example, when using a computer with a Pentium E5400 processor, calculations by method 1 will take several tens of hours.

Since expression (2), as noted above, can be calculated using the FFT, so method 2 requires less the computation time t_2 which does not independent on the number of scatterers M and is determined as follows

$$t_2 \sim N \times N \times \log_2 N. \tag{5}$$

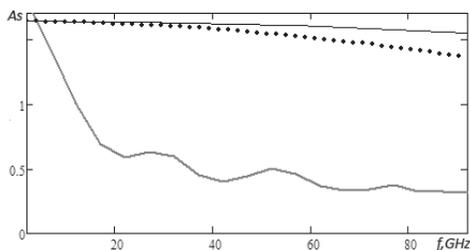


Fig. 9. The frequency dependences of the amplitude in the center of the hologram computed by methods 1 (dotted line), 2 (dotted line) and 3 (solid line).

Method 3 requires computation time t_3 , which is N_p longer than method 2:

$$t_3 \sim N \times N \times N_p \times \log_2 N. \quad (6)$$

If the number of scatterers making up the volumetric object under study is equal to $M \gg (N_p \times \log_2 N)$, then it follows from a comparison of (4) with (6) that method 3, which is a method of virtual sources on a set of parallel planes, significantly reduces the computation time as compared to the known method 1.

7. CONCLUSION

In this paper, a method for calculating holograms for volumetric objects based on the representation of objects in the form of ensembles of virtual point sources distributed on a set of parallel planes has been proposed. The proposed method is the development of the known method in which objects are represented as ensemble of real point scatterers.

The possibilities of the proposed method are demonstrated by the example of calculating a hologram of a fragment of a sphere, on which 1000 points are randomly selected and those scatter radiation emanating from the center of the sphere. The choice of a fragment of a sphere as an object under study is due to the fact that when calculating its hologram, phase errors inherent in approximate calculations, are most pronounced.

The calculations were performed for the frequency range of 2...100 GHz, the sphere radius of 0.5 m, a two-dimensional hologram size of 0.65×0.65 m, and a pixel count of 512×512. It is shown that, in comparison with the known method, the proposed method makes it

possible to calculate the amplitude of a hologram with satisfactory accuracy if virtual sources are placed on parallel planes in an amount of more than 64 pieces. In the case of objects that require representation in the form of an ensemble of point scatterers in the amount of more than 1000 pieces, the calculation of their holograms by the proposed method turns out to be much more efficient than the known method.

REFERENCES

1. Zverev VA. *Radiooptika* [Radiooptics]. Moscow, Sovetskoe radio Publ., 1975, 304 p.
2. Razevig VV. Modelirovanie processa registratsii radiogogramm ob'ektov slozhnoy formy radiolokatorami maloy i sverhmaloy dal'nosti [Simulation of recording the microwave holograms of complex objects by the near range radars]. *Nauka i obrazovanie*, 2014, 6:336-349 (in Russ.).
3. Razevig VV, Bugaev AS, Chapurskiy VV. Sravnitel'nyy analiz fokusirovki klassicheskikh i mul'tistaticheskikh radiogogramm [The comparative analysis of classical and multistatic microwave holograms focusing]. *Radiotekhnika*, 2013, 8:8-17 (in Russ.).
4. Krainy VI, Semenov AN, Chapurskiy VV. Fokusirovka mnogochastotnykh mul'tistaticheskikh radiogogramm metodom neekvidistantnogo BPF [Multi-static multi-frequency image reconstruction based on non-uniform FFT]. *Nauka i obrazovanie*, 2015, 11:292-300 (in Russ.).
5. Razevig VV, Chizh MA, Chapurskiy VV, Ivashov SI, Zhuravlev AV. Numerical

- comparison of mono-static and multi-static array performance in personnel screening systems. *Proc. PIERS-2016*, Shanghai, China, 8-11 Aug, 2016, N.Y., IEEE, 2016:2137-2141.
6. Suhanov DYa, Zav'yalova KV. Beskontaktnaya radiotomografiya cherez neploskuyu granitsu razdela sred v priblizhenii fazovogo ekrana [Non-contact radiotomography across a non-planar interface in the phase screen approximation]. *Estestvennye i matematicheskie nauki v sovremennom mire*, 2014, 18:81-90 (in Russ.).

DOI: 10.17725/rensit.2020.12.437

Optimization of a Polarization Nephelometer

Viktor G. Oshlakov, Anatoly P. Shcherbakov

V.E. Zuev Institute of Atmospheric Optics SB RAS, <http://www.iao.ru/>

Tomsk 634055, Russian Federation

E-mail: oshlakov@iao.ru, molnija2@inbox.ru

Yaroslav A. Ilyushin

Lomonosov Moscow State University, <https://www.msu.ru/>

Moscow 119991, Russian Federation

E-mail: ilyushin@physics.msu.ru

Received June 01, 2020; peer-reviewed June 29, 2020; accepted July 06, 2020

Abstract: The effect of parameters of a polarization nephelometer on its accuracy characteristic is analyzed. Errors in approximation of the actual scattering volume and actual optical beam by the elementary scattering volume and elementary beam are considered. A five-wave monochromatic source of radiation with the high spectral intensity of $0.15 \div 0.6$ W is described. The design of polarization units is demonstrated.

Keywords: polarization, nephelometer, non-isotropic scatterer, scattering matrix, elementary volume, elementary optical beam, approximation error, LED, encoder, microprocessor, precision position control, pulse width modulation, PID controller

UDC 535.36+530.1

Acknowledgements: The authors are thankful to N.N. Bakin, the head of division of Scientific Research Institute of High-Precision Devices, for consultation concerning application of modern LEDs. This study has been accomplished within the framework of State Assignment of IAO SB RAS.

For citation: Viktor G. Oshlakov, Anatoly P. Shcherbakov, Yaroslav A. Ilyushin. Optimization of a Polarization Nephelometer. *RENSIT*, 2020, 12(4):437-450. DOI: 10.17725/rensit.2020.12.437.

CONTENTS

1. INTRODUCTION (437)
 2. OPTIMIZATION OF DIMENSIONS OF A POLARIZATION NEPHELOMETER (438)
 3. OPTIMAL METER OF THE MATRIX $\mathbf{D}(\varphi, L_s, L)$ AND ITS CALIBRATION (442)
 4. SOURCE OF RADIATION (447)
 5. CONCLUSION (450)
- REFERENCES (450)

1. INTRODUCTION

The scattering matrix (D_{mn}) , $m, n = \overline{1,4}$, relates the Stokes vector of the incident radiation \mathbf{S}_i to the Stokes vector \mathbf{S} of the radiation scattered by a small volume dv at an angle φ with respect to the incident radiation (Fig. 1). The scattering matrix for radiation of the optical wavelength range is used in problems of optical wave propagation, laser sensing, polarization filtering, and others. In the general case of a nonisotropic medium, it

is impossible to predetermine identical and zero elements of the scattering matrix. That is why all 16 elements of the matrix (D_{mn}) , $m, n = \overline{1,4}$, are to be determined. The importance and complexity

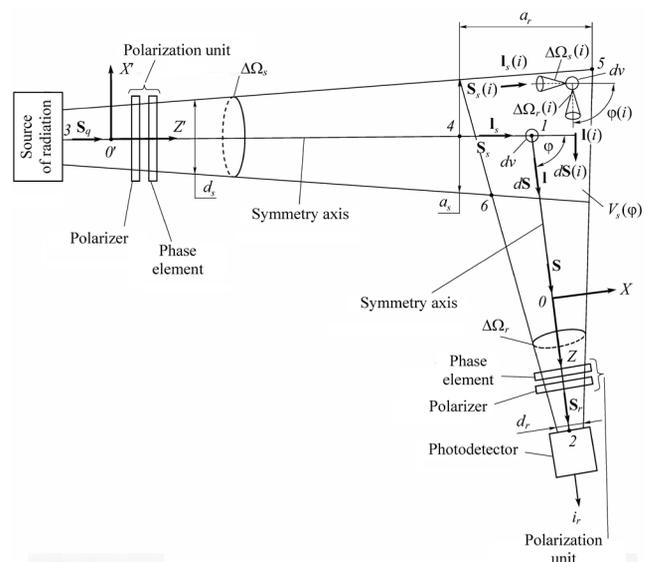


Fig. 1. Block diagram of goniometer of the polarization nephelometer.

of making a device that automatically measures the matrix (D_{mn}) , $m, n = \overline{1, 4}$, were indicated in [1]. This problem is solved methodologically in [2, 3]. An optimal meter of the scattering matrix needs in the proper calibration. The modulation method of measuring the scattering matrix, which was used, for example, in [4, 5], cannot be applied in the laser sensing. Theoretical studies have shown that in the case of an isotropic medium, identical and zero elements of the matrix D can be predetermined. The measurements of the matrix D in [4, 5] employ these results. Consequently, the technique [4, 5] is applicable only in isotropic media. The technique for measuring the matrix of any medium that was proposed in [2, 3] is applicable to laser sensing as well. The achieved level of development of microprocessor technology makes it possible to accomplish completely the task specified in [1].

The scattering volume of a polarization nephelometer and an optical beam incident on the scattering volume should be elementary. However, only qualitative definitions of these terms are known. That is why the paper considers the approximation of the actual scattering volume and the actual optical beam by the elementary ones.

A monochromatic source with high spectral intensity required for polarization measurements can be created using high-brightness LEDs. The analysis of parameters of the polarization nephelometer will allow improvement of its accuracy and size characteristics.

2. OPTIMIZATION OF DIMENSIONS OF A POLARIZATION NEPHELOMETER

Write the Stokes vector \mathbf{S} of radiation in the form

$$d\mathbf{S} = (IQUV)^T, \tag{1}$$

where T is the transposition sign. Then, the Stokes vector $d\mathbf{S} = (dI(i)dQ(i)dU(i)dV(i))^T$ of radiation scattered by the small volume dv (Fig. 1) being at some point i of the scattering

medium at the observation point spaced by $r(i)$ from the scattering volume is [1]

$$d\mathbf{S}(i) = \frac{1}{r^2(i)} D(\varphi(i), \mathbf{I}_s(i), \mathbf{I}(i)) \mathbf{S}_s(i) dv, \tag{2}$$

where $\mathbf{S}_s(i) = (I_s(i)Q_s(i)U_s(i)V_s(i))^T$ is the Stokes vector of the optical beam incident on the volume dv ; this beam propagates within the small solid angle $d\Omega_s(i)$ and its axis is described by the direction vector $\mathbf{I}_s(i)$; $\mathbf{S}(i)$ is the Stokes vector of the optical beam of scattered radiation; it propagates within the small solid angle $\Delta\Omega(i)$ and its axis is described by the direction vector $\mathbf{I}(i)$; $\varphi(i)$ is the angle between the vectors $\mathbf{I}_s(i)$ and $\mathbf{I}(i)$. We adhere to the notation of the parameters I, Q, U, V of the Stokes vector \mathbf{S} as in [2,3].

In the case of a non-isotropic medium, the elements of the matrix $D(\varphi(i), \mathbf{I}_s(i), \mathbf{I}(i))$ are functions of the position of the scattering plane described by the vectors $\mathbf{I}_s(i)$ and $\mathbf{I}(i)$ and the angle $\varphi(i)$. To measure the matrix $D(\varphi, \mathbf{I}_s, \mathbf{I})$ in the open range $\varphi = (0^\circ, 180^\circ)$, the polarization nephelometer constructed by the goniometer scheme [4,5] (Fig. 1) is used. The radiation source generates an optical beam with the diameter d_s (diameter is understood as the maximal cross section size) and divergence $\Delta\Omega_s$. In the course of propagation, the radiation passes successively the polarization unit (PU) of the radiation source. The polarization unit consists of a polarizer and a phase element. The radiation should be maximally monochromatic for the optical path difference of its orthogonal components at the exit of the phase element to be constant. A photodetector with the entrance pupil diameter d_r and the field-of-view divergence $\Delta\Omega_r$ receives the radiation upon the successive passage of the PU phase element and polarizer. The symmetry axes of the optical beam emitted by the source and the photodetector field of view are given by the direction vectors \mathbf{I}_s and \mathbf{I} , respectively. The symmetry axes intersect at point 1. The photodetector should turn around the axis passing through this point and perpendicular

to the plane defined by the vectors \mathbf{I}_s and \mathbf{I} . The angle φ between the vectors \mathbf{I}_s and \mathbf{I} defines the scattering angle of the matrix $D(\varphi, \mathbf{I}_s, \mathbf{I})$, while the vectors themselves define the scattering plane. In [6,7], a single photodetector rotating around the axis is replaced with five photodetectors installed at different angles φ .

The sum of the distances between the points 1, 2 (\mathcal{A}) and 1, 3 (\mathcal{A}) is called the nephelometer base \mathcal{A} .

Examples of some nephelometers used in experiments in the isotropic atmosphere are given in **Table 1**.

As can be seen from Table 1, the nephelometers differ widely in the values of the parameters \mathcal{A} , d_s , and d_r . The ranges of applicability of these nephelometers with respect to atmospheric parameters are discussed insufficiently. The issue of optimization of nephelometer parameters from the viewpoint of improvement of its accuracy characteristics has received insufficient attention as well.

The descriptions of nephelometers often include the concepts of the elementary volume $V_{se1}(\varphi)$ and the elementary optical beam, which are some idealizations of the actual volume $V_s(\varphi)$ and the actual beam. However, the accuracy of approximation of the actual parameters by the elementary volume $V_{se1}(\varphi)$ and the elementary beam was not discussed yet.

The scattering volume $V_s(\varphi)$ (Fig. 1) is bounded by the lateral surfaces of the optical beam emitted by the radiation source and the photodetector field of view. It is characterized by the cross-section diameters of the optical beam a_s and the photodetector field of view a_r .

For the small volume dV at point 1, Eq. (2) has the form

$$d\mathbf{S} = \frac{1}{A_r^2} D(\varphi, \mathbf{I}_s, \mathbf{I}) \mathbf{S}_s dV. \tag{3}$$

Introduce the right coordinate systems XYZ and $X'Y'Z'$ with the axes X, Z and X', Z' lying in the scattering plane determined by the vectors \mathbf{I}_s and \mathbf{I} , which are at the same time the direction vectors of the axes Z' and Z , respectively. The vectors \mathbf{S}_s and $d\mathbf{S}$ will be determined relative to the axes X' and X , respectively.

The main task of a polarization nephelometer is to determine matrix $D(\varphi, \mathbf{I}_s, \mathbf{I})$ elements (in absolute units) satisfying Eq. (3). The volume should be large enough for we can believe that it includes the complete set of all particles characteristic of this medium. In the case of well mixed “clear” air containing aerosol particles with the size of few micrometers near the Earth’s surface, the volume can be taken equal to few cubic centimeters. The same is also true for moderately dense, but stable stratiform clouds. However, in the case of dense cumulus clouds, the volume of 1 cm³ is sufficient to determine reliably the size distribution function and concentration of particle [8]. The small volume is a volume of the scattering medium containing the general set of particles.

Let the volume $V_s(\varphi)$ contain N small volumes dV and let every i -th volume dV have some point i , the right coordinate systems $X(i)Y(i)Z(i)$ and $X'(i)Y'(i)Z'(i)$ are anchored to. The axes $Y(i)$ and $Y'(i)$ are perpendicular to the scattering plane at the point i of the i -th volume dV , which is determined by the vectors $\mathbf{I}_s(i)$ and $\mathbf{I}(i)$ coinciding with the axes of optical beams of the incident and scattered radiation of the i -th volume dV , respectively. The directions of the axes $Z(i)$ and $Z'(i)$ coincide with those of the vectors $\mathbf{I}(i)$ and $\mathbf{I}_s(i)$. The vectors $\mathbf{S}_s(i)$ and $d\mathbf{S}(i)$ of, respectively, the incident and scattered radiation of the i -th volume dV are determined relative to the axes $X'(i)$ and $X(i)$.

The distance A_r should always exceed (no less than fivefold) the diameter of any cross

Table 1

Parameters of nephelometers

Parameter	Reference		
	[4]	[6]	[7]
Base \mathcal{A} , m	8	0.58	0.58
d_s , m	0.2	0.03	0.03
d_r , m	0.1	0.03	0.03
Divergence of optical beam emitted by source	40'	40'	40'
Divergence of photodetector field of view	2÷4°	3°	3°

section of the volume $V_s(\varphi)$ by the plane perpendicular to the axis of the photodetector field of view [12]. In addition, the photodetector PU should not fall within the volume $V_s(\varphi)$. The requirement for the distance between points 1 and 6 of the volume $V_s(\varphi)$ to be no larger than A_r at the angle

$$\varphi = \begin{cases} \varphi_{\min}, & \text{at } \varphi_{\min} < 10^\circ, \\ 10^\circ, & \text{at } \varphi_{\min} > 10^\circ, \end{cases}$$

where $[\varphi_{\min}, \varphi_{\max}]$ is the angular range of measurement of the matrix D , satisfies this condition. Consequently, the following condition should be true:

$$A_r > \frac{dr}{2} \operatorname{ctg} \frac{\varphi}{2}, \quad (3a)$$

where

$$\varphi = \begin{cases} \varphi_{\min}, & \text{at } \varphi_{\min} < 10^\circ, \\ 10^\circ, & \text{at } \varphi_{\min} > 10^\circ, \end{cases}$$

It is assumed that for $V_s(\varphi) \leq V_{se1}(\varphi)$ at the distance A_r fulfilling Eq. (3a) with allowance made for Eq. (3), we can write with a small error

$$S = \frac{1}{A_r^2} \int_{V_s(\varphi)} D(\varphi, \mathbf{I}_s, \mathbf{I}) \mathbf{S}_s(i) dv \cong \frac{1}{A_r^2} D(\varphi, \mathbf{I}_s, \mathbf{I}) \mathbf{S}_s V_s(\varphi), \quad (4)$$

where $D(\varphi, \mathbf{I}_s, \mathbf{I})$ is the scattering matrix of the volume dv , $\mathbf{S}_s(i)$ is the Stokes vector of the optical beam incident onto this volume, and \mathbf{S}_s is the Stokes vector of the optical beam irradiating the volume $V_s(\varphi)$.

Equation (4) is valid when the intensity of the radiation incident on the volume $V_s(\varphi)$ remains practically constant within this volume. In the volume $V_{se1}(\varphi)$, the amount of the absorbed and scattered radiation is proportional to the volume [9]. The propagation of radiation in a scattering medium is accompanied by multiple scattering. In the radiation incident on the volume $V_s(\varphi)$, the fraction of the multiply scattered radiation should be negligibly small in comparison with the direct radiation. An isotropic medium can be characterized by the extinction coefficient $k_\lambda =$

$\alpha + \sigma$, where α is the absorption coefficient, σ is the scattering coefficient. In this case, the optical thickness $\tau(A_s) = k_\lambda A_s$ and $\tau(a) = k_\lambda a_r$ should not exceed 5 to 6 [1].

Consider the approximation of the volume $V_s(\varphi)$ of an isotropic medium by the volume $V_{se1}(\varphi)$.

The optical flux Φ_s of unit power upon the propagation of the path l decreases by Φ_{ext} equal to

$$\begin{aligned} \Phi_{ext} &= 1 - e^{-\tau(l)} = \\ &= \tau(l) - \frac{\tau(l)^2}{2!} + \frac{\tau(l)^3}{3!} - \dots - \frac{(-\tau(l))^n}{n!} = \\ &= \tau(l) - \tilde{\tau}(l), \end{aligned} \quad (5)$$

where

$$\begin{aligned} \tau(l) &= k_\lambda l, \\ \tilde{\tau}(l) &= \frac{\tau(l)^2}{2!} - \frac{\tau(l)^3}{3!} + \dots + \frac{(-\tau(l))^n}{n!}. \end{aligned}$$

It follows from Eq. (5) that if the optical thickness $\tau(a_r)$ (Fig. 1) of the volume $V_s(\varphi)$ fulfills the condition

$$\tau(a_r) \gg \tilde{\tau}(a_r), \quad (6)$$

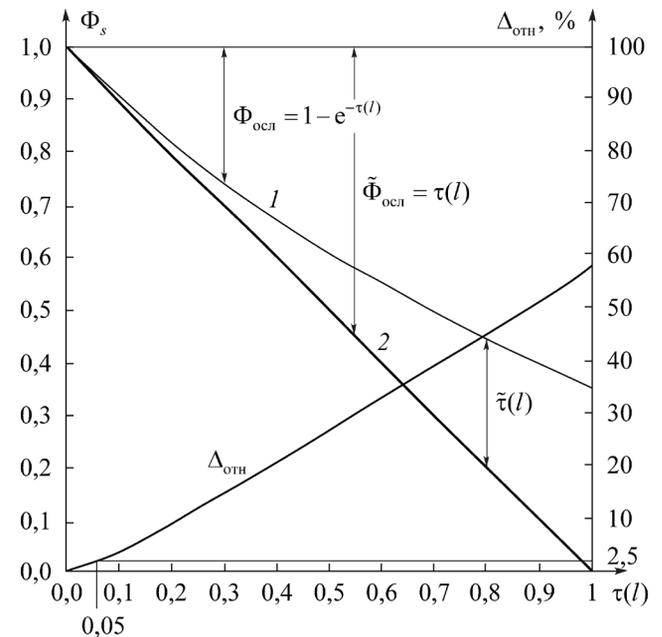


Fig. 2. Extinction of optical beam of unit power by an actual volume $V_s(\varphi)$ (1) and an elementary volume $V_{se1}(\varphi)$ (2), relative error $\Delta_{rel}(\%)$ of approximation of the actual volume by the elementary volume.

then it can be considered as the volume $V_{se1}(\varphi)$ with some error. **Fig. 2** demonstrates the extinction of the optical beam Φ_s of unit power by the volume $V_s(\varphi)$ (line 1) and by its approximation by the volume $V_{se1}(\varphi)$ (line 2). The relative error of this approximation Δ_{rel} can be determined as

$$\Delta_{rel} = \frac{\tilde{\Phi}_{ext} - \Phi_{ext}}{\Phi_{ext}}, \quad (7)$$

where $\tilde{\Phi}_{ext} = \tau(l)$ is the extinction of unit-power flux by the volume $V_{se1}(\varphi)$.

The thickness a_r of the volume $V_s(\varphi)$, whose optical thickness is $\tau(a_r) = 0.05$, can be taken as that of the volume $V_{se1}(\varphi)$ with the error $\Delta_{rel} = 2.5\%$.

This criterion allows us to determine $a_r = 0.05/k_\lambda$. The volume, whose thickness a_r for the optical beam is

$$a_r < 0.05/k_\lambda, \quad (8)$$

can be taken as the volume $V_{se1}(\varphi)$.

Practically any volume $V_s(\varphi)$ of the sufficiently homogeneous medium contains the large number of scattering and absorbing centers. Therefore, the condition of statistical representativeness of an elementary volume, whose optical thickness is $\tau(a_r) = 0.05$, is always fulfilled with high accuracy. That is why the thickness a_r of the volume dV can be much smaller than $0.05/k_\lambda$.

Using the Koschmieder formula, we can rewrite Eq. (8) as

$$a_r = 0.013L_{MVR}, \quad (9)$$

where L_{MVR} is the meteorological visibility range.

ASSUMPTION 1. Matrix \tilde{D}_{mn} , $m, n = \overline{1,4}$, measured by the polarization nephelometer is the matrix (D_{mn}) , $m, n = \overline{1,4}$, of the medium, if in the volume $V_s(\varphi)$ the scattering planes of all small volumes dV determined by the vectors $\mathbf{I}_s(z)$ and $\mathbf{I}(z)$ are parallel and the angles $\varphi(z)$ between them are identical.

Assumption 1 is satisfied, if Items 1 or 2 are satisfied.

1. The scattering volume $V_s(\varphi)$ is contracted to the volume dV at point 1.

2. The optical beam emitted by the radiation source at the length a_r can be approximated by the elementary beam with a small error, and the photodetector 2 field-of-view angle α_r in the scattering plane is small.

The optical beam with the small value of relative change in the brightness at the distance a_r can be substituted by the elementary optical beam with a small error. If the optical beam from the radiation source propagates in the solid angle $\Delta\Omega_s$, then the brightness of the optical beam L_s irradiating the volume $V_s(\varphi)$ at point 4 is

$$L_s = \frac{E^\perp}{\Delta\Omega_s} = \frac{4I}{\pi a_s^2}, \quad (10)$$

where E^\perp is the illuminance in the cross section of the beam, I is the luminous intensity.

The relative change of brightness δL_s at the distance a_r is

$$\delta L_s = \frac{L_s - L_s(a_r)}{L_s(a_r)} = \frac{L_s}{L_s(a_r)} - 1, \quad (11)$$

where $L_s(a_r)$ is brightness of the optical beam at the distance a_r from point 4.

With Eq. (10), we can write a decrease in brightness $\delta = L_s/L_s(a_r)$ in the form

$$\begin{aligned} \delta &= \frac{L_s}{L_s(a_r)} = \frac{2(a_s/2 + a_r \operatorname{tg} \alpha_s)^2}{a_s^2} = \\ &= 1 + \frac{2a_r \operatorname{tg} \alpha_s}{a_s} \left(1 + \frac{2a_r \operatorname{tg} \alpha_s}{a_s} \right), \end{aligned} \quad (12)$$

where $2L_s$ is the divergence of the optical beam.

With allowance made for Eq. (12), the value of δL_s is

$$\delta L_s = \frac{2a_r \operatorname{tg} \alpha_s}{a_s} \left(1 + \frac{2a_r \operatorname{tg} \alpha_s}{a_s} \right). \quad (13)$$

The value of δL_s characterizes the relative change in the brightness L_s of the optical beam at the distance a_r and is determined by the parameter $\tilde{\delta L}_s = \frac{a_r}{a_s} \operatorname{tg} \alpha_s$. The parallel optical beam is an ideal representative of an elementary beam, and $\tilde{\delta L}_s = 0$ in it. At the small distance a_r , δL_s of the beam can be neglected, and the beam itself at this distance can be considered

as an elementary optical beam. Thus, we can believe that the error of approximation of the light beam of a polarization nephelometer with the length a_r by the elementary optical beam is determined by the value of the parameter

$$\widetilde{\delta L}_s = \frac{a_r}{a_s} \operatorname{tg} \alpha_s \quad (14)$$

3. OPTIMAL METER OF THE MATRIX $D(\varphi, \mathbf{I}_s, \mathbf{I})$ AND ITS CALIBRATION

The method employed in the optimal meter of the matrix $D(\varphi, \mathbf{I}_s, \mathbf{I})$ allows all elements of the matrix to be determined from separate measurements of the signal i_r at the exit of the photodetector without additional transformations of the signal i_r [2, 3]. The algorithm for control of polarization elements in PUs of the radiation source and the photodetector is optimal, because it provides the smallest errors in measuring the elements of the matrix $D(\varphi, \mathbf{I}_s, \mathbf{I})$ due to inaccuracies in manufacture and installation of polarization elements. At the same time it is easy to implement, because only the position of the fast axis (FA) of the phase elements in PUs of the source and the photodetector is controlled, the discrete number is minimal, and the positions of the fast axes of the phase elements in PUs of the source and the photodetector are independent of elements of the matrix $D(\varphi, \mathbf{I}_s, \mathbf{I})$.

Let the volume $V_s(\varphi)$ satisfy Assumption 1 with a small error, and we can take $(\widetilde{D}_{mn}) \cong (D_{mn})$, $m, n = \overline{1, 4}$.

If the distance $A_r \gg a_s$, but it influence on the value of the vector \mathbf{S} is negligibly small, then Eq. (4) can be written in the form

$$\mathbf{S} = D(\varphi, \mathbf{I}_s, \mathbf{I}) \mathbf{S}_s V_s(\varphi), \quad (15)$$

where $\mathbf{S}_s = (I_s Q_s U_s V_s)^T$, $\mathbf{S} = (IQUV)^T$.

The measurements in the polarization nephelometer should satisfy Eq. (15). To measure 16 elements of the matrix $D(\varphi, \mathbf{I}_s, \mathbf{I})$, 16 independent equations are sufficient. Every vector \mathbf{S}_s corresponds to four equations in Eq. (15), each containing one parameter of the vector \mathbf{S} . Consequently, to determine 16 elements of the matrix, it is sufficient for the source to

generate four types of polarization described by the vectors $\mathbf{S}_{sj} = (I_{sj} Q_{sj} U_{sj} V_{sj})^T, j = \overline{1, 4}$ [2,3].

These 16 equations can be written in the form

$$\mathbf{S}_j = (I_j Q_j U_j V_j)^T = D(\varphi, \mathbf{I}_s, \mathbf{I}) \mathbf{S}_{sj} V_s(\varphi), j = \overline{1, 4} \quad (16)$$

The vector \mathbf{S}_{sj} determined by the vector \mathbf{S}_q and the Mueller matrices of the polarizer and phase element of PU of the radiation source can be written in the form

$$\mathbf{S}_{sj} = E \begin{pmatrix} 1 \\ \cos 2\alpha'_j \cos 2(\alpha'_j - \beta') + \cos \tau' \sin 2\alpha'_j \sin 2(\alpha'_j - \beta') \\ \sin 2\alpha'_j \cos 2(\alpha'_j - \beta') - \cos \tau' \cos 2\alpha'_j \sin 2(\alpha'_j - \beta') \\ \sin \tau' \sin 2(\alpha'_j - \beta') \end{pmatrix} \quad (17)$$

where $E = (1/2)(I_0 + Q_0 \cos 2\beta' + U_0 \sin 2\beta')$, and I_0, Q_0, U_0 are the parameters of the vector \mathbf{S}_0 of generated radiation (in the presence of a depolarizer, $Q_0 = U_0 = 0$), β' is the angle of orientation of the polarizer transmission plane (TP), α'_j is the angle of orientation of FA of the phase element, which is measured relative to the axis X' (α'_j corresponds to the j -th position of FA, $j = 1, 4$); τ' is the phase shift of orthogonal components of the phase element.

Separate the normalized vector \mathbf{S}_{ij} from the vector \mathbf{S}_{sj} given by Eq. (17)

$$\mathfrak{S}_{sj} = \begin{pmatrix} 1 \\ Q_{sj} \\ U_{sj} \\ V_{sj} \end{pmatrix} = \begin{pmatrix} 1 \\ \cos 2\alpha'_j \cos 2(\alpha'_j - \beta') + \cos \tau' \sin 2\alpha'_j \sin 2(\alpha'_j - \beta') \\ \sin 2\alpha'_j \cos 2(\alpha'_j - \beta') - \cos \tau' \cos 2\alpha'_j \sin 2(\alpha'_j - \beta') \\ \sin \tau' \sin 2(\alpha'_j - \beta') \end{pmatrix} \quad (18)$$

and rewrite Eq. (16) in the form

$$\mathbf{S}_j = D'(\varphi, \mathbf{I}_s, \mathbf{I}) \mathbf{S}_{sj}, j = \overline{1, 4} \quad (19)$$

where

$$D'(\varphi, \mathbf{I}_s, \mathbf{I}) = E V_s(\varphi) D(\varphi, \mathbf{I}_s, \mathbf{I}).$$

We use the parameters of the vectors $\mathbf{S}_{sj}, j = \overline{1, 4}$ to form the matrix W nondegenerate at any α'_j

$$W = \begin{pmatrix} 1 & Q_{s1} & U_{s1} & V_{s1} \\ 1 & Q_{s2} & U_{s2} & V_{s2} \\ 1 & Q_{s3} & U_{s3} & V_{s3} \\ 1 & Q_{s4} & U_{s4} & V_{s4} \end{pmatrix} \quad (20)$$

and then group system (19) into four systems, each determining a line of the matrix $D'(\varphi, \mathbf{I}_s, \mathbf{I})$:

$$W\mathbf{D}'_1 = \mathbf{I}_w, \quad W\mathbf{D}'_2 = \mathbf{Q}_w, \quad W\mathbf{D}'_3 = \mathbf{U}_w, \quad W\mathbf{D}'_4 = \mathbf{V}_w, \quad (21)$$

where $\mathbf{D}'_m = EV_s(\varphi)(D_{m1}D_{m2}D_{m3}D_{m4})^T$, $m = \overline{1,4}$; $(D_{m1}D_{m2}D_{m3}D_{m4})$ – is the m -th line of the matrix $D(\varphi, \mathbf{I}_s, \mathbf{I})$; $\mathbf{I}_w = (I_1I_2I_3I_4)^T$, $\mathbf{Q}_w = (Q_1Q_2Q_3Q_4)^T$, $\mathbf{U}_w = (U_1U_2U_3U_4)^T$, $\mathbf{V}_w = (V_1V_2V_3V_4)^T$ – are the vectors composed of the parameters of the vectors $\mathbf{S}_j = (IQU_jV_j)^T, j = \overline{1,4}$.

For systems (21) $W, \mathbf{I}_w, \mathbf{Q}_w, \mathbf{U}_w$, and \mathbf{V}_w are input data in the meter of the vectors $\mathbf{D}'_1, \mathbf{D}'_2, \mathbf{D}'_3$ and \mathbf{D}'_4 . The inaccuracy in determination of the input data in Eq. (21) has the smallest effect on the accuracy of determination of the vectors $\mathbf{D}'_i, i = \overline{1,4}$, when the conditioning number $CondW$ of the matrix W is minimal [10]. The conditioning number of the nondegenerate matrix can be determined as

$$CondW = \|W\| \|W^{-1}\|,$$

where $\|...\|$ denotes the norm of the matrix; W^{-1} is the matrix inverse to the matrix W .

The choice of some or other particular norm is dictated by the requirements imposed on the accuracy of solution. The choice of the Euclidian norm corresponds to the criterion of smallness of the root-mean-square error.

In the optimal meter, the minimal conditioning number $CondW$ of the matrix W obeying the Euclidian norm of the matrix is equal to 4.4722 at the position of FA of the phase element determined by the angles $\alpha'_1 = 38.54^\circ$; $\alpha'_2 = 75.14^\circ$; $\alpha'_3 = 105.38^\circ$; $\alpha'_4 = 141.857^\circ$, the phase shift $\tau' = 131.795^\circ$, and the angle of orientation of the polarizer TP $\beta' = 90^\circ$. The parameters $\tau', \alpha'_j, j = \overline{1,4}$, and β' should be measured with high accuracy, and the Mueller matrices of polarization elements should differ only slightly from the Mueller matrices of ideal polarization elements.

The photodetector can measure only the parameter I_r of the vector $\mathbf{S}_r = (IQ_rU_rV_r)^T$ through measurement of the signal $i_r = \psi I_r$, where ψ is the photodetector sensitivity. The vector \mathbf{S}_r is determined by the vector \mathbf{S} and the Mueller matrices of the phase element and the polarizer of the photodetector PU.

To measure the parameters $IQUV$ of the vector \mathbf{S} , four independent equations obtained through the measurement of the parameter I_r at each of the four positions of FA of the phase element of photodetector PU are sufficient. These equations have the form

$$I_{ri} = \frac{1}{2}(m_{i1}I + m_{i2}Q + m_{i3}Y + m_{i4}V), \quad i = \overline{1,4}, \quad (22)$$

where

$$m_{i1} = 1; \quad m_{i2} = \cos 2\alpha_i \cos 2(\alpha_i - \beta) + \cos \tau \sin 2\alpha_i \sin 2(\alpha_i - \beta);$$

$$m_{i3} = \sin 2\alpha_i \cos 2(\alpha_i - \beta) - \cos \tau \cos 2\alpha_i \sin 2(\alpha_i - \beta);$$

$$m_{i4} = -\sin \tau \sin 2(\alpha_i - \beta);$$

i is the serial number of the position of FA of the phase element; α_i is the angle of orientation of FA of the phase element corresponding to the i -th position, β is the angle of orientation of TP of the polarizer of photodetector PU relative to the axis X ; τ is the phase shift of orthogonal components of the phase element.

The matrix $(m_{in}), i, n = \overline{1,4}$, of the system of equation (22) is denoted as M . This matrix differs from the matrix W given by Eq. (20) by the sign of the elements in the fourth column. Therefore, the minimal conditioning numbers of these matrices are $CondM = CondW = 4.4722$.

In the meter of the vector \mathbf{S} , the input data are $I_r, i = \overline{1,4}$, and the matrix M . In the optimal meter of the vector \mathbf{S} , the minimal value $CondM \leq 1.53$ is achieved at the position of FA of the phase element determined by the angles $\alpha'_1 = 38.54^\circ$; $\alpha'_2 = 75.14^\circ$; $\alpha'_3 = 105.38^\circ$; $\alpha'_4 = 141.857^\circ$, the phase shift $\tau' = 131.795^\circ$, and angle of orientation of the polarizer TP $\beta' = 90^\circ$. The parameters $\tau', \alpha'_j, j = \overline{1,4}$, and β' should be measured with high accuracy, and the Mueller matrices of the polarization elements should be

close to the Mueller matrices of ideal polarization elements.

Let I_{rj} be the value of the parameter I_r of the vector \mathbf{S}_r at the j -th position of FA of the phase element of PU of the radiation source and i -th position of FA of the phase element of the photodetector PU. Then, it follows from Eq. (22) that

$$\mathbf{S}_j = 2M^{-1}\mathbf{I}_{rj}, \quad i = \overline{1,4}, \quad (23)$$

where $\mathbf{I}_{rj} = (I_{r1j} I_{r2j} I_{r3j} I_{r4j})^T$, $M^{-1} = (\tilde{\mathbf{m}}_n)$, $i, n = \overline{1,4}$ is the matrix inverse to the matrix M .

Introduce the matrix E , whose elements are determined by the photodetector signal $i_{rj} = \psi I_{rj}$

$$E = \begin{pmatrix} i_{r11} & i_{r21} & i_{r31} & i_{r41} \\ i_{r12} & i_{r22} & i_{r32} & i_{r42} \\ i_{r13} & i_{r23} & i_{r33} & i_{r43} \\ i_{r14} & i_{r24} & i_{r34} & i_{r44} \end{pmatrix}. \quad (24)$$

Then, with allowance made for Eqs. (21), (23), and (24), we have

$$\begin{aligned} \mathbf{D}'_1 &= W^{-1}\mathbf{I}_W = \frac{2}{\psi} W^{-1} E \tilde{\mathbf{m}}_{1n}, \quad n = \overline{1,4}, \\ \mathbf{D}'_2 &= W^{-1}\mathbf{Q}_W = \frac{2}{\psi} W^{-1} E \tilde{\mathbf{m}}_{2n}, \quad n = \overline{1,4}, \\ \mathbf{D}'_3 &= W^{-1}\mathbf{U}_W = \frac{2}{\psi} W^{-1} E \tilde{\mathbf{m}}_{3n}, \quad n = \overline{1,4}, \\ \mathbf{D}'_4 &= W^{-1}\mathbf{V}_W = \frac{2}{\psi} W^{-1} E \tilde{\mathbf{m}}_{4n}, \quad n = \overline{1,4}, \end{aligned} \quad (25)$$

where $\tilde{\mathbf{m}}_n = (\tilde{m}_{n1} \tilde{m}_{n2} \tilde{m}_{n3} \tilde{m}_{n4})^T$; $(\tilde{m}_{i1} \tilde{m}_{i2} \tilde{m}_{i3} \tilde{m}_{i4})$ is the i -th line of the matrix M^{-1} .

In the general case, Eq. (25) can be written as

$$\mathbf{D}_i = \frac{1}{K(\varphi)} 2W^{-1} E \tilde{\mathbf{m}}_{in}, \quad n = \overline{1,4}, \quad (26)$$

where $\mathbf{D}_i^T = (D_{i1} D_{i2} D_{i3} D_{i4})$; $(D_{i1} D_{i2} D_{i3} D_{i4})$

is the i -th line of the matrix $D(\varphi, \mathbf{I}_s, \mathbf{I})$;

$\tilde{\mathbf{m}}_{in}^T = (\tilde{m}_{i1} \tilde{m}_{i2} \tilde{m}_{i3} \tilde{m}_{i4})$; $(\tilde{m}_{i1} \tilde{m}_{i2} \tilde{m}_{i3} \tilde{m}_{i4})$ is the i -th line of the matrix M^{-1} ; $K(\varphi) = EV_s(\varphi)\psi$ is the calibration coefficient.

The control over the discrete positions of the phase elements can be the following:

1. FA of the phase element of PU of the radiation source is successively set to the position $j = \overline{1,4}$.
2. At each position j of FA of the phase element of PU of the radiation source, the signal i_{rj} , $i = \overline{1,4}$, is measured through discrete setting FA of the phase element of the photodetector to the position $i = \overline{1,4}$.

Thus, at each position of FA of the phase element of PU of the radiation source, the j -th line of the matrix E is determined. Execution of Items 1 and 2 and Eq. (25) determines the matrix $\psi D'(\varphi, \mathbf{I}_s, \mathbf{I}) = K(\varphi)D(\varphi, \mathbf{I}_s, \mathbf{I})$.

Using Eq. (25), we can write the element $EV_s(\varphi)\psi D_{11}$ necessary for the calibration in the form $EV_s(\varphi)\psi D_{11} = \tilde{D}_{11}$, where

$$\begin{aligned} \tilde{D}_{11} &= 2[\tilde{w}_{11}(i_{r11}\tilde{m}_{11} + i_{r21}\tilde{m}_{12} + i_{r31}\tilde{m}_{13} + i_{r41}\tilde{m}_{14}) + \\ &+ \tilde{w}_{12}(i_{r12}\tilde{m}_{11} + i_{r22}\tilde{m}_{12} + i_{r32}\tilde{m}_{13} + i_{r42}\tilde{m}_{14}) + \\ &+ \tilde{w}_{13}(i_{r13}\tilde{m}_{11} + i_{r23}\tilde{m}_{12} + i_{r33}\tilde{m}_{13} + i_{r43}\tilde{m}_{14}) + \\ &+ \tilde{w}_{14}(i_{r14}\tilde{m}_{11} + i_{r24}\tilde{m}_{12} + i_{r34}\tilde{m}_{13} + i_{r44}\tilde{m}_{14})] \end{aligned} \quad (27)$$

is the relative value of the element D_{11} of matrix $D(\varphi, \mathbf{I}_s, \mathbf{I})$.

The matrix D can be calculated from one system of 16 equations, whose right-hand sides are the parameters $I_{rj} = i_{rj}/\psi$ [3]. Let us form the vector

$$\mathbf{I}_{rj} = (I_{r11} \dots I_{r14} I_{r21} \dots I_{r24} I_{r31} \dots I_{r34} I_{r41} \dots I_{r44})^T.$$

These systems of equations in the matrix form can be represented as:

$B\mathbf{D}_{mn} = \mathbf{I}_{rj}$,
where $D_{mn} = (D_{11} \dots D_{14} D_{21} \dots D_{24} D_{31} \dots D_{34} D_{41} \dots D_{44})^T$ and D_{ij} are elements of the matrix $D(\varphi, \mathbf{I}_s, \mathbf{I})$; $B = M \otimes W$ is the Kronecker (direct) multiplication of the matrices M and W .

Using the equation [10]:

$$\|M\|_E = \sqrt{S_p M^T M},$$

where $\|M\|_E$ is the Euclidian norm of the matrix M and $S_p M^T M$ is the spur (trace) of the matrix $M^T M$, we can obtain

$$\text{Cond} B = \|B\|_E \|B^{-1}\|_E = \text{Cond} M \text{Cond} W. \quad (28)$$

Consequently, $CondB$ achieves the minimum when $CondM$ and $CondW$ determined from the positions of FA of the phase elements of PUs of the source and the photodetector achieve the minima. The values of $CondM$ and $CondW$, which provide for the minimum, are independent of the elements of the matrix $D(\varphi, \mathbf{I}_s, \mathbf{I})$ and absolutely optimal.

The aim of the calibration is to determine the calibration coefficient $K(\varphi) = EV_s(\varphi)\Psi$.

The method of calibration with the use of a diffusively reflecting (diffusively transmitting) screen in the non-polarized radiation from a source is described in [11]. This method is most correct, because it takes into account the geometric configuration of the volume $V_s(\varphi)$ and irregular illuminance of its points. In the optimal meter of the matrix $D(\varphi, \mathbf{I}_s, \mathbf{I})$, the source radiation is fully polarized. Therefore, the method [11] should be modified to be used in the optimal meter of the matrix $D(\varphi, \mathbf{I}_s, \mathbf{I})$.

The ideal diffusively reflecting (diffusively transmitting) screen has the following properties:

- (a) all the incident radiation is reflected (or transmitted) by the screen surface;
- (b) the screen brightness is identical in all directions;
- (c) the reflected (transmitted) radiation is non-polarized in all directions.

If the surface of an actual diffusively reflecting (diffusively transmitting) screen has the illuminance E , then its brightness L regardless of polarization of the incident radiation is

$$L = \beta \frac{E}{\pi},$$

where β is the brightness coefficient of the diffusively reflecting (diffusively transmitting) screen. The brightness coefficient β is understood as a ratio of the brightness of this screen in some direction to the brightness of an ideal screen being under the same illuminance conditions. If a surface coated by an MgO layer is illuminated normally and observed at the angle $\xi \approx 30^\circ$ to the normal, then $\beta = 1$ and the brightness is $L = E/\pi$ (Fig. 3) [12].

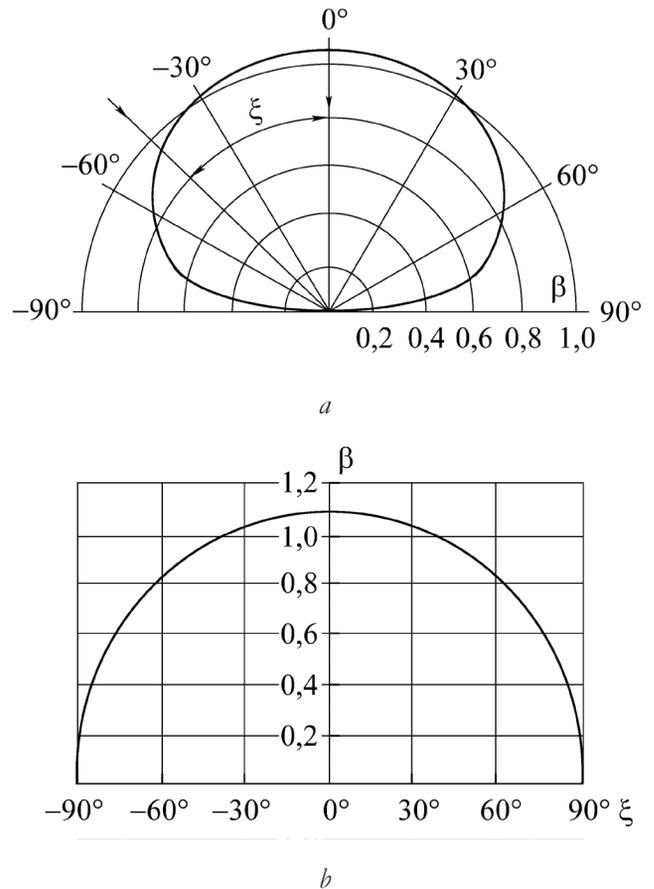


Fig. 3. Brightness coefficients in the scattering MgO surface in polar (a) and orthogonal (b) coordinates.

Assume that FA of the phase element lies in TP of the polarizer in PUs of the radiation source and the photodetector, which means $\alpha'_j = \beta'$ in Eq. (17) and $\alpha_i = \beta$ in Eq. (22), and remains at this position as the polarizers rotate in the process of calibration. The Mueller matrix of PU with FA of the phase element lying in TP of the polarizer is equal to the Mueller matrix of the polarizer [14].

Let the angle φ be obtuse and the screen move from point 5 to point 6 (Fig. 1) along the axis Z . Set the angle $\beta' = \alpha'_j = 90^\circ$ in PU of the radiation source. Then the optical beam of the radiation source has the vector \mathbf{S}_s given by Eq. (17) equal to

$$\mathbf{S}_s = (E, -E, 0, 0)^T \tag{29}$$

for the radiation linearly polarized along the axis Y' .

The light flux $d\Phi_r^c$ incident on the photosensitive element of the photodetector

from the element dA of the screen surface \mathcal{A} irradiated by the optical beam with the vector \mathbf{S}_s given by Eq. (29) is equal to

$$d\Phi_r^c = C_1 \beta(\gamma'') E \cos \gamma' \cos \gamma'' dA / 2\pi, \quad (30)$$

where C_1 is the coefficient equal to zero for the screen points not belonging to the volume $V_s(\varphi)$, γ' is the angle of incidence of the optical beam on the screen, γ'' is the angle between the normal to the surface \mathcal{A} and the axis of the photodetector field of view, $\beta(\gamma'')$ is the coefficient β (Fig. 3) at the angle $\xi = \gamma''$, the coefficient $1/2$ is equal to the attenuation of the nonpolarized radiation by the polarizer of the photodetector PU. At the angle $\gamma'' = 0$, the coefficient β is minimally dependent on the angle γ'' , and the angle $\gamma' = \pi - \varphi$, and therefore the angle $\gamma'' = 0$ is optimal.

The light flux $d\Phi_r^c$ gives rise to the signal di_r^c at the photodetector output

$$di_r^c = C_2 C_1 \beta(\gamma'') \cos \gamma' \cos \gamma'' E dA / 2\pi,$$

where C_2 is the coefficient of photodetector sensitivity to the light flux from the source at the point of the element dA .

The integral signal i_r^c from the entire luminous surface \mathcal{A}

$$i_r^c = \frac{\beta(\gamma'') \cos \gamma' \cos \gamma''}{2\pi} \iint_{\mathcal{A}} C_1 C_2 E dA.$$

As the screen moves from point 5 to point 6 along the axis Z , the integral signal W_r^c is calculated as

$$W_r^c = \int_5^6 i_r^c dz = \frac{\beta(\gamma'') \cos \gamma' \cos \gamma''}{2\pi} \iiint_{V_s(\varphi)} C_1 C_2 E dv, \quad (31)$$

where $dv = dA dz$.

The analogous conclusions can be also drawn for the diffusively transmitting screen (thin fluoroplastic sheet) at the angle $0^\circ \leq \varphi \leq 10^\circ$ [13].

If we remove the screen and irradiate the volume $V_s(\varphi)$ by the beam with the vector \mathbf{S}_s given by Eq. (29), then the vector $d\mathbf{S}(i)$ of radiation scattered by the volume $dv(i)$ is

$$d\mathbf{S}(i) = \frac{E}{r(i)^2} (D_{11} - D_{12}, D_{21} - D_{22}, D_{31} - D_{32}, D_{41} - D_{42})^T dv(i),$$

and the vector $d\mathbf{S}_r(i)$ of radiation after photodetector PU at $\alpha_1 = \beta = 0$ (polarizer TP is directed along the axis X) is

$$d\mathbf{S}_r(i) = \frac{E}{2r(i)^2} (D_{11} - D_{12} + D_{21} - D_{22}, D_{11} - D_{12} - D_{21} - D_{22}, 0, 0)^T dv(i).$$

The Stokes vector $d\mathbf{S}_r(i)$ corresponds to the luminous intensity

$$dH_r(i) = \frac{E}{2} (D_{11} - D_{12} + D_{21} - D_{22}) dv(i),$$

and the light flux $d\Phi_r(i)$ incident on the sensitive element of the photodetector is

$$d\Phi_r(i) = C_1 d\vec{I}_r(i) = C_1 \frac{E}{2} (D_{11} - D_{12} + D_{21} - D_{22}) dv(i). \quad (32)$$

The light flux $d\Phi_r(i)$ described by Eq. (32) gives rise to the signal $i_r \uparrow \rightarrow (i)$ at the photodetector output

$$i_r \uparrow \rightarrow (i) = \frac{C_1 C_2}{2} E (D_{11} - D_{12} + D_{21} - D_{22}) dv(i). \quad (33)$$

The Stokes vector $d\mathbf{S}_r(i)$ of the radiation scattered by the volume element $dv(i)$ exposed to the incident radiation with the vector \mathbf{S}_s described by Eq. (29) after PU of the photodetector with the polarizer TP along the axis Y is

$$d\mathbf{S}_r(i) = \frac{E}{2r(i)^2} (D_{11} - D_{12} - D_{21} + D_{22}, -D_{11} + D_{12} + D_{21} - D_{22}, 0, 0)^T dv(i).$$

This vector corresponds to the signal $i_r \uparrow \uparrow (i)$ at the photodetector output equal to

$$i_r \uparrow \uparrow (i) = \frac{C_1 C_2 E}{2} (D_{11} - D_{12} - D_{21} + D_{22}) dv(i). \quad (34)$$

If we irradiate the volume $V_s(\varphi)$ by the optical beam with the Stokes vector $\mathbf{S}_s = (E, E, 0, 0)^T$ (the radiation is linearly polarized long the axis X') at the angle $\beta' = \alpha_1' = 0^\circ$, then the Stokes vector $d\mathbf{S}(i)$ of the radiation scattered by the element $dv(i)$ is

$$d\mathbf{S}(i) = \frac{E}{r(i)^2} (D_{11} + D_{12}, D_{21} + D_{22}, D_{31} + D_{32}, D_{41} + D_{42})^T dv(i),$$

and the Stokes vector $d\mathbf{S}_r(i)$ of radiation after the photodetector PU with the polarizer TP along the axis X is

$$d\mathbf{S}_r(i) = \frac{E}{2r(i)^2} (D_{11} + D_{12} + D_{21} + D_{22}, D_{11} + D_{12} + D_{21} + D_{22}, 0, 0)^T dv(i).$$

The radiation with this Stokes vector corresponds to the signal $i_{r \rightarrow}(i)$ at the photodetector output equal to

$$i_{r \rightarrow}(i) = \frac{C_1 C_2 E}{2} (D_{11} + D_{12} + D_{21} + D_{22}) dv(i). \quad (35)$$

The Stokes vector $d\mathbf{S}_r(i)$ of radiation scattered by the element $dv(i)$ after the photodetector PU with the polarizer TP along the axis Y is

$$d\mathbf{S}_r(i) = \frac{E}{2r(i)^2} (D_{11} + D_{12} - D_{21} - D_{22}, -D_{11} - D_{12} + D_{21} + D_{22}, 0, 0)^T dv(i).$$

The radiation with this Stokes vector corresponds to the signal $i_{r \rightarrow \uparrow}(i)$ at the photodetector output equal to

$$i_{r \rightarrow \uparrow}(i) = \frac{C_1 C_2 E}{2} (D_{11} + D_{12} - D_{21} - D_{22})^T dv(i). \quad (36)$$

Summing up the signals $i_{r \rightarrow \uparrow}(i)$ (33), $i_{r \rightarrow \uparrow}(i)$ (34), $i_{r \rightarrow}(i)$ (35) and $i_{r \rightarrow \uparrow}(i)$ (36) from all elements $dv(i)$ of the volume $V_s(\varphi)$, we obtain the signal $i_{r\Sigma}$ equal to

$$i_{r\Sigma} = \iiint_{V_s(\varphi)} (i_{r \rightarrow \uparrow} + i_{r \rightarrow \uparrow} + i_{r \rightarrow} + i_{r \rightarrow \uparrow}) dv = \frac{D_{11}}{2} \iiint_{V_s(\varphi)} C_1 C_2 E dv. \quad (37)$$

Taking into account Eqs. (31) and (37), we obtain

$$D_{11} = \frac{i_{rD} \beta(\gamma'') \cos \gamma' \cos \gamma''}{\pi W_r^c}. \quad (38)$$

Substitute Eq. (38) into Eq. (27) and obtain the calibration coefficient

$$K(\varphi) = \frac{\pi W_r^c \tilde{D}_{11}}{i_{r\Sigma} \beta(\gamma'') \cos \gamma' \cos \gamma''}. \quad (39)$$

4. SOURCE OF RADIATION

The ideal source of radiation for polarization measurements should be monochromatic, have

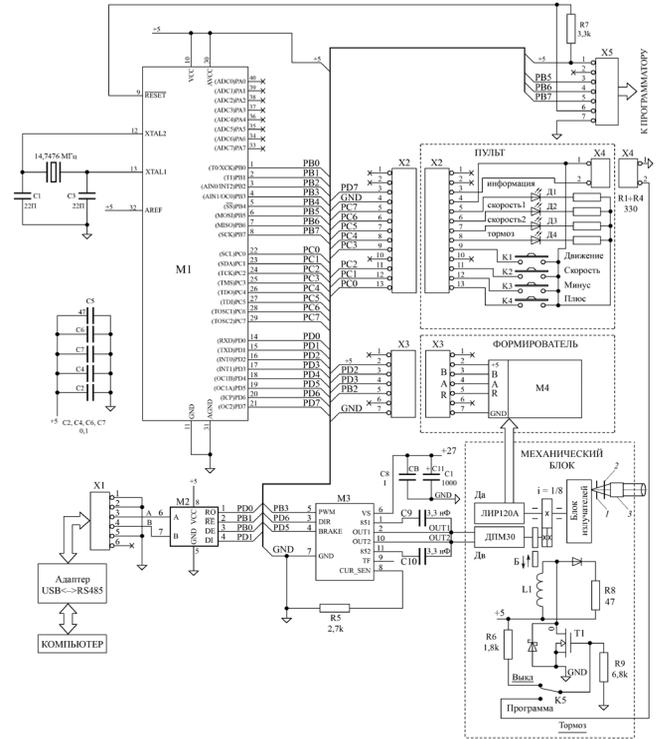


Fig. 4. Block diagram of the source of radiation: Ø 0.5 diaphragm 1, depolarizer 2, GELIOS-44-2 objective 3, T1 – RLU120, M1– ATmega16, M2 – MAX 485, M3 – LMD18200.

high spectral intensity, cover a wide spectral range, and have high directivity of radiation. **Fig. 4** shows the block diagram of a five-wave source of quasi-monochromatic radiation of high spectral intensity and directivity. The requirements for an ideal light source for polarization measurements can be met by a set of high-brightness LEDs with small size of luminous elements. High-brightness LEDs, whose main parameters are given in **Table 2**, are used as generators of radiation.

In the radiation source, LED with the diameter of the luminous element dle of 1 mm is set at a distance of 1 mm in front of the diaphragm 0.5

Table 2

Main LED parameters.

LED type	Color	λ_{rad} , nm	P _{rad} , W	U _{power} , V	I _{max} , mA	I _{av} , mA	$\delta\lambda$, nm/°C	d _{lum} , mm
IPR169A9S	blue	455	0.6	3.5	700	350	0.17	1
IPR169A9L	green	525	0.3	3.5	700	350	0.17	1
IPR169A9Zh	yellow	555	0.15	2	700	350	0.17	1
IPR169A9Zh	orange	590	0.15	2	700	350	0.17	1
IPR169A9K	red	625	0.28	2	700	350	0.17	1

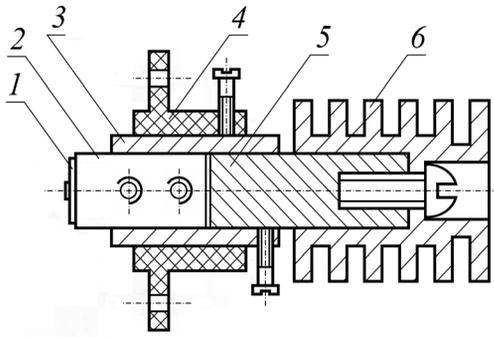


Fig. 5. Block diagram of the emitter: LED 1, heatsink 2, cylinder 3, insulator 4, rod 5, and radiator 6.

mm in diameter. The GELIOS-44-2 objective allows us to obtain the optical beam 25 mm in diameter with the divergence of 0.3°.

The minimal volume dv of clear air is equal to few cubic centimeters. We can take dr equal to 25 mm. Then the minimal volume is $V_s(90^\circ)$ in our case. It is possible to obtain more than 10 cm^3 . The nephelometer base $A = 1.2$ m and $A_r = 0.6$ m can provide the measurement of the matrix D at $\varphi_{min} = 1.8^\circ$. The optical beam parameter characterizing the error of approximation of the optical beam by the elementary optical beam varies from 0.0025 at $\varphi = 90^\circ$ to 0.013 at $\varphi = 180^\circ$.

The emitter arrangement is shown in Fig. 5. LED 1 is soldered to heat sink 2 (copper plate 0.5 mm thick), which is clamped with screws in the groove of copper rod 5. Rod 5 moves longitudinally in copper cylinder 3. Cylinder 3 moves longitudinally in caprolon insulator 4. Radiator 6 increases the heat removal from LED. The power source of the LED D2 is an emitter voltage regulator on the transistor T1 (Fig. 6). As low voltage is applied to the input

of the microcircuit M1, the voltage determined by the divider at resistors R1 and R3 is formed at the emitter of the transistor T1. Resistor R2 provides the negative feedback for stabilization of LED current. Resistor R2 is a constantan wire wound on an MLT (metal-film varnished heat-resistant)-type resistor.

The unit of emitters is a disk 150 mm in diameter turning around the axis. It comprises of five emitters placed in a circle (Fig. 5) and power sources (Fig. 6) blown from below by a micro-fan (Fig. 7). The required LED is placed against diaphragm 1 (Fig. 4) by the motor DV through a back-lash-free gearbox. Once the LED is set, the rotation of the disk with emitters and power supplies can be slowed down by the stop of the slider B on the rubberized disk. The angle of disk rotation around the axis is measured by the LIR120A transducer, which reflects it by the number of pulses with respect to the pulse R. Pulses B, A, and R from LIR120A come to the M4 shaper on the 555IP11 microcircuit and then through the X3 connector to the M1 AEmega16 microprocessor. The motor is controlled by the M3 LMD18200 driver. The program to the M1 microprocessor is saved through the X5 connector. The software of the M1 microcontroller accomplishes the function of a servo drive, in which the M3 LIR120A angle converter acts as a sensor, and the LDM18200 driver acts as a power bridge. The servo feedback operates by the principle of a PID

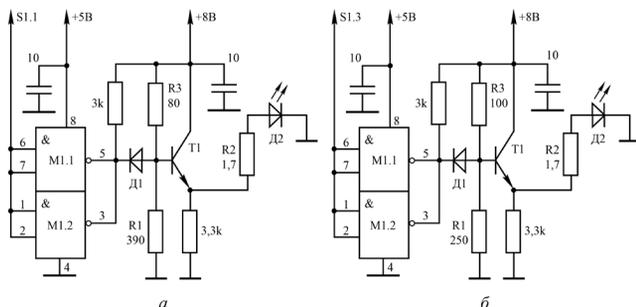


Fig. 6. Power supply: a – 3.5 V; b – 2 V; M1 – 155L A18; T1 – SC3807; D1 – 1N4148.

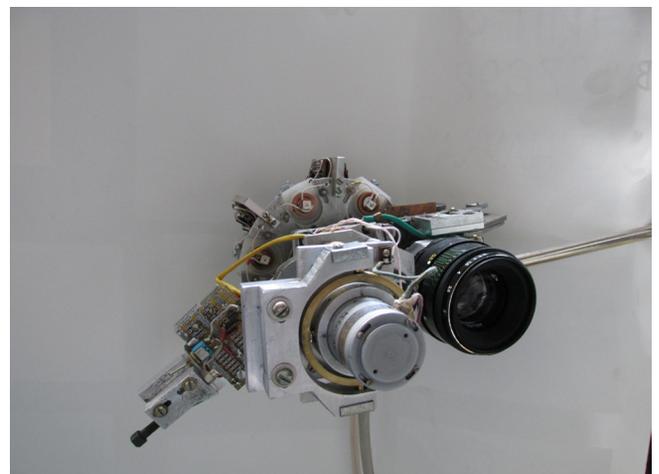


Fig. 7. Source of radiation.

(Proportional - Differential - Integral) controller, which calculates the speed and direction of movement depending on the difference between the current coordinate and the required one. In its turn, the required coordinate or, more exactly, the variable associated with it changes by the mathematical law during execution of the movement command. This ensures a smooth start of the electric drive and its smooth deceleration to zero speed at the finish point. The value of acceleration and deceleration of the electric drive is set by a separate variable in the microcontroller program. Smooth acceleration and deceleration, as well as coordinate feedback with the PID controller, are provided by standard numerical control (CNC) mathematics.

The radiation source can be controlled both from a personal computer and from the control panel. Control commands from a personal computer are transferred through the X1 connector and the M2 MAX485 microcircuit.

The control panel allows the position of the emitter to be changed smoothly with three gradations of speed. High-speed gradation allows the drive to turn quickly to the needed angle. The low-speed gradation allows accurate adjustment of the emitter position. There is also a command for the initial setting of the emitter coordinate system with respect to the reference pulse of the output R of the LIR120A angle converter. The control panel also allows us to save the coordinate of a current position of the drive to the internal memory of the microcontroller to be able to return to this point at any time. The number of recorded positions can be up to five. With commands from the control panel, it is possible to move the drive to any of these recorded positions.

The PC control program has the same capabilities as the built-in control panel. In addition, with commands from the computer, it is possible to set the position of the emitter directly to the needed angle with the accuracy provided by the LIR120A angle converter (0.009°).



Fig. 8. Polarization unit.

A polarization unit (PU) can be connected in place of the mechanical unit (Fig. 8). The position of FA of the phase element of PU is controlled in the way similar to that described above.

The computer program allows up to five devices similar to the mechanical unit and PU to be controlled. It has three dialog tabs "Operation," "Settings," and "Port" (in Fig. 9). Two last tabs include various settings, in particular, the communication port, positioning resolution (still identical for all axes of motion), and limiting speed. Figure 9 shows the dialog box of motion commands. In this window, one can see five horizontal arrays of keys and text fields arranged in one line. Each group of elements is associated with the control of one of five devices.

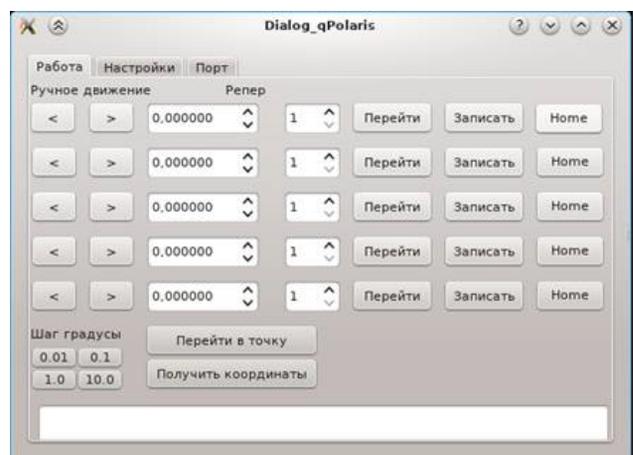


Fig. 9. Example of operation of the software for control of five emitters.

The computer program is written in C++ with the QtCreator version 4 open source software. This tool includes the free GCC-4.7 compiler and the popular Qt free interface library version 4. This solution opens the feasibility of not only writing programs with a modern user interface, but also compiling them for any platform, including Windows and Linux, for which there is a compiled Qt library on the Internet. Thus, the program can be compiled without changes even for most so-called embedded industrial computers, if they have the Qt-embedded 4.0 or 5.0 library.

5. CONCLUSIONS

The material presented in the paper allows one to determine the main parameters of a quasi-monochromatic radiation source of high spectral brightness being a component of a polarization nephelometer that measures matrix D in the range $\varphi = [\varphi_{\min}, \varphi_{\max}]$, as well as its design and operating principle. The external view of the polarization unit demonstrates the structure of the unit for high-precision control of the phase element.

REFERENCES

1. Rozenberg GV (ed.). *Searchlight Beam in the Atmosphere*. Moscow, AN SSSR Publ., 1960, pp. 146-174.
2. Oshlakov VG. Optimal measurer of the scattering phase matrix. *Optika atmosfery i okeana* [Atmospheric and oceanic optics], 1992, 5(11):767-769 (in Russ.).
3. Oshlakov VG, Borkov YuG. Numerical analysis of the instrumental matrix of a polarization meter. *Optika atmosfery i okeana* [Atmospheric and oceanic optics], 2002, 15(07):575-581 (in Russ.).
4. Gorchakov GI, Rozenberg GV. Measurement of the radiation scattering matrix in the atmospheric surface layer. *Izvestiya AN SSSR. Fizika atmosfery i okeana*, 1965, 1(12):1279-1287 (in Russ.).
5. Gorchakov GI. Matrices of radiation scattering by near-surface air. *Izvestiya AN SSSR. Fizika atmosfery i okeana*, 1966, 11(6):593-605 (in Russ.).
6. Sidorov VN. Flow-through polarization nephelometer. *Izvestiya AN SSSR. Fizika atmosfery i okeana*, 1979, 15(7):763-768 (in Russ.).
7. Lukshin VV, Isakov AA. Optical characteristics of smoke aerosols. *Izvestiya AN SSSR. Fizika atmosfery i okeana*, 1988, 24(3):250-257 (in Russ.).
8. Deirmendjian D. *Electromagnetic Scattering on Spherical Polydispersions*. RAND Corporation, Santa Monica, California, American Elsevier Publishing Company, New York, 1969.
9. Ivanov AP. *Optics of Scattering Media*. Moscow, Nauka i tekhnika Publ., 1969, 592 p.
10. Voevodin VV, Kuznetsov YuA. *Matrices and Calculations*. Moscow, Nauka Publ., 1984, 318 p.
11. Pritchard BS, Elliott WG. Two instruments for atmospheric optics measurements. *J. Opt. Soc. Am.*, 1960, 50(3):111.
12. Gurevich MM. *Photometry: Theory, Methods, and Devices*. Leningrad, Energoatomizdat Publ., 1983, 269 p.
13. Isakov AA. Study of atmospheric aerosol by the methods of radiation scattering at small angles. *Ph.D. Thesis*, Moscow, IFA AN Publ., 1979.
14. Gerrard A, Burch JM. *Introduction to Matrix Methods in Optics*. New York, Willey, 1975, 341 p.
15. Born M, Wolf E. *Principles of Optics*. London, Pergamon Press, 1959, 719 p.

DOI: 10.17725/rensit.2020.12.451

Using of Stopping and Range of Ions in Matter Code to Study of Radiation Damage in Materials

Hiwa Mohammad Qadr

University of Raparin, College of Science, Physics Department, <https://www.uor.edu.krd/>
Main Street, Sulaimanyah, Ranya, 46012 Iraq

E-mail: hiwa.physics@uor.edu.krd

Ari Maghdid Hamad

Koya University, Faculty of Science and Health, Physics Department, <http://fsch.koyauniversity.org/>
Koya KOY45, Kurdistan Region, F.R. Iraq

E-mail: ari.physics@koyauniversity.org

Received April 28, 2020; peer reviewed May 18, 2020; accepted May 20, 2020

Abstract: The aim of this work to investigate the impact of the radiation damage in the materials by the proton energy irradiation. The damage parameter used in the evaluation is displacement per atom (DPA) in material as a function of proton energy. Stopping and Range of Ions in Matter (SRIM) code was used to calculate the total vacancy and the number of atomic displacements based on the Norgett-Robinson-Torrens model by difference energies for proton irradiation damage. The option of this code was calculated by using Ion Distribution and Quick Calculation of Damage (Kinchin-Pease) for Fe and Cu target and also Full damage cascade (F-C) was chosen for only Fe. The result is that, the prediction of the F-C model are higher than the K-P calculation. Comparisons has been made with an international standard definition of DPA.

Keywords: Proton Energy Irradiation, Radiation Damage, Displacement per Atom (DPA), Stopping and Range of Ions in Matter (SRIM) Code, Primary Knock-on Atom (PKA)

PACS: 34.50.Bw

For citation: Hiwa Mohammad Qadr, Ari Maghdid Hamad. Using of Stopping and Range of Ions in Matter Code to Study of Radiation Damage in Materials. *RENSIT*, 2020, 12(4):451-456. DOI: 10.17725/rensit.2020.12.451.

Contents

1. Introduction (451)
2. Displacement Damage (452)
 - 2.1. Displacements Produced by a PKA (452)
 - 2.2. Displacement Per Atom (452)
3. SRIM/TRIM Analysis (453)
 - 3.1. Vacancies and damage calculation (453)
 - 3.2. Comparison between K-P and F-C (453)
 - 3.3. DPA Calculation (453)
 - 3.4. Damage Calculations and Comparison in SRIM (455)
4. Conclusion (455)
- References (455)

1. INTRODUCTION

A Monte Carlo simulation code is developed for the study of neutron and proton-induced radiation damage in the materials which results from nuclear collision as well as reactions that create energetic

recoil atoms of the host material or reaction creates. The development of codes for the radiation damage method by neutrons and protons can be highly useful in technology of advanced nuclear systems, for energy and nuclear fusion reactors [1].

There are several codes used that are related to radiation damage [2-6], because radiation damage can be measured as a function of displacement per atom (DPA), which is one of the critical issues for high intensity beams, particularly, for protons. So, for the measurement of the displacement per atom values of the material, there is The Stopping and Range of Ions in Matter Code, in a short, SRIM or TRIM for ion transport in matter which is a one of the types of code for various physical quantities related to ion implantation, energy deposition and other effects of interaction of ions with matter [7-9]. It is also used to calculate radiation damage calculation method exposure unit known as displacement per atom.

Thus, DPA has been used as a standard measure for computing proton induced radiation damage production from different radiation sources [10]. There are two basic types of damage calculation: Quick Calculation of damage (Kinchin-Pease model) (K-P) and Full damage cascade (F-C) [11,12]. Accordingly, this paper is mainly to investigate the effect of radiation damage in materials and demonstrate the displacement per atom in the results of the advanced Monte Carlo transport simulation code.

2. DISPLACEMENT DAMAGE

2.1. DISPLACEMENTS PRODUCED BY A PKA

Displacement cascade, in a material can be visualised as a series of elastic collisions initiated with the lattice atoms, where the lattice atom hit by a high energy particle. The initially-bombard a target is named Primary Knock-on Atom (PKA) [13]. Displacement cascade can be created by PKAs, in turn when an energetic incident particle, transferred to the PKA, is high enough, $E \gg E_p$, the PKA will be able to continue the PKA process to displace other atoms of the crystal, creating secondary recoil atom displacement. Thus, it can depend on the amount of the energy and mass of the recoil [14]. **Fig. 1** shows collision cascade or displacement cascade by a PKA.

If a lattice atom is hit by a PKA, it must receive energy in the collision in order to be displaced atom from its lattice site. This minimum energy is called the displacement threshold energy. It obviously depends on the initial direction. The lattice atom in collision receives energy that is less than the displacement threshold energy, the atom can be knocked out of its position in the crystal but will not be displaced [13].

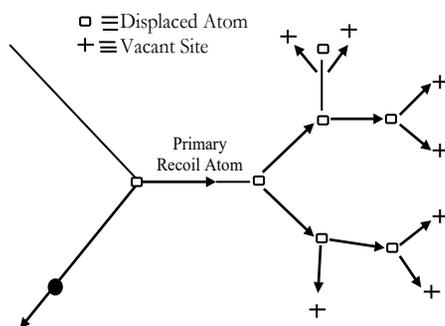


Fig. 1. Schematic representation of the formation of displacement cascade by a PKA.

The average number of displacements for a PKA energy E can be estimated by the following derivation given by Kinchin-Pease damage function [15]:

$$N_d(E) = \begin{cases} 0 & \text{if } E > E_C \\ 1 & \text{if } E_d < E < 2E_d \\ \frac{E}{2E_d} & \text{if } 2E_d < E < E_C \\ \frac{E}{2E_d} & \text{if } E > E_C \end{cases}, \quad (1)$$

where E_d is the threshold displacement energy and represented the energy required to generate a stable Frankel pair and E_C is the cut-off energy of discrimination of elastic and inelastic.

2.2. DISPLACEMENT PER ATOM (DPA)

The most common measure of the amount of radiation damage for displacement damage in a different type of particles is displacement per atom [16]. E521 ASTM standard particle for neutron radiation damage simulation by charged-particle irradiation recommends the use of the Norgett-Robinson-Torrens secondary displacement model that allows for calculating irradiation damage and it also allows DPA correlations from neutron damage [17].

There are many different ways that have been developed to measure the displacement per atom. One of them is the Kinchin and Pease (K-P) model; they were the first people to come up with a great process. In this model, the actual number of point defects produced by an ion implantation can be derived from the energy, which is transferred from an ion to an atom of the target materials [18].

The majority of the authors in this field have tried to establish a new process for measuring DPA. The Norgett-Robinson-Torrens model was the most successful one. It developed a method in 1975 for calculating DPA [14]. Based on this model, the number of displaced atoms is given by the following equation:

$$N_d(E) = \frac{\beta T_{dam}}{2E_d}, \quad (2)$$

where $\beta = 0.8$ factor is the displacement efficiency which was determined from binary collision formulas to account for realistic scattering anisotropy in the screened interaction potential. T_{dam} is the damage

energy and represents the portion of the PKA energy which is lost by elastic collisions with the target atoms.

Stoller et al. [19] have discussed another way for calculating DPA which is based on the number of vacancies produced. Simply summing incoming ion and the target atom from the *vacancy.txt* file provides the average number of vacancies produced.

Unfortunately, the NRT-DPA is not always a good correlation parameter available for all the irradiation correlations since only the actual number of atomic displacement did not control all the radiation effects. Particularly, the NRT-DPA is used merely as an irradiation dose more than an appropriate correlation parameter. The concept of the NRT-DPA is not capable of including spatial fluctuation distribution [20].

3. SRIM/TRIM ANALYSIS

3.1. VACANCIES AND DAMAGE CALCULATION

SRIM-2013 code has used to calculate the total vacancy by difference energies for proton irradiation damage. The option of this code was calculated by using Ion Distribution and Quick Calculation of Damage (Kinchin-Pease model). Calculations were for H ion implantation into Fe and Cu for a range of H ion energies using 0.5, 1.0 and 9.0 MeV of PKA. The threshold displacement energy used in the calculations was 40 eV for Fe and 30 eV for Cu based on ASTM E521 Standard and the lattice binding energy was set at 0.0 [17,21]. Then, the incident ion was fixed for 5000 ions. At the end, the data has found in the “*vacancy.txt*” output files, which were detailing the number of vacancies created for ions and recoils.

Fig. 2 shows SRIM calculation of vacancy creation cause H into Fe and Cu using 0.5, 1.0 and 9.0 MeV. In the case of 0.5 MeV H into Fe and Cu beams, the prediction of the Cu is greater than the Fe with the total number of vacancies per ion being around 12/Ion for Cu and 8/Ion for Fe. For 1.0 MeV, the total number of vacancies per ion is about 10/Ion for Fe and 16/Ion for Cu. And it is also for 9.0 MeV that the average number of vacancies per ion is 31 for Fe and 48 for Cu. From Fig. 2, the number of target vacancies increase, which is dependent on the ion energy. For example, for 1.0MeV H PKA in Fe is greater damage than the 0.5MeV H PKA in Fe

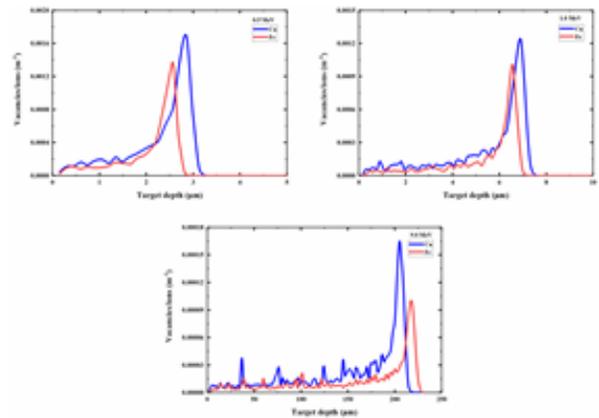


Fig. 2. Number of vacancies per incident ion versus depth for 0.5, 1.0 and 9.0 MeV H PKA into Fe and Cu.

because the ion with higher energy can transfer more energy to the recoiling atom.

3.2. COMPARISON BETWEEN K-P AND F-C

In this section, SRIM-2013 code has used to calculate the total vacancy. The option of this code was selected by using *Full damage cascade* (F-C) and compared with *Quick Calculation of Damage* (Kinchin-Pease model) (K-P). Calculations were carried out for H ion implantation into only Fe for a range of H ion energies using 0.5 and 1.0 MeV of PKA. Then, the incident ion was also fixed for 5000 ions. At the end, the data has found by summing the second column and the third column from the “*vacancy.txt*” output files. The results of the difference between the K-P and F-C models for Fig. 3 is that, the prediction of the full cascade (F-C) model for 0.5 and 1.0MeV are higher than the Kinchin-Pease (K-P) calculation which is independent of the electronic to nuclear stopping ratio.

3.3. DPA CALCULATION

Read and de Oliveira [19] has also argued about the damage of radiation energy into SRIM as a result of many different ways, and presented a different mechanism for measuring the amount

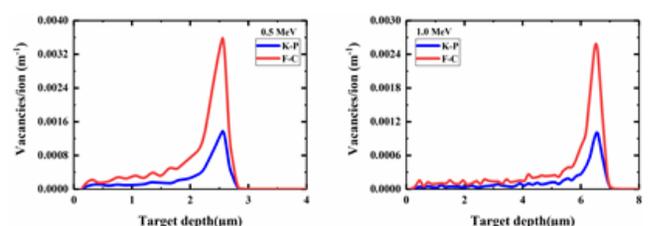


Fig. 3. SRIM calculation of vacancy creation cause H PKA into Fe and Cu using 0.5 and 1.0 MeV ion.

of displacement. According to Equation (3), they have illustrated that T_{dam} is able to be calculated by SRIM.

$$T_{dam} = E_{ions}^0 - E_{ions}^I - E_{target}^I \tag{3}$$

In this relation, E_{ions}^0 can be the incident ion, on the other hand beam energy, E_{ions}^I can be the beam energy loss to ionization and E_{target}^I is the atom energy target loss to ionization. These values ought to be found in the “*ioniz.txt*” output file.

Furthermore, for determine T_{dam} , the values must be found in the “*phonons.txt*” based on this equation:

$$T_{dam} = E_{ions}^P + E_{target}^P \tag{4}$$

where E_{ions}^P is the beam energy lost to phonons which is given by the second column (in the TRIM output folder) and E_{target}^P is the target atom energy lost to phonons which is given by the third column.

The option of this code was calculated using Ion Distribution and Quick Calculation of Damage. The damage calculated with this option is according to the Kinchin-Pease (K-P) formalism which was used in the NRT standard model. This option is currently used for comparison with proton damage. The ion used was Hydrogen, with ions of energies of 0.5 1.0, 5.0 and 9.0 MeV H PKA into Fe and copper Cu for 5000 incident ion. Then, data has found in the “*phonons.txt*” and obtaining T_{dam} by Eq. (4) and solving in Eq. (2) (NRT equation). The figures show the target depth of the radiation induced damage in the units of displacements per atoms (DPA). **Fig. 4** shows the displacements at Bragg peak can increase,

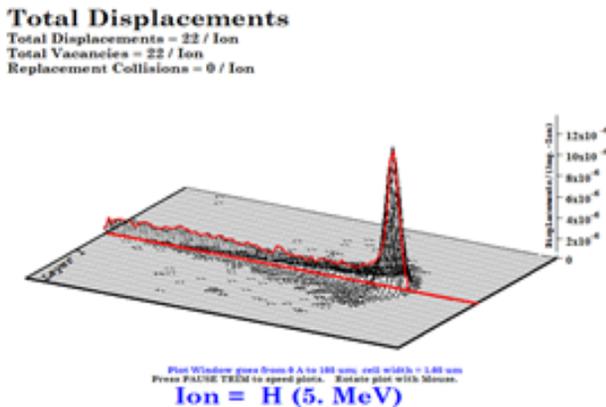


Fig. 4. Shows the Bragg peak of proton due to H PKA into Fe by using 5.0 Me.

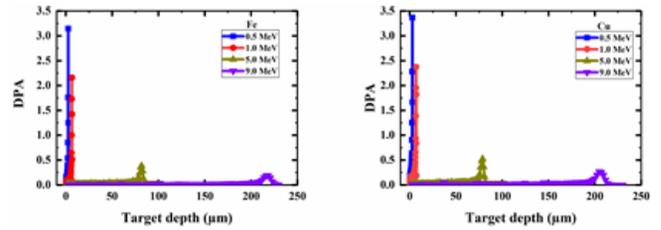


Fig. 5. Shows TRIM calculation of DPA production due to H PKA into Fe and Cu by using 0.5, 1.0, 5.0 and 9.0 MeV.

therefore protons lose their energy on arriving at Bragg peak.

It can be seen from **Fig. 5** that the target depth change with increase DPA a 0.5, 1.0, 5.0 and 9.0 MeV of PKA energy for Fe and Cu. The results were pretty satisfactory as the DPA behaviour is clear. For lower energy there are taller Bragg peaks which do not distribute their energy through the material easily and so have less DPA an average. For higher energy of 9.0 MeV in Fe have higher displacement damage, which is greater than those for 0.5, 1.0 and 5.0 MeV. The maximum damage level is about 2.85 μm depth, 7.2 μm depth, 87.2 μm depth and 230 μm depth for four energies.

It clearly shows in Fig. 5 hat compared with damage level at 0.5 MeV for copper is lower than those which is at a lower energy around 3.3 μm depth and at higher energy around 2190 μm depth. So, it is clear that the penetration depth varies strongly with its energy. The average displacement per atoms were 0.649 at an ion energy of 0.5 MeV, 0.351 at an ion energy of 1.0 MeV, 0.0639 at an ion energy of 5.0 MeV and 0.035 at an ion energy of 9.0 MeV.

As mentioned before, the factor T_d in Eq. 2 can be called the damage energy as well as can be E_{PKA} . **Fig. 6** shows the plots of damage energy and the

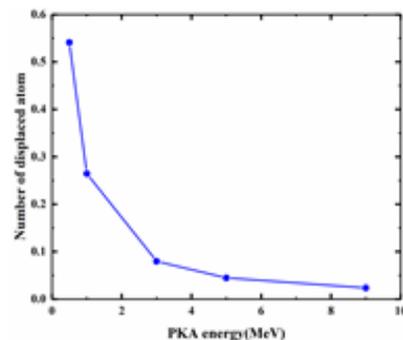


Fig. 6. Shows the plot of PKA energy with the number of displaced atoms for H into Fe.

Table 1

DPA of Fe target for two cases

Incident energy (MeV)	DPA - first case	DPA - second case
0.5	2.71E-04	1.46E-03
1	1.38E-04	7.33E-04
5.0	2.50E-05	1.30E-04

average number of displaced atoms (N_d) produced by a PKA of a given energy. It was calculated for iron using a standard method.

3.4. DAMAGE CALCULATIONS AND COMPARISON IN SRIM

In this section, comparisons has been made with an international standard that the results of DPA, it was calculated for two cases with SRIM model for 5000 incident ions. The first case, the threshold displacement energy used in the calculations was 40 eV for iron and 30 eV for copper based on ASTM E521 Standard and the lattice binding energy was set at zero. The second case, the threshold displacement energy was 25 eV for iron and copper as well as set the lattice binding energy was 3 eV. The data has also found in the “*vacancy.txt*” output files.

Table 1 was compared with DPA by difference in the threshold displacement energy and the lattice binding energy for 0.5, 1.0 and 5.0 MeV H PKA into Fe. The average number of DPA for second case result is greater than DPA for the first case result, which was produced by PKA directly. This is because the Coulomb scattering for the second case is much bigger than that of the first case.

Table 2 shows DPA of Cu target for two cases, the average number of DPA for the second case is also bigger than the number of DPA for the first case.

In addition, when comparing Table 1 and Table 2, the average number of DPA of copper are higher than those of Fe for the both cases. And as is also shown, the number of DPA an decrease with increasing incident ion energy for the both cases.

Table 2

DPA of Cu target for two cases

Incident energy (MeV)	DPA - first case	DPA - second case
0.5	3.75E-04	1.40E-03
1	2.09E-04	7.79E-04
5.0	3.93E-05	8.11E-05

4. CONCLUSION

SRIM-2013 code was used in this work to calculate the total vacancy and the number of displacement based on NRT by difference energies for proton irradiation damage. First, the option K-P model was chosen, and data taken from *vacancy.txt file*. It can be observed that the number of target vacancies increase, which is dependent on the ion energy. Thus, the prediction of the Cu for 0.5, 1.0 and 9.0 MeV were greater than the Fe with the total number of vacancies per ion. Second, using F-C option, and compared with the K-P option. It was found that the results of the F-C model for 0.5 and 1.0 MeV into Fe were greater than the K-P model. So, the F-C option could be more accurate.

Calculations were carried out for SRIM code based on the NRT for determining the number of displacement, which was preferred in the *phonons.txt file* for difference energies into Fe and Cu target when the K-P option was used. It can be noticed that for lower energy there were taller Bragg peaks which did not distribute their energy through the material easily and so had less DPA an average. So, it is clear that the target depth also varies strongly with increasing ion energy. In addition, comparison have been made for SRIM code based on two cases. It was found that the average, the number of DPA of Cu results were bigger than DPA. This is because the Coulomb scattering is much bigger.

REFERENCES

1. Ledingham K, McKenna P, Singhal R. Applications for nuclear phenomena generated by ultra-intense lasers. *Science*, 2003, 300(5622):1107-1111.
2. Trachenko K, Dove M, Salje E, Todorov I, Smith W, Pruneda M, Artacho E. Radiation damage in the bulk and at the surface. *Molecular Simulation*, 2005, 31(5):355-359.
3. Ferrari A, Sala PR, Fasso A, Ranft J. *FLUKA: a multi-particle transport code*. Stanford Linear Accelerator Center (SLAC), 2005.
4. Waters LS. *MCNPX user's manual*. Los Alamos National Laboratory, 2002, pp.
5. Niita K, Sato T, Iwase H, Nose H, Nakashima H, Sihver L. PHITS - a particle and heavy ion transport code system. *Radiation measurements*, 2006, 41(9-10):1080-1090.

6. Mokhov N, Striganov S. MARS15 overview. *AIP Conference Proceedings*, American Institute of Physics, 2007, pp. 50-60.
7. Strašik I, Pavlovič M. Improvements to the SRIM simulations. *Radiation Effects & Defects in Solids*, 2009, 164(7-8):470-476.
8. Qadr HM. Effect of ion irradiation on the hardness properties of Zirconium alloy. *Annals of the University of Craiova, Physics*, 2019, 29:68-76.
9. Hiwa M, Ari M. Investigation of long and short term irradiation hardening of P91 and P92 ferritic/martensitic steels. *Problems of Atomic Science and Technology, Series Thermonuclear Fusion*, 2019, 42(2):81-88.
10. Ziegler JF, Biersack JP, Ziegler MD. *SRIM: the stopping and range of ions in matter*. Cadence Design Systems 2008.
11. Chen S, Bernard D. On the calculation of atomic displacements using damage energy. *Results in Physics*, 2020, 16:102835.
12. Crocombette J-P, Wambeke C. Quick calculation of damage for ion irradiation: implementation in Iradina and comparisons to SRIM. *EPJ Nuclear Sci. Technol.*, 2019, 5:1-9.
13. Linn B. *Radiation Damage Effects in Materials: Damage Formation Models*, 2013.
14. Stoller RE. Primary radiation damage formation. In: *Comprehensive Nuclear Materials*, Konings RJM, Allen TR, Stoller RE, and Yamanaka S, Editors, Elsevier Ltd., Amsterdam, 2012, pp. 293-332
15. Fasso A, Vlachoudis V, Smirnov G, Ferrari A, Sommerer F. FLUKA realistic modeling of radiation induced damage. *Prog. Nucl. Sci. Tech.*, 2011, 2:769-775.
16. Iwamoto Y, Niita K, Sawai T, Ronningen R, Baumann T. Displacement damage calculations in PHITS for copper irradiated with charged particles and neutrons. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 2013, 303:120-124.
17. ASTM, E521-96 (2009). *Standard Practice for Neutron Radiation Damage Simulation by Charged-Particle Irradiation*, ASTM International, West Conshohocken, PA, 2009.
18. Nordlund K, Zinkle SJ, Sand AE, Granberg F, Averbach RS, Stoller RE, Suzudo T, Malerba L, Banhart F, Weber WJ. Primary radiation damage: a review of current understanding and models. *Journal of Nuclear Materials*, 2018, 512:450-479.
19. Stoller RE, Toloczko MB, Was GS, Certain AG, Dwaraknath S, Garner FA. On the use of SRIM for computing radiation damage exposure. *Nuclear instruments and methods in physics research section B: beam interactions with materials and atoms*, 2013, 310:75-80.
20. Norgett M, Robinson M, Torrens I. A proposed method of calculating displacement dose rates. *Nuclear engineering and design*, 1975 33(1):50-54.
21. Lucasson P. The production of frenkel defects in metals. *Fundamental aspects of radiation damage in metals*. US: Oak Ridge National Laboratory, 1975, pp. 42-65.

DOI: 10.17725/rensit.2020.12.457

Nanoparticles Surface as a Spheroidal Defect in the Parent Material

Sergey P. Gubin, Yulia V. Ioni, Elena Yu. Buslaeva

N.S. Kurnakov Institute of General and Inorganic Chemistry, Russian Academy of Sciences, <http://igic.ras.ru/>
Moscow 119991, Russian Federation

E-mail: gubin@igic.ras.ru, acidladj@mail.ru, eyubuslaeva@inbox.ru

Received November 09, 2020, peer-reviewed November 13, 2020, accepted November 16, 2020

Abstract: The review considers a general theoretical model of nanoparticles similar to the structure of a "green walnut", consisting of a core, a surface, and a layer of ligands. Due to the peculiarities of such surface defect structure in nanoparticles, the phenomenon of high-temperature ferromagnetism in the nonmagnetic parent structure is possible (it is considered by the example of semiconductor metal oxides, noble metals, and carbon nanostructures). It is shown that this phenomenon is not of an artifact nature, but it is possible for any materials in a highly dispersed state. The main reasons are proposed which can determine the cooperative magnetic effect in the system of defect-induced magnetic moments. The review also discusses the ferron model, which is successfully used to analyze the properties of modern magnetic nanomaterials.

Keywords: carbon materials, nanoparticles, defects, defect-induced magnetism, magnetism, ferromagnetism

UDC 546.05+537.6

Acknowledgements: The authors express their deep gratitude to Dr Sci Phys&Math, Prof. Yuri A. Koksharov for valuable advice and comments during the preparation of the publication. This work was supported by the Grant of the President of the Russian Federation MK-893.2020.8 and the Russian Science Foundation 20-19-00395.

For citation: Sergey P. Gubin, Yulia V. Ioni, Elena Yu. Buslaeva. Nanoparticles Surface as a Spheroidal Defect in the Parent Material. *RENSIT*, 2020, 12(4):457-470. DOI: 10.17725/rensit.2020.12.457.

CONTENTS

1. INTRODUCTION (457)
 2. NANOPARTICLES SURFACE AS A DEFECT (458)
 3. LIGAND SHELL (460)
 4. IS IT POSSIBLE TO DIRECTLY INFLUENCE THE SURFACE MAGNETIC EXCHANGE, TO ENHANCE IT? (463)
 5. THE ROLE OF BASS SIZE (464)
 6. THEORETICAL MODELS OF HTFM IN LF AND NM (464)
 7. CONCLUSION (465)
- REFERENCES (465)

1. INTRODUCTION

Nanostructuring made possible to reveal a number of fundamentally new properties of materials; this was most clearly manifested by the example of various types of nanomaterials (NM), consisting

of isolated (by various methods) nanoparticles (NPs) with an average diameter of 2-10 nm [1-6].

Numerous examples [7-12] have shown that a "simple" reduction of a material up to nano sizes (2-10 nm) makes them highly defective, which radically changes their physical and chemical properties and the range of applications.

Most real materials are defective to some extent. Such defects are well-studied as vacancies - anionic and cationic; impurity atoms of substitution or insertion; dangling bonds, altered crystal lattice parameters; element-wise or structural separation (internal nanoclusters), etc. [13] In a number of cases, a relationship has been established between structural defects and material properties. Obviously, materials consisting of nanoparticles (NPs) 2-10 nm in size are unconditionally defective, but it is difficult to use the above types

of defects for describing their properties. NPs and NMs based on them contain specific defects characteristic only of them, which determine their unique properties [14,15].

2. NANOPARTICLES SURFACE AS A DEFECT

Let us consider the NP surface as an extended defect in the structure of the “parent” nanomaterial, which determines its behavior under the action of external electromagnetic radiation or a magnetic field, which differs from the properties of the initial parent material. Thus, let us assume that the surface of nanoparticles is a special type of defects in nanomaterials.

In order to understand how the physical properties of a material (including magnetic) can (should) change during the transition from a compact (for example, basically diamagnetic) material to a NP system of the same composition, we will conduct a mental experiment.

Let us consider the NP surface as an extended defect in the structure of the “parent” nanomaterial, which determines its behavior under the action of external electromagnetic radiation or a magnetic field, which differs from the properties of the initial parent material. Thus, let us assume that the surface of nanoparticles is a special type of defects in nanomaterials.

In order to understand how the physical properties of a material (including magnetic) can (should) change during the transition from a compact (for example, basically diamagnetic) material to a NP system of the same composition, we will conduct a mental experiment.

Let's take a sample with an arbitrary crystal structure (**Fig. 1A**) (in this case, for simplicity,

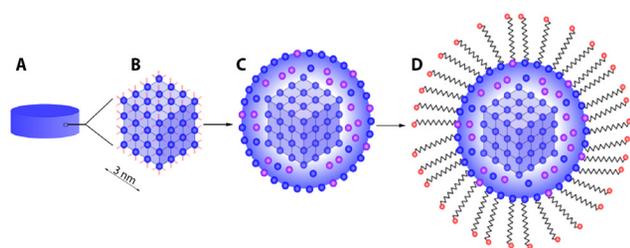


Fig. 1. Diagram of NP formation from compact material.

a simple cubic lattice was chosen). The sample is in equilibrium at room temperature, normal atmospheric pressure. If you remove a cube with an edge of 3 nm from the middle of this crystal, you will get an object (**Fig. 1B**) along the edges, vertices and sides of which there are exposed atoms (ions) that were previously included in the crystal lattice of the parent material. The broken ties sticking out are clearly visible. These can be vacant or filled orbitals (a pair of electrons can be in an orbital, and a vacancy in a matrix). When they interact with each other, an electron cloud appears on the surface - a plasmon, which is easily polarized under the influence of electromagnetic radiation [16].

But, for example, for the occurrence of magnetism when removing a cube from a "non-magnetic" matrix, the most important is the homolytic breaking of bonds inside the crystal, as a result of which orthogonal surfaces appear on the surface of the removed cube - orbitals with unpaired electrons (spins). As a result of exchange interactions between them, a magnetic domain appears, covering the NP surface, and it becomes magnetic [17,18]. At the band approach, it is assumed that an electron from the valence or conduction band gets to the surface level appearing due to the broken bond (ie, a defect). The electron captured by the defect, first, stabilizes the ferromagnetic state closed to itself, and, second, it self-traps in this cavity. Thus, a ferromagnetic region appears on the surface near the defect, called ferron [19–22].

Plasmon is located along, ferron - orthogonal to the NP surface. This orbital configuration usually does not hinder spin-spin exchange.

Extracting the cube from the matrix will remove the external pressure of the crystal structure of the compact sample, which leads to the transformation of the cube into a spheroid - NP (**Fig. 1C**), consisting of a core, in which the crystal structure of the parent material and its physical properties are mainly preserved, and consisting of a shell of 1-2 atomic layers, in which the crystal structure of the parent sample is significantly distorted-destroyed [23,24]. Due to

surface decompensation of metal polyhedra, the structure degraded, bond lengths, bond angles, and coordination polyhedra changed. A significant number of vacancies - structural defects and impurities - are concentrated in such a shell; there is often structural disorder and similar changes. Thus, the surface (1–2 atomic layers) of spheroidal NPs (2–10 nm) is a specific defect in the structure of the parent material, which is responsible for the appearance of properties uncharacteristic of a compact material [25,26].

As the size decreases, the curvature of the NP surface arises, which leads not only to an increase in the distance between surface atoms and a weakening of electron – electron interactions, but also enhances the exchange interactions between the spins of dangling bonds [27,28]. The system of such exchange interactions covering the entire NP surface determines the behavior of the NP in a magnetic field.

It is generally accepted that NPs consist of a core and two shells (Fig. 1D) (a “green walnut” structure) [28,29]. Each of the 3 parts of a NP (**Fig. 2**) behaves in its own way in an external magnetic field, so that the actually measured magnetic quantities are always a superposition of at least three components.

For small particles, the contribution of the surface layer to the determination of physical properties can be more than 50%, but with an increase in the NP size, the effect of the surface layer begins to rapidly decrease, and for particles > 100 nm, its contribution is usually neglected [30,31]. The set of such NPs is the most

rigorous and well-studied part of the entire huge array of objects which are considered today as nanomaterials.

NPs can be isolated from each other in gaseous, liquid, or solid media; in powders consisting ONLY of NPs and composite materials, where individual NPs are dispersed in matrices - inorganic or polymer [32,33]. Methods for obtaining NPs are well developed; they are obtained both by dispersing compact materials by various methods [34,35] and using various chemical reactions in solutions [36]. In the first case, at the destruction of the sample in the core (but not in the shell) of the formed NPs, the structure of the initial material is mainly retained, while in the second case, the composition and structure of NPs are reconstructed from individual atoms or ions [37]. Such materials based on NPs are universal among other disperse systems, in the sense that a highly defect state can be created in them not only by dispersing macrosamples, but also "construct" them from subatomic fragments, i.e. to one degree or another to manage the defectiveness of the created materials [38,39]. Methods for the synthesis of NPs of a certain size, composition, and structure, with a narrow size distribution, covered with a certain type of ligands, are well developed [40]. In other words, the type of defects and the degree of defectiveness (and hence the physical properties) of such materials can be finely controlled. However, when NPs are compacted, nanomaterials are obtained in which the idealized structure considered above can be distorted: the shells are deformed, penetrate into each other, are destroyed, the cores become larger, etc.

NPs have high surface energy. In a real medium - gas, liquid or solid - the NP surface (2-10 nm) is ALWAYS covered with a layer of light atoms (L in Fig. 2) - specially introduced ligands. The components of the real environment are also potential ligands for the nanoparticle surface. [41] Interaction with ligands makes it possible to partially or completely fill vacant coordination sites (structural defects) and makes the particle significantly more stable. However, such

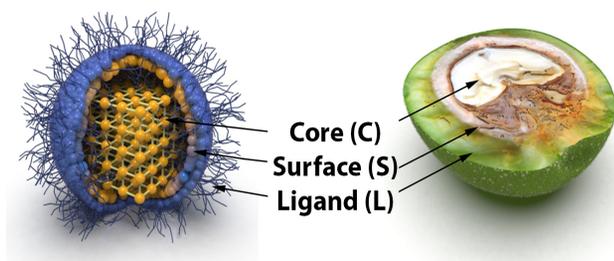


Fig. 2. NP model representing the structure of a "green walnut".

coordination can hinder the spin exchange in the surface layer up to complete destruction.

Thus, any NP with a size of 2-10 nm consists of at least three parts: 1 - a core with a slightly changed structure of the parent sample; 2 - the first shell (1-2 nm), which has the same composition as the particle core, but a significantly distorted structure; 3 - ligand shell, usually consisting of light atoms (O, N, S, etc.). Each of the three components behaves differently under the action of electromagnetic radiation or an external magnetic field and contributes to the observed physicochemical parameters of NPs.

NPs and NMs based on them are a typical example of defective materials, the physical properties of which, along with chemical reactivity, are determined not so much by the composition and crystal structure as by the type and number of defects in the surface layer of NPs.

3. LIGAND SHELL

The NPs of the sizes considered in this article are never "naked"; on their surface there are always either individual atoms from the environment, or specially introduced molecules (CO, P(Ph)₃, etc.). Polymers – organic or inorganic – are in use as specific stabilizers. [42] All these groups are called ligands, meaning that their main function is to interact with dangling surface bonds and fill vacant positions in the coordination sphere of surface atoms of NPs. The ligand shell is an integral part of NPs and has a significant effect both on the magnetic interactions in the particle itself and between them. [43-46] Interaction with the environment also often leads to a change in the chemical composition of the surface (in the most common variant – to oxidation), which inevitably affects the magnetic properties.

As a successful application of the above approach, let us consider NPs of "nonmagnetic" materials, the features of magnetism of which are described in detail in the review [47]. Magnetism still remains one of the most mysterious properties of matter. No less mysterious gravity differs in that its manifestations are universal – gravitational forces do not show any selectivity at the acting

on a material, regardless of its composition and electronic structure. [48,49] In contrast to this, the result of the action of a magnetic field on a material is highly dependent on its composition, structure, shape and other factors. Materials due to their behavior in an external magnetic field are roughly divided into two unequal groups: magnetic (their minority) and all the rest – "non-magnetic". [50,51] The latter, being as a rule diamagnets, are not completely indifferent to the action of an external magnetic field, although the general physical theory of magnetism does not contain reasons for differences.

Modern theoretical concepts of the nature of magnetism, as a rule, describe quite successfully the properties of most traditional magnetic materials, leaving aside "non-magnetic" materials [52,53].

However, as it often happened before, "the experiment turned out to be ahead of theory": it turned out quite unexpectedly that ferromagnetism is found in some samples of semiconductor nanomaterials which do not contain magnetic dopants [54,55]. The number of examples grew rapidly, it became clear that these are not isolated artifacts. In 2005, the term "d⁰ magnetism" was introduced to determine this class of phenomena [56]. Later, systematic studies, reviews [57-61] appeared. Later, reports appeared devoted to high-temperature ferromagnetism (HTFM) of carbon NPs (quantum dots), graphene, and other highly defect structures [59,62-64]. It became clear that the observed phenomenon is of a general nature and requires consideration of the entire set of facts from a unified standpoint before it will be possible to start developing the theory of HTFM "nonmagnetic nanomaterials".

Various types of structural inhomogeneity, for example, vacancies and impurity atoms, can contribute to the appearance of HTFM in nanoparticles of materials whose volume analogs do not exhibit such magnetic properties [11,12,58,60,65].

In compact semiconductor NMs, a continuous layer L is usually absent on each NP; its role is most often played by the surface of adjacent neighboring NPs. The interactions arising here

are random and difficult to formalize. To stabilize semiconductor NPs as ligands, mercaptans with long hydrocarbon radicals are most often used [28,66]. It has been shown that the covering of ZnO NPs with mercapto ligands can significantly change their magnetic properties: they can be diamagnetic, have Pauli paramagnetism, or become ferromagnetic [27]. It is believed that although the surface is covered with mercapto ligands, the main contribution to magnetism comes from Zn-O, not Zn-S bonds [67].

The brief review given in Ref. [68] clearly indicates that the behavior of “nonmagnetic” materials in a magnetic field directly depends on the degree of their dispersion and other types of defectiveness. The listed experimental results cover a wide range of materials and make it possible to assert *that a ferro (antiferro) magnetic state can be realized in any material by converting it into a nanodispersed state with a sharp increase in the degree of structural imperfection.*

Based on the above, it is now necessary to expand this conclusion.

SURFACE FERROMAGNETISM AT ROOM TEMPERATURE PRESENTS TO ANY MATERIALS IN A HIGH-DISPERSION NANOSTATE. The exception is organic, polymer and biomaterials.

This is most clearly manifested in nanomaterials, namely, nanopowders consisting of nanoparticles isolated from each other. Defective structures responsible for the appearance of magnetism in such NMs can be regulated within certain limits. Their characteristics are reproducible and can work as a basis for theoretical developments. However,

there are still few such materials [69]. The bulk of nanopowders are formed by simple compaction of nanopowders. In this case, the “ideal” structure of individual NPs in them is violated (**Fig. 3**). In addition, for the implementation of collective magnetic effects, it is necessary to ensure the interaction between the magnetic shells (often severely destroyed) of individual NPs. All this is difficult to control. Therefore, the magnetic characteristics of such materials are generally poorly reproducible. Despite this, the general trends can be traced quite clearly.

At the first stage of the development of these works, it was believed that the surface HTFM is a characteristic of only macroform diamagnetic nanomaterials of a certain type, such as ZnO, Al₂O₃, MgO, GaN, and CdSe semiconductors [66,70-73]. Within this type of objects, models were built and calculations were carried out. For example, in [74-77], Thus, for the first time, the following proposition appears, which later became generally accepted for some time: for the manifestation of cooperative magnetic phenomena in some materials, it is necessary (but not sufficient) that uncompensated magnetic moments (spins) exist in them in a significant amount (this is a requirement of cooperative effects), and so that between these moments there were so-called “exchange” interactions, thanks to which their ordered spatial arrangement would be energetically favorable. The proposed model of “localized magnetic moments” well (at least qualitatively) explained the fact that in the overwhelming majority of materials known at that time (including important from a practical point of view), exhibiting collective forms of magnetism, uncompensated magnetic moments are due *to empty electrons d and f of the inner shells of transition metal atoms. The electrons of the outer valence s and p shells, as a rule, participate in the formation of a chemical bond and therefore their magnetic moments are compensated.* It was this phenomenon that later became known as “d⁰ magnetism” [56], i.e. magnetism in the absence of d-metals.

The physical models proposed to explain the HTFM of NPs of “nonmagnetic” materials were based on the above-mentioned

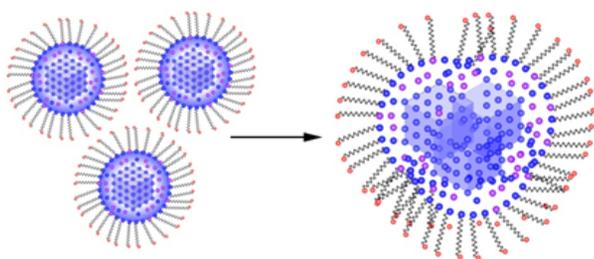


Fig. 3. Violation of the “ideal” structure of NPs when they are compacted.

well-known concepts. In addition, the concept of uncompensated magnetic sublattices, various types of spin subsystems, the emergence of a new crystal structure that promotes magnetism, etc. [78-82] were attracted. These approaches were suitable for describing the properties of a narrow range of materials and were difficult to use when considering systems other than those based on such approaches. This situation could not satisfy anyone.

It was noted that the realization of d^0 magnetism requires a substantial defectiveness of the structure, which is more often found in low-dimensional systems of NPs or thin (only a few nm) films [83-86].

Gradually, an opinion began to form about the critical importance of structural defects (of various types) for obtaining HTFM in systems exhibiting d^0 magnetism. This fact was reflected in the appearance in 2008 of the term “defect-induced” magnetism [56,83,87].

However, there are many defects, and it was not obvious which ones and how they determine magnetism. Many types of defects in NM did not lead to the appearance of HTFM. Often talked about artifact in such experiments, but this idea was quickly refuted in a number of works [27,85,88,89].

Each of these studies has been fairly successful in describing the properties (magnetic behavior) of a selected narrow set of materials. For example, when discussing the HTFM of a hafnium oxide nanofilm, the authors noted that, taking into account the conditions of sample preparation, a significant number of oxygen vacancies can be expected. In such oxides, which have a high dielectric constant, unpaired electrons tend to form an impurity band, where they can be localized due to correlation effects and local potential fluctuations [90]. Such an impurity band mixes with vacant $5d$ states, which in turn polarizes the impurity band and provides the necessary ferromagnetic coupling. The observed anisotropy suggests a large $5d$ -orbital contribution to the spin-orbit coupling. However, such approaches are

hardly suitable for describing the phenomenon as a whole.

It became clear that the existing understanding of magnetism in solids is insufficient to explain these results. However, there are no attempts to consider the entire set of available facts from a unified standpoint.

The clearly expressed HTFM of individual “pure” carbon nanostructures that have nothing other than s and p orbitals has received a definite theoretical consideration [62,91-93]. But the question arises, what is common between a 2D nanofilm of hafnium oxide and graphene, if they behave identically in a magnetic field? Realizing this, a number of authors began to consider structural disturbances – defects as a possible universal reason for the manifestation of HTFM in such materials. [94-98] Gradually they began to pay attention to edge effects, distorted surface states, various structural defects [99].

Initially, it looks paradoxical that in order to obtain the required magnetic properties in the systems under consideration, it is necessary not to get rid of defects, but to learn how to control them, *to turn a defect into an effect*. In this case, magnetism remains an intrinsic property of the system, as well as a significant degree of defectiveness, which is fundamentally unavoidable, as, for example, for NPs [14]. A well-developed system of defects (with the exception of defects arising during the growth of crystals – dislocations, disclinations, etc., which are rigidly embedded in the crystal structure – they are not considered here) can provide a sufficiently large number of exchange magnetic bonds necessary for the emergence of a cooperative magnetic effect. However, it immediately became clear that not every defect creates uncompensated magnetic moments. And not every system of defect-induced magnetic moments can turn into a cooperative state.

In review [100], an attempt was made not only to analyze and systematize the experimental results, but also to generalize the methods which make it possible to control the intensity of ferromagnetic exchange interaction in systems of NPs of nonmagnetic (in the volume state) materials. It

should be noted that the introduction of 3D metal ions, which is often used in the creation of magnetic semiconductors, is nothing more than one of the methods for realizing defect structures [83,101]. Such defect centers can act as mediators of magnetic exchange [102].

4. IS IT POSSIBLE TO DIRECTLY INFLUENCE THE SURFACE MAGNETIC EXCHANGE, TO ENHANCE IT?

The reasons which can determine the cooperative magnetic effect in the system of defect-induced magnetic moments can be reduced to the following:

- 1) it is believed that a decrease in the distances between magnetic centers by careful compaction of NPs enhances exchange. On the other hand, compaction destroys the integrity of the surface layer of NPs, which can lead to unpredictable changes in the magnetic characteristics. [86,103];
- 2) surface modification by removing (replacing) adsorbents (ligands) that hinder the magnetic exchange [44,46,104];
- 3) adsorption of ligands – mediators of magnetic exchange [105].

From the data presented, it follows without doubt that the described phenomenon is of a general nature. A distinctive feature of all considered objects is the presence in them of a large number of defects of various nature. Purely phenomenologically, it is reasonable, in the first approximation, to associate the appearance of this or that type of magnetic ordering with the presence of defects of one nature or another in the object. But a number of difficult questions immediately arise, exhaustive answers to which have not yet been found.

There are direct, indirect, double indirect exchange, the interaction of Ruderman-Kittel-Kasuya-Yosida (RKKY) and some others [106]. The concept of exchange interactions is of a model nature (in fact, the form of the Hamiltonian is postulated), this explains its limitations [107]. But what about those cases when there are no d or f electrons, and exchange interactions are present and cooperative effects take place?

HTFM of NPs of “nonmagnetic” materials is a widespread phenomenon, for which there are still no unified phenomenological approaches, the role of defects is recognized, but which defects and how they affect HTFM is not specified.

It is believed that the surface of spheroidal nanoparticles is a defect that determines and causes HTFM. NPs isolated from each other in the matrix behave like pseudo-atoms with large (often giant) collective magnetic moments. As the distance between particles decreases, exchange interactions arise and intensify. The results of intraparticle exchanges are summed up with interparticle ones and together determine the behavior of an NM in a magnetic field.

Based on the thought experiment proposed at the beginning, as a result of the rupture of the structure of the original crystal, three types of orbitals can appear on the surface of the removed cube. (Fig. 4): two (A and B) as a result of heterolytic rupture of bonds and C - after homolysis of bonds. This is a simplified diagram designed to demonstrate the electronic consequences of bond breaking using the p-orbital as an example. In the case of d or f orbitals, the picture is more complicated, but in the end the electronic result of bond breaking is the same as in the diagram.

During the transformation of a cube into a spheroidal NP, these orbitals behave differently: orbitals with unpaired electrons remain directed orthogonally to the NP surface and cannot participate in the formation of direct valence bonds with each other in the surface layer. At the

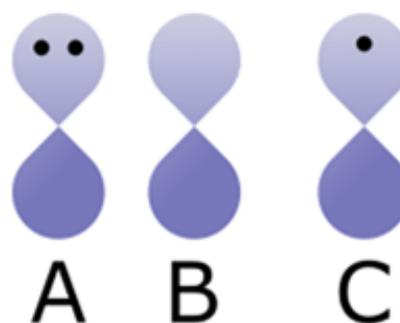


Fig. 4. Three types of bond breaking in NM.

same time, such a configuration does not prevent exchange interactions between parallel spins; as a result, a "magnetic layer" is formed, covering the entire surface of the NP and determining its behavior in a magnetic field.

The rest of the electrons and vacant orbitals formed as a result of breaking bonds in the macrostructure form an "electronic layer" also located on the NP surface and determining its behavior in an electromagnetic field, the interaction with which manifests itself in the spectra in the form of plasmon resonance [108].

These two subsystems "magnetic" and "electronic" coexist in the surface layer of the NP and determine its properties. The model considers idealized spheroidal NPs located separately from each other.

In real powders and other NMs, individual NPs behave like giant magnetic pseudo-atoms. In this case, interparticle magnetic interactions (exchange) occur along with intraparticle interactions. The convergence of NPs over short distances destroys the ideal picture, which often affects the reproducibility of the measurement results and complicates their interpretation.

The above "ideal" scheme stops working. In such materials (and there are most of them), intraparticle and interparticle exchange interactions, as a rule, cannot be separated and formalized. In the case of non-spheroidal morphology of planar 2D nanosystems (in nanofilms, graphene, etc.), "linear" defects are added to a spheroidal defect (shell) that has become flat – end and along internal breaks in the structure.

Orbitals with unpaired spins lie in the plane and their exchange interactions can occur only with neighbors at the edge of the plane.

There are two subsystems: electronic, due to the overlap of orbitals filled with electrons, and spin, due to exchange interactions between unpaired electrons. These two subsystems are orthogonal to each other. The electronic subsystem is polarized by an external electromagnetic field (light quanta); spin – by an external magnetic field. The model conceptually reflects the presence of

two subsystems, as exemplified by spheroidal and flat defects. [100]

5. THE ROLE OF NP SIZE

When NPs grow larger, surface ferromagnetism does not disappear anywhere, just the fraction of surface atoms sharply decreases with increasing NP size, and their contribution to the magnetic behavior of the sample is beyond the sensitivity limit of the method [103].

The proposed model makes it possible to qualitatively generalize and classify the available information on the HTFM of NPs of "nonmagnetic" materials. Thus, as a result of the above-described virtual operation, the diamagnetic material is transformed into a system of NPs with uncompensated surface spins; their exchange interactions can lead to the manifestation of various forms of magnetic ordering, from para up to ferro (less often antiferro) magnetism.

6. THEORETICAL MODELS OF HTFM IN LF AND NM

The experimental detection of the HTFM phenomenon in NPs and other nanostructures was accompanied by attempts to interpret it theoretically (see reviews [100,107,109] and references therein). Ferromagnetism is more common in metal-containing compounds; therefore, the band approach dominates in the theory of magnetism as a whole. At the same time, the HTFM phenomenon is observed both in metals and in dielectrics and semiconductors, i.e. claims to be universal [110]. Therefore, it is not surprising that there is currently no unified explanation of this phenomenon for all materials where it was observed.

Attempts to theoretically explain HTFM in carbon nanostructures, with the exception of graphene, on the whole, do not go beyond the model of defect-induced band or localized magnetism used for most inorganic nanoparticles [62]. A theoretical study of HTFM in graphene "nanoislands" of triangular and hexagonal shape with zigzag edges was carried out in the Hubbard mean field model and using the DFT methods [111]. It is shown that electronic states with

nonzero spin actually appear in the structures. Attention was drawn to the dimensional features of HTFM in graphene in review [112]. Thus, carbon nanomaterials are an integral part of the family of materials with HTFM.

7. CONCLUSION

The nanostate of matter presented us with another surprise – the fundamental possibility of obtaining magnetic materials from almost any substance by fairly simple, sometimes (as in the case of ZnO) primitive methods. At the same time, theoretical concepts and deep understanding of this phenomenon (as often happened before) are still far from complete. The task of the authors of the review was to show that the material presented is not a collection of individual random facts taken from numerous experimental and theoretical studies, but is a phenomenon (a general pattern) that requires deep and theoretical analysis from a unified standpoint. The future will show to what extent we have managed to do this. Summarizing the state of the problem of the theoretical description of HTFM, we can say that, on the one hand, its satisfactory solution is still hampered by the diversity and incompleteness of experimental studies. On the other hand, the unusualness of the detected magnetic properties, apparently, should stimulate theorists to critically analyze the existing ideas about the magnetic properties of materials, model approaches, and computational techniques.

The discovery of the HTFM phenomenon occurred in the process of searching for materials in which the electrical and magnetic properties are closely interrelated. Such materials are necessary for spintronics which is a modern branch of electronics, which uses not only the electric charge of current carriers, but also their spin, i.e. magnetic properties [113,114]. In the theoretical models of HTFM, in particular, in the ferron theory of E.L. Nagaev, the electrical and magnetic properties of the material are closely interrelated on the NP surface. Therefore, there is no doubt that with further study of HTFM, it, in addition to undoubted fundamental scientific interest, will find important practical applications [115]. The most probable closest practical use of the HTFM

phenomenon is a magnetoresistance-based device. Although the magnitude of the magnetoresistance in magnetized materials is not too high [22,116], targeted control of the defect structure of the NP surface [117] can correct this drawback.

It can be assumed that the widespread involvement of magnetic methods in the study of NPs and NMs will not only lead to the discovery of new examples of the manifestation of magnetism in materials, but will also expand our understanding of magnetism as a fundamental property of matter. The authors hope that the approach presented in the article will become a reason not only for discussions, but also serve as a basis for a deeper consideration of the HTFM problem by specialists.

REFERENCES

1. Gubin SP, Kataeva NA, Khomutov GB. Promising areas of nanoscience: chemistry of nanoparticles of semiconductor materials. *Izv. akad. nauk. Ser. Khim.*, 2005, 4:811-836 (in Russ.).
2. Jung DR, Kim J, Nahm C, Choi H, Nam S., Park B. Review paper: Semiconductor nanoparticles with surface passivation and surface plasmon. *Electronic Materials Letters*, 2011, 7 (3):185-194.
3. Gubin SP, Koksharov YA, Khomutov GB, Yurkov GYu. Magnetic nanoparticles: methods of preparation, structure and properties. *Advances in Chemistry*, 2005, 74 (6):539-574.
4. Faivre D, Bennet M. Magnetic nanoparticles line up. *Nature*, 2016, 535 (7611):235-236.
5. Gawande MB, Goswami A, Felpin F-X, Asefa T, Huang X, Silva R, Zou X, Zboril R, Varma RS. Cu and Cu-based nanoparticles: synthesis and applications in catalysis. *Chem. Rev.*, 2016, 116:3722-3811.
6. Chaudhuri RG, Paria S. Core/Shell Nanoparticles: Classes, Properties, Synthesis Mechanisms, Characterization, and Applications. *Chem. Rev.*, 2012, 112:2373-2433.
7. Ilka M, Bera S, Kwon S-H. Influence of Surface Defects and Size on Photochemical Properties of SnO₂ Nanoparticles. *Materials*, 2018, 11 (6):904.

8. Bera S, Khan H, Biswas I, Jana S. Polyaniline hybridized surface defective ZnO nanorods with long-term stable photoelectrochemical activity. *Appl. Surf. Sci.*, 2016, 383:165-176.
9. Anuchai S, Phanichphant S, Tantraviwat D, Pluengphon P, Bovornratanaraks T, Inceesungvorn B. Low temperature preparation of oxygen-deficient tin dioxide nanocrystals and a role of oxygen vacancy in photocatalytic activity improvement. *J. Colloid Interface Sci.*, 2018, 512:105-114.
10. Martínez B, Roig A, Molins E, González-Carreño T, Serna CJ. Magnetic Characterization of γ -Fe₂O₃ Nanoparticles Fabricated by Aerosol Pyrolysis. *J. Appl. Phys.*, 1998 83:3256-3262.
11. Singh R. Unexpected magnetism in nanomaterials. *JMMM*, 2013, 346:58-73.
12. Crespo P, de la Presa P, Marín P, Multigner M, Alonso JM, Rivero G, Yndurain F, González-Calbet JM, Hernando A. Magnetism in nanoparticles: tuning properties with coatings. *J. Phys.: Condens. Matter.*, 2013, 25:484006-4840027.
13. Tilley RD. *Defects in solids*. Wiley, Weinheim, 2008, 552 p.
14. Gubicza J. *Defect Structure and Properties of Nanomaterials: Second and Extended Edition*. Cambridge, Woodhead Publishing, 2017, 410 p.
15. Wu Z, Cao S, Zhang C, Piao L. Effects of bulk and surface defects on the photocatalytic performance of size-controlled TiO₂ nanoparticles. *Nanotechnology*, 2017, 28:275706.
16. Pi X, Delerue C. Tight-binding calculations of the optical response of optimally P-doped Si nanocrystals: a model for localized surface plasmon resonance. *Phys. Rev. Lett.*, 2013, 111 (17):177402-5.
17. Schnack J. Large magnetic molecules and what we learn from them. *Contemp. Physics*, 2019, 60:127-145.
18. Gómez Vitoria M, Weick G, Wienmann D, Jalabert RA. Orbital magnetism in ensembles of gold nanoparticles. *Phys. Rev. B.*, 2018, 98 (19):195417-38.
19. Nagaev EL. Ground state and anomalous magnetic moment of conduction electrons in an antiferromagnetic semiconductor. *JETP Letters*, 1967, 6:484.
20. Nagaev EL. Ferromagnetic microregions in a semiconductor antiferromagnet. *JETP*, 1968, 54: 228.
21. Kasuya T. Mobility of the antiferromagnetic large polaron. *Solid State Commun.*, 1970, 8:1635.
22. Nagaev EL. Magnetic polarons (ferrons) of complicated structure. *JETP Letters*, 2001, 74:431.
23. Hamdy Makhlof AS, Barhoum A. *Fundamentals of Nanoparticles: Classifications, Synthesis Methods, Properties and Characterization*. Elsevier, 2018, 666 p.
24. de Mello Donega C. *Nanoparticles*. Springer-Verlag Berlin Heidelberg, 2014, 299 p.
25. Nealon GL, Donnio B, Greget R, Kappler J-P, Terazzi E, Gallani JL. Magnetism in gold nanoparticles. *Nanoscale*, 2012, 4:5244-52258.
26. Agrachev M, Antonello S, Dainese T, Ruzzi M, Zoleo A, Aprá E, Govind N, Fortunelli A, Sementa L, Maran F. Magnetic Ordering in Gold Nanoclusters. *ACS Omega*, 2017, 2(6):2607-2617
27. Garcia MA, Merino JM, Pinel EF, Quesada A, de la Venta J, Gonzalez MLR, Castro GR, Crespo P, Llopis J, Gonzalez-Calbet JM, Hernando A. Magnetic Properties of ZnO Nanoparticles. *Nano Lett.*, 2007,7(6):1489-1494.
28. Guglieri C, Laguna-Marco MA, García MA, Carmona N., Céspedes, Garsia-Hernández M., Espinosa A., Chaboy J. XMCD Proof of Ferromagnetic Behavior in ZnO Nanoparticles. *J. Phys. Chem. C.*, 2012, 116 (11):6608-6614.
29. Gubin SP, Koksharov YA. Preparation, structure and properties of magnetic materials based on Co-containing nanoparticles. *Neorg. Mat.*, 2002, 38(11):1287-1304.
30. Yamamoto Y, Hori H. Direct Observation of the Ferromagnetic Spin Polarization in Gold Nanoparticles: A Review. *Rev. Adv. Mater. Sci.*, 2006, 12:23-32.

31. Fiorani D. *Surface Effects in Magnetic Nanoparticles*. New York, Springer, 2005, 300 p.
32. Wang XQ, Mujumdar AS. A review on nanofluids — part I: theoretical and numerical investigations. *Brazilian J of Chem. Eng.*, 2008, 25(4):613-630.
33. Yu W, Xie H. A Review on Nanofluids: Preparation, Stability Mechanisms, and Applications. *J.Nanomaterials*, 2012, Article ID 435873, 1 p.
34. Merkel TJ, Herlihy KP, Nunes J, Orgel RM, Rolland JP, DeSimone JM. Scalable, Shape-specific, Top-down Fabrication Methods for the Synthesis of Engineered Colloidal Particles. *Langmuir*, 2010, 26(16):13086-13096.
35. Bulychev NA, Kazaryan MA, Kirichenko MN, Ivanov AV Study of Acoustoplasma Discharge as a Technique for Synthesis of Optically Active Materials. *Int. J. Nanotec.*, 2019, 16(1/2/3):34-41.
36. Lo CH, Tsung TI, Chen LC. Shape-controlled synthesis of Cu-based nanofluid using submerged arc nanoparticle synthesis system (SANSS). *J. Crystal. Growth*, 2005, 277(1-4):636-642.
37. Sato-Berrú R, Redón R, Vázquez-Olmos A, Saniger J. Silver nanoparticles synthesized by direct photoreduction of metal salts. Application in surface-enhanced Raman spectroscopy. *J. Raman Spectrosc.*, 2009, 40:376-380.
38. Jin R, Cao YC, Hao E, Metraux GS, Schatz GC, Mirkin C. Controlling anisotropic nanoparticle growth through plasmon excitation. *Nature* 2003, 425:487-490.
39. Kim S, Yoo B, Chun K, Kang W, Choo J, Gong M-S, Joo SW. Catalytic effect of laser ablated Ni nanoparticles in the oxidative addition reaction for a coupling reagent of benzylchloride and bromoacetonitrile. *J. Mol. Catal. A: Chem.*, 2005, 226(2):231-234.
40. Heuer-Jungemann A, Feliu N, Bakaimi I, Hamaly M, Alkilany A, Chakraborty I, Masood A, Casula MF, Kostopoulou A, Oh E, Susumu K, Stewart MH, Medintz IL, Stratakis E, Parak WJ, Kanaras AG. The Role of Ligands in the Chemical Synthesis and Applications of Inorganic Nanoparticles. *Chem. Rev.*, 2019, 119(8):4819-4880.
41. Ling D, Hackett MJ, Hyeon T. Surface ligands in synthesis, modification, assembly and biomedical applications of nanoparticles. *Nano Today*, 2014, 9(4):457-477.
42. Kate K, Singh K, Khanna PK. Microwave formation of Polypyrrole/Ag nano-composite based on interfacial polymerization by use of AgNO₃. *Synth. React. Inorg Met-Org Chem.*, 2011, 41:199-202.
43. Angappane S, Park J, Jang Y, Hyeon T, Park JG. Magnetic Pd nanoparticles: Effects of surface atoms. *J.Phys.: Condens. Matter.*, 2008, 20:295209-295215.
44. Guerrero E, Muñoz-Márquez MA, García MA, Crespo P, Fernández-Pinel E, Hernando A, Fernández A. Surface plasmon resonance and magnetism of thiol-capped gold nanoparticles. *Nanotechnology*, 2008, 19:175701-175706.
45. Ionita P, Caragheorgheopo A, Gilbert B, Chechik V. Dipole-Dipole Interactions in Spin-Labeled Au Nanoparticles as a Measure of Interspin Distances. *J. Phys. Chem. B*, 2005, 109:3734-3742.
46. Tu W, Fukui K, Miyazaki A, Enoki T. Magnetic and Electronic Properties of Palladium Nanoparticles Coated with π -Conjugated Tetrathiafulvalenes Derivative. *J. Phys. Chem. B*, 2006, 110:20895-20900.
47. Gubin SP, Koksharov YA, Ioni YV. Magnetism of nanoparticles of "non-magnetic" materials; the role of defects. *Zhurnal neorg. khimii*, 2021, 66(1):1-27 (in Russ.).
48. Cai L, Zhu J, Hou Y, Tong M, Kim H. Influence of gravity on transport and retention of representative engineered nanoparticles in quartz sand. *J. of Cont. Hydr.*, 2015, 181:153-160.
49. Palma-Jiménez M, Corrales Ureña Y, Villalobos Bermúdez C, Vega Baudrit JR. Microgravity and Nanomaterials. *Int. J of Biophysics*, 2017, 7(4):60-68.

50. Vonsovsky SV. *Magnetism. Magnetic properties of dia-, para-, ferro-, antiferro- and ferrimagnets.* Moscow, Nauka Publ., 1971, 1032 p.
51. Hurd CM. Varieties of Magnetic Order in Solids. *Contempt. Phys.*, 1982, 23:469-493.
52. Veciana J. *π -Electron Magnetism from Molecules to Magnetic Materials.* Springer, 2001, 208 p.
53. Mattis D.C. *The theory of magnetism made simple: an introduction to physical concepts and to some useful mathematical methods.* L.: World Scientific, 2006, 580 p.
54. Norberg NS, Kittilstved KR, Amonette JE, Kukkadapu RK, Schwartz DA, Gamelin DR. Synthesis of Colloidal Mn²⁺: ZnO Quantum Dots and High-TC Ferromagnetic Nanocrystalline Thin Films. *JACS*, 2004, 126(30):9387-9398.
55. Kittilstved KR, Norberg NS, Gamelin DR. Chemical Manipulation of High-TC Ferromagnetism in ZnO Diluted Magnetic Semiconductors. *Phys. Rev. Lett.*, 2005, 94(14):147209-147213.
56. Coey JMD. d⁰ ferromagnetism. *Solid State Sciences*, 2005, 7:660-667.
57. Droghetti A, Pemmaraju CD, Sanvito S. Predicting d⁰ magnetism: Self-interaction correction scheme. *Phys.Rev.B*, 2008, 78:140404 (R)-4.
58. Sundaresan A, Rao CNR. Ferromagnetism as universal feature of inorganic nanoparticles. *Nano today*, 2009, 4:96-106.
59. Hernando A, Crespo P, García MA, Coey M, Ayuela A, Echenique PM. Revisiting magnetism of capped Au and ZnO nanoparticles: Surface band structure and atomic orbital with giant magnetic moment. *Phys. Status Solidi B*, 248:2352-2360.
60. Trudel S. Unexpected magnetism in gold nanostructures: making gold even more attractive. *Gold Bull.*, 2011, 44:3-13.
61. Ackland K, Coey JMD. Room temperature magnetism in CeO₂ — A review. *Physics Rep.*, 2018, 746:1-39.
62. Makarova T, Palacio F. *Carbon-based Magnetism. An Overview of the Magnetism of Metal Free Carbon-Based Compounds and Materials.* Amsterdam, Elsevier, 2006, 576 p.
63. Wang Y, Huang Y, Song Y, Zhang X, Ma Y, Liang J, Chen Y. Room-Temperature Ferromagnetism of Graphene. *Nano Lett.*, 2009,9:220-224.
64. Matte HSSR, Subrahmanyam KS, Rao CNR. Novel Magnetic Properties of Graphene: Presence of Both Ferromagnetic and Antiferromagnetic Features and Other Aspects. *J. Phys. Chem. C*, 2009, 113:9982-9985.
65. Morozavska AN, Eliseev EA, Glinchuk MD, Blink R. Surface-induced magnetism in the solids with impurities and vacancies. *Physics B*, 2011, 406:1673-1688.
66. Singh SB, Limaye MV, Date SK, Kulkarni SK. Room temperature ferromagnetism in thiol-capped CdSe and CdSe: Cu nanoparticles. *Chem. Phys. Lett.*, 2008, 464(4-6):208-210.
67. Chaboy J, Boada R, Piquer C, Laguna-Marco MA, García-Hernández M, Carmona N, Llopis J, Ruíz-González ML, González-Calbet J, Fernández JF, García MA. Evidence of intrinsic magnetism in capped ZnO nanoparticles. *Phys. Rev. B.*, 2010, 82:064411-9.
68. Dietl T. A ten-year perspective on dilute magnetic semiconductors and oxides. *Nature Materials*, 2010, 9:965-974.
69. Sundaresan A, Bhargavi R, Rangarajan N, Siddesh U, Rao CNR. Ferromagnetism as a universal feature of nanoparticles of the otherwise nonmagnetic oxides. *Phys. Rev. B*, 2006, 74:161306 (R)-6.
70. Sharma P, Gupta A, Rao KV, Owens FJ, Sharma R, Ahuja R, Guillen JMO, Johansson B, Gehring GA. Ferromagnetism above room temperature in bulk and transparent thin films of Mn-doped ZnO. *Nat. Mater.*, 2003, 2:673-679.
71. Hong NH, Sakai J, Huong NT, Brizé V. Room temperature ferromagnetism in laser ablated Ni-doped In₂O₃ thin films. *Appl Phys. Lett.*, 2005, 87(10):102505-102507.
72. Hong NH, Sakai J, Poirot N, Brizé V. Room-temperature ferromagnetism observed in

- undoped semiconducting and insulating oxide thin films. *Phys. Rev. B*, 2006, 73(13):132404-132407.
73. Seehra MS, Dutta P, Neeleshwar S, Chen Y-Y, Chen CL, Chou SW, Chen CC, Dong C-L, Chang C-L. Size-Controlled Ex-nihilo Ferromagnetism in Capped CdSe Quantum Dots. *Adv. Mater.*, 2008, 20:1656-1660.
74. Moriah T. Recent advances in the theory of magnetism of itinerant electrons. *UFN*, 1981, 135:117-170.
75. Moriah T. *Spin fluctuations in magnets with itinerant electrons*. Moscow, Mir Publ., 1988, 287 p.
76. Izyumov YA, Katsnelson MI, Skryabin YUN. *The magnetism of collectivized electrons*. Moscow, Fizmatlit Publ., 1994, 368 p.
77. Donath M., Nolting W. *Local-Moment Ferromagnets*. Springer, Berlin-Heidelberg, 2005, 329 p.
78. Aguilera-Granja F, Rodríguez-López JL, Michaelian K, Berlanga-Ramírez EO, Vega A. Structure and magnetism of small rhodium clusters. *Phys. Rev. B*, 2002, 66(22):224410-224419.
79. Gómez Vitoria M, Weick G, Wienmann D, Jalabert R. Orbital magnetism in ensembles of gold nanoparticles. *Phys. Rev. B*, 2018, 98:195417-195425.
80. Grèget R, Nealon GL, Vilenó B, Turek P, Mèny C, Ott F, Derory A, Voirin E, Rivière E, Rogalev A, Wilhelm F, Joly L, Knafo W, Ballon G, Terazzi E, Kappler JP, Donnio B, Gallani JL. Magnetic Properties of Gold Nanoparticles: A Room-Temperature Quantum Effect. *Chem. Phys. Chem.*, 2012, 13:3092-3097.
81. Hernando A, Crespo P, García MA. Origin of Orbital Ferromagnetism and Giant Magnetic Anisotropy at the Nanoscale. *Phys. Rev. Lett.*, 2006, 96(17):057206-057209.
82. Jeon YT, Lee GH. Magnetism of the fcc Rh and Pd nanoparticles. *J. Appl. Phys. B*, 2008, 103:094313-094317.
83. Coey JMD, Chambers SA. Oxide Dilute Magnetic Semiconductors - Fact or Fiction? *MRS Bull.*, 2008, 33:1053-1058.
84. Ivanovskiy AL. Magnetic effects in nonmagnetic sp-materials, induced by sp-impurities and defects. *UFN*, 2007, 177:1083-1105.
85. Tuboltsev V, Savin A, Pirojenko A, Räisänen J. Twist of Magnetism in Nanocrystalline Gold. *ACS Nano*, 2013, 7:6691-6699.
86. Suda M, Kameyama N, Ikegami A, Suzuki M, Kawamura N, Einaga Y. Size-reduction induced ferromagnetism and photo-magnetic effects in azobenzene-thiol-passivated gold nanoparticles. *Polyhedron*, 2009, 28:1868-1874.
87. Esquinazi P, Hergert W, Spemann D, Setzer A, Ernst A. Defect-induced magnetism in solids. *IEEE Transactions on Magnetics*, 2013, 49:4668-4674.
88. Abraham DW, Frank MM, Guha S. Abraham, D. W., Frank, M. M., & Guha, S. (2005). Absence of magnetism in hafnium oxide films. *Applied Physics Letters*, 2005, 87(25):252502-252504.
89. Coey JMD, Ackland K, Venkatesan M, Sen S. Collective magnetic response of CeO₂ nanoparticles. *Nat. Phys.*, 2016, 12: 694-699.
90. Venkatesan M, Fitzgerald CB, Coey JMD. Unexpected magnetism in a dielectric oxide. *Nature*, 2004, 430:630.
91. Sheka EF. Odd Electrons in Molecular Chemistry, Surface Science, and Solid State Magnetism. *Int. J. Quant. Chem.*, 2007, 107:2935-2955.
92. Ray S.C. *Magnetism and Spintronics in Carbon and Carbon Nanostructured Materials*, 1st Ed. Elsevier, 2020, 240 p.
93. Yaziev O. Emergence of magnetism in graphene materials and nanostructures. *Rep. Prog. Phys.*, 2010, 73:056601-056616.
94. Lee H, Son Y-W, Park N, Han S, Yu J. Magnetic ordering at the edges of graphitic fragments: Magnetic tail interactions between the edge-localized states. *Phys. Rev. B*, 2005, 72(17):174431-174438.
95. Okada S, Oshiyama A. Magnetic Ordering in Hexagonally Bonded Sheets with First-Row Elements. *Phys. Rev. B*, 2001, 64(14):146803-146806.

96. Scopel WL, Paz WS, Freitas JCC. Interaction between single vacancies in graphene sheet: An ab initio calculation. *Solid State Commun.*, 2016, 250:5-10.
97. Urita K, Suenaga K, Sugai T, Shinohara H, Iijima S. In Situ Observation of Thermal Relaxation of Interstitial-Vacancy Pair Defects in a Graphite Gap. *Phys. Rev. Lett.*, 2005, 94(15):155502-155505.
98. Zhou Q, Fu Z, Wang C, Tang Y, Zhang H, Yuan L, Yang X. The electronic and magnetic properties of B-doping Stone – Wales defected graphene decorated with transition-metal atoms. *Physica E*, 2015, 73:257-261.
99. Shukla V. Observation of critical magnetic behavior in 2D carbon based composites. *Nanoscale Adv.*, 2020, 2:962-995.
100. Coey JMD. Magnetism in d^0 oxides. *Nat. Mater.*, 2019, 18:770.
101. Coey JMD. Dilute magnetic oxides. *Current Opinion in Solid State and Materials Science*, 2006, 10:83-92.
102. Kikoin K. Ferromagnetic ordering in dilute magnetic dielectrics with and without free carriers. *J. Magn. Magn. Mater.*, 2009, 321:702-705.
103. Hori H, Yamamoto Y, Iwamoto T, Miura T, Teranishi T, Miyake M. *Phys. Rev. B.*, 2004, 16:174411-174415.
104. Crespo P, Litrán R, Rojas TC, Multigner M, de la Fuente JM, Sánchez-López JC, García MA, Hernando A, Penadés S, Fernández A. Permanent Magnetism, Magnetic Anisotropy, and Hysteresis of Niol -pped Gold Capped. *Phys. Rev. B*, 2004, 93(8):087204-087207.
105. Tang N. *In Spintronic 2D Materials: Fundamentals and Applications*. (Eds. Liu W, Xu Y). Amsterdam, Elsevier, 2005, pp. 137-161.
106. Rakitin YuV, Kalinnikov TV. *Sovremennaya magnetokhimiya* [Modern magnetochemistry]. Sankt-Petersburg, Nauka Publ., 1994, 287 p.
107. Zunger A, Lany S, Raebiger H. The quest for dilute ferromagnetism in semiconductors. *Guides and misguides for theory. Physics*, 2010 3:53-62.
108. Boken J, Khurana P, Thatai S, Kumar D, Prasad S. Plasmonic nanoparticles and their analytical applications: A review. *Applied Spectroscopy Reviews*, 2017, 52(9):774-820.
109. Stoneham M. The strange magnetism of oxides and carbons. *J. Phys.: Condens. Matter*, 2010, 22(7):074211-074217.
110. Sundaresan A, Rao CNR. Ferromagnetism as universal feature of inorganic nanoparticles. *Nano today*, 2009, 4:96-106.
111. Fernández-Rossier J, Palacios JJ. Magnetism in Graphene Nanoislands. *Physical Review Letters*, 2007, 99(17):177204-177207.
112. Sheka EA, Popova NA, Popova VA. Physics and chemistry of graphene. Emergence, magnetism, mechanophysics and mechanochemistry. *UFN*, 2018, 188:720-772.
113. Wolf SA, Treger D. Spintronics: a new paradigm for electronics for the new millennium. *IEEE Trans. On Magn.*, 2000, 36:2748-2751.
114. Saha SK, Baskey M, Majumdar D. Graphene Quantum Sheets: A New Material for Spintronic Applications. *Adv. Matter*, 2010, 22:5531-5536.
115. Hu J, Li W, Hou Y. Two-Dimensional Magnetic Nanostructures. *Trends in Chemistry*, 2020, 2:163-174.
116. Nagaev EL. *Izbrannyye proizvedeniya* [Selected Works]. Moscow, Fizmatlit Publ., 2004, 320 p.
117. Ziese M. Extrinsic magnetotransport phenomena in ferromagnetic oxides. *Rep.Prog. Phys.*, 2002, 65(2):143-249.

DOI: 10.17725/rensit.2020.12.471

Highly Conductive Polymer PEDOT:PSS - Application in Biomedical and Bioelectrochemical Systems

Anatoly N. Reshetilov, Anna E. Kitova, Sergey E. Tarasov, Yulia V. Plekhanova, Alexander G. Bykov

Federal research Center "Pushchino Biological Research Center of the Russian Academy of Sciences", G.K. Skryabin Institute of Biochemistry and Physiology of Microorganisms of RAS, <http://www.ibpm.ru/>

Pushchino 142290, Moscow Region, Russian Federation

E-mail: anatol@ibpm.pushchino.ru, kitova@ibpm.pushchino.ru, setar25@gmail.com, plekhanova@ibpm.pushchino.ru, agbykov@rambler.ru

Ashok K. Sundramoorthy

SRM Institute of Science and Technology, <https://www.srmist.edu.in/>

Kattankulathur 603203, India

E-mail: ashokkumar.sun@ktr.srmuniv.ac.in

Iren E. Kuznetsova, Vladimir V. Kolesov

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

E-mail: kuziren@yandex.ru, kv@cplire.ru

Pavel M. Gotovtsev

National Research Center "Kurchatov Institute", <http://www.nrcki.ru/>

Moscow 123182, Russian Federation

E-mail: gotovtsevp@yandex.ru

Received September 14, 2020, peer-reviewed September 28, 2020, accepted October 05, 2020

Abstract: This review deals with the use of the highly conductive polymer PEDOT:PSS in biomedical and bioelectrochemical systems. The examples of toxic effects on living cells, positive effects of PEDOT:PSS on the viability of cells and tissues are given. The properties of the polymer, methods of increasing its electrical conductivity by its modification with various nanoparticles and nanomaterials are discussed. Examples of using PEDOT and its composites in bioelectrochemical devices, such as biosensors and biofuel cells, are considered. Changes in the characteristics of biosensors and biofuel cells under the influence of PEDOT are discussed.

Keywords: highly conductive polymer PEDOT:PSS; biomedical and bioelectrochemical systems; application of PEDOT:PSS; biosensors; biofuel cells

UDC 543.07

Acknowledgements: The research was carried out with the financial support of the Russian Foundation for Basic Research and DNT within the framework of the scientific project No. 19-58-45011 (sections 1-2) and with the financial support of the RFBR within the framework of the scientific projects No. 18-29-23042 (section 3) and 18-29-23024 (sections 4, 5).

For citation: Anatoly N. Reshetilov, Anna E. Kitova, Sergey E. Tarasov, Yulia V. Plekhanova, Alexander G. Bykov, Ashok K. Sundramoorthy, Iren E. Kuznetsova, Vladimir V. Kolesov, Pavel M. Gotovtsev. Highly Conductive Polymer PEDOT:PSS - Application in Biomedical and Bioelectrochemical Systems. *RENSIT*, 2020, 12(4):471-482. DOI: 10.17725/rensit.2020.12.471.

CONTENTS

1. INTRODUCTION (472)
 2. ALTERATION OF PHYSICAL-CHEMICAL PROPERTIES OF PEDOT:PSS AFTER ITS MODIFICATION WITH DIFFERENT MATERIALS (472)
 3. STIMULATING AND TOXIC EFFECTS OF PEDOT:PSS ON CELLS (474)
 4. APPLICATION OF PEDOT:PSS IN BIOELECTROCHEMICAL SYSTEMS: (476)
 - 4.1. USE FOR FORMATION OF BIOSENSOR RECEPTORS (476)
 - 4.2. APPLICATION IN BIOFUEL CELLS (477)
 5. CONCLUSION (479)
- REFERENCES (479)

1. INTRODUCTION

Currently, much attention is paid to the creation of various composite materials and mixtures for effective integration and conjunction with biological material; further use of such compositions can lead to the creation of analytical devices based on them – biosensors; the creation of energy-producing system – biofuel cells (BFCs); the creation of systems with the biomedical perspective. Recent studies show that polymer compounds such as chitosan, polyethylene glycol, polyaniline, polyvinyl alcohol, and so forth play an important role in composite materials. Some of the promising polymers that draw attention are poly(3,4-ethylenedioxythiophene) – PEDOT and poly(3,4-ethylenedioxythiophene): polystyrenesulfonic acid – PEDOT:PSS. In the oxidized state, PEDOT is one of the most stable polymers known. When combined with polystyrene sulfonic acid (PSS) as a counter ion, PEDOT forms a polyelectrolyte complex that has the properties of a stable dispersion that can be produced on an industrial scale. PEDOT:PSS is a polymer

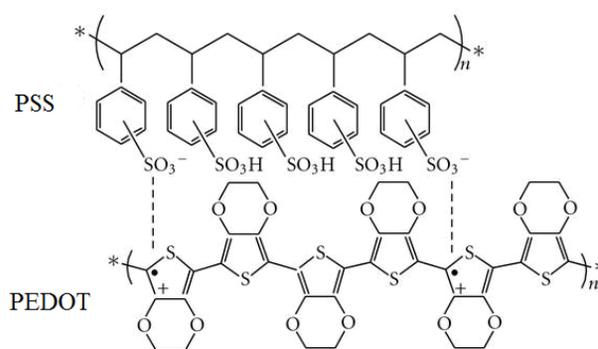


Fig. 1. The chemical formula of PEDOT:PSS.

that is formed by two basic molecules – EDOT (3,4-ethylenedioxythiophene) and PSS (polystyrenesulfonic acid); the chemical formula of PEDOT:PSS is given in Fig. 1. Some of the sulfonyl groups of PSS are deprotonated and carry a negative charge. PEDOT is a conjugated polymer and carries positive charges. The charged macromolecules form a macromolecular salt.

Although the practical importance of the said polymers is highlighted, there remains much to learn about PEDOT:PSS and its properties.

In this review, main characteristics of PEDOT/PEDOT:PSS are described, they are important from a practical perspective enabling the use of this polymer as the basis of analytical and technical systems.

2. ALTERATION OF PHYSICAL-CHEMICAL PROPERTIES OF PEDOT:PSS AFTER ITS MODIFICATION WITH DIFFERENT MATERIALS

An important characteristic of polymers in any bioelectrochemical devices is their conductivity. PEDOT:PSS is considered to be one of the most promising conductive materials and it is used in device production in various fields: in displays, touch panels, polymer and photoelectrochemical cells, organic light-emitting diodes [1], as a coating

in electroluminescent layers [2], as a sensitive element in photo-voltaic devices [3], in memory cells [4], resistive switches [5] and others. Electron transfer in the PEDOT polymer is carried out by a system of conjugated bonds due to electron exchange reactions between neighboring redox sites and is accompanied by conjugated movement of dopant anions along the polymer chain. Thus, PEDOT and PEDOT:PSS have a mixed electron-ion type of conductivity. Addition of metal nanoparticles that actively inject electrons, into such a polymer matrix can contribute to the production of nanostructures with the increased conductivity. Researchers are constantly searching for new ways to improve the conductive properties of PEDOT-based films. One of the most common treatment for improving the conductivity of PEDOT:PSS is to include various nanomaterials and nanoparticles in the composition of the polymer. For instance, in [6] authors sought to investigate the conductivity of PEDOT:PSS films containing gold nanoparticles and inferred that the conductivity of these films is due to hopping charge transport. They examined the absorption spectra in the visible region and volt-ampere characteristics in a wide range of electric fields at the macro (in planar structures) and micro levels (using a conducting atomic force microscope) in the films on the basis of electroactive polymer PEDOT:PSS and gold/silver nanoparticles. It was shown that the behavior of the current-voltage characteristics of nanocomposite films depends on the magnitude of the electric field. It was also found that addition of the gold nanoparticles to PEDOT:PSS in weak fields leads to an increase in volume conductivity by almost two orders of magnitude (due to donor-acceptor interactions), a 50% decrease

in the activation energy of conductivity, and an increase in sensitivity to adsorbed oxygen.

The PEDOT:PSS with conductivity of about $3000 \text{ Cm} \times \text{cm}^{-1}$ was obtained quite recently. The main strategy to increase the conductivity is to change the properties of the polymer when it interacts with various solvents, such as DMSO (dimethylsulfoxide). In addition to the effect of increased conductivity, the authors [7] found that the range of potential, within which the PEDOT:PSS demonstrates highly conductive properties, expands to the region of negative potentials. This characteristic can be used later on when using PEDOT:PSS as part of chemical sensors where high material conductivity is required at particular potentials.

In [8], the authors propose a method for manufacturing flexible conductive structures of PEDOT:PSS polymer and reduced graphene with a multilayer structure. The resulting capacitance value of such composites was about 193 F/g at a current density of 500 mA/g. The obtained material was highly stable – the capacitive component remained at the level of 90% after 1000 cycles, which indicates the prospect of using this approach to manufacture flexible energy storage devices. The works [9,10] demonstrate the possibility of multi-stage formation of composite transparent coatings with increased conductivity based on the carbon nanotube system - conductive PEDOT:PSS polymer. The produced coatings are characterized by a combination of low surface resistance (89 Ohm) and high optical transparency (about 81%). The main advantage of the obtained coatings is their mechanical stability to bending deformations.

Yan et al. [11] were the first to discover that completely isolated halloysite nanotubes

(HNTs) improve the electrical conductivity of PEDOT:PSS films by several orders of magnitude in the case of simple mixing. Due to this effect a highly porous and highly conductive composition of PEDOT:PSS-HNTs was created. By modifying the mixing process, the authors created a flexible conducting hybrid structure with a high specific surface. They propose a mechanism according to which colloidal particles with size of several dozens of nanometers are tightly packed in channels formed by PEDOT:PSS particles. In this case, the resulting conductivity is several orders of magnitude higher than that of the PEDOT:PSS located outside the nanotubes.

The formation of "sandwich structures" based on the successive repetition of procedures of metal oxide layers and polymer deposition led to the creation of composite materials $\{\text{PEDOT}/\text{MnO}_2\}_x$ [12] and $\{\text{PEDOT}/\text{NiO}\}_x$ [13]. Such structures make it possible to increase the capacitance characteristics of materials while maintaining high charge-discharge rates, which is used in supercapacitors.

Cross-linking of PEDOT:PSS with diglycidyl ether of bisphenol A results in production of the electrically conductive films with increased water resistance. Moreover, there is no evidence of decrease in the electrical properties of the film as thermal stability increases. The researchers proposed a promising technique of creation of electrodes that can be used in biological environments where contact with water is often unavoidable.

3. STIMULATING AND TOXIC EFFECTS OF PEDOT:PSS ON CELLS

One of the basic properties of any polymer for potential use in bioelectrochemical

devices is the biocompatibility of the material. This term is usually used for the description of the implantable devices. The materials used in implants should be chemically and biologically inert, non-allergenic and non-toxic. Several studies have indicated that the PEDOT:PSS polymer is used in biomedical studies showing its good compatibility with the cells of different tissues. The review [15] emphasized a growing interest in the development of neural prosthetic devices from organic nanomaterials which was driven by recent advances in nanotechnology. An ideal material should have the properties that allow seamless binding and long lifetime. As a result, a great number of tests show that lots of nanomaterials, synthesized primarily for various purposes, are able to detect neural signals and stimulate the neurons. It is worthy of note that PEDOT is among the polymer materials, which exhibit the said properties. This polymer has extremely high conductivity, chemical stability in the oxidized state and great biocompatibility. The applications of polymers during exploration of tissue regeneration may give us an answer to questions regarding peculiarities of the organism response induced by the insertion of the stent (tensile elastic construction in the form of a cylindrical frame which is inserted in the clearance of hollow organs, providing expansion of the site reduced by pathological process). The scientists [16] explored the potential application of the PEDOT thin films for the stents with the aim to elucidate the extent of their effects on cell adhesion and proliferation, to determine their role in physical-chemical processes. As a matter of fact, this research focused on study of the stimulating effects of PEDOT. The authors demonstrated that all tested PEDOT films were cytocompatible,

contributed to adsorption of serum albumin and increased cell survival, and PEDOT:PSS promoted cell proliferation. Finally, the authors have concluded that the applications of nanomaterials in medicine may improve the relationship between medicine and bioelectronics; the use of PEDOT:PSS and PEDOT:TOS polymers (PEDOT doped with anions such as tosylate) stimulates regeneration processes during implantation of cardiovascular devices.

The work [17] reported the development of a dual component coating that could help stabilize a cochlear implant (a medical device which is intended to compensate for hearing loss). This coating is a combination of an arginine-glycine-aspartic acid, alginate hydrogel and the PEDOT polymer. The use of this coating significantly improved implant performance, lowered electrode impedance, improved charge delivery and locally enhanced the level of a trophic factor. In addition, the authors wrote that this coating is biocompatible with the organism tissues.

The studies [18-21] provide the results of the experiments on the development of the layer around nerve cells with the use of the conductive polymer PEDOT:PSS to obtain biomaterial with high electrical conductivity for direct electrical contact between electrically active cells such as heart cells, neurons, cells of the skeleton muscle. The authors described a technique which can be used for this procedure. It was found that nerve cells can perceive the concentration of the monomer EDOT (10 mM) for 72 hours retaining 80% of the survived cells. PEDOT could be delivered to the neuronal cells which are present in the tissue sample using 0.5-1.0 mA/mm² electrodes. Living cells kept viability in the polymer for 120 hours. The

"PEDOT-neuron" composite yielded lower impedance by 1-1.5 order of the magnitude; a significant increase in the capacity of charge delivery was also observed. The application of PEDOT for the development of these hybrid electrode layers is effective to prevent negative effects of short circuits at the electrode-tissue interface.

Meanwhile, several studies have indicated that the PEDOT:PSS films show antibacterial activity. For instance, in the work [22] the biohybrid film (BHF) was prepared by a combination of PEDOT, chitosan and polyvinyl alcohol. To obtain BHF a set of samples were selected by varying the concentrations of chitosan. Poly(vinyl alcohol) is the compound acting as a glue, thereby, improving mechanical properties such as strength of the film. The resultant biohybrid film presented high antibacterial activity in respect to Gram positive bacteria (*Staphylococcus aureus*). The researchers [23] reported a simple strategy for creating an efficient photo thermal nanocomposite based on PEDOT:PSS and agarose. This nanocomposite shows high antibacterial activity against both Gram-positive and Gram-negative bacteria. The exposure of the composition to IR light is related to photo thermal conversions, leading to efficient death of approximately 100% of pathogenic bacteria within 2 minutes. PEDOT with nanoparticles of fluoro hydroxyapatite, distributed uniformly on its matrix to cover orthopedic implants, has also antibacterial effects against Gram-positive and Gram-negative bacteria [24].

To improve the conductivity, PEDOT is frequently modified, as is shown above, by different nanomaterials. Notably, these compositions may have antimicrobial properties. For instance, in the work

[25], nanohybrid coatings on the basis of PEDOT with inclusions of the particles of graphene oxide and modified additionally with poly-diallyldimethylammonium chloride demonstrate (1) low roughness of the surface that prevents bacterial adhesion to surfaces and (2) positive charge that may be efficient in killing the bacteria. These properties are useful in biomedical devices to create cardiovascular stents and surgical devices.

Thus, an addition of different compounds into the PEDOT solution influences the properties in respect to different living cells. The question concerning biocompatibility of the polymer PEDOT:PSS relative to bacterial cells is yet to be answered and needs further exploration.

4. APPLICATIONS OF PEDOT:PSS FILMS IN BIOELECTROCHEMICAL SYSTEMS

4.1. APPLICATION OF PEDOT:PSS FILMS FOR THE FORMATION OF BIOSENSOR RECEPTORS

Let us consider examples of the use of PEDOT:PSS films for the creation of bioelectrochemical devices. Most frequently, publications focus on the use of PEDOT as a component for the amperometric enzyme-based biosensor. Researchers in [26] suggested a biosensor on the basis of glucose oxidase entrapped in PEDOT-based microcuvettes. A signal was recorded during detection of hydrogen peroxide resulting from the enzymatic reaction. A biosensor had high sensitivity with a rapid signal generation occurring within 20 s. This procedure is universal and can be applied to any enzyme.

The common concept of the field-effect transistor is based around the idea that it is constructed of ceramic and semiconductor materials coated with metal. The study [27] shows that it is possible to develop all-organic

field-effect transistors which include elements – channel, source, drain, gate – made from PEDOT:PSS. The combination of PEDOT:PSS and ferrocene, as a mediator, allowed measurement of glucose in the range from 0.001 to 100 mM. The biosensor was successfully applied to detect glucose levels in saliva. In [28], a similar transistor was used to determine both glucose and lactate. For immobilization of the enzymes, a simple, fast and reproduction procedure has been performed based on one-step electrochemical co-deposition of the glucose oxidase enzyme (GOx) or lactate dehydrogenase (LDH) and the particles of *Ni-Al* Layered double hydroxide (**Fig. 2**).

The studies [29,30] reported that graphene was utilized to modify PEDOT:PSS-based glucose biosensors, and now it is not necessary to use artificial mediators for electron transfer.

PEDOT:PSS is used not only as a component of glucose biosensors, but also in combination with other enzymes. For instance, the study [31] reported on a highly sensitive disposable amperometric biosensor

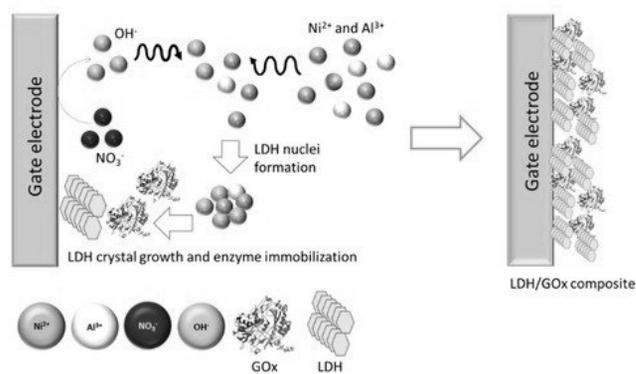


Fig. 2 Processes occurring at a PEDOT:PSS-based thin film gate electrode; nitrate recovery contributes to an increase in OH level leading to the deposition of the composite of a dual hydroxide/enzyme. Taken out from Gualandi et al, 2020 [28], under a Creative Commons Attribution 4.0 International License. Published by Multidisciplinary Digital Publishing Institute (MDPI).

for detecting triglycerides based on Au/PEDOT:PSS nanocomposite deposited by inkjet-printing on screen-printed graphite electrodes and co-immobilization of lipase, glycerol kinase and glycerol-3-phosphate oxidase. The designed biosensor had a wide detection limit (0–531 mg/dl), its sensitivity was $2.66 \mu\text{A}/\text{mM}$, a response time was 30 s, and detection limit was as low as 7.88 mg/dL. The work [32] presented an amperometric biosensor based on PEDOT and multi-wall carbon nanotubes for detection of organophosphorus compounds using the enzyme acetylcholinesterase. Detection limit of malathion was 1fM in linear range from 1 fM to $1 \mu\text{M}$. The study [33] described the formation of a biosensor based on the principle of electrochemiluminescence (ECL) imaging for ethyl alcohol identification. The biosensor was based on the enzyme (alcohol dehydrogenase) immobilized in the matrix from PEDOT:PSS and graphene. Biosensor was applied to detect the concentration of ethanol in real beverages; detection limit was $2.5 \cdot 10^{-6} \text{ M}$.

Nafion was used to prepare biocompatible medium for biologically active compounds, to improve adhesion and binding force between the PEDOT:PSS film and the electrode interface, as well as to prevent exfoliation of enzymatic molecules for PEDOT modification in [34]. Ascorbate oxidase was used as a basis of the biosensor. PEDOT modification had pronounced water resistance, preventing sensitive element to be exfoliated from the electrode that is important while carrying out experiments in aqueous media. The biosensor (sensitivity of $158.1 \mu\text{A}/\text{mM}$ and detection limit of $0.193 \mu\text{M}$) was applied to determine the ascorbic acid content of juices.

PEDOT:PSS is applied not only in enzymatic biosensors, but in the other types of sensors as well, for example, immunosensors. An example of the use of PEDOT:PSS, as a component of immunosensors, is an amperometric device which measures the content of clenbuterol, medical preparation, stimulating the growth in milk [35]. The authors used a competitive scheme of analysis, where horseradish peroxidase enzyme was the label of an antigen. The detection limit for clenbuterol was 0.196 ng/ml. The article [36] described the aptasensor based on electrochemical impedance spectroscopy. The paper electrode was modified with PEDOT:PSS and graphene to provide conductivity and sensitivity, the surface was then functionalized with amino- and carboxyl groups, and then with aptamers. This biosensor was sensitive to carcinoembryonal antigens in the linear range from 0.77 to $14 \text{ ng}/\text{ml}^{-1}$; it would be able to assist in the diagnosis of cancer at an early stage.

Also, the use of matrices based on PEDOT seems to be effective for DNA exploration. Electrochemical DNA biosensors have advantages over conventional techniques due to rapid, precise, highly sensitive and selective signal responses. The work [37] demonstrated that the graphite electrode modified with chitosan and PEDOT was applied to electrochemically detect interaction of DNA with anticancer medical preparation (mitomycin C is an antitumor antibiotics and a strong agent, which is able to crosslink DNA).

4.2. APPLICATIONS IN BIOFUEL CELLS

The use of PEDOT:PSS polymer in biofuel cells (BFC) would facilitate electron transport in the system and increase the overall

efficiency of these devices. So far, only a few variants of BFCs based on PEDOT are described in literature. For example, the study [38] reported that the anode for BFC was produced by successive deposition of layers of carbon, para-benzoquinone, glucose oxidase, and PEDOT. Such an electrode could work both as a glucose biosensor and as a BFC anode. The maximum specific power of such a device depended on the temperature and was equal to 18.9 and 22.5 $\mu\text{W}/\text{cm}^2$ at 25° and 37°C, respectively. PEDOT- and graphene-based anode served as the basis for the immobilization of *Escherichia coli* [39]. A compact biofilm was formed on the hybrid anode due to the electrostatic interaction between negatively charged bacteria and positively charged PEDOT. The maximum power generated by such a BFC was 873 mW/m². Researchers [40] suggested to develop a PEDOT:PSS - and sulphonated graphene oxide-based biocompatible BFC anode. Glucose oxidase was used as a biocatalyst. It was shown that similar structure of the anode can be used both in glucose biosensor and in BFC whereas the element generates a high power density of $27 \pm 2 \text{ mA} \cdot \text{cm}^{-2}$.

In [41], the creation of miniature BFCs that have higher output of electrical energy and shorter time to reach steady state (about 1 hour) is reviewed. Appliance of PEDOT polymer and an anode chamber with a volume of 12 μl is the distinctive feature of such BFC. BFC develops a power of about $423 \text{ mW} \times \text{cm}^{-3}$ and uses unadapted *Shewanella oneidensis* MR-1 bacteria as a model biocatalyst. The paper [42] presents BFCs based on microbial communities of activated sludge, which anodes are covered with chemically polymerized PEDOT films. The use of polymer made it possible to achieve

power of 3.5 A/m² and increase the degree of wastewater treatment from 51% to 86%.

PEDOT can be used not only for bioanode modification, but also in BFC cathodes. Researchers [43] demonstrated that PEDOT was used to immobilize laccase at the BFC cathode. It was shown that electropolymerization of PEDOT with various alloying additives significantly affects the structural features and morphology of PEDOT films, increasing the active surface of the electrode, and guarantees effective fuel mass transfer through the matrix and, as a result it is a crucial step that determines the cell capacity.

The use of PEDOT enables the creation of BFCs based on flexible electrodes. In [44] PEDOT:PSS was the basis for microbial fuel cell (MFC) made of flexible stretchable fabric. The flexible *Pseudomonas aeruginosa*-based MFC generated the maximum power and current density of 1.0 $\mu\text{W} \times \text{cm}^2$ and 6.3 $\mu\text{A}/\text{cm}^2$, respectively, which is comparable to or exceeds similar parameters of a flexible paper-based BFC. Such a device can be easily integrated into the next-generation of flexible electronics to implement low-power and autonomous systems. Such flexible electrodes can be used to create wearable biosensors and BFCs for monitoring human health. Researchers [45] designed a wearable BFC based on *Staphylococcus epidermidis*, *Staphylococcus capitis* and *Micrococcus luteus* that exist on human skin and eat sweat as a substrate. Prussian blue was used as a mediator of electronic transport, and PEDOT:PSS polymer and dimethylsulfoxide were used to ensure the conductivity and fixation of bacteria, as well as to collect electricity.

Biological fluids can be the basis of power for biofuel cells, which in its turn can be a power source for low-power

devices that monitor human health (for example, continuous glucose monitoring) or perform functions that support human life (for example, cardiac pacemaker, neurostimulators, etc.). For example, the study [46] reported that human urine served as a power source for BFC. The BFC anode was a carbon film coated with PEDOT:PSS. The electrodes were inoculated with a mixture of sludge and pure human urine in a periodic mode. The BFC maintained its functionality for more than 90 days generating average stable power of 283.5 μ W. Moreover, it was shown that PEDOT:PSS not only improves the electrochemical properties of the electrodes, but also contributes to the growth of the biofilm and, consequently, increases the overall energy characteristics and long-term functionality of the BFC.

5. CONCLUSION

In the present study, we show that the main property of the PEDOT:PSS polymer is low electrical resistance and the majority of the components involved in conjugation with this polymer get it. One of the prior sphere of practical application of PEDOT:PSS is medicine. It is, therefore, not surprisingly, that there is increasing attention paid to the combination of PEDOT:PSS and glucose oxidase – an enzyme used to create biosensors for glucose detection. Notably, combinations of PEDOT:PSS and nanomaterials significantly increase the detection sensitivity and enhance the stability of the electrodes. Applications of different modifications of PEDOT allows for the development of flexible bioelectrochemical devices that can be recommended both for medical and other industrial areas.

The question is still unanswered with regard to the biocompatibility and toxicity

of PEDOT. The effects of PEDOT:PSS on neurons in the human brain have been explored and the biocompatibility of the polymer in the cells of the human brain have been shown. Additionally, PEDOT:PSS is used for fabrication of films with high antibacterial activity against *Staphylococcus aureus* and *Escherichia coli*.

On the whole, the overall direction to the use of the properties of the PEDOT:PSS polymer aims to improve the quality of life, increase the production efficiency while monitoring the environmental indicators, in other words we give the outlook of this polymer for possible applications in easy-to-follow and positively oriented trends.

REFERENCES

1. Hecht DS, Hu LB, Irvin G. Emerging transparent electrodes based on thin films of carbon nanotubes, graphene, and metallic nanostructures. *Adv. Mater.*, 2013, 23:1482-1513.
2. Chin BD. Role of the polymeric hole injection layer on the efficiency and stability of organic light emitting diodes with small molecular emitters. *J. Phys. D: Appl. Phys.*, 2008, 41(21):215104. doi: 10.1088/0022-3727/41/21/215104.
3. Scharber MC, Muehler D, Koppe M, Denk P, Waldauf C, Heeger AJ, Brabec CJ. Design Rules for Donors in Bulk Heterojunction Solar Cells-Towards 10% Energy Conversion Efficiency. *Adv. Mater.*, 2006, 18:789-794.
4. Scott JC, Bozano LD. Nonvolatile Memory Elements Based on Organic Material. *Adv. Mater.*, 2007, 19(11):1452-1463.
5. Pettersson LA, Ghosh S, Inganäs O. Optical anisotropy in thin films of poly(3,4-ethylenedioxythiophene)-poly(4-styrenesulfonate). *Org. Electron.*, 2002, 3(3-4):143-148.
6. Kukhto AV, Pochtenny AE, Misevich AV, Kukhto IN, Semenova EM, Vorobyova SA,

- Sarantopoulou E. Optical and electrophysical properties of nanocomposites based on PEDOT:PSS and gold/silver nanoparticles. *Phys. Solid State*, 2014, 56(4):827-834.
7. Park HS, Ko SJ, Park JS, Kim JY, Song HK. Redox-active charge carriers of conducting polymers as tuner of conductivity and its potential window. *Sci. Rep.*, 2013, 3(2454):1-6.
 8. Yang W, Zhao Y, He X, Chen Y, Xu J, Li S, Jiang Y. Flexible conducting polymer/reduced graphene oxide films: synthesis, characterization, and electrochemical performance. *Nanoscale Res. Lett.*, 2015, 10(1). doi: 10.1186/s11671-015-0932-1.
 9. Voronin AS, Simunin MM, Ivanchenko FS, Shiverskii AV, Fadeev YV, Tambasov IA, Nemtsev IV, Matsynin AA, Khartov SV. Preparation and investigation of composite transparent electrodes of poly(3,4-ethylenedioxythiophene) polystyrene sulfonate/single-wall carbon nanotubes. *Tech. Phys. Lett.*, 2017, 43(9):783-786.
 10. Ventura IA, Zhou J, Lubineau G. Investigating the Inter-Tube Conduction Mechanism in Polycarbonate Nanocomposites Prepared with Conductive Polymer-Coated Carbon Nanotubes. *Nanoscale Res. Lett.*, 2015, 10:485.
 11. Yan H, Zhang P, Li J, Zhao XL, Zhang K, Zhang B. PEDOT/PSS-Halloysite Nanotubes (HNTs) Hybrid Films: Insulating HNTs Enhance Conductivity of the PEDOT/PSS Films. *Sci. Rep.*, 2015, 5:18641.
 12. Nizhegorodova AO, Kondratiev VV. Synthesis and electrochemical properties of composite materials based on poly-3,4-ethylenedioxythiophene with manganese dioxide inclusions. *Russ. J. Electrochem.*, 2014, 50(12):1157-1163.
 13. Nizhegorodova AO, Apraksin RV, Kondratiev VV. Electrochemical properties of composite materials based on poly-3,4-Ethylenedioxythiophene with nickel oxide inclusions. *Russ. J. Electrochem.*, 2015, 51(10):908-915.
 14. Park MU, Lee SM, Chung D. Model system of cross-linked PEDOT:PSS adaptable to an application for an electrode with enhanced water stability. *Synth. Met.*, 2019, 258:116195.
 15. Fattahi P, Yang G, Kim G, Abidian MR. A Review of Organic and Inorganic Biomaterials for Neural Interfaces. *Adv Mater*, 2014, 26(12):1846-1885.
 16. Karagkiozaki V, Karagiannidis PG, Gioti M, Kavatzikidou P, Georgiou D, Georgeraki E, Logothetidis S. Bioelectronics meets nanomedicine for cardiovascular implants: PEDOT-based nanocoatings for tissue regeneration. *Biochim. et Biophysica Acta*, 2013, 1830:4294-4304.
 17. Chikar JA, Hendricks JL, Richardson-Burns SM, Raphael Y, Pfingst BE, Martin DC. The use of a dual PEDOT and RGD-functionalized alginate hydrogel coating to provide sustained drug delivery and improved cochlear implant function. *Biomaterials*, 2012, 33:1982-1990.
 18. Richardson-Burns SM, Hendricks JL, Foster B, Povlich LK, Kim DH, Martin DC. Polymerization of the conducting polymer poly(3,4-ethylenedioxythiophene) (PEDOT) around living neural cells. *Biomaterials*, 2007, 28:1539-1552.
 19. Venkatraman S, Hendricks JL, Richardson-Burns SM, Jan E, Martin D, Carmena JM. PEDOT coated microelectrode arrays for chronic neural recording and stimulation. 2009, *4th International IEEE/EMBS Conference on Neural Engineering*. doi: 10.1109/ner.2009.5109313.
 20. Wilks S. Poly(3,4-ethylene dioxythiophene) (PEDOT) as a micro-neural interface material for electrostimulation. *Front Neuroeng.*, 2009, 2. doi: 10.3389/neuro.16.007.2009.
 21. Richardson-Burns SM, Hendricks JL, Martin DC. Electrochemical polymerization of conducting polymers in living neural tissue. *J. Neural Eng.*, 2007, 4(2):6-13.
 22. Khan S, Narula AK. Bio-hybrid blended transparent and conductive films

- PEDOT:PSS:Chitosan exhibiting electroactive and antibacterial properties. *Eur. Polym. J.*, 2016, 81:161-172.
23. Ko Y, Kim J, Jeong HY, Goomin Kwon, Kim D, Ku M, Yang J, Yamauchi Y, Kim HY, Lee C, You J. Antibacterial poly(3,4-ethylenedioxythiophene):poly(styrene-sulfonate)/agarose nanocomposite hydrogels with thermoprocessability and selfhealing. *Carbohydr. Polym.*, 2019, 203:26-34.
24. Madhan Kumar A, Adesina AY, Hussein MA, Ramakrishna S, Al-Aqeeli N, Aktar S, Saravanan S. PEDOT/FHA nanocomposite coatings on newly developed Ti-Nb-Zr implants: Biocompatibility and surface protection against corrosion and bacterial infections. *Mater. Sci. Eng. C*, 2019, doi: 10.1016/j.msec.2019.01.012.
25. Hsu CC, Cheng YW, Liu CC, Peng XY, Yung MC, Liu TY. Anti-Bacterial and Anti-Fouling Capabilities of Poly(3,4-Ethylenedioxythiophene) Derivative Nanohybrid Coatings on SUS316L Stainless Steel by Electrochemical Polymerization. *Polymers*, 2020, 12(7):1467; doi: 10.3390/polym12071467.
26. Park J, Kim HK, Son Y. Glucose biosensor constructed from capped conducting microtubules of PEDOT. *Sens Actuators B Chem*, 2008, 133:244-250.
27. Shim NY, Bernards DA, Macaya DJ, DeFranco JA, Róisín MN, Owens M, Malliaras GG. All-Plastic Electrochemical Transistor for Glucose Sensing Using a Ferrocene Mediator. *Sensors*, 2009, 9:9896-9902.
28. Gualandi I, Tessarolo M, Mariani F, Arcangeli D, Possanzini L, Tonelli D, Scavetta E. Layered Double Hydroxide-Modified Organic Electrochemical Transistor for Glucose and Lactate Biosensing. *Sensors*, 2020, 20(12):3453; doi: 10.3390/s20123453.
29. Wisitsoraat A, Pakapongpan S, Sriprachuabwong C, Phokharatkul D, Sritongkham P, Lomas T, Tuantranont A. Graphene-PEDOT:PSS on screen printed carbon electrode for enzymatic biosensing. *J. Electroanal. Chem.*, 2013, 704:208-213.
30. David M, Barsan MM, Brett CMA, Monica Florescu. Improved glucose label-free biosensor with layer-by-layer architecture and conducting polymer poly(3,4-ethylenedioxythiophene). *Sens Actuators B Chem.*, 2018, 255 (3):3227-3234.
31. Phongphut A, Sriprachuabwong C, Wisitsoraat A, Tuantranont A, Prichanont S, Sritongkham P. A disposable amperometric biosensor based on inkjet-printed Au/PEDOT-PSS nanocomposite for triglyceride determination. *Sens Actuators B Chem*, 2013, 178:501-507.
32. Kaur N, Thakur H, Prabhakar N. Conducting polymer and multi-walled carbon nanotubes nanocomposites based amperometric biosensor for detection of organophosphate. *J. Electroanal. Chem.*, 2016, 775:121-128.
33. Gao Y, Li J, Yang X, Xiang Q, Wang K. Electrochemiluminescence Biosensor Based on PEDOT:PSS-Graphene Functionalized ITO Electrode. *Electroanalysis*, 2014, 26:382-388.
34. Wen Y, Xu J, Li D, Liu M, Kong F, He H. A novel electrochemical biosensing platform based on poly(3,4-ethylenedioxythiophene):poly(styrenesulfonate) composites. *Synth. Met.*, 2012, 162:1308-1314.
35. Talib NAA, Salam F, Sulaiman Y. Development of highly sensitive immunosensor for Clenbuterol detection by using poly(3,4-ethylenedioxythiophene)/graphene oxide modified screen-printed carbon electrode. *Sensors*, 2018, 18:4324-4336.
36. Yen YK, Chao CH, Yeh YS. A Graphene-PEDOT:PSS Modified Paper-Based Aptasensor for Electrochemical Impedance Spectroscopy Detection of Tumor Marker. *Sensors*, 2020, 20:1372.
37. Kuralay F, Demirci S, Kiristi M, Oksuz L, Oksuz AU. Poly(3,4-ethylenedioxythiophene) coated chitosan modified disposable electrodes for DNA and DNA-drug

- interaction sensing. *Colloids Surf B Biointerfaces*, 2014, 123: 825-830.
38. Nien PC, Wang JY, Chen PY, Chen LC, Ho KC. Encapsulating benzoquinone and glucose oxidase with a PEDOT film: Application to oxygen-independent glucose sensors and glucose/O₂ biofuel cells. *Bioresour. Technol.*, 2010, 101(14):5480-5486.
39. Wang Y, Zhao C, Sun D, Zhang JR, Zhu JJ. A Graphene/Poly(3,4-ethylenedioxythiophene) Hybrid as an Anode for High-Performance Microbial Fuel Cells. *ChemPlusChem*, 2013, 78:823-829.
40. Inamuddin B, Ahamed MI, Asiri AM, AlAmry KA. Biocompatible mediated bioanode prepared by using poly(3,4-ethylene dioxy thiophene) poly(styrene sulfonate) (PEDOT:PSS) and sulfonated graphene oxide integrated enzyme for biofuel cells applications. **Mat. Sci. for En. Tech.**, 2018, 1(1):63-69.
41. Jiang H, Halverson LJ, Dong L. A miniature microbial fuel cell with conducting nanofibers-based 3D porous biofilm. *J. Micromech. Microeng.*, 2015, 25:125017.
42. Kang YL, Ibrahim S, Pichiah S. Synergetic effect of conductive polymer poly(3,4-ethylenedioxythiophene) with different structural configuration of anode for microbial fuel cell application. *Bioresour. Technol.*, 2015, 189:364-369.
43. Wang X, Sjöberg-Eerola P, Eriksson JE, Bobacka J, Bergelin M. The effect of counter ions and substrate material on the growth and morphology of poly(3,4-ethylenedioxythiophene) films: Towards the application of enzyme electrode construction in biofuel cells. *Synth. Met.*, 2010, 160(13-14):1373-1381.
44. Pang S, Gao Y, Choi S. Flexible and stretchable microbial fuel cells with modified conductive and hydrophilic textile. *Biosens. Bioelectron.*, 2018, 100:504-511.
45. Mohammadifar M, Tahernia M, Yang JH, Koh A, Choi S. Biopower-on-Skin: Electricity generation from sweat-eating bacteria for self-powered E-Skins. *Nano Energy*, 2020, 104994; doi: 10.1016/j.nanoen.2020.104994.
46. Salar-Garcia MJ, Montilla F, Quijada C, Morallon E, Ieropoulos I. Improving the power performance of urine-fed microbial fuel cells using PEDOT-PSS modified anodes. *Appl. Energy*, 2020, 278:115528; doi: 10.1016/j.apenergy.2020.115528.

DOI: 10.17725/rensit.2020.12.483

Interferometric Processing of Acoustic Information by Using Extended Antennas in Dispersing Media

Venedikt M. Kuz'kin

A.M. Prokhorov General Physics Institute, Russian Academy of Sciences, <http://www.gpi.ru/>
Moscow 119991, Russian Federation

E-mail: kumiov@yandex.ru

Sergey A. Pereselkov, Sergey A. Tkachenko, Ilia V. Kaznacheev

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

Voronezh 394006, Russian Federation

E-mail: pereselkov@yandex.ru; sega-tk@mail.ru; kaznacheev.ilya@gmail.com

Received May 05, 2020, peer-reviewed May 26, 2020, accepted June 01, 2020

Abstract: The theory of interferometric processing of acoustic information by using extended antennas in media with frequency dispersion is presented. The dependence between the two-dimensional spectral density of the two-fold Fourier transform of the interference pattern formed by a moving noise source and the aperture and angular structure of the received field is analyzed. The gain factor, directivity characteristic, and noise immunity of processing are estimated. Depending between the input signal/noise ratio on the antenna element and the maximal range of the noise source is obtained. This maximal range allows stable detection and estimation of direction, radial velocity, range and depth are close to real values. Numerical simulation results are presented and discussed.

Keywords: dispersion, interferogram, hologram, noise source, extended antennas, gain, directional characteristic, noise immunity, limiting offset, numerical simulation

UDC 004.052.34

Acknowledgements: The work was supported by grants from the Russian Foundation for Basic Research (projects No. 19-08-00941, No. 19-29-06075). The work of I.V. Kaznacheev was supported by the grant of the President of the Russian Federation MK-933.2019.8.

For citation: Venedikt M. Kuz'kin, Sergey A. Pereselkov, Sergey A. Tkachenko, Ilia V. Kaznacheev. Interferometric Processing of Acoustic Information by Using Extended Antennas in Dispersing Media. *RENSIT*, 2020, 12(4)483-494. DOI: 10.17725/rensit.2020.12.483.

CONTENTS

1. INTRODUCTION (483)
 2. ALGORITHM OF SIGNAL PROCESSING (484)
 - 2.1 HORIZONTAL LINEAR ANTENNA (484)
 - 2.2 VERTICAL LINEAR ANTENNA (488)
 3. NOISE IMMUNITY OF SIGNAL PROCESSING (488)
 4. SIMULATION RESULT (489)
 5. CONCLUSION (493)
- REFERENCES (494)

1. INTRODUCTION

A typical particular case of media with frequency dispersion is an oceanic waveguide with pronounced waveguide dispersion, on which we will focus our attention. One of the poorly developed problems in the processing of hydroacoustic information is to provide noise immunity in the conditions of spatial and temporal variability of the oceanic environment. Difficulties in approaching its solution arise

from the a priori uncertainty of knowledge about the propagation medium (including random), as well as the weak value of the useful signal against the background of intense localized interference and surrounding noise. It is this situation that is of practical interest. In this regard, an urgent task seems to be the development of processing, which combined high noise immunity with adaptability to changing conditions of wave field propagation, i.e. with a decrease in the requirements for the volume of a priori information about the transfer function of the waveguide.

Theoretical studies, computational and field experiments have shown that such a combination can be provided by information technology for processing an interference pattern (interferogram) [1–5]. It is based on the formation mechanism of the interferogram of a broadband source caused by waveguide dispersion and multimode propagation [6]. Interferometric processing implements the coherent accumulation of spectral intensity along localized bands, which is then subjected to a 2D Fourier transformation. The converted spectral density (which we will conventionally call a hologram) is concentrated in a small area, providing high noise immunity.

At present, most of the main provisions of the theory, practical issues and potential possibilities of interferometric processing using single vector-scalar receivers have been sufficiently worked out. This made it possible to radically solve the problem of localization of low-noise sound sources [1–5] and outline the areas of such new applications as restoration of the transfer function of an unperturbed waveguide and diagnostics of hydrodynamic disturbances

[7,8]. If the current development of interferometric processing with the use of single vector-scalar receivers undoubtedly confirms its fruitfulness, then the questions remain, what results can be obtained by applying it in relation to extended receiving antennas.

This article is devoted to this issue, in which the results of the theory [1,2] are generalized for the case of using horizontal and vertical receiving linear antennas. The results of numerical simulation are presented.

2. ALGORITHM OF SIGNAL PROCESSING

The waveguide is assumed to be horizontally uniform. Let the number of elements receiving antenna is B , $b = \overline{1, B}$, interelement distance – d . The fields from each element of the antenna are summed, and an interferogram is formed at the output, to which a 2D Fourier transformation is applied.

2.1. HORIZONTAL LINEAR ANTENNA

Let's designate the distance from the element Q_b to the source S as r_b (Fig. 1). Let us choose the first element as a reference Q_1 . The aperture $L = (B - 1)d$ is assumed to be much less than the distance to the source, $L \ll r_b$, then $r_b = r_1 - (b - 1)d \sin \theta$, where θ – the bearing. Antenna compensation is provided in the direction

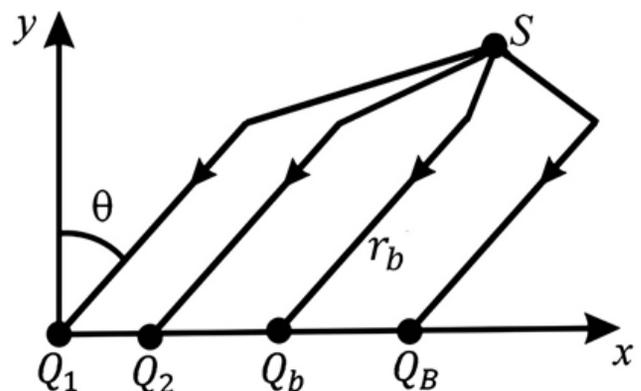


Fig. 1. Problem geometry.

of the angle θ_* . Since we have to compensate for the difference in the distances from the source to the various elements of the antenna, we multiply the field of the b -th element by $\exp[ih_*(\omega_0)(b-1)d\sin\theta_*]$, where $h_*(\omega_0)$ – where is the selected horizontal wavenumber at the middle frequency of the source spectrum $\omega_0 = 2\pi f_0$.

The field at the output of the b -th element is written as a sum of modes [9]

$$p_b(\omega, r_b) = \sum_m A_m(\omega, r_b) \exp\{i[h_m(\omega)r_1 - 2(b-1)(h_m(\omega) - h_*(\omega_0)\eta_*)]\}, \quad (1)$$

where

$$\eta = d\sin\theta/2, \quad \eta_* = d\sin\theta_*/2. \quad (2)$$

Here A_m and h_m – are the amplitude and horizontal wavenumber of the m -mode. Cylindrical field divergence, modal attenuation, source z_s and z_b antenna element depths are formally taken into account by the amplitude dependence of the modes. At the antenna output $p_{an}(\omega, r_1)$, the field, neglecting the dependence of the amplitude on the distance $A_m(\omega, r_b) \approx A_m(\omega, r_1)$, after simple transformations can be represented as

$$p_{an}(\omega, r_1) = \sum_m A_m(\omega, r_1) \exp\{i[h_m(\omega)r_1 - (B-1)(h_m(\omega)\eta - h_*(\omega_0)\eta_*)]\} I_m, \quad (3)$$

$$I_m = \frac{\sin[B(h_m(\omega)\eta - h_*(\omega_0)\eta_*)]}{\sin[(h_m(\omega)\eta - h_*(\omega_0)\eta_*)]}. \quad (4)$$

The antenna interferogram $P_{an}(\omega, r_1) = |p_{an}(\omega, r_1)|^2$, according to (3), is

$$P_{an}(\omega, r_1) = \sum_m \sum_n P_{mn}^{(an)}(\omega, r_1), \quad (5)$$

where

$$P_{mn}^{(an)}(\omega, r_1) = A_m(\omega, r_1) A_n^*(\omega, r_1) \exp[ih_{mn}(\omega)(r_1 - (B-1)\eta)] I_{mn}, \quad (6)$$

$$I_{mn} = I_m I_n.$$

Here $h_{mn}(\omega) = h_m(\omega) - h_n(\omega)$. Consider the case of a moving source with a constant radial velocity w (the projection of the velocity in

the direction to the antenna). We assume that the distance r_1 corresponds to the initial moment of time $t_0 = 0$. Further, in the interferogram (5), we pass from the distance r_1 variable to the time variable t and carry out a 2D Fourier transformation. At the output of the integral transformation, the spectral density is determined by the expression

$$F_{an}(\tilde{\nu}, \tau) = \int_0^{\Delta t} \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} P_{an}(\omega, t) \exp[i(\tilde{\nu}t - \omega\tau)] dt d\omega = \sum_m \sum_n F_{mn}(\tilde{\nu}, \tau), \quad (7)$$

where $\tilde{\nu} = 2\pi\nu$ and τ – are the frequency and time of the hologram, Δt and $\Delta\omega$ – are the observation time and spectrum width. Using the approach for obtaining a hologram of a single receiver [1], we obtain

$$F_{mn}(\tilde{\nu}, \tau) = A_m(\omega_0, r_1) A_n^*(\omega_0, r_1) I_{mn}(\omega_0, B, \eta, \eta_*) w \Delta\omega \Delta t \times \exp\left[i\left(\frac{\tilde{\nu}\Delta t}{2} - \tau\omega_0\right)\right] \times \exp\left\{i\left[(m-n)\alpha\left(\frac{\Delta t}{2}w + r_1 - (B-1)\eta\right) + (r_1 - (B-1)\eta)(\tilde{\nu}/w)\right]\right\} \times \frac{\sin\left\{\left[r_1 - (B-1)\eta + w\tau\right](m-n)\frac{d\alpha}{d\omega} - \tau\right\} \frac{\Delta\omega}{2}}{\left[r_1 - (B-1)\eta + w\tau\right](m-n)\frac{d\alpha}{d\omega} - \tau} \frac{\Delta\omega}{2} \frac{\sin\left\{[w(m-n)\alpha + \tilde{\nu}]\frac{\Delta t}{2}\right\}}{[w(m-n)\alpha + \tilde{\nu}]\frac{\Delta t}{2}}. \quad (8)$$

Here $\alpha = dh(\omega_0)/dl = h_{l+1}(\omega_0) - h_l(\omega_0)$, l – is the number of the mode, in the vicinity of which the modes are in phase, t_* the selected time moment in the observation interval Δt , $0 < t_* < \Delta t$. The introduction of the quantity is useful in interpreting a hologram. In fact $\alpha(m-n) = h_{mn}(\omega_0)$, $(d\alpha/d\omega)(m-n) = dh_{mn}(\omega_0)/d\omega$. If we put $B = 1$, then $I_{mn} = 1$, and relation (8) turns into an expression for a single receiver [1].

As in the case of a single receiver, the spectral density of the antenna hologram is localized on a plane $(\tau, \tilde{\nu})$ in a narrow strip in the form of focal spots, mirror-inverted relative to the origin. This feature is due to the symmetry of function (8) with respect to the permutation of the numbers of interfering

modes: $F_{mn}(\tilde{\nu}, \tau) = F_{nm}(-\tilde{\nu}, -\tau)$. Focal spots are located in the first and third quadrants if the radial velocity $w < 0$ (the source is approaching the antenna) and in the second and fourth quadrants when the source is moving away from the antenna ($w > 0$). The localization region contains $(M - 1)$ the main maxima with coordinates $(\tau_\mu, \tilde{\nu}_\mu)$, where $M - 1$ is the number of modes that form the field, $\mu = \overline{1, M - 1}$ is the number of the focal spot. The peak closest to the origin is due to the interference of neighboring modes and is located at a point $(\tau_1, \tilde{\nu}_1)$. The coordinates of the adjacent peak caused by the interference of the mode numbers, $(m, m + 2)$, $-(\tau_2, \tilde{\nu}_2)$. Coordinates of the most distant peak from the origin, due to the interference of the first and last modes $(\tau_{M-1}, \tilde{\nu}_{M-1})$. The $(M - \mu)$ main peaks are summed at the points with coordinates $(\tau_\mu, \tilde{\nu}_\mu)$.

The coordinates of the main maximums $(\tau_\mu, \tilde{\nu}_\mu)$ are located on a straight line $\tilde{\nu} = \tilde{\varepsilon}\tau$ with a slope factor

$$\tilde{\varepsilon} = -\frac{wh_{mn}(\omega_0)}{r_1(dh_{mn}(\omega_0)/d\omega)}, \quad (9)$$

and occupy a band between the values $\tau = \pm 2\pi\Delta\omega$ and $\tilde{\nu} = \pm 2\pi/\Delta t$. Outside these bands, the spectral density of the hologram is practically suppressed. The angular coefficients of the interference fringes of the interferogram and the line of the location of the main maxima of the spectral density are related by the relation $\tilde{\varepsilon} = -\Delta\Omega/\Delta t$, where $\Delta\Omega$ is the frequency shift of the interference maxima of the wave field over time. The frequency shift expresses the condition for maintaining the phase between constructively interfering modes caused by a change in the distance between the source and the antenna [6].

Under the condition $r_1 \gg |(B - 1)\eta - wt_*|$, as follows from (8), the estimates of the distance and radial velocity of the source

$$\dot{r}_1 = \kappa_{r\mu}\tau_\mu, \quad \dot{w} = -\kappa_{w\mu}\tilde{\nu}_\mu \quad (10)$$

coincide with expressions for a single receiver [1]. Here

$$\begin{aligned} \kappa_{r\mu} &= |\overline{dh_{m(m+\mu)}(\omega_0)/d\omega}|, \\ \kappa_{w\mu} &= |\overline{h_{m(m+\mu)}(\omega_0)}|^{-1} \end{aligned} \quad (11)$$

– coefficients that determine the frequency and spatial scales of variability of the transfer function of the waveguide [9]. The source parameter estimates obtained as a result of measurements, in contrast to their true values, are indicated by a dot above. The bar above denotes averaging over mode numbers. Adaptive methods of interferometric processing that allow estimating the source distance and radial velocity in the absence of information about the transfer function, i.e. without knowing the coefficients (11), are presented in [3].

Factor

$$I_{mn} = I_m(\omega_0, B, \eta, \eta_*)I_n(\omega_0, B, \eta, \eta_*), \quad (12)$$

determined from formulas (4), (6), characterizes the distribution of the spectral density of the antenna hologram in relation to the hologram of a single receiver. Peculiarities of dependence (12) on bearing θ will be considered by the example of the compensation angle $\theta_* = 0$. The main maxima of the functions $I_{m,n}$, $\max I_{m,n} = B$, correspond to the values

$$\sin \theta = \pm 2k\pi \frac{1}{h_{m,n}(\omega_0)d}, \quad (13)$$

where $k = 0, 1, \dots$ – is the order of the spectrum. In the direction of bearing $\theta = 0$, $k = 0$, a coherent addition of complex

amplitudes of all mode numbers is carried out; $k = 1, 2, \dots$ the position of the maxima depends on the mode numbers. The zeros of the function $I_{m,n}$ correspond to the values of the angles determined by the expression

$$\sin \theta = \pm 2j\pi \frac{1}{Bh_{m,n}(\omega_0)d}, \quad (14)$$

where j – are integers, except for $B, 2B, \dots$. Zeros are located about B times more often than the main highs. The secondary maxima $I_{m,n}$ lie approximately in the middle between two adjacent zeros, i.e. fall on values

$$\sin \theta \cong \pm(2g+1)\pi \frac{1}{Bh_{m,n}(\omega_0)d}, \quad (15)$$

where g – are integers. Secondary maximum levels do not exceed 0.2 of the B value [10].

Thus, the positions of the main highs do not depend on the number of elements B , between each two main maximums there are $(B - 1)$ zeros and secondary maximums $(B - 2)$. As the parameter $Bd\omega_0$ increases, i.e. the larger the aperture and the higher the frequency range, as follows from (14), the sharpness of the main maxima increases (the width decreases) and the number of bearings at which the spectral density is equal to zero increases. The distance between the major maxima for a particular wavelength is determined by the inter-element distance.

We put the interelement distance, $d = \lambda(n + 1)/2$, where λ is the wavelength for which we will take the value $\lambda = 2\pi/h_*(\omega_0)$. Then, as follows from (13), if $n = 0$, then only the main maximum of order zero exists ($\theta = 0, \pi$), if $n = 1$ – two main maxima of the zero and first orders ($\theta = 0, \pi/2, \pi, 3\pi/2$), etc. The number of orders k is determined by the condition $2k \leq n + 1$. With an increase in the interelement distance and a decrease in the wavelength, the number of main maxima increases. The

width of the main maxima, according to (14), does not depend on their spectrum order and is equal to $\Delta\theta = \lambda/Bd$. By changing the frequency range (at a given inter-element distance), you can adjust the positions of the main maxima. The multimode propagation regime, obviously, leads to the "blurring" of the zeros of the spectral density of the hologram, a shift in the positions of the main maxima (for angles $\theta \neq 0$) determined by relation (13), and an increase in their width. The introduction of the compensation angle θ_* leads to the appearance of a term in the right-hand sides of relations (13)-(15) $h_*(\omega_0) \sin\theta_*/h_{m,n}(\omega_0)$.

With respect to a single receiver, the effectiveness of interferometric processing using an antenna is characterized by the gain

$$\chi = |G_{an}(B, \theta, \theta_*)| / |G_r|, \quad (16)$$

where

$$G_{an}(B, \theta, \theta_*) = \int \int_U |F_{an}(\tau, \tilde{\nu})| d\tau d\tilde{\nu}, \quad (17)$$

$$G_r = \int \int_U |F_r(\tau, \tilde{\nu})| d\tau d\tilde{\nu} \quad (18)$$

and antenna directivity

$$D(B, \theta, \theta_*) = G_{an}(B, \theta, \theta_*) / \max G_{an}. \quad (19)$$

Here, the subscript « n » refers to a single receiver, U – is the spectral density localization area.

From (8), (12) and (16) it follows that $\chi_{\max} = B^2$.

Thus, in the case of a horizontal antenna, the processing allows the direction finding of the source both by scanning the directional characteristic (19) and based on processing the antenna hologram for various vector-scalar field components, as it was proposed

with respect to a single vector-scalar receiver [5].

2.2. VERTICAL LINEAR ANTENNA

The field at the output of the *b*-th element, according to [9], is represented as

$$p_b(\omega, r) = \sum_m \Psi_m(z_b) A_m(\omega, r) \exp[ih_m(\omega)r], \quad (20)$$

where $\Psi_m(z)$ is the eigenfunction of the *m*-th mode; z_b – is the depth of the *b*-th element. As above, the cylindrical field divergence, modal attenuation, and source depth z_s are formally taken into account by the mode amplitude. In (20), the slow change in the eigenfunction of frequency is neglected. At the antenna output, the field is

$$P_{an}(\omega, r) = \sum_b \sum_m \Psi_m(z_b) A_m(\omega, r) \exp[ih_m(\omega)r]. \quad (21)$$

The interferogram $P_{an}(\omega, r)$, according to (20), (21), takes the form

$$P_{an}(\omega, r) = \sum_b \sum_a \sum_m \sum_n P_{mn}^{(ba)}(\omega, r), \quad (22)$$

where

$$P_{mn}^{(ba)}(\omega, r) = \Psi_m(z_b) \Psi_n^*(z_a) A_m(\omega, r) A_n^*(\omega, r) \exp[ih_{mn}(\omega)r]. \quad (23)$$

Let the source move with radial velocity *w* and assume that the distance *r* corresponds to the initial moment of time $t_0 = 0$. Passing in interferogram (22) from the distance variable *r* to the time variable *t* and applying the 2D Fourier transformation to it, we obtain

$$F_{an}(\tilde{\nu}, \tau) = \int_0^{\Delta t} \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} P_{an}(\omega, t) \exp[i(\tilde{\nu}t - \omega\tau)] dt d\omega = \sum_b \sum_a \sum_m \sum_n F_{mn}^{(ba)}(\tilde{\nu}, \tau), \quad (24)$$

where

$$F_{mn}^{(ba)}(\tilde{\nu}, \tau) = A_m(\omega_0, r) A_n^*(\omega_0, r) \Psi_m(z_b) \Psi_n^*(z_a) w \Delta\omega \Delta t \times \exp\left[i\left(\frac{\tilde{\nu}, \Delta t}{2} \tau \omega_0\right)\right] \exp\left\{i\left[(m-n)\alpha\left(\frac{\Delta t}{2} w+r\right)+t(\tilde{\nu}/w)\right]\right\} \times \frac{\sin\left\{\left[(r_1+wt_s)(m-n)\frac{d\alpha}{d\omega}-\tau\right]\frac{\Delta\omega}{2}\right\}}{\left[(r_1+wt_s)(m-n)\frac{d\alpha}{d\omega}-\tau\right]\frac{\Delta\omega}{2}} \frac{\sin\left\{[w(m-n)\alpha+\tilde{\nu}]\frac{\Delta t}{2}\right\}}{[w(m-n)\alpha+\tilde{\nu}]\frac{\Delta t}{2}}. \quad (25)$$

If we put $B = 1$, then $z_b = z_a$, and formula (25) turns into the corresponding formula for a single receiver [1].

The qualitative picture of the localization of the two-dimensional spectral density (24) of a vertical antenna, as in the case of a horizontal antenna (7), is similar to a single receiver. When the condition $r \gg |w|t_s$ is satisfied, as follows from (25), the distance and radial velocity of the source are determined by expression (10).

The gain χ is given by expression (16), from which, according to (25), one should expect $\chi \approx B^2$. The equality is fulfilled when the values of the eigenfunctions of the modes of the antenna elements at different depths are equal to each other $\Psi_m(z_b) = \Psi_m(z_a)$, $z_b \neq z_a$. The source direction finding is carried out on the basis of an algorithm as proposed in relation to a single vector-scalar receiver [5].

3. NOISE IMMUNITY OF SIGNAL PROCESSING

It is proposed to characterize the noise immunity of interferometric processing using a single receiver by the limiting (minimum) input signal-to-noise ratio (*s/n*) $q_{lim}^{(r)}$, when stable detection and estimates of bearing, radial velocity, distance and depth are close to real ones for *s/n* values $q \gg q_{lim}^{(r)}$ [1]. In the case of isotropic interference for the scalar field component (pressure) of a noise source, $q_{lim}^{(r)} \approx 1.5J^2$, where $J = \Delta t / (T + \delta t)$

is the number of time intervals (counts), on which the coherent accumulation of spectral maxima of the wave field along the interference fringes is realized [2]. Here T is the duration of the noise realization, δt - the interval between samples. Despite the fact that the estimate $q_{\text{lim}}^{(r)} = 1.5J^2$ was established on the basis of a number of physical considerations, and not derived from any more general principles, it was found correctly and repeatedly verified on the results of numerical and field experiments. This made it possible to construct a theory of noise immunity of interferometric processing using a single receiver. Let us generalize the obtained estimate for extended antennas.

Suppose that the noise signal and interference are statistically unrelated random processes, at the input of the antenna elements the interference is not correlated. To satisfy the second condition, it is sufficient to require the fulfillment of the inequality $d \gg \lambda/2$. Then the limiting input s/n ratio at the antenna element is estimated as

$$q_{\text{lim}}^{(an)} = Bq_{\text{lim}}^{(r)}/\chi. \tag{26}$$

In the case of a horizontal linear antenna, the highest processing immunity is achieved with a bearing equal to the compensation angle $\theta = \theta_*$, $\min q_{\text{lim}}^{(an)} = q_{\text{lim}}^{(r)}/B$. For the limiting input s/n ratio of the vertical antenna, we have $q_{\text{lim}}^{(an)} \approx q_{\text{lim}}^{(r)}/B$.

For the limiting (maximum) removal r_{lim} of the noise source, when the operability of the interferometric processing is preserved, using the same approach as in the case of a single receiver [3], we obtain

$$r_{\text{lim}} = \frac{1}{5\sqrt{1.5}} \sqrt{\frac{q\chi}{B}} \frac{\Delta t}{|dh_{1M}(\omega_0)/d\omega|}, \tag{27}$$

With an increase in the input value of s/n at the antenna element, observation time Δt , antenna elements B , gain χ , and center frequency ω_0 , the maximum range of the noise source increases. The latter is due to the fact that, with increasing frequency, the group velocities of the modes asymptotically tend to a constant value, which does not depend on the mode number [9]. A characteristic feature of relation (27) is the fact that it includes the parameters of the transfer function of the waveguide and therefore the value of the limiting distance r_{lim} is different depending on the choice of the water area. If we put $B = 1$, then formula (27) turns into the corresponding formula for a single receiver [3].

Note that if at each b -th receiver we first perform interferometric processing and then sum up the spectral densities of the holograms at the antenna output, then there will be no gain in noise immunity with respect to a single receiver.

4. SIMULATION RESULT

Numerical simulation was performed for a horizontally uniform waveguide with a depth of $H = 70$ m. The distribution of the sound velocity over depth is shown in **Fig. 2**. Parameters of an absorbing liquid homogeneous bottom: the ratio of the density of bottom and water $\rho = 1.8$, a complex refractive index $n = 0.85(1 + i0.02)$.

The number of antenna elements $B = 21$. The elements of the horizontal antenna are located at the bottom, $z_b = 70$ m, the elements of the vertical antenna are at depths of $z_b = 10 + 2.5(b - 1)$ m, $b = \overline{1, 21}$. The inter-element distance is $d = 2.5$ m, which is approximately equal to half the wavelength λ at a frequency of $f_0 = 310$ Hz. The noise source, located at a depth of $z_s = 30$ m, was moving away from

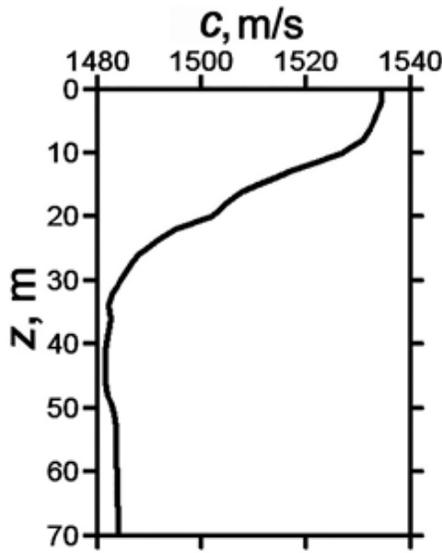


Fig. 2. Sound speed.

the antennas at a speed of $w = 3$ m/s. At the moment in time $t = 0$, the distance of the source from the support element of the horizontal antenna and the vertical antenna was $r = 10$ km. The accumulation time $\Delta t = 60$ s, the duration of the random realization $T = 2$ s, the time interval between samples $\delta t = 0.5$ s, so the number of samples $J = 12$. The limiting input ratio s/n at the antenna elements, taking $q_{\text{lim}}^{(am)} = q_{\text{lim}}^{(r)}/B$, is equal to $q_{\text{lim}}^{(am)} = 4.9603 \cdot 10^{-4}$.

Highlighted horizontal wavenumber $b_*(\omega_0) = b_1(\omega_0)$ (Table 1). If the wavenumbers $b_*(\omega_0)$ of other modes are used as the quality, then the calculation results practically do not change. The results of numerical simulations for horizontal and vertical antennas are shown in Fig. 3-6 and Fig. 7-10, respectively. In order to increase the contrast and information content, average values are cut out on interferograms and holograms.

Fig. 3 shows interferometric processing when received on an antenna support element. A contrasting interference pattern is observed (Fig. 3a), the spectral density of the hologram (Fig. 3b) is concentrated in six focal spots. The focal spots of the numbers $\mu = 1, 2$ partially overlap. The directivity characteristic (Fig. 3c), naturally, has circular symmetry.

In Fig. 4 shows the results of interferometric processing at the antenna output in the absence of compensation. Compared to a single receiver (Fig. 3), the contrast of the interferogram decreases (Fig. 4a) and the topology of the spectral density of the hologram changes (Fig. 4b). This is due to the incoherent summation of the fields on the antenna elements. In this case, the positions of the maxima of the focal spots are retained. The directional characteristic (Fig. 4c) has one main maximum corresponding to the zero order of the spectrum. Its width is half the spectral density $\Delta\theta \approx 9.5^\circ$.

Directional characteristics at different compensation angles are shown in Fig. 5. With an increase in the compensation angle θ_* , the width of the main maximum increases. The largest width occurs at an angle $\theta_* = 90^\circ$. These features of the behavior of the directivity characteristic, as evidenced by other modeling results (not reflected in this work), take place with an increase in the wave size d/λ while maintaining the number of elements or with a decrease in the number of antenna elements with a constant wave size.

In Fig. 6 shows the dependences of the normalized gain $\hat{\chi} = \chi/\chi_{\text{max}}$ (16) on the bearing θ for different values of the

Table 1

Horizontal wave numbers of modes h_m and their derivatives $dh_m/d\omega$ at a frequency of $f_0 = 300$ Hz.

Modes Numbers, m	1	2	3	4	5	6	7
h_m, m^{-1}	1.2787	1.2737	1.2670	1.2593	1.2504	1.2412	1.2325
$d_{hm}/d\omega, 1/\text{m}\cdot\text{Hz}$	$0.6798 \cdot 10^{-3}$	$0.6807 \cdot 10^{-3}$	$0.6820 \cdot 10^{-3}$	$0.6838 \cdot 10^{-3}$	$0.6853 \cdot 10^{-3}$	$0.6847 \cdot 10^{-3}$	$0.6861 \cdot 10^{-3}$

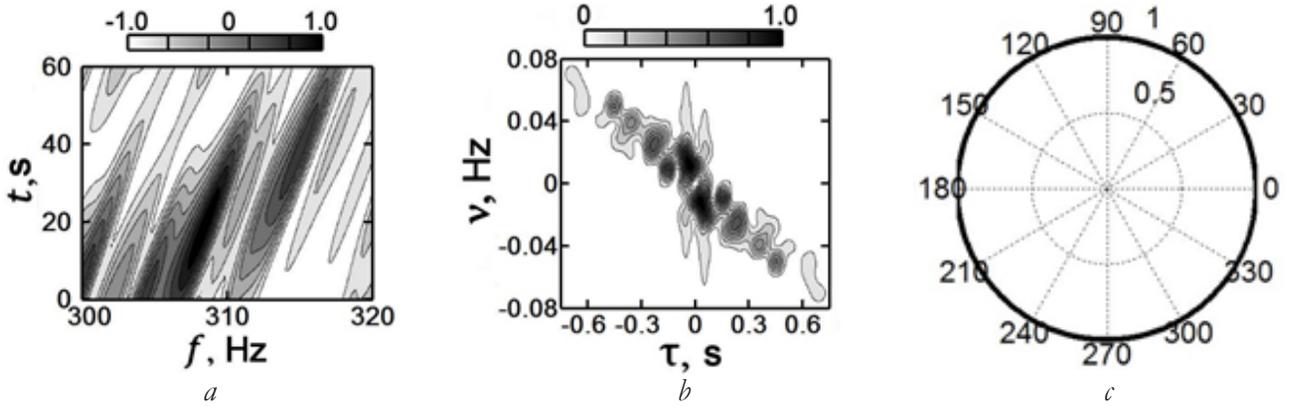


Fig. 3. Normalized interferogram (a), hologram (b) and directional characteristic (c) of the antenna reference element.

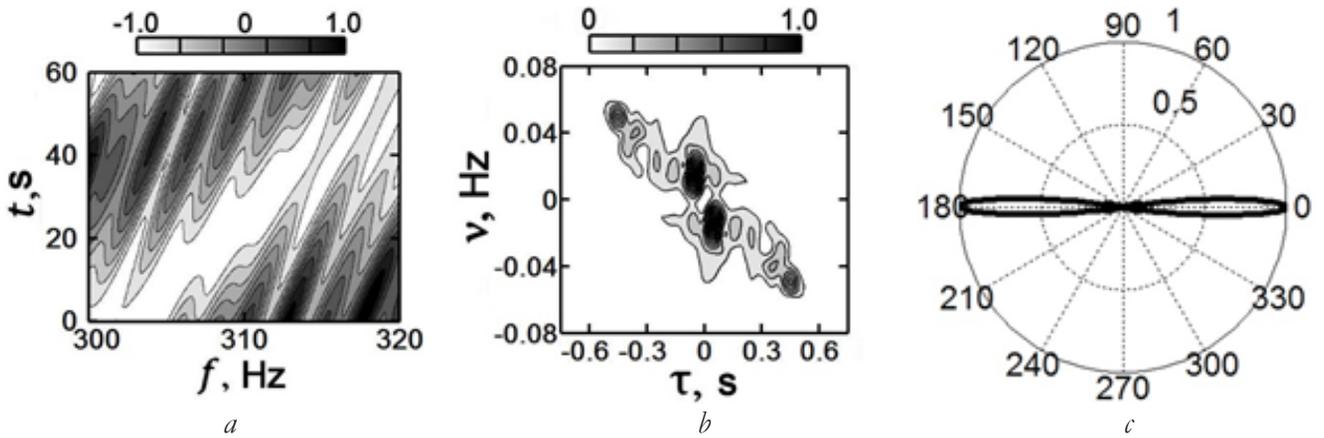


Fig. 4. Normalized interferogram (a), hologram (b) and directional characteristic (c) of the antenna at the absence of compensation.

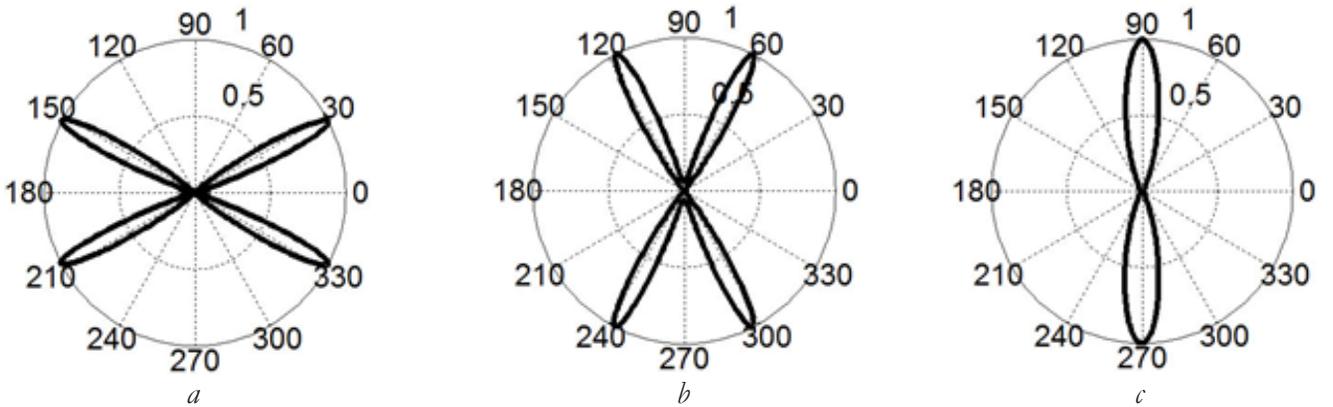


Fig. 5. Antenna directivity characteristics for compensation angles: a) $\theta_* = 30^\circ$, (b) $\theta_* = 60^\circ$, (c) $\theta_* = 90^\circ$.

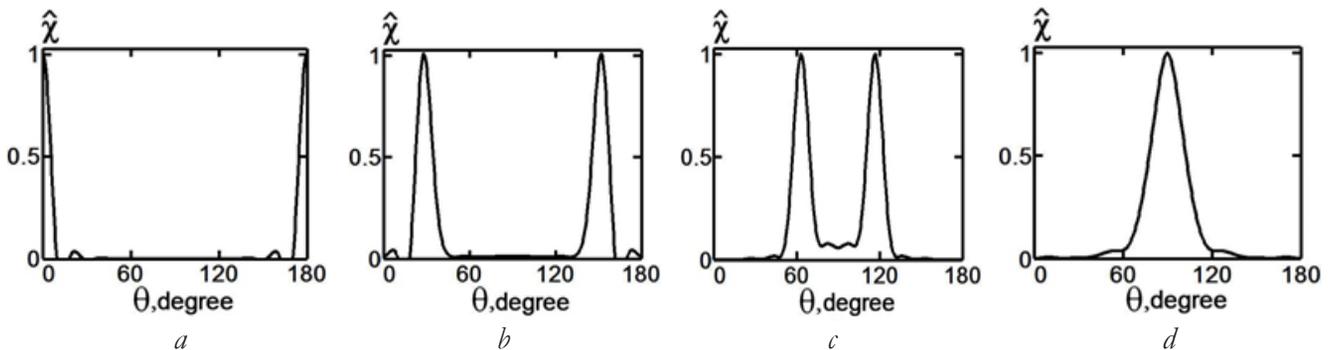


Fig. 6. Dependence of the normalized gain $\hat{\chi}$ on the bearing θ for different values of the compensation angle θ_* : (a) $\theta_* = 0^\circ$, (b) $\theta_* = 30^\circ$, (c) $\theta_* = 60^\circ$, (d) $\theta_* = 90^\circ$.

compensation angle θ_* . With an increase in the compensation angle, the width of the maxima increases. These dependencies are essentially another form of representing dependencies in Fig. 5. For this reason, the main maximums in Figs. 5, 6 have the same width. Normalized value of the coefficient $\chi_{\max} = 384$. The maximum gain B^2 , approximately equal, occurs in the direction of the compensation angle.

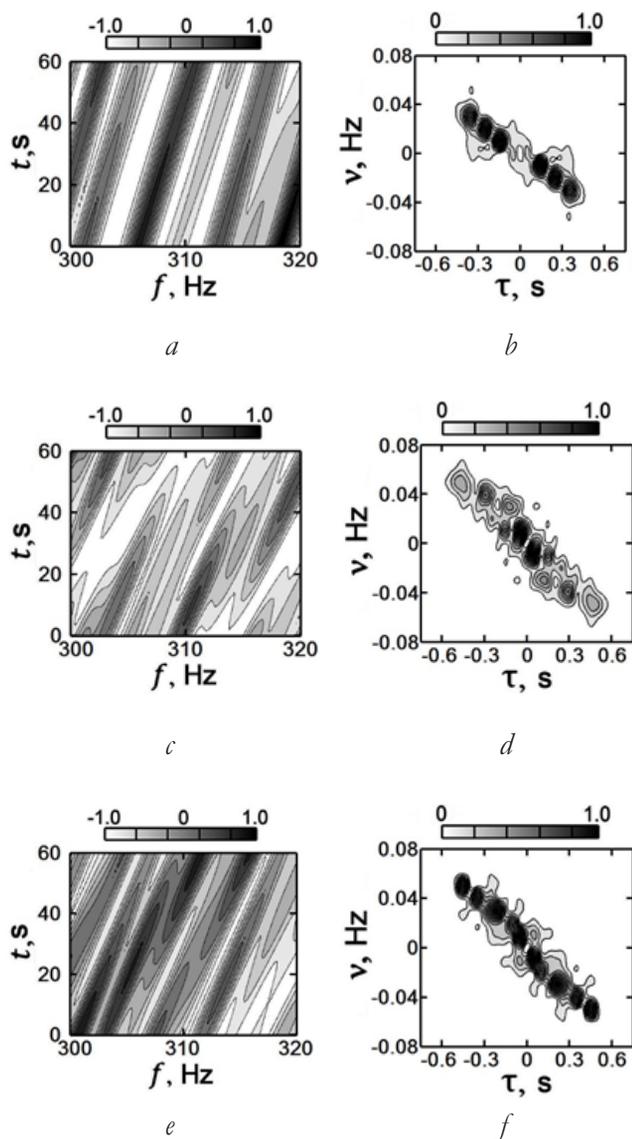


Fig. 7. Normalized interferograms (a, c, e) and holograms (b, d, f) of antenna elements located at depths of $z_b = 10, 37.5, 60$ m, respectively.

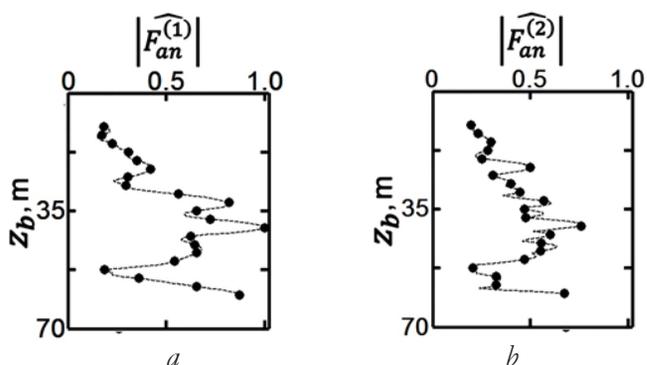


Fig. 8. Dependences of the normalized spectral maxima of the holograms of the antenna elements on their depth: (a) the first maximum $|F_b^{(1)}|$; (b) the second maximum, $|F_b^{(2)}|$.

In Fig. 7 shows interferograms and holograms of three antenna elements. The configuration of the spectral density distribution regions is different, despite the coherent addition of the fields on the antenna elements. a shift in the position of the maxima of the focal spots is observed. This difference is explained by the different values of the eigenfunctions at the depths of the elements.

This difference is most clearly illustrated in Fig. 8, which shows the dependences of two normalized spectral maxima $|F_b^{(1,2)}| = F_b^{(1,2)} / F_{\max}^{(1,2)}$ from the antenna element depth. The regions of localization of the spectral density of the first two focal spots in Fig. 7b. Numerical calculations are marked with dots. The distribution of the focal spot maxima has an oscillating form due to

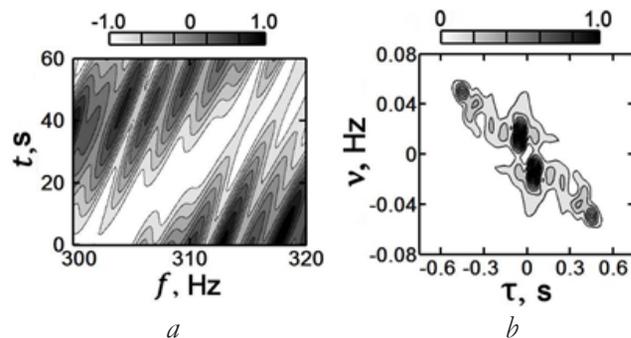


Fig. 9. Normalized interferogram (a) and hologram (b) of the antenna.

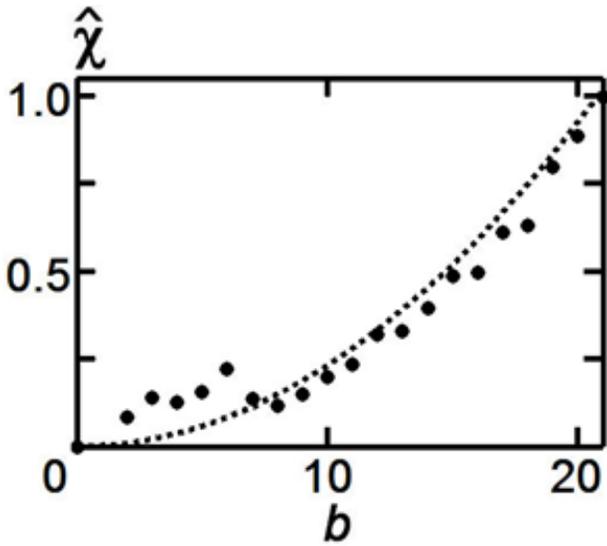


Fig. 10. Dependence of the normalized gain $\hat{\chi}$ on the number b of antenna elements. Points – calculated values, dotted line – quadratic dependence.

the different values of the mode excitation coefficients (eigenfunctions) at the reception horizons. The nature of the change in the oscillations with a change in the position of the focal spot is associated with the interference of different numbers of modes that determine the location of localized regions. The ratio of normalized coefficients $\beta = F_{\max}^{(1)} / F_{\max}^{(2)} = 1.53$.

Fig. 9 shows the behavior of the normalized interferogram and the antenna hologram. Spectral density is predominantly localized in the region of the first focal spot.

In **Fig. 10** shows, obtained by numerical simulation (points), the dependence of the normalized gain $\hat{\chi} = \chi(b) / \chi_{\max}$ (16) on the number of elements b . The calculated

values fit well with the estimated quadratic dependence $\hat{\chi} = b^2 / \chi_{\max}$ (dotted line). The value of the coefficient $\chi_{\max} = 362$ is obtained from the condition for normalizing the calculated values.

Table 2 shows the values of the maximum ranges of a noise source using a single receiver and a linear antenna, depending on the input ratio s/n q , calculated by formula (27). Observation time $\Delta t = 60$ s, value $|dh_{1M}(\omega_0)/d\omega| = 6.3 \cdot 10^{-6}$ s/m (see Table 1), parameter $\chi/B = 17.62$.

5. CONCLUSION

The intensive introduction of interferometric processing in hydroacoustics over the past few years has already made it possible to obtain a number of new results using single receivers, forcing to revise the previously established classical processing methods (matched-field processing), in short, all those areas where wave interference plays a role. These considerations, as well as, of course, to a large extent the importance of physical and applied problems that can be solved using antennas, stimulated the consideration of interferometric processing using extended linear antennas. An expression is obtained for the distribution of the spectral density of the hologram, which determines the gain and directivity characteristic. The noise immunity and the limiting removal of the noise source are estimated. The results obtained significantly expand the field of application of interferometric processing. The material is illustrated by numerical calculations for the low-frequency region of a noise source, which make it possible to clearly understand the efficiency of interferometric processing when working with multi-element antennas.

Table 2

Limiting removal of a noise source depending on the input s/n ratio.

Ratio $s/n, q$	Single receiver	Linear Antenna, Elements count $B = 21$
	Limit of noise source distance $r_{\text{lim}}, \text{ km}$	
10^{-3}	49.2	206.6
10^{-4}	15.5	65.0
10^{-5}	4.9	20.5
10^{-6}	1.5	6.3

REFERENCES

1. Kuznetsov GN, Kuzkin VM, Pereselkov SA. Spectrogram and localization of the sound source in the shallow sea. *Acoustical physics*, 2017, 63(4):406-418.
2. Kaznacheev IV, Kuznetsov GN., Kuzkin VM, Pereselkov SA. Interferometric method for detecting a moving sound source by a vector-scalar receiver. *Acoustical physics*, 2018, 64(1):33-45.
3. 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.
4. Kuz'kin VM, Kuznetsov GN, Pereselkov SA, Grigor'ev VA. Resolving power of the interferometric method of source localization. *Phys. Wave Phenom.*, 2018, 26(2):150-159.
5. Kuznetsov GN, Kuz'kin VM, Lyakhov GA, Pereselkov SA, Prosovetskiy DYU. Direction finding of a noise sound source. *Phys. Wave Phenom.*, 2019, 27(3):237-241.
6. Chuprov SD The interference structure of the sound field in a layered ocean. In: *Acoustics of Ocean. The current situation.* Moscow, Nauka Publ., 1982, p. 71-82.
7. Kuz'kin VM, Pereselkov SA, Zvyagin VG, Malykhin AYU, Prosovetskiy DYU. Intense internal waves and their manifestation in interference patterns of received signals on oceanic shelf. *Phys. Wave Phenom.*, 2018, 26(2):160-167.
8. Badiy M, Kuz'kin VM, Lyakhov GA, Pereselkov SA, Prosovetskiy DYU, Tkachenko SA. Intense internal waves and their manifestation in the interference patterns of received signals on oceanic shelf. Part II. *Phys. Wave Phenom.*, 2019, 27(4):313-319.
9. Brekhovskikh LM, Lysanov YuP. *Theoretical foundations of ocean acoustics.* Leningrad, Gidrometeoizdat Publ., 1982, 264 p.
10. Landsberg GS. *Optics.* Moscow, Nauka Publ, 1976, 92 p.

DOI: 10.17725/rensit.2020.12.495

Restoring Missing Fragments of a Distorted Image Due to Defocusing or Blur of a Known Hardware Function

Alexander V. Kokoshkin, Evgeny P. Novichikhin

Kotel'nikov Institute of Radioengineering and Electronics of RAS, Fryazino Branch, <http://fire.relarn.ru/>
Fryazino 141190, Moscow Region, Russian Federation

E-mail: shvarts65@mail.ru, epnov@mail.ru

Received April 30, 2020, peer-reviewed May 18, 2020, accepted May 29, 2020

Abstract: In this paper, a comparative analysis of the methods for recovering images distorted by defocusing or blurring from incomplete data is performed using examples. Incomplete data means the absence of any image fragments that were retouched using different types of interpolation - linear, spline and the interpolation method for the sequential calculation of the Fourier spectrum (IMSCS) developed by us. Then, the famous deconvolution method, the Wiener Filter (WF), was applied to the entire image. Analysis of the quality of restoration, carried out on the example of aerospace images, suggests that using IMSCS to fill in missing fragments (gaps) is either preferred or no less competitive than alternative methods. This is a consequence of the fact that IMSCS does not just retouch the gap, but also tries to reconstruct the lost data.

Keywords: distorted images, remote sensing of the earth's surface, blurring, defocusing, missing fragments, interpolation, restoration

UDC 621.369

Acknowledgements: The work was carried out within the framework of the state assignment of Kotel'nikov institute of radio engineering and electronics of RAS.

For citation: Alexander V. Kokoshkin, Evgeny P. Novichikhin. Restoring Missing Fragments of a Distorted Image Due to Defocusing or Blur of a Known Hardware Function. *RENSIT*, 2020, 12(4):495-506. DOI: 10.17725/rensit.2020.12.495.

CONTENTS

1. INTRODUCTION (495)
 2. BASIC ASSUMPTIONS (496)
 3. HARDWARE DEFOCUS FUNCTION (497)
 4. HARDWARE BLURRING FUNCTION (500)
 5. FINAL EXAMPLE (502)
 6. CONCLUSIONS (503)
- REFERENCES (504)

1. INTRODUCTION

The development of methods for remote sensing of the Earth's surface and other planets is a topical trend in science and technology. And one of the key tasks of digital image processing is the restoration of missing parts (gaps). In aerospace imagery, data loss can arise due to the peculiarities of the trajectories of the aircraft, cloud

shading or technical failures of the recording equipment. In some cases, incompleteness of data in published images may mean deliberate removal of fragments in order to hide information. Thus, the improvement of methods for reconstructing gaps is one of the most important directions.

Defocusing and blurring can be distinguished among the most common image distortions, which can be described on the basis of the convolution equation [1,2] with the corresponding distorting hardware function (HF). In [3-5], the possibility of restoring images distorted by blur and defocusing from incomplete data was considered. The gaps in such images were retouched using linear interpolation. Then the well-known deconvolution methods

were applied to the entire image. In [6-13], the possibilities of filling in the missing parts of the image using wavelets, various types of interpolation (bilinear, spline, trigonometric, polynomial) are shown in order to reconstruct the distortions caused by a known HF. However, the assumed non-rectangular shape of the gap and the appearance of additional artifacts after deconvolution caused by such retouching limit the possibilities of using these methods.

In [14], we have developed a method for interpolation of a sequentially calculated Fourier spectrum (MIPVS), which is capable of filling in gaps of arbitrary shape. In works [15-19], possible areas of application of MIPVS are shown and a comparative analysis of the methods used to fill gaps in images was carried out.

This work is devoted to a comparative analysis of how the methods of interpolation of missing fragments distorted due to defocusing or blurring of images, with a known instrumental function, affect the final result of reconstruction. MIPVS is compared with linear interpolation and spline interpolation described in [20,21]. The physical meaning of spline interpolation is that for an arbitrary set of reference points (nodes) a system of linear equations is solved that simulates the behavior of a curved elastic plate. At the same time, it is possible to take into account when calculating the spline the border around the gap with a width of one to several nearest pixels. This approach has a certain versatility and can be applied for comparative analysis with the interpolation method of a sequentially calculated Fourier spectrum. Comparing MIPVS with simpler methods of interpolation, for example, linear, will not be entirely correct. Nevertheless, linear interpolation will be used by us in this

work, since with the instrumental function of linear blur it exactly corresponds to the distortion. Thus, provided that the orientation of the blur and the interpolation line match exactly, all other methods of filling the gap are in a more difficult position compared to linear interpolation.

2. BASIC ASSUMPTIONS

As test we use aerospace images taken from publicly available Internet resources. There are a large number of methods for determining the distorting hardware function from the original image and methods of "blind" deconvolution [1,22-26], but any inaccuracy in determining the HF parameters will inevitably lead to a deterioration in the quality of reconstruction. Therefore, in order to eliminate additional interfering factors, this work assumes that the hardware functions are accurately known. The gap is filled using linear interpolation, spline and MIPVS. The missing fragment is square and varies in size from experience to experience. Its shape was chosen to be square in order to put all the considered interpolation methods into equal conditions, since the MIPVS (in contrast to alternative methods) has an arbitrary lacuna shape that does not cause difficulties. Reconstruction of the image with the reconstructed lacuna is performed using the Wiener Filter (WF) [1]. No correction of artifacts after deconvolution is performed. An objective criterion for the quality of restoration will be the relative values of the standard deviation of the result of restoration with an interpolated gap from the result of restoration of an image without a gap taken as "ideal".

In [4] it was shown that the irreplaceable area is determined by the relative size of the gap (in [4] of the shading object) and the

amount of defocusing of the image. Thus, in our work, the size of the defocusing spot is chosen large enough for reasons of the fact that with a small size of the HF, or as they sometimes say - the point blur function (PBF), the shading effects are hardly noticeable. According to the results of [4], it is obvious that inside the gap one can only hope for partial recovery of the lost data, and only within the radius of the blur spot from the edge of the gap. **Fig. 1** shows the gap in black, gray circles denote the HF coverage area at different sizes of the gap and HF relative to each other.

Thus, with a known HF with a radius of $S = 10$ pixels, the forecast for the reconstruction of a lacuna with a size of no more than 20 by 20 pixels is positive. In the proposed work, we vary the size of the gap from 10 by 10 pixels to 30 by 30 pixels. Those, we begin the study of the reconstruction quality from the zone of sufficiently confident restoration (a gap of minimum size – Fig. 1a) and end in the zone of obviously problematic restoration of the center of a gap of maximum size – Fig. 1c. In what follows, all figures in this work show the results of image reconstruction for the most difficult of the selected cases – a gap size of 30 by 30 pixels. This is done so that readers can visually (expertly) assess the quality of the restoration.

Similar reasoning was used to select the relative sizes of the gap and the length of the HF of rectilinear blur. When modeling the instrumental blur function, we assume that in the process of shooting, distortion occurs due to the steady progressive motion of the recording system (presumably an aircraft) relative to the image scene. Thus, we take a horizontal line of length $SL = 21$ pixels for the hardware blur function.

3. HARDWARE DEFOCUS FUNCTION

The sequence of actions required to conduct a comparative analysis of the effect of interpolation methods for missing fragments on the final result of restoration is as follows:

Fig. 2a shows an aerospace image of Rostov Veliky, 512 by 512 pixels. In **Fig. 2b** image of **Fig. 2a** is shown after defocusing with the known HF with a radius of $S = 10$ pixels. **Fig. 2c** is figure 2b with a missing fragment measuring 30 by 30 pixels (black). **Fig. 2d** – restoration using the Wiener filter of **Fig. 2c** with a gap filled using linear interpolation. **Fig. 2e** – restoration using the Wiener filter of **Fig. 2c** with a gap filled using spline interpolation. **Fig. 2f** - restoration using the Wiener filter of **Fig. 2c** with a gap filled with MIPVS.

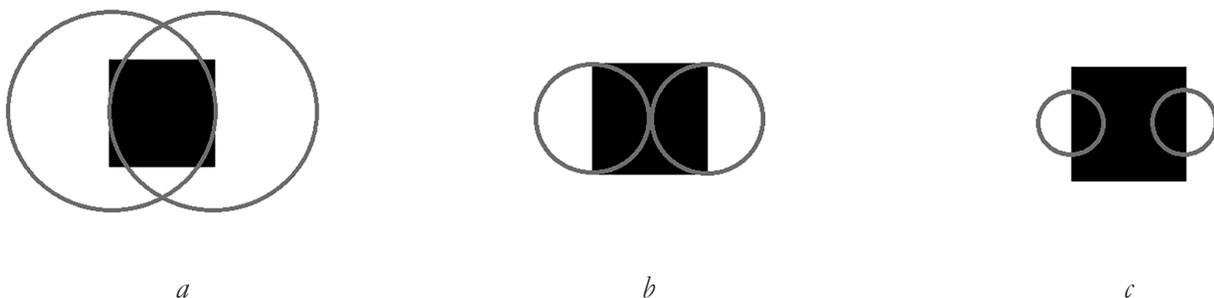


Fig. 1. The coverage area of the HF with different sizes of the gap (black square) and HF "round spot" (gray circles): (a) – HF is much wider than the gap; (b) – the HF radius is equal to the half-width of the square lacuna; (c) HF is much smaller than the gap.

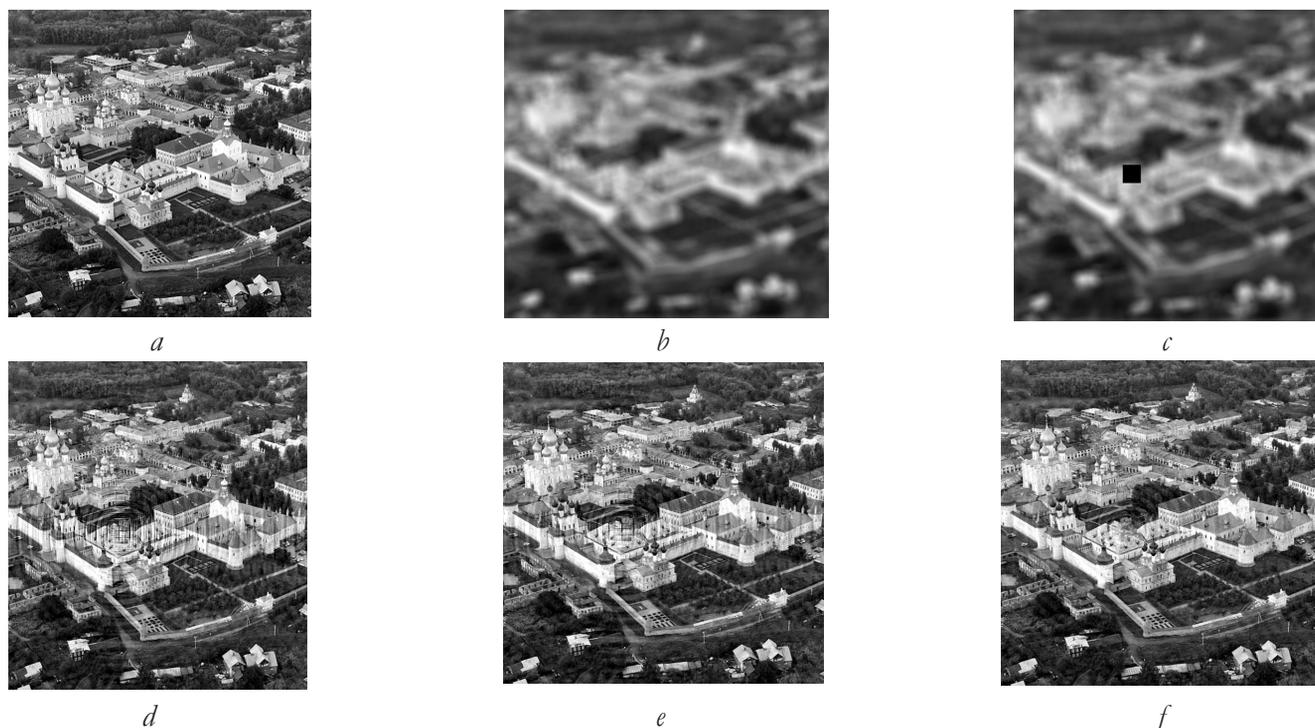


Fig. 2. (a) - original test aerospace image; (b) - defocused image in Fig. 2a ($S = 10$ pixels); (c) - image of Fig. 2b with the missing fragment 30 by 30 pixels (black); (d) - restoration using the Wiener filter of Fig. 2c with a gap filled using linear interpolation; (e) - restoration using the Wiener filter of Fig. 2c with the gap filled with the help of spline interpolation; (f) - restoration using the Wiener filter of Fig. 2c with a gap filled with MIPVS.

Fig. 3 is a fragment of Fig. 2. It shows a 90 by 90 pixel patch with a 30 by 30 pixel gap inside. Each of Fig. 3 (from a to e) is a fragment of the corresponding Fig. 2 (from a to e).

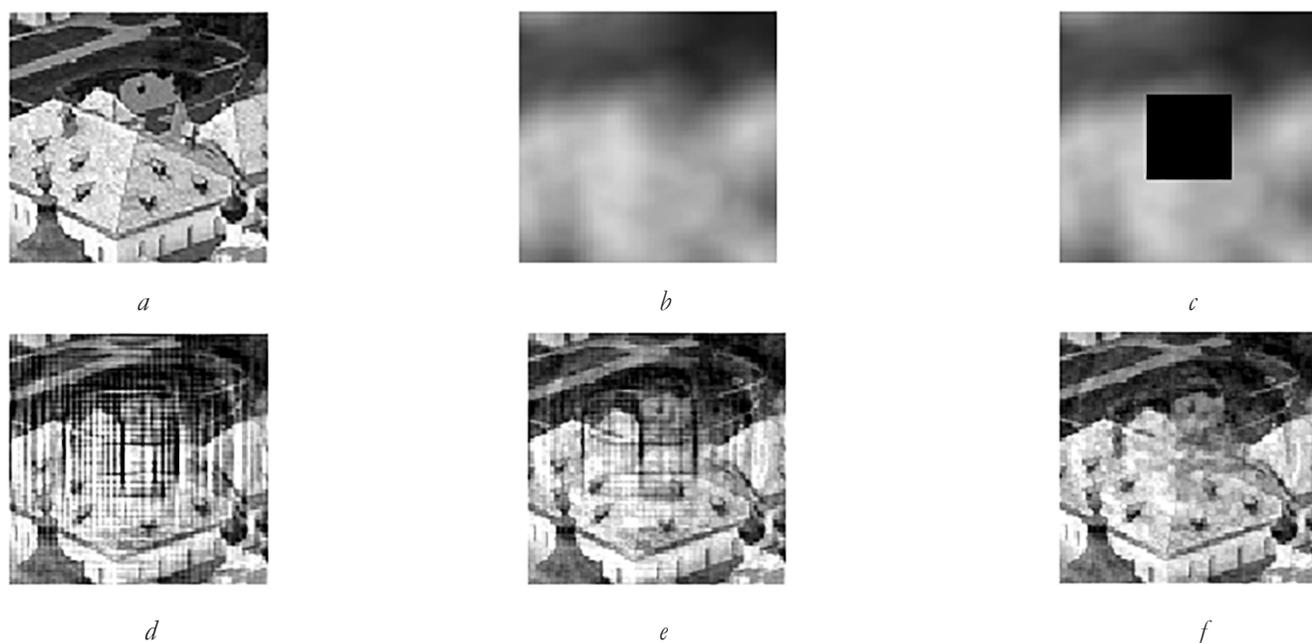


Fig. 3. Fragments of 90 by 90 pixels: (a) - the original test aerospace image; (b) - defocused image of Fig. 2a ($S = 10$ pixels); (c) - images 2b with a missing fragment of 30 by 30 pixels (black); (d) - Fig. 2c, reconstructed using the Wiener Filter, with a gap filled with linear interpolation; (e) - Fig. 2c reconstructed using the Wiener Filter with a gap filled with spline interpolation; (f) - Fig. 2c, reconstructed using the Wiener Filter, with a gap filled with MIPVS.

A close examination of the images in Fig. 2*d,e,f* it can be seen that on the reconstructed images the intensity of artifacts over the entire area gradually decreases from filling the gap using linear interpolation to filling the gap using MIPVS. This is confirmed by the data in Fig. 4*a*, which shows the dependence of the standard deviation over the entire image of the result of reconstruction with an interpolated gap on the reconstruction of an image without

a gap, taken as "ideal", along the horizontal gap size (DL). The standard deviation for the corresponding interpolation is indicated as follows: SKOL - linear, SKOS - spline, SKOF - MIPVS. In Fig. 4*b* shows similar dependences of the RMS on the size of the gap, calculated only inside the gap itself. The standard deviation for the corresponding interpolation is indicated as follows: SKOLm - linear, SKOSm - spline, SKOFm - MIPVS. Figs 4*a* and 4*b* show a

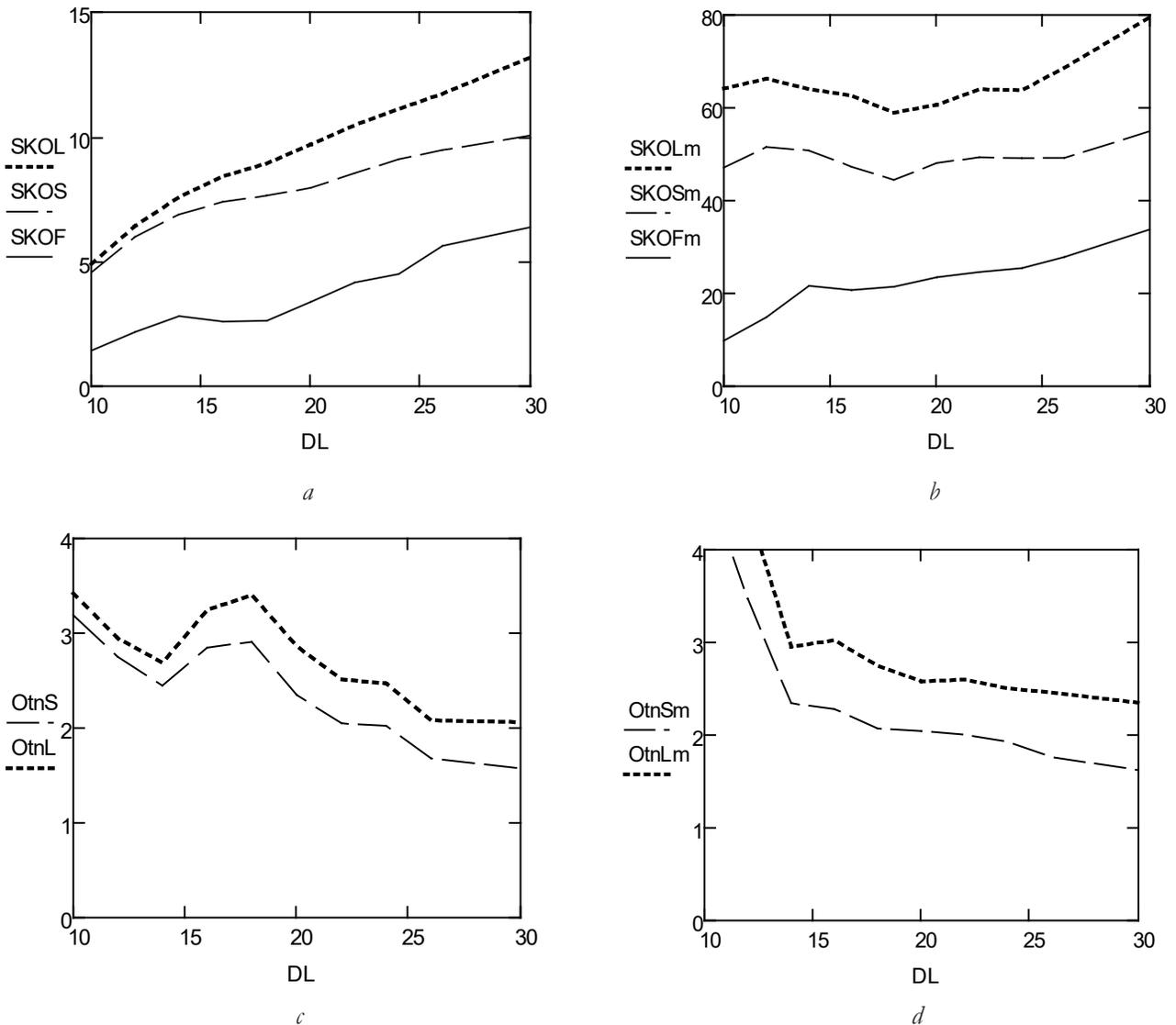


Fig. 4. Case of AF defocusing. The standard deviation of the result of reconstruction with an interpolated lacuna from the reconstruction of an image without a lacuna taken as "ideal": (a) - over the entire image (interpolation designations: SKOL - linear, SKOS - spline, SKOF - MIPVS); (b) - only by the area of the gap (notations of interpolations: SKOLm - linear, SKOSm - spline, SKOFm - MIPVS); (c) - relations $OtnS = SKOS/SKOF$, $OtnL = SKOL/SKOF$; (d) - ratios $OtnSm = SKOSm/SKOFm$, $OtnLm = SKOLm/SKOFm$.

significant advantage of MIPVS (solid line) compared to linear (dashed line) and spline (dashed line) interpolations. Fig. 4c depicts the dependences $OtnS = SKOS/SKOF$, $OtnL = SKOL/SKOF$ on the size of the gap DL. They show how many times the RMSD of restoration over the entire image with linear (OtnL) or spline (OtnS) interpolations is greater than the RMSD of restoration with MIPVS. Similar to Fig. 4d illustrates dependencies $OtnSm = SKOSm/SKOFm$, $OtnLm = SKOLm/SKOFm$ referring to RMS only inside the reconstructed lacuna. All dependencies indicate significantly better performance of MIPVS compared to competing methods.

4. HARDWARE BLUR FUNCTION

We assume that linear interpolation occurs exactly along the blur line (length $SL = 21$ pixels). In this case, other types of filling the gap are in a more difficult position. The

initial test is the image in Fig. 2a. In order to save space, here (in Fig. 5) we present only fragments of the test image measuring 90 by 90 pixels with a gap of 30 by 30 pixels in the middle. In fact, as with HF defocus, restoration was performed over the entire 512 by 512 pixel image. That is, all actions for HF blur were carried out as for HF defocusing.

Fig. 6 is analogous to Fig. 4, for HF blur only. The designations are the same. The behavior of the curves in Fig. 6 indicates that, according to the objective criterion of standard deviation from the "ideal", none of the investigated interpolations has an advantage over the alternative ones. This is especially evident in Fig. 6c and Fig. 6d, which, respectively demonstrate the behavior of the dependence $OtnS = SKOS/SKOF$, $OtnL = SKOL/SKOF$ (for RMS over the entire image) on the size of the gap DL and the dependence $OtnSm = SKOSm/$

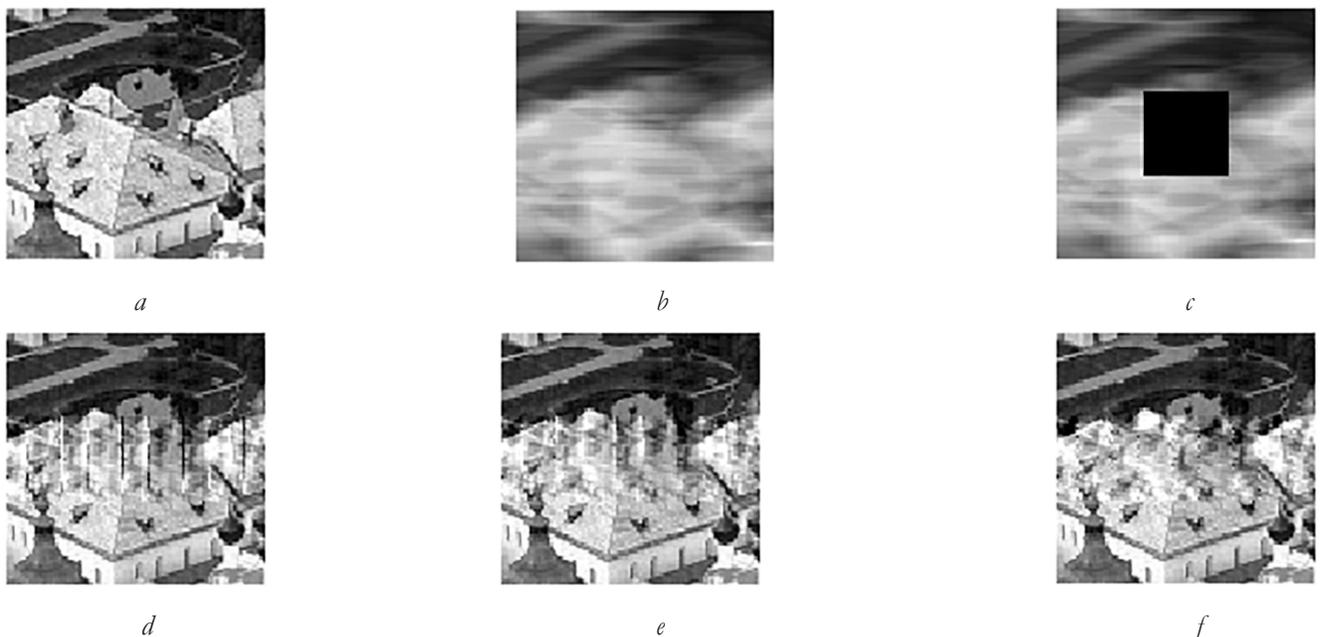


Fig. 5. Fragments of 90 by 90 pixels: (a) - the original test aerospace image; (b) - image distorted by blur in Fig. 2a ($SL = 21$ pixels); (c) - images 5b with a missing fragment of 30 by 30 pixels (black); (d) - Fig. 5c reconstructed using the Wiener Filter with a gap filled using linear interpolation; (e) - Fig. 5c reconstructed using the Wiener Filter with a gap filled with spline interpolation; (f) - Fig. 5c reconstructed using the Wiener Filter with a gap filled with MIPVS.

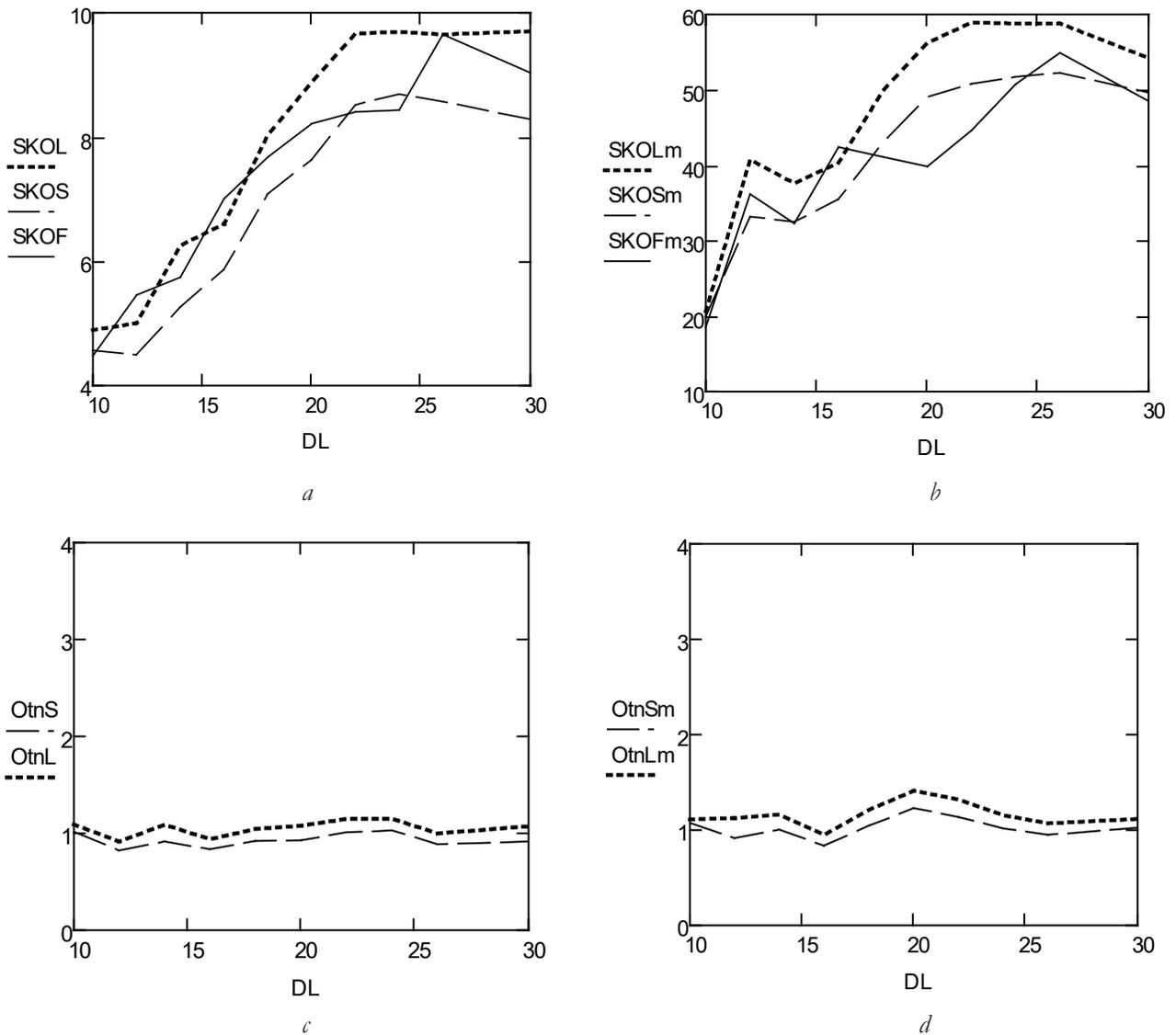


Fig. 6. A case of HF blur. The standard deviation of the result of reconstruction with an interpolated lacuna from the reconstruction of an image without a lacuna taken as "ideal": (a) - over the entire image (interpolation symbols: SKOL - linear, SKOS - spline, SKOF - MIPVS); (b) - only by the area of the gap (notation for interpolations: SKOLm - linear, SKOSm - spline, SKOFm - MIPVS); (c) - relations $OtnS = SKOS/SKOF$, $OtnL = SKOL/SKOF$; (d) - ratios $OtnSm = SKOSm/SKOFm$, $OtnLm = SKOLm/SKOFm$.

SKOFm, $OtnLm = SKOLm/SKOFm$ (for standard deviation, only inside the restored gap). In Fig. 6c and Fig. 6d standard deviation of restoration over the entire image (or standard deviation only in the gap) with linear or spline interpolation is practically equal to standard deviation of restoration with MIPVS, i.e. their ratio is close to unity for all sizes of the gap. At the same time, for HF defocusing (Fig. 4c and Fig. 4d), the reconstruction after MIPVS showed

much better results than the reconstruction after retouching the gap using competing interpolation methods. Note that the expert assessment of the restoration results in Fig. 5d, Fig. 5e, 5f, due to the visually lesser artifacts, indicates a preference for using MIPVS. However, according to the objective criterion of standard deviation from "ideal", none of the interpolation methods has any advantages.

5. FINAL EXAMPLE

In conclusion, we will give an example for the well-known aerospace image "the position of the Egyptian air defense S-75 in the desert"

Fig. 7. Original photo:

<https://picryl.com/media/an-aerial-view-of-soviet-built-sa-2b-guideline-surface-to-air-missiles-positioned-016058>.

For this image, a similar sequence of actions for modeling, processing and analysis was carried out as for the previous photo (in the case of HF defocusing). The lacuna was created in the area of the lower left rocket launcher. **Fig. 8a** is a fragment of the image in Fig. 7. The sequence of Figures 8 is analogous to Fig. 3 90 by 90 pixels with a gap of 30 by 30 pixels inside the resulting sequence of actions: defocusing - Fig. 8b, creating a gap in Fig. 8c, reconstruction using the WF with linear interpolation of the gap - Fig. 8d, restoration with the help of WF with spline interpolation of the gap - Fig. 8e, restoration with the help of WF



Fig. 7. The original image of the C-75 in the desert.

during lacuna interpolation using MIPVS - Fig. 8f.

The behavior of the curved lines in **Fig. 9** in general terms corresponds to similar graphs presented in Fig. 4. I.e. indicates an advantage when using MIPVS for retouching missing fragments over alternative interpolation methods. We believe that this happens because MIPVS, working in the frequency domain, uses all available data (original image), while spline or linear

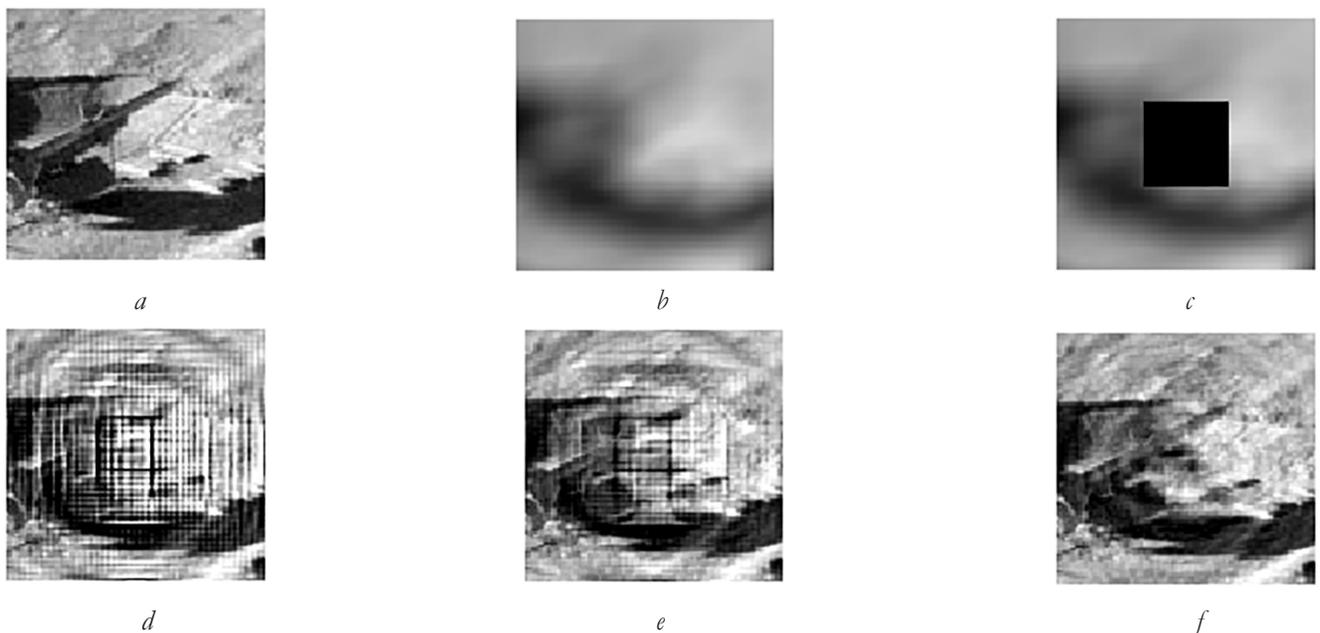


Fig. 8. Fragments of 90 by 90 pixels: (a) - the original test aerospace image Fig. 7; (b) - defocused image of Fig. 7 ($S = 10$ pixels); (c) - images 8b with a missing fragment of 30 by 30 pixels (black); (d) - Fig. 8c reconstructed using the Wiener Filter with a gap filled with linear interpolation; (e) - Fig. 8c, reconstructed using the Wiener Filter, with a gap filled using spline interpolation; (f) - Fig. 8c, reconstructed using the Wiener Filter, with a gap filled with MIPVS.

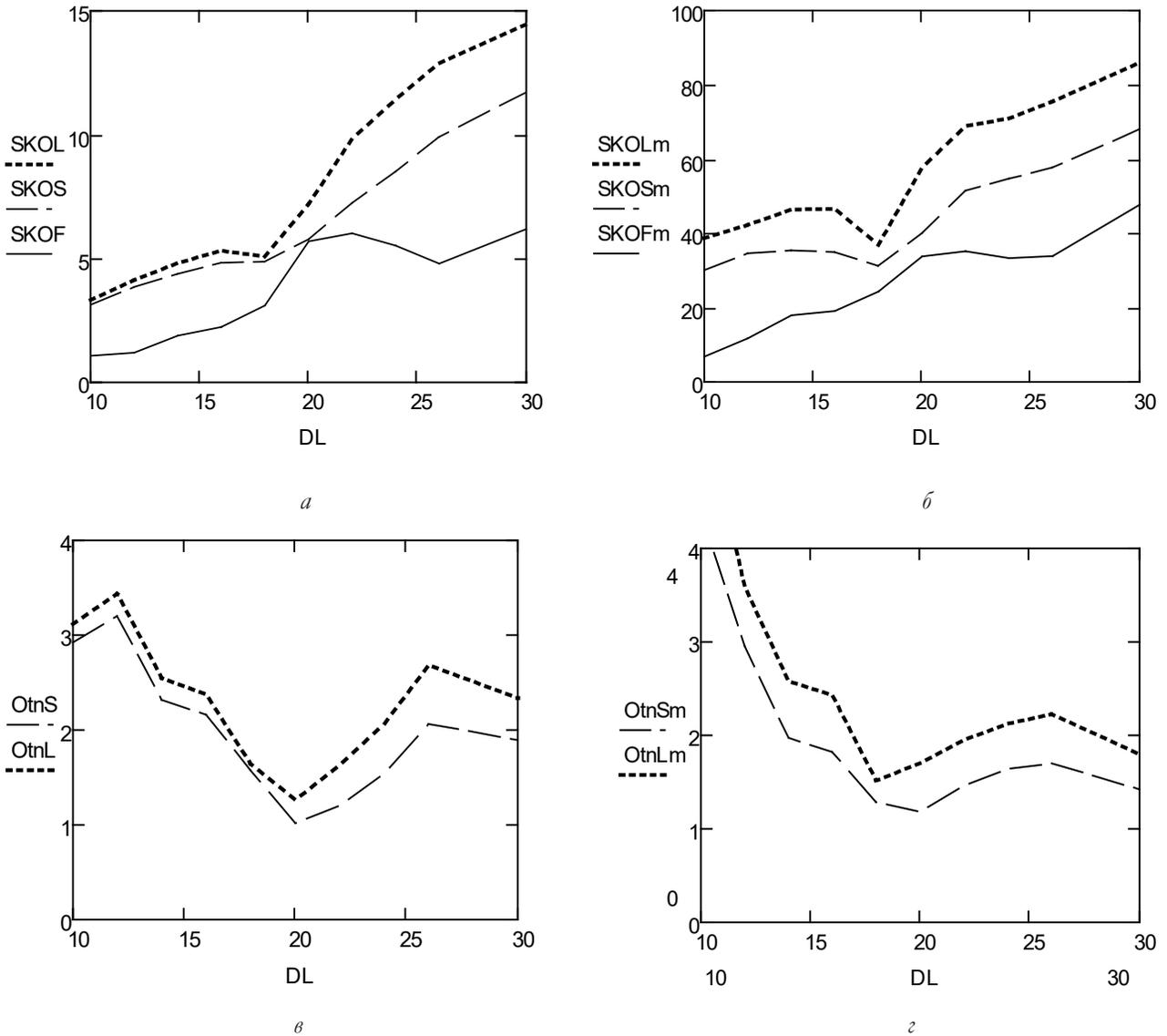


Fig. 9. For HF defocusing, Fig. 7. The standard deviation of the restoration result with an interpolated gap from the restoration of an image without a gap taken as "ideal": (a) - over the entire image (interpolation symbols: SKOL - linear, SKOS - spline, SKOF - MIPVS); (b) - only by the area of the gap (designations of interpolations: SKOLm - linear, SKOSm - spline, SKOFm - MIPVS); (c) - relations $OtnS = SKOS/SKOF$, $OtnL = SKOL/SKOF$; (d) - ratios $OtnSm = SKOSm/SKOFm$, $OtnLm = SKOLm/SKOFm$.

interpolation, performing actions in the spatial domain, operates with information about the edges of the gap. Thus, MIPVS not only retouches, but also tries to reconstruct the lost data [14,15,16,19]. Some fluctuations in the curves in Fig. 9 take place due to the specifics of the selected image and, among other things, depend on the location of the lacuna. Obviously, if the gap is located on a "flat" place (clear sky or smooth desert), then there is not much difference which of the

interpolation methods fill the place of the lost data. It is quite another matter if there is no fragment of a highly informative image (saturated with various elements).

6. CONCLUSIONS

Comparative analysis, both by expert judgment and by the objective standard deviation criterion from the "ideal", shows that when solving the problems of restoring missing fragments of images distorted due

to defocusing with a known instrumental function, the use of MIPVS is preferable compared to linear and spline interpolations. At the same time, despite the seemingly favorable for MIPVS expert assessment with HF blur, standard deviation does not show any advantages for any of the methods of lacuna retouching considered here. Thus, the MIPVS proposed by us in each of the considered HF options is either preferable or no less competitive than alternative methods. We believe that this happens because MIPVS, working in the frequency domain, uses all available data (the entire original image), while spline or linear interpolation, performing actions in the spatial domain, operates with information only about the edges of the gap. Thus, MIPVS not only retouches, but also tries to reconstruct the lost data [14,15,16,19].

REFERENCES

1. Gonzalez RC and Woods RE. *Digital Image Processing*. PrenticeHall, Upper Saddle River, NJ, 2002; Moscow, Tekhnosfera Publ., 2006:1071.
2. Gulyaev YuV, Zrazhevsky AYu, Kokoshkin AV, Korotkov VA, Cherepenin VA. Correction of the spacial spectrum distorted by the optical system using the reference image method. Part 2. Adaptive reference image method (in Russ.). *Journal of Radioelectronics* [electronic journal], 2013, №12. Available at: <http://jre.cplire.ru/jre/dec13/2/text.html>.
3. Kokoshkin AV, Korotkov VA, Novichihin EP. Effects of partially shaded in reconstructing the image, distorted by blurring (in Russ.). *Journal of Radioelectronics* [electronic journal], 2014, №9. Available at: <http://jre.cplire.ru/jre/sep14/3/text.html>.
4. Zrazhevsky AYu, Kokoshkin AV, Korotkov VA, Korotkov KV. Recovery of defocusing partially shaded image (in Russ.). *Journal of Radioelectronics* [electronic journal], 2014, №10. Available at: <http://jre.cplire.ru/jre/oct14/9/text.html>.
5. Kokoshkin AV, Korotkov VA, Korotkov KV, Novichihin EP. Comparison of objective methods of assessing quality of digital images (in Russ.). *Journal of Radioelectronics* [electronic journal], 2015, №6. Available at: <http://jre.cplire.ru/jre/jun15/15/text.html>.
6. Jong-Keuk Lee, Ji-Hong Kim, Jin-Seok Seo. Adaptive Recovery of Image Blocks Using Spline Approach. *IJCSNS International Journal of Computer Science and Network Security*, 2011, 11(2):213-217.
7. Jiho Park, Dong-Chul Park, Mark RJ, El-Sharkawi MA. Block loss recovery in DCT image encoding using POCS. *Proc. 2002 IEEE International Symposium on Circuits and Systems (ISCAS)*, Vol. 5. DOI: 10.1109/ISCAS.2002.1010686.
8. Strohmer T. Computationally attractive reconstruction of bandlimited images from irregular samples. *IEEE Trans. on Image Processing*, 1997, 6(4):540-548.
9. Chen Chen, Tramel Eric W, Fowler James E. Compressed-Sensing Recovery of Images and Video Using Multihypothesis Predictions. *Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Pacific Grove, CA, 2011, pp. 1193-1198, doi: 10.1109/ACSSC.2011.6190204.
10. Seung Hwa Hyun, Sang Soo Kim, Byoung Chul Kim, Il Kyu Eom, Yoo Shin Kim. Efficient Directional Interpolation for Block Recovery Using Difference Values of Border Pixels. *Congress on Image and*

- Signal Processing*, 2008, (CISP '08), vol. 3:565-568, doi: 10.1109/CISP.2008.659.
11. Hsieh Ching-Tang, Chen Yen-Liang and Hsu Chih-Hsu. Fast Image Restoration Method Based on the Multi-Resolution Layer. *Tamkang Journal of Science and Engineering*, 2009, 12(4):439-448.
 12. Park Jiho, Park Dong-Chul, Marks Robert J, Fellow, El-Sharkawi Mohamed. Recovery of image blocks using the method of alternating projections. *IEEE Transactions on Image Processing*, 2005, 14(4):461-474. DOI: 10.1109/TIP.2004.842354.
 13. Hyuna Seung Hwa, Eomb Il Kyu, Kim Yoo Shin. Directional Filtering for Block Recovery Using Wavelet Features. *Proc. SPIE 5960, Visual Communications and Image Processing*, 2005, 59600Z (31 July 2006); doi: 10.1117/12.631414.
 14. Kokoshkin AV, Korotkov VA, Korotkov KV, Novichihin EP. Using Fourier spectrum for retouching and restoration missing parts of the image which were deformed by instrumental function (in Russ.). *Journal of Radioelectronics* [electronic journal], 2016, №7. Available at: <http://jre.cplire.ru/jre/jul16/4/text.html>.
 15. Kokoshkin AV, Korotkov VA, Korotkov KV, Novichihin EP. Application of digital image processing methods for the goal of restoration of fine art objects (in Russ.). *Journal of Radioelectronics* [electronic journal], 2018, № 9. Available at: <http://jre.cplire.ru/jre/sep18/16/text.pdf>. DOI 10.30898/1684-1719.2018.9.16.
 16. Kokoshkin AV, Korotkov VA, Korotkov KV, Novichihin EP. Comparison of the results of filling the gaps in the images by interpolation by sequential calculation of the Fourier spectrum and spline interpolation. *Reports XII All-Russian Scientific and Technical Conference "Radar and Radio Communication"*, November 26-28, 2018, Moscow, Kotelnikov IRE RAS, 2018:14-18.
 17. Kokoshkin AV, Korotkov VA, Novichihin EP. Comparison of interpolation methods when achieving superresolution of images based on the analysis of several frames. *Radioelektronika, Nanosistemy, Informacionnye Tehnologii (RENSIT)*, 2019, 11(1):85-91; doi: 10.17725/rensit.2019.11.085.
 18. Kokoshkin AV, Korotkov VA, Korotkov KV, Novichihin EP. Choosing an interpolation method for implementing superresolution of images. *Proceedings XIII All-Russian Scientific and Technical Conference "Radar and Radio Communication"*, November 25-27, 2019, Moscow, IRE RAS, pp. 200-204.
 19. Kokoshkin AV, Korotkov VA, Korotkov KV, Novichihin EP. Retouching and restoration of missing image fragments by means of the iterative calculation of their spectra. *Computer Optics*, 2019; 43(6):1030-1040. DOI: 10.18287/2412-6179-2019-43-6-1030-1040.
 20. Ashkenazy AV. *Spline surfaces. Fundamentals of the theory and computational algorithms: Proc. allowance* (in Russ.). Tver, Publishing House of Tver State University, 2003, p. 82.
 21. Nesterenko EA. The ability to use spline surfaces for constructing surfaces based on the results of shooting (in Russ.). *Notes of the Mining Institute*, 2013. 204:127-133.
 22. Cherepenin VA, Zhuravlev AV, Chizh MA, Kokoshkin AV, Korotkov VA, Korotkov KV, Novichihin EP. Reconstruction of Subsurface Radio Holograms Fully and Partially Measured by Different Methods. *Journal of Communications Technology and Electronics*, 2017, 62(7):780-787. DOI: 10.1134/S1064226917070038.

23. Kokoshkin AV, Korotkov VA, Korotkov KV, Novichihin EP. Blind image restoration distorted by smear and defocusing when the shape and parameters of the AF are unknown (in Russ.). *Journal of Radioelectronics* [electronic journal], 2014, №9. Available at: <http://jre.cplire.ru/jre/sep14/8/text.html>.
24. Zrazhevsky AYu, Kokoshkin AV, Korotkov VA, Korotkov KV, Novichihin EP. The universal reference spectrum and its use for finding the hardware function of distortion and image restoration. *Proc. VIII All-Russian Scientific and Technical Conference "Radar and Radio Communication"*, November 24-26, 2014, Moscow, Kotelnikov IRE RAS, 2014, pp. 191-195.
25. Sizikov VS, Stepanov AB, Mezhenin AV, Burlov DI, Eksemplyarov RA. Determining image-distortion parameters by spectral means when processing pictures of the earth's surface obtained from satellites and aircraft. *Journal of Optical Technology*, 85(4):203-210.
26. Kokoshkin AV, Korotkov VA, Korotkov KV, Novichikhin EP. Reconstruction of images distorted by defocusing and blurring without determining the type and parameters of the hardware function. *J. of Communications Technology and Electronics*, 2019, 64(6):569-580. DOI: 10.1134/S1064226919060044.

DOI: 10.17725/rensit.2020.12.507

Analyze and Evaluate the Performance Velocity Control in DC Motor

Hazim M. Alkargole, Abbas S. Hassan, Raof T. Hussein

Mustansiriyah University, <https://uomustansiriyah.edu.iq/>

Palestine Street, P.O. Box: 14022, Baghdad, Iraq

E-mail: h.mohammed2@uomustansiriyah.edu.iq, abassalmhyanj@uomustansiriyah.edu.iq, raoofatal@uomustansiriyah.edu.iq

Received April 05, 2020, peer-reviewed May 18, 2020, accepted May 29, 2020

Abstract: A mathematical model of controlling the DC motor has been applied in this paper. There are many and different types of controllers have been used with purpose of analyzing and evaluating the performance of the of DC motor which are, Fuzzy Logic Controller (FLC), Linear Quadratic Regulator (LQR), Fuzzy Proportional Derivative (FPD), Proportional Integral Derivative (PID), Fuzzy Proportional Derivative with integral (FPD plus I), and Fuzzy Proportional Integral (FPI) with membership functions of 3*3, 5*5, and 7*7 rule bases. The results show that the (FLC) controller with 5*5 rule base provides the best results among all the other controllers to design the DC motor controller.

Keywords: PID controller, LQR controller, Fuzzy controller, DC Motor

UDC 681.5.01:519.6

Acknowledgment: The authors thank Mustansiriya University for its full support of this work. It is one of the Iraqi public universities named after the Mustansiriya school, which was founded in the time of the Abbasids in Baghdad in 1233 by the Caliph Al-Mustansir Billah. It was an important scientific and cultural center. Located in the capital Baghdad. Includes 13 faculties. Founded in 1963. www.uomustansiriyah.edu.iq.

For citation: Hazim M. Alkargole, Abbas S. Hassan, Raof T. Hussein. Analyze and Evaluate the Performance Velocity Control in DC Motor. *RENSIT*, 2020, 12(4):507-516. DOI: 10.17725/rensit.2020.12.507.

CONTENTS

- | | |
|---|--|
| <ol style="list-style-type: none"> 1. INTRODUCTION (508) <ol style="list-style-type: none"> 1.1. STATEMENT OF PROBLEM AND METHODOLOGY OF SOLUTION (508) 2. CONTROLLERS DESIGN (508) <ol style="list-style-type: none"> 2.1. CLASSICAL PID (508) <ol style="list-style-type: none"> 2.1.1. PROPORTIONAL CONTROLLER (P-CONTROLLER) (508) 2.1.2. PROPORTIONAL PLUS DERIVATIVE CONTROLLER (PD-CONTROLLER) (508) 2.1.3. PROPORTIONAL PLUS INTEGRAL CONTROLLER (PI-CONTROLLER) (509) 2.1.4. PROPORTIONAL PLUS INTEGRAL PLUS DERIVATIVE CONTROLLER (PID CONTROLLER) (509) <ol style="list-style-type: none"> 2.1.4.1. TUNING OF PID-CONTROLLER (509) 2.2. LINEAR QUADRATIC REGULATOR (LQR) CONTROLLER (510) 2.3. FUZZY LOGIC CONTROLLER DESIGN (510) <ol style="list-style-type: none"> 2.3.1. MEMBERSHIP FUNCTION OF 7*7 RULE BASE (511) | <ol style="list-style-type: none"> 3. RESULTS AND DISCUSSION (512) <ol style="list-style-type: none"> 3.1. RESPONSE OF DC MOTOR (512) 3.2. TUNING USING PID CONTROLLER RESULTS (512) 3.3. LQR CONTROLLER RESULTS (512) 3.4. FUZZY LOGIC CONTROLLER RESULTS (513) 3.5. PI-FUZZY LOGIC CONTROLLER RESULTS (513) 3.6. PD FUZZY LOGIC CONTROLLER RESULTS (513) <ol style="list-style-type: none"> 3.6.1. MEMBERSHIP FUNCTION OF 7*7 RULE BASE RESULTS (513) 3.7. PID FUZZY LOGIC CONTROLLER RESULTS (514) <ol style="list-style-type: none"> 3.7.1. MEMBERSHIP FUNCTIONS OZ. 5*5, AND 7*7 RULES BASE RESULTS (514) 4. DISCUSSION (514) 5. CONCLUSION (515) 6. REFERENCES (515) |
|---|--|

1. INTRODUCTION

Last decades, due to system instability, many controllers have been explored, such as PID controller, which refers to the family of controllers with various configurations of the Proportional, Integral and Derivative terms. The most common controller that has been used in industry is Conventional PID controllers due to its simplicity in tuning and effectiveness and design for general linear systems with convenient low cost and implementation. The Linear Quadratic Regulator (LQR) is a method that runs optimally controlled response in order to enable the high-performance design and the closed-loop systems [1,2]. Recently, fuzzy logic is becoming one of the most important for producing and developing a control system. Fuzzy logic has solved many complex problems as it is simple, easily maintained, and inexpensive. The mathematical model in this case is very important to build a good controller for controlling any complex system. The differential equations are the most common system that is used for discrete time systems or continuous time systems. As it's known that the nonlinear physical system could be designed based on the collected data and system identification methods but practically is difficult and challenging due to complexity particularly for conventional control design [3,4]. Furthermore. The quantitative and qualitative information could be used by the Fuzzy controller. Qualitative information has been collected from common knowledge and the expert operator strategy [5,6]. The fuzzy logic control has not been used for linear systems as most of the them based on misconceptions as mentioned above. However, the linear controllers such as PID have been able to solve any kind of control problem with less cost, effort and time. Therefore, the PID has to be tried first [2]. The main characteristics of fuzzy controllers is:

- 1) a fuzzy controller is cheaper to be developed rather than developing the other controllers such as model-based;
- 2) the fuzzy controllers are covering a wider range of operating conditions rather than LQR and PID;
- 3) the fuzzy controller is easy to be designed as it is not complicated and easy to understand [7].

1.1. STATEMENT OF PROBLEM AND METHODOLOGY OF SOLUTION

The DC motor speed may change due to disturbance present surrounding it. This will make the desired speed sometimes change and will be not maintain. By using classical Proportional Integral and Derivative (PID), Linear Quadratic Control (LQR), and soft computing methods, such Fuzzy Logic Controller (FLC), PI-Fuzzy, PID-Fuzzy and PD-Fuzzy controllers, the speed could be minimized. The main aim of this paper is analyzing and evaluating the speed control performance of DC motor with various controllers. These controllers are Fuzzy-PID, PID controller, LQR controller, Fuzzy Logic controller, Fuzzy-PI, Fuzzy-PD controllers. A comparison is made among these controllers in order to see which one among them give best performance [8].

2. CONTROLLERS DESIGN

2.1. CLASSICAL PID

The classical PID controller could have different configurations which are integral, proportional and others. The most common controller that is used in the industries is conventional PID controller because it is very simple in tuning and design and effectiveness. As is known that the using of P-controller is to reduce input and output phase shift signal, and detection of tracking error [1,2,4].

2.1.1. PROPORTIONAL CONTROLLER (P-CONTROLLER)

It is one of the linear feedback control system. this system is not for an on/off system which is simpler than the PID control system. The signal of this controller is proportional to the error signal as it is the difference between the process set point and variable (the proportional controller is the multiplication of the proportional and the error signal. The following formula is how the P-controller calculated mathematically [5,6]:

$$P_{out} = K_p e(t), \quad (1)$$

$$e(t) = SP - PV, \quad (2)$$

where P_{out} – the proportional controller output, K_p – the gain of the proportional, $e(t)$ – immediate process error at time 't', SP – Set point, PV – the variable of the process.

2.1.2. PROPORTIONAL PLUS DERIVATIVE CONTROLLER (PD-CONTROLLER)

The PD-regulator can lessen the greatest overshoot yet may hold a consistent state following blunder. The utilization of subsidiary control is constrained

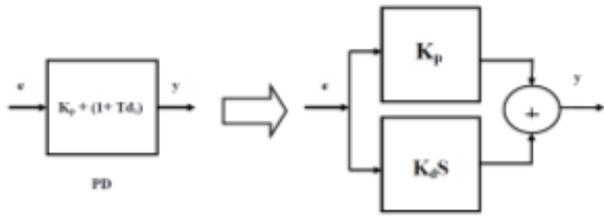


Fig. 1. PD controller [1].

from the start, subordinate control looks appealing as appeared in Fig. 1. It assists with lessening the time required to balance out a blunder. In any case, it won't expel counterbalance. The control signal from subsidiary activity stops when the blunder quits changing, which won't really be at the set point. Essentially it is likewise constrained to slow acting procedures When its utilizing for quick acting procedure, for example, stream. Control flags because of subsidiary activity will frequently drive the control valve to limits following very little however steep (enormous de dt changes in input). Mathematically proportional plus derivative (PD) control is expressed as [1,2,4]:

$$m = K_p \left(e + TD \frac{d_e}{d_t} \right) + b, \quad (3)$$

where m – controller signal, K_p – controller gain, TD – derivative time, e – error, b – constant.

2.1.3. PROPORTIONAL PLUS INTEGRAL CONTROLLER (PI-CONTROLLER)

A proportional plus integral (PI) controller contains the transfer function: $Gc(s) = K_p + K_i/s$ as shown in Fig. 2. The task of this controller is to tune the control parameters K_p and K_i to achieve better control. By combining the P and I controllers, the system performance will be better since there are two parameters to tune [5,6].

2.1.4. PROPORTIONAL PLUS INTEGRAL PLUS DERIVATIVE CONTROLLER (PID CONTROLLER):

This controller is fast and settling in time, and no steady state error as shown in Fig. 3 below. Despite the fact that PID controller could be logically designed and pre-tuned for provided lower-order

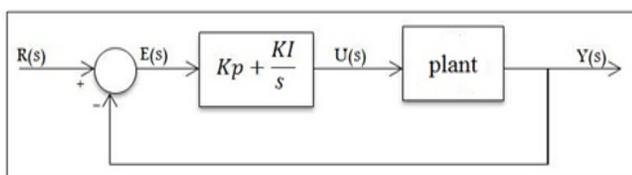


Fig. 2. PI Controller in unity Feedback [1].

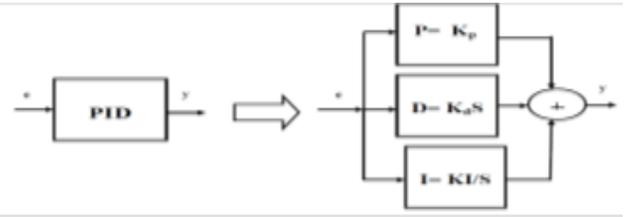


Fig. 3. PID Controller Structure [1].

linear system. Also, it could be physically worked for most systems that include higher-order system, for example, Nonlinearity and vulnerabilities., Ziegler-Nichols and Cohen-Coon of the Taylor Instrument Company started, well-known heuristic guidelines for experimental design and tuning strategies, have been utilized so as to have powerful controllers. However, the tuning of the system is consistently a test in the best in class of PID controller structure. This issue turns out to be more significant and basic, especially, when issues including stability, specifications and performance that are considered [1,2].

2.1.4.1. TUNING OF PID-CONTROLLER

In order to take advantage of this feature, this console systematically introduces the design process based on Ziegler and Nichols' approach. In the Ziegler-Nichols method, parameter setting is based on the stability limits of the system. The derivative and complementary terms are initially taken out of the system and the relative gain is increased to the critical oscillation point. A desirable function of industrial automation is to modify PID control parameters, which mainly includes control gain and possibly also some measurement parameters used in the controller, depending on changes to the systems (installations, process) and their working environments [1,2].

• TUNING WITH ZIEGLER-NICHOLS METHODS (THE FIRST METHOD)

This method of tuning PID controller by obtaining experimentally the step response of the position control of a DC Motor as shown in Fig. 4. The curve is characterized by two constants, the time constant

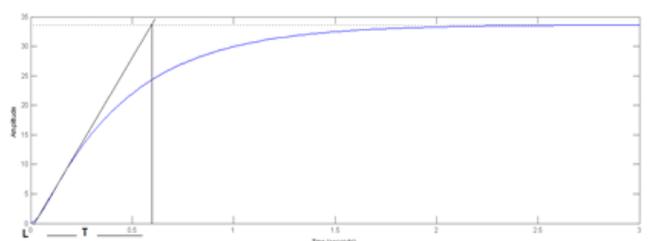


Fig. 4. The step response by using Ziegler-Nichols methods.

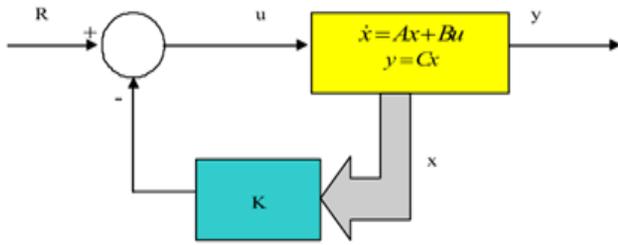


Fig. 5. LQR system design [1].

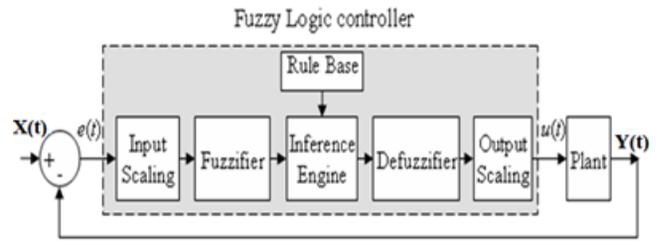


Fig. 6. Fuzzy logic controller structure (feedback control system) [10].

T and the delay time L , the time constant and the delay time are determined by drawing a tangent line to the point of incidence of the curve and determining the intersection of the tangent line with the time axis and the line $c(t) = k$ as shown in Fig. 5. The transfer function $c(s)/u(s)$ can be approximated by a first-order system with the transfer shift as follows: $C(s)/u(s) = ke^{-Ls}/Ts + 1$, Ziegler-Nichols suggested assigning values K_p , T_i and T_d according to the formula given in the Table 1 below [9]:

2.2. LINEAR QUADRATIC REGULATOR (LQR) CONTROLLER:

LQR controller is classified as an ideal control system. Fig. 6 below shows the configuration of the designed LQR. LQR is the best theory of pole placement; In theory, pole placement method involves determining the desired location of the electrodes and relocating the system poles position to the desired location of the electrodes to achieve the desired response of the system. The LQR algorithm determines the optimal location of columns based on two cost functions. To find the optimum gain, we must first determine the optimum performance index, and then solve the Riccati algebraic equation. LQR has no specific solution for defining the cost function to achieve the optimum gain and the cost function must be determined iteratively. The state space representation of a linear system is [1]:

$$\dot{X} = Ax + Bu, \quad y = Cx, \tag{4}$$

where x – state vector, y – output vector, u – input vector, A – state matrix, B – input matrix, and C – output matrix.

Table 1

Ziegler-Nichols values of K_p , T_i and T_d

The controller type	K_p	T_i	T_d
PI	$0.9T/$	$L/$	0
P	T/L	∞	0
PID	$1.2T/L$	$2L$	$0.5L$

The feedback control equation, that is minimizes the cost value [1]:

$$u = -k x(t). \tag{5}$$

To find the gain values, $K = K1$ and $K2$ only because the system is a second order, it must solve the Riccati equation below:

$$A^T X + X A - X B R^{-1} B^T X + Q = 0. \tag{6}$$

The design procedures can be found from [2]. The matrices Q and R has been selected by trial and error. Some text books and literatures use $Q = C^*C$ and $R = 0.1$ or 0.2 .

2.3. FUZZY LOGIC CONTROLLER DESIGN

Fluffy Logic has been effectively applied to many applications. The most normally utilized controller is the PID controller. The PID controller has been provided by the fuzzy logic system in order to have better controlling [10,11,12]. This is better for control systems because it is not easy to model the fuzzy logic. Recently, the fuzzy logic and PID controller becomes one of the most efficient systems for developing advanced control systems. Furthermore, any other requirements can be executed in controllers that are simple, inexpensive and easy to maintain [13,14,15]. The fuzzy control uses only a small portion of the available fuzzy mathematics. This part is fairly simple mathematically and conceptually easy to understand. This article introduces some basic concepts, terminology, arithmetic for ambiguous combinations, and fuzzy logic. The fuzzy controller as shown in Fig. 6 consists of four main mechanisms: a – Rule Base, b – Inference Mechanisms, c – Fuzzification Interface, D – Defuzzification Interface [16,17,18,19].

In general, two signals should be considered as input signals, which are change of error ($\Delta e(s)$) and error ($e(s)$) signals. These signals represent the PD gain. The change of error signal could be obtained by multiplying the error signal with the delay signal and then subtracting it from the original error signal

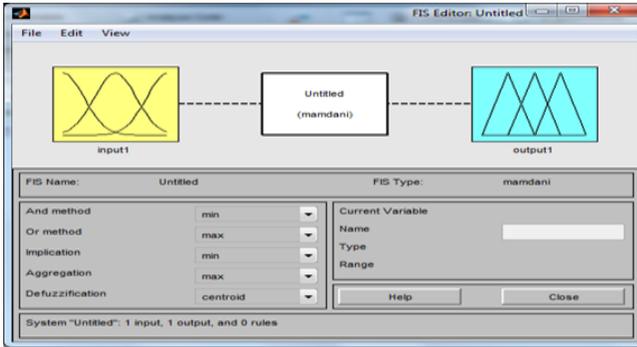


Fig. 7. Input/output Fuzzy Sets [10].

[20,21]. While, the error signal could be obtained by subtract the plant (Y) output from the output ($Y(t)$). Fig. 7 shows the simulation of symmetric triangular and the singleton fuzzy sets [22,23]. Additionally, in this article the table of 9, 25 and 29 rules have been used with the group of 3, 5, 7, as discussed later in the next section.

The fuzzy logic controller action be expressed with membership function and simple "if-then" rules to the position control of an DC Motor with a 3×3 , 5×5 and 7×7 rule base in chapter three.

2.3.1. MEMBERSHIP FUNCTION OF 7*7 RULE BASE

In this design 49 rules have been used based on the seven triangular membership, Table 2 shows a 7×7 rules base for the purpose of developing the DC motor speed system.

The membership function of 7×7 rules is shown in Fig. 8 below

The response from applying fuzzy logic controller to the speed control of a DC motor for the 7×7 rule base will be shown in chapter three.

2.4. DC MOTOR

Generally, the motors have been used to convert the electrical energy into mechanical energy. This conversion has been done by two very interactive magnetic fields, which are stator and rotor. The DC motor has the ability of providing very high

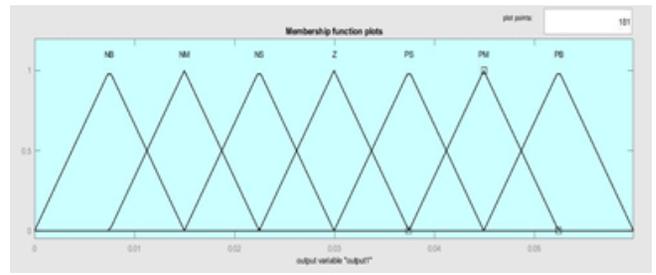
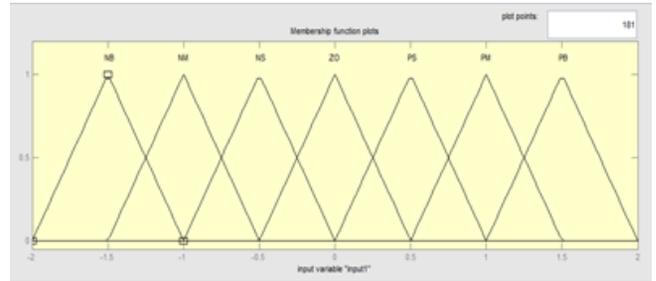


Fig. 8. The Membership function of 7×7 (a) Inputs variables, (b) Output Variable.

torque, and easy to miniaturize. The electromagnetic induction machines have been discovered by Faraday, Gauss, and Oersted in early 1800's [24]. Moreover, two kind of converting machine have been used recently, the machine that converts the mechanical to electrical called generator, and the machine that converts the electrical to mechanical called motor. Fig. 9 shows the DC motor with its equivalent circuit [25].

Table 2

Fuzzy Rule Base [10].

e/de	NB	NM	NS	Z	PS	PM	PB
NB	NB	NB	NM	NM	NS	NS	Z
NM	NB	NM	NM	NS	NS	PS	PS
NS	NM	NM	NS	NS	Z	PS	PS
Z	NS	NS	NS	Z	PS	PS	PM
PS	NS	Z	PS	PM	PS	PM	PM
PM	NS	Z	PS	PS	PM	PM	PB
PB	Z	PS	PS	PM	PM	PB	PB

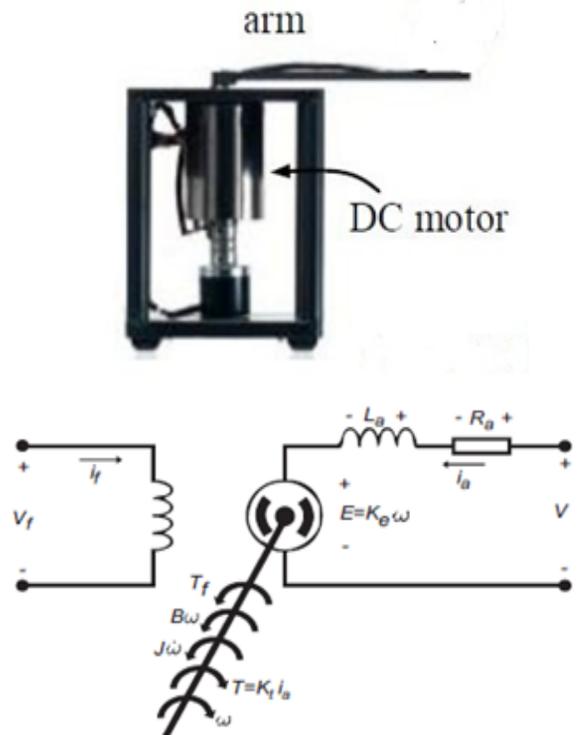


Fig. 9. DC motor with its equivalent circuit [25].

According to the Kirchhoff's voltage law, consequently the dynamic of DC motor are expressed by the following equations [8]:

$$V = Ra La + La \frac{di}{dt} + ke w, \tag{7}$$

$$Kt ia = J \frac{dw}{dt} + B w + Tf. \tag{8}$$

The actual parameters of a DC motor from [8] are: Ke – the back emf constant = $14.7e 3 N.m/A$, Tf – the load torque = 0, Ra – the armature resistance = 4.67Ω , La – the armature inductance = $170 e-3 H$, w – the rotor angular speed, Kt = the torque constant = $14.7e 3 V.sec/rad$, B – the viscous (damping) friction = $47.3e 6 N.m/rad/sec$.

And by substituting equation (8) into equation (7) and taking Laplace transformation, the transfer function of speed to voltage is

$$\frac{W(s)}{V(s)} = \frac{2030}{S^2} + 28.58S + 60.34, \tag{9}$$

where $W(s)$ and $V(s)$ are the output and the input of the system respectively. simply it has obvious that the denominator roots are equal to -26.2843 and -2.2957, that means the system is stable but it will be shown in section 3.2 the system suffers from oscillation with overshoot and steady state error.

3. RESULTS AND DISCUSSION

This chapter contains all the result of the graduation project which can be obtained by using Mat lab toolbox to design fuzzy logic system like (PI- Fuzzy, Fuzzy, PD-Fuzzy and PID-Fuzzy) controllers and applying these controllers for a DC motor speed control and compares the results with the classical PID and LQR controllers.

3.1. RESPONSE OF DC MOTOR

By taking the DC motor closed-loop transfer function as in equation (9) with unity feedback and step input the results are show in Fig. 10.

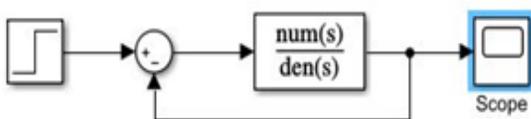


Fig. 10a. DC motor in unity feedback.



Fig. 10b. DC motor in unity feedback.

Table 3

Result of applying the Ziegler-Nichols (The First Method)

K_p	T_i	$K_i = K_p/T_i$	T_d	$K_d = K_p * T_d$
17.38	0.077	225.714	0.019	0.33

The system suffers from oscillation with maximum overshoot, $M_p = 31.7\%$, Rising Time = 0.028 sec., Settling Time = 0.245 sec., Steady State Error (es.s) after oscillation = 0.04.

3.2. TUNING USING PID CONTROLLER RESULTS

In this section the tuning of classical PID controller include with Ziegler-Nichols (The first method). The result in table (3) was obtained by applying the Ziegler-Nichols (The first Method) to tuning DC motor speed based on classical PID controller.

The system information are $T_r = 0.027$ sec., $T_s = 0.24$ sec., $M_p = 14.3 \%$, $es.s = 0.01$.

3.3. LQR CONTROLLER RESULTS

It has been shown that from section 2.4 to achieve LQR controller it must solve the Riccati equation (6). Riccati equation can be easily programmed for a computer, or solved using MATLAB function lqr, that is:

$$[k_s, e] = lqr(A, B, Q, R). \tag{10}$$

The matrices A , B , and C in Riccati equation must be written in a Jacobian matrices [2], from equation (10) as follows:

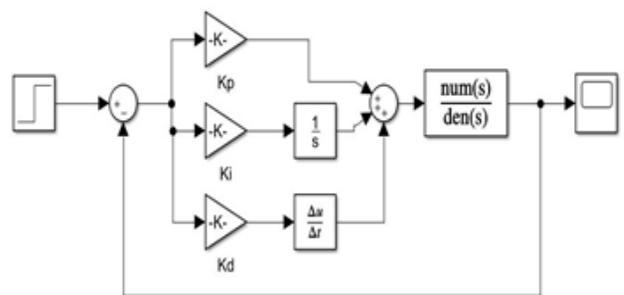


Fig. 11a. PID controller.

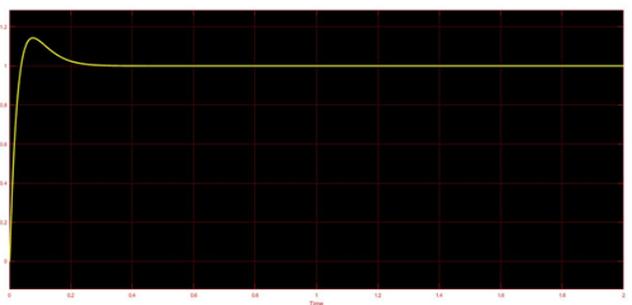


Fig. 11b. The response of PID controller.

$$A = \begin{bmatrix} 0 & 1 \\ -60.34 & -28.58 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 2030 \end{bmatrix}, C = [1 \ 0].$$

The state-feedback gain is K , LQR returns the solution S of the associated Riccati equation and e is the eigen values. This type of controller needs prefilter to remove the offset between input and output. By simulating the Matlab program for the DC motor system with LQR controller, the optimal values of gains by adding the pre-filter are: $k_1 = 0.9707$ and $k_2 = 0.3616$. The response will be as follows:

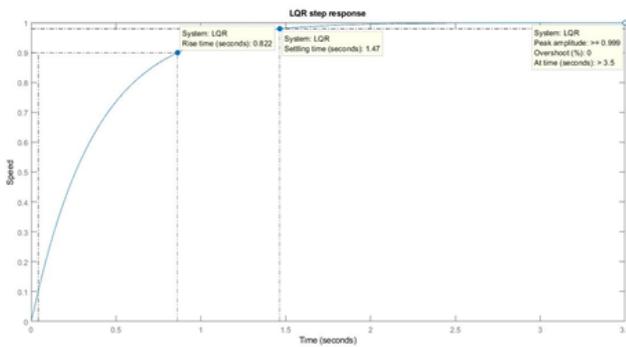


Fig. 12. The response of LQR controller.

The system information are $T_r = 0.822$ sec., $T_s = 1.47$ sec., $M_p = 0\%$, $es.s = 0.01$.

3.4. FUZZY LOGIC CONTROLLER RESULTS

The fuzzy logic controller action can be expressed with membership function and simple "if-then" rules to the speed control of DC motor which will implemented with a 3*3, 5*5 and 7*7 rule base as shown in Fig. 12.2 below:

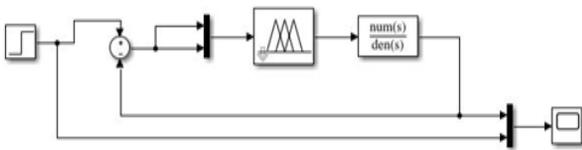


Fig. 12.2. Fuzzy logic controller.

3.4.1. MEMBERSHIP FUNCTION OF 7*7 RULE BASE RESULTS:

By using group of seven triangular membership functions input/output rule an variables table of 49 rules that are used in this design a 7*7 rule base as shown in Table 2 in section two to develop the speed control of DC motor system we obtained this response.

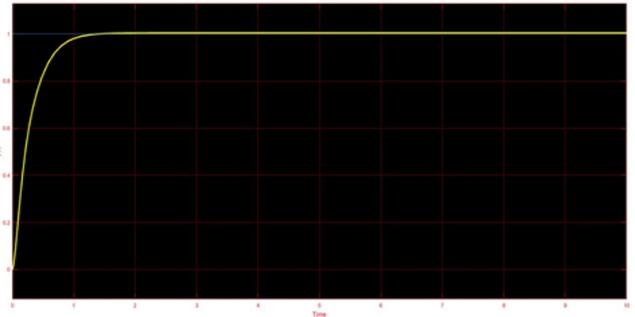


Fig. 13. The response of Fuzzy Logic Controller 7*7 rule base.

The system information are $T_r = 0.58$ sec., $T_s = 1.32$ sec., $M_p = 0\%$, $es.s = 0$.

3.5. PI-FUZZY LOGIC CONTROLLER RESULTS

In this section applying PI controller with fuzzy logic controller to the speed control of DC motor which will implemented with a 3*3, 5*5 and 7*7 rule base as shown in Fig. 14 below, where the values of K_p and K_I were obtained from optimal PID controller, section 3.3:

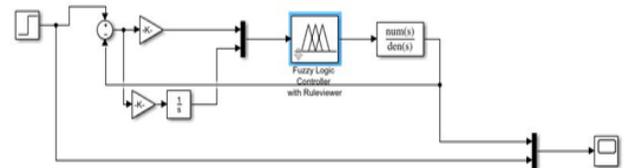


Fig. 14. PI-Fuzzy logic controller.

3.6. PD FUZZY LOGIC CONTROLLER RESULTS

The fuzzy logic controllers action can be expressed with membership function and simple "if-then" rules to the speed control of DC motor which will implemented with a 3*3, 5*5 and 7*7 rule base as shown in Fig. 15 below, where the values of K_p and K_d were obtained from optimal PID controller, section 3.3:

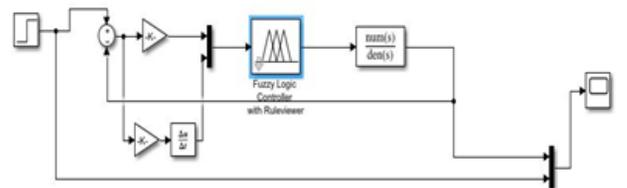


Fig. 15. PD-Fuzzy logic controller.

3.6.1. MEMBERSHIP FUNCTION OF 7*7 RULE BASE RESULTS:

By using group of seven triangular membership functions input/output variables and rule table of 49 rules were used in this design a 7*7 rule base was defined in table (2) in chapter two to develop the speed control of DC motor system we obtained this response:

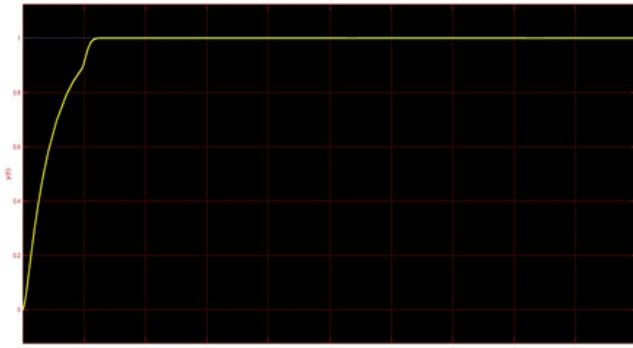


Fig. 16. The Response of PD-Fuzzy Logic Controller 7*7 rule base.

The system information are $T_r = 0.899$ sec., $T_s = 1.8$ sec., $M_p = 0\%$, $es.s = 0$.

3.7. PID FUZZY LOGIC CONTROLLER RESULTS

In this section applying I controller as assistant to PD controller with fuzzy logic controller to the DC motor speed control which will implemented with a, 5*5 and 7*7 rule base as shown in **Fig. 17** below, where the values of K_p , K_d , and K_I were obtained from optimal PID controller, section 3.3:

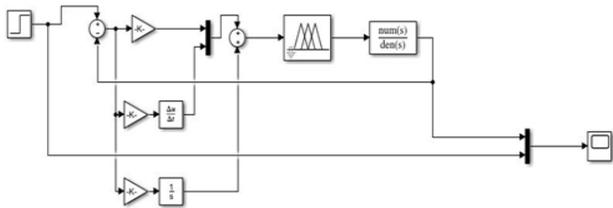


Fig. 17. PID-Fuzzy controller.

3.7.1. MEMBERSHIP FUNCTIONS OZ. 5*5, AND 7*7 RULES BASE RESULTS

The response from applying PID-Fuzzy Logic controller to the speed control of DC motor for the 3*3, 5*5, and 7*7 rules base were defined in table (2, 3 and 4) in the previous sections to develop the speed control of a DC motor system we obtained the same response for all rules, that as shown in **Fig. 18** below:



Fig. 18. The response of PID-Fuzzy Logic controller 5*5, 7*7 rules base.

The system information for all rules are the same and are as follows $T_r = 0.948$ sec., $T_s = 2$ sec., $M_p = 0\%$, $es.s = 0$.

4. DISCUSSION

To discuss the results which are obtained by using MATLAB toolbox to design fuzzy logic system controllers by applying these controllers for speed control of a DC motor, it is necessary to compare the results of the PID and LQR controllers with fuzzy logic controllers. The response of the DC motor system has approximately the same rise time (T_r) and settling time (T_s) and equal to 0.028 sec. and 0.24 sec. respectively before and after applying PID controller, but the maximum overshoot (M_p) decrease from 31.7% to 14.3% with minimum steady state error ($es.s$) with PID controller. LQR controller has larger $T_r = 0.822$ sec. and $T_s = 1.47$ sec. compared with the PID controller but it removed the maximum overshoot (M_p).

For the Fuzzy Logic Controller (FLC), it has been shown that from the table below both of PI-Fuzzy, PD-Fuzzy, and PID-Fuzzy controllers have the same $T_r = 0.9$ sec., $T_s =$ between 1.5 to 2 sec., $M_p = 0$, and $es.s = 0$, but the Fuzzy Logic Controller (FLC) with

Table 4

Discussion of result				
	Rise Time (sec.)	Settling Time (sec.)	Overshoo (%)	Steady State Error
DC motor	0.028	0.24	31.7	0.01
PID Tuning With Ziegler-Nichols	0.027	0.24	14.3	0.01
LQR	0.822	1.47	0	0
Fuzzy logic controller 3*3	0.46	0.9	0	0
Fuzzy logic controller 5*5	0.38	0.7	0	0
Fuzzy logic controller 7*7	0.58	1.32	0	0
PI-Fuzzy logic Controller 3*3, 5*5, and 7*7	0.94	2	0	0
PD-Fuzzy logic Controller 3*3, 5*5, and 7*7	0.93	1.5 to 1.8	0	0
PID-Fuzzy logic Controller 3*3, 5*5, and 7*7	0.94	2	0	0

5*5 rule base is better than others, which has $T_r = 0.3$ sec., $T_s = 0.7$ sec., $M_p = 0\%$, and $e_{s.s} = 0$, so the later controller is better than LQR controller.

5. CONCLUSION

This paper presents the analysis and performance evaluation of speed control of a DC motor as follows:

1. Classical PID controller provides a higher execution for DC motor, but the PID controller does not remove the maximum overshoot compared with other controllers.
2. PI-Fuzzy, PD-Fuzzy, and PID-Fuzzy controllers provide a better control performance by removing maximum overshoot and steady state error, but with larger rise time and settling time compared with PID controller.
3. Fuzzy Logic Controller with 5*5 rule base improved the system much more than PID, PI-Fuzzy, PD-Fuzzy, and PID-Fuzzy controllers by increasing the speed of system and decrease time delay.

REFERENCES

1. Westphalia L. *Handbook of Control Systems Engineering*. John-Wiley Inc.,1995.
2. D.K.Anand, R.B.Zmood. *Introduction to Control Systems*. 3rd Edition, Butterworth Heinemann Ltd, 1995.
3. Fakhruddin H. Ali and Hussein Mohammed Mahmood. LABVIEW FPGA Implementation of PID Controller for DC Motor Speed Control. *1st International Conference on Energy, Power and Control (EPC-IQ)*, pp.1-6. IEEE, 2010.
4. Guishan S. Chen. Particle Swarm Optimization for PID Controllers with Robust Testing. *International Conference on Machine Learning and Cybernetics*, Hong Kong, 2007, pp. 956-961. IEEE, 2007.
5. Sujit Kumar and V. Kumar. Speed Control DC Motor using Fractional Fuzzy based PID Controller. *2012 IEEE International Conference on Engineering and Technology (ICETECH)*, Coimbatore, pp. 1-6, IEEE, 2012.
6. P. vas and M.J.E. Bulale. Design and Implementation of Fuzzy based PID Controller. *Industrial Technology, IEEE ICIT '02. 2007 IEEE International Conference on*, vol. 1, pp. 220-225, IEEE, 2007.

7. Wang Xiao kan and Sun Zhong liang. Design and Research based Fuzzy-PID Parameters Self-Tuning Controller with MATLAB. *International Conference on Advanced Computer Theory and Engineering, ICACTE '08*, pp.996-999, IEEE, 2008.
8. Munaz Cesar and A. Abdelrahman. Speed Control of DC Motor using Fuzzy Logic Controller. *2008 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE)*, Khartoum, pp.1-8, IEEE, 2008.
9. RG Kanojiya and PM Meshram. Tuning of PID Controller using Ziegler-Nichols Method for Speed Control of DC Motor. *IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM-2012)*, Nagapattinam, Tamil Nadu, 2012, pp. 117-122, IEEE, 2012.
10. E. Cox. Fuzzy Fundamentals. *IEEE spectrum*, 29(10), October 1992.
11. H. Ying. *Fuzzy Control and Modeling, Analytical Foundations and Applications*. Institute of Electrical and Electronic Engineers Inc., USA, 2000.
12. KM Passino and Stephen Yurkovich. *Fuzzy Control*. Addison-Wesley Longman Inc., USA, 1998.
13. *Math works. Fuzzy Logic Toolbox User's Guide*. Math works, Inc., 1999.
14. Y Huang and S Yasunobu. Design of Neuro-Fuzzy Controller. *9th International Symposium on Communications and Information Technology*, Icheon, pp. 808-813, IEEE, 2009.
15. MM Kamal, L Mathew and S Chatterji. Speed Control of Brushless DC Motor using Genetic Algorithm. *2014 Students Conference on Engineering and Systems*, Allahabad, pp. 1-5, IEEE, 2014.
16. Zhang Xiao Ming and Long Shi Liu. Simulation Study of Fuzzy-PID Controller for DC Motor based on DSP. *2012 International Conference on Industrial Control and Electronics Engineering*. pp.3401-3405, IEEE, 2012.
17. SG Kadwane and S Gupta. Real Time Based PI-Fuzzy Controller for DC Servomotor. *2006 International Conference on Power Electronic, Drives and Energy Systems*, New Delhi, 2006, pp.1-4.
18. Zhao Yongjuan and Pan Yutitan. The Design and Simulation of Hybrid Fuzzy-PID Controller. *International Forum on Information Technology and Applications (IFITA)*, 2010, vol. 3, pp. 95-98, IEEE, 2010.

19. JP Rey and MM Dalvi. A Numerical-based Fuzzy-PID Controller applied to a DC Drive. *IEEE International Symposium on Industrial Electronics, Symposium Proceedings, ISIE'*, pp. 429-434, IEEE, 2009.
20. M Surya Kalawati and Mitra R. Comparative Analysis of Robustness of Optimally Offline Tuned PID Controller and Fuzzy Supervised PID Controller. *2014 Recent Advances in Engineering and Computational Sciences (RAECS)*, pp.1-6, 2014.
21. Salim Jyoti Ohri and Naveen. Speed Control of DC Motor using Fuzzy Controller in LABVIEW WSEAS. *IEEE Conference on Transactions on Systems and Control*, Vol. 10, pp. 2224-2856. IEEE, 2015.
22. Vishal Mehra and V Kala. Design and Simulation of FOPID Composite Parameters Controller with MATLAB. *2010 International Conference on Computer Design and Applications (ICCCA)*, vol. 4, pp. V4-308-V4-311, IEEE, 2010.
23. Ahamadraza Ahmadi and Kumari SU. Optimal Fuzzy-PID Controller. *International Conference on Power Signals Control and Computations (EPSCICON)*, pp. 1-6, IEEE, 2012.
24. Austin Hughes. *Electric Motors and Drives Fundamentals, Types and Applications*. 3rd Edition, Linacre House, Jordan Hill, Oxford OX2 8DP, 2006.
25. Mohammed Salah, Mohammed Abdelati. Parameters Identification of a Permanent Magnet DC Motor. *International Conference on Modelling Identification and Control (MIC 2010)* At: Aust.

DOI: 10.17725/rensit.2020.12.517

Information Processing Algorithms in Aviation-Based Radioelectronic Surveillance Systems

Ivan D. Biryukov, Pavel V. Buchuchan, Polina I. Timoshenko

A.I. Berg Central Research Radio Engineering Institute, <http://cnirti.ru/>
Moscow 107078, Russian Federation

E-mail: biryukov68@gmail.com, p.v.buchuchan@yandex.ru, polina.timoshenko.2011@yandex.ru

Received September 23, 2020, peer-reviewed October 12, 2020, accepted October 26, 2020

Abstract: This article is dedicated to development of algorithms on information processing in aviation-based radioelectronic surveillance systems, aimed to solve problems of detection, location and identification of radiosignal-emitting sources.

Key words: radioelectronic surveillance system, the algorithm of information processing, detection, location, identification, radiosignal-emitting source

UDC 621.396.96

For citation: Ivan D. Biryukov, Pavel V. Buchuchan, Polina I. Timoshenko. Information Processing Algorithms in Aviation-Based Radioelectronic Surveillance Systems. *RENSIT*, 2020, 12(4):517-528. DOI: 10.17725/rensit.2020.12.517.

CONTENTS

1. INTRODUCTION (518)
 2. INCREASING THE EFFICIENCY OF DETECTING SIGNALS FROM RADIO EMISSION SOURCES OF A KNOWN TYPE IN AVIATION-BASED RADIO ELECTRONIC SURVEILLANCE SYSTEMS (AB RESS) (518)
 - 2.1. CHARACTERISTICS OF THE DETECTION OF SIGNALS FROM RADIO SOURCES IN THE RESS PLANT (518)
 - 2.2. FILTER IMPLEMENTATION METHODS (519)
 - 2.2.1. IMPLEMENTATION METHOD OF THE EQUALIZING FILTER (519)
 - 2.2.2. IMPLEMENTATION METHOD OF A FILTER MATCHED TO THE TYPE OF RSES (520)
 - 2.3. EVALUATION OF THE EFFECTIVENESS OF THE PROPOSED METHOD FOR INCREASING THE EFFICIENCY OF DETECTING SIGNALS OF RSES (521)
 3. IMPROVING THE ACCURACY OF DETERMINING THE LOCATION OF THE RSES IN THE AB RESS (521)
 - 3.1. CHARACTERISTIC OF DETERMINING THE LOCATION OF RSES IN THE AB RESS (521)
 - 3.2. ANALYSIS OF THE SYSTEMATIC ERROR IN DETERMINING THE POSITION OF THE RSES IN THE AB RESS (522)
 - 3.3. ALGORITHM FOR ELIMINATING THE SYSTEMATIC ERROR IN DETERMINING THE POSITION OF THE RSES IN THE AB RESS (523)
 - 3.4. EVALUATION OF THE EFFICIENCY OF THE ALGORITHM FOR ELIMINATING THE SYSTEMATIC ERROR IN DETERMINING THE POSITION OF THE RSES IN THE AB RESS (524)
 4. IMPROVING THE EFFICIENCY OF RECOGNITION OF RSES (525)
 - 4.1 CHARACTERISTICS OF THE SOLUTION TO THE PROBLEM OF RECOGNITION OF RSES (525)
 - 4.2 ANALYSIS OF METHODS FOR STORING THE DIGITIZED SIGNAL OF RSES IN THE RESS SYSTEM (526)
 - 4.3. ALGORITHM FOR AUTOMATIC STORAGE OF THE REQUIRED DIGITIZED SIGNALS OF RSES(526)
 5. CONCLUSION (528)
- REFERENCES (528)

1. INTRODUCTION

Aviation-based radioelectronic surveillance systems (AB RESS), being passive systems, efficiently in a large spatial area provide the solution of tasks: detection, estimation of parameters of radio emission sources, including coordinates, and their recognition. In view of the dependence of the solutions of the listed problems, the highest efficiency of aviation electronic surveillance systems can be achieved with their joint optimal solution. However, in practice, due to the complexity of the synthesized joint optimal algorithms, which are difficult to implement even at the modern level of computing technology, they decompose and proceed to the construction of quasi-optimal algorithms. Decomposition is carried out according to the tasks being solved, allowing to increase both the efficiency of solving each of them and the efficiency of information processing in the AB RESS as a whole, while maintaining the main interdependence of solutions.

The structure of the information processing algorithm in AB RESS depends on the category of the source of radio emission. This paper deals with a source of radio emission of the primary category - a radio electronic device. In the case of such a source of radio emission, the information processing algorithm can be conventionally represented in the form of processing stages. At the first stage, information about the received signal of the radio emission source is processed. At the second stage - about the source of radio emission of the received signals.

The urgency of increasing the efficiency of information processing in AB RESS is due to the constant counteraction to them and the development of radio emission sources, the complication of the electronic situation in the observation area.

The work consists of three parts, each of which is respectively devoted to increasing the efficiency of signal detection, determination of coordinates and recognition of radio emission sources. They substantiate and describe

the developed algorithms, and assess their effectiveness.

2 IMPROVING THE EFFICIENCY OF DETECTING SIGNALS FROM RADIO EMISSION SOURCES OF A KNOWN TYPE IN AB RESS

2.1 CHARACTERISTICS OF THE DETECTION OF SIGNALS FROM SOURCES OF RADIO EMISSION IN AB RESS

Modern AB RESS are built on a multichannel principle to ensure the observation of radiosignal-emitting sources (RSES) in a wide frequency range [1]. In view of the fact that RSESs can emit radio signals in a wide range of possible values of their radio technical parameters (RTP), in particular, frequency and spectrum width, two cases of signal reception and detection are possible:

1. The bandwidth of the signal is greater than the bandwidth of the frequency channel.
2. The bandwidth of the signal is less than the bandwidth of the frequency channel.

In the first case, the broadband signal is fed to several frequency channels at once. In the REW aviation system, each channel receiving a given radio signal introduces its own distortions, which significantly complicate the solution of the signal detection problem due to the unevenness and differences in their amplitude-frequency characteristics (AFC). Reducing the level of these distortions due to the optimal construction of the onboard analog part of the RESS aviation system is not always possible due to its excessive complexity. Therefore, it is advisable to assign the task of signal processing to the on-board special computer in order to correct the distortions introduced by each channel of the RESS aviation system.

In the second case, several narrow-band signals can enter one frequency channel. An increase in the number of signals that are simultaneously within the bandwidth of the receiving path leads to the need to solve the

problem of signal resolution, which, as a rule, is solved by means of frequency filtering.

At the same time, in contrast to radar systems, in which, on the basis of knowledge of the emitted signal, optimal methods of receiving reflected radio signals from objects and the underlying surface are used, in modern aviation RESS systems, due to the lack of such knowledge, non-optimal methods of receiving radio signals from RSES are used. However, the RESS aviation systems contain information about the received signals, which is used in solving the recognition problem and is contained in the catalog of types of radioactive sources. In it, each type is assigned a description in space of radio technical parameters, such as carrier frequency, pulse duration, spectrum width, and others. Based on these data, it seems possible to form matched filters with each type from the RSES catalog [2].

Obviously, not all RSES, and some are not fully represented in the catalog of types of RSES, especially during the period of active opposition of RSES to the RESS aviation systems. Therefore, the proposed method for increasing the efficiency of detecting signals of radiation sources should be considered as an addition to the existing methods. At the same time, in the course of a regular RESS, one should expect a decrease in the number of such RSESs and, consequently, an increase in the value of the proposed method in RESS aviation systems.

The consistent use of a correcting and matched filter with the type of RSES will increase the efficiency of detecting signals from these RSESs due to a significant increase in the signal-to-noise ratio at the detector input.

2.2 METHODS FOR IMPLEMENTING FILTERS

2.2.1 IMPLEMENTATION METHOD OF THE EQUALIZATION FILTER

The proposed method for the implementation of a correcting filter in digital form makes it

possible to adjust its parameters individually for each channel of the receiving path. It is based on the determination of the values of the correcting filter on a finite set of frequency samples using the Parks-McClellan algorithm. They are calculated using the *firpm(...)* function in the MATLAB software package, which allows you to obtain filter coefficients with a finite impulse response.

Let the initial frequency response $\tilde{S}(n)$ of the receiving channel, obtained at $n = 1, 2, \dots, N$ frequency samples, be known, which distorts the received signal. Then the frequency response of the correcting filter that compensates for these distortions at each frequency point must satisfy the condition:

$$\frac{c_a}{\tilde{S}(n)} - \frac{\varepsilon_a}{\tilde{S}(n)} \leq |\dot{K}_\kappa(n)| \leq \frac{c_a}{\tilde{S}(n)} + \frac{\varepsilon_a}{\tilde{S}(n)}, \quad (1)$$

where c_a – some coefficients to which the correction occurs; ε_a – permissible deviation of the frequency response from the values to which the correction occurs; $|\dot{K}_\kappa(n)|$ – AFC of the correcting filter.

As an example, **Fig. 1** shows the original normalized frequency response of the receiving channel approximated by line segments, where a solid line indicates its part in the used frequency band, and dashed lines - its part in unused frequency bands. For this frequency response of the receiving channel at $\varepsilon_a = 0.02$, in accordance with condition (1), the frequency response of the

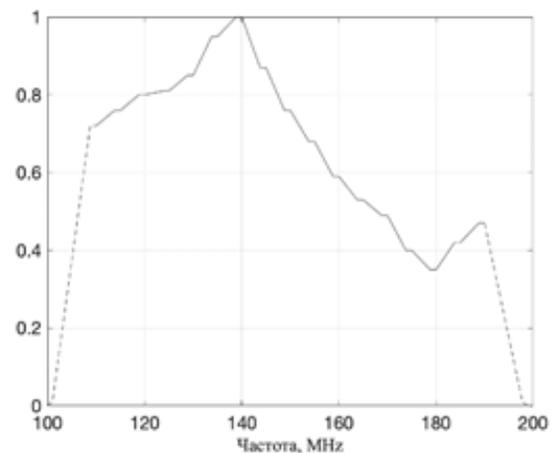


Fig. 1. Initial normalized AFC of the receiving channel.

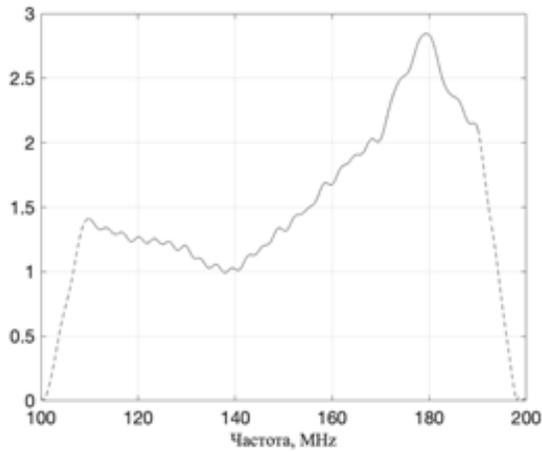


Fig. 2. Frequency response of the correcting filter.

correcting filter is determined, which is shown in Fig. 2.

The ideal, initial AFC of the receiving channel and the AFC of the receiving channel with a compensation filter are shown in Fig. 3. The root-mean-square deviation of the initial frequency response of the receiving channel from the ideal is $1.4 \cdot 10^{-1}$, and the standard deviation of the frequency response of the receiving channel, taking into account the compensation filter, from the ideal is $1.9 \cdot 10^{-4}$.

2.2.2 IMPLEMENTATION METHOD OF A FILTER MATCHED TO THE TYPE OF RSES

The description of the types of RSESs in the space of radio technical parameters depends on the completeness and inaccuracy of a priori information about them. The specified description is determined by the alphabet of

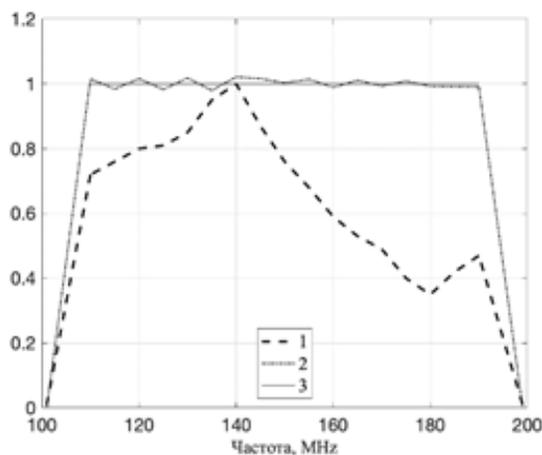


Fig. 3. AFC: (1) - initial receiving channel; (2) - a receiving channel with a compensation filter; (3) - perfect.

the features of the types and the density of the probability distribution of the values of the features for each type.

The method of implementing the filter, matched with the type of RSES, is reduced to the definition of the operator L of transition from the parametric description of the s -type RSES to the expression of the impulse response $\overline{h^{[s]}}$ of the filter matched with it, where $s = \overline{1, N_b}$, N_b is the number of types in the catalog.

In the case of an unknown distribution law of RTP values, the description of the RSES types is, as a rule, specified in the known boundaries $\lambda_{\max}^{[s]}, \lambda_{\min}^{[s]}$ in the form of a multidimensional uniform probability density of RTP values. The composition of vectors $\lambda_{\max}^{[s]}, \lambda_{\min}^{[s]}$ of the catalog of types of RSES in the RESS:

$$\lambda_{\max}^{[s]} = [f_{0\max}^{[s]}, \tau_{\max}^{[s]}, T_{\max}^{[s]}, \Delta F_{\max}^{[s]}]^T,$$

$$\lambda_{\min}^{[s]} = [f_{0\min}^{[s]}, \tau_{\min}^{[s]}, T_{\min}^{[s]}, \Delta F_{\min}^{[s]}]^T.$$

Then the procedure for determining the impulse response of a filter matched to the type of RSES can be written in the form

$$\overline{h^{[s]}} = L\{\lambda_{\max}^{[s]}, \lambda_{\min}^{[s]}\}.$$

The calculation of filters agreed with the types of RSES is performed once before the RESS. On the basis of the calculated coefficients in the digital processing unit, the corresponding digital filters are formed, which are used in the RESS process when detecting RSES signals [5].

If in the catalog of RSES types their description is limited by the carrier frequency and the signal spectrum width, then $\lambda_{\max}^{[s]} = [f_{0\max}^{[s]}, \Delta F_{\max}^{[s]}]^T$, $\lambda_{\min}^{[s]} = [f_{0\min}^{[s]}, \Delta F_{\min}^{[s]}]^T$ and the implementation of a filter consistent with the s -th type of RSES, is reduced to the formation of a band-pass digital filter. In this case, the operator L is the *firpm(...)* function of the MATLAB batch application, which calculates the coefficients of the desired filter with a finite impulse response. The values $(f_{0\max} + \Delta F_{\max})$ and

$(f_{0\min} - \Delta F_{\max})$. are specified as the bandwidth frequencies in the calculation.

It should be expected that in the case of a description of the types of RSES by a large number of RTPs and the known law of their distribution, the calculation of filter coefficients matched with the type of RSES will become more complicated, and the signal-to-noise ratio at their output will increase.

2.3 EVALUATION OF THE EFFECTIVENESS OF THE PROPOSED METHOD FOR INCREASING THE EFFICIENCY OF DETECTING SIGNALS OF RSES

Evaluation of the effectiveness of the proposed method for increasing the efficiency of detecting signals of RSES is based on comparing the efficiency of an autocorrelation detector of one receiving channel with and without the use of correcting and matched with the type of RSES filters.

Fig. 4 shows a diagram of an autocorrelation detector of one receiving channel using a correction filter $K_k(n)$ and a filter $K_c(n)$, matched with the s -type of RSES.

Efficiency assessment was carried out in the MATLAB programming environment.

The signal is generated in accordance with the following description of the RSES type:

[Pulse sequence, $\tau_{\min} = 3 \mu s$, $\tau_{\max} = 5 \mu s$, $T_{\min} = 10 \mu s$, $T_{\max} = 15 \mu s$, $f_{0\min} = 170 \text{ MHz}$, $f_{0\max} = 180 \text{ MHz}$, $\Delta F_{\min} = 200 \text{ kHz}$, $\Delta F_{\max} = 333 \text{ kHz}$].

The signal is in white Gaussian noise. The normalized frequency response of the receiver input circuits corresponds to the frequency response shown in Fig. 2.

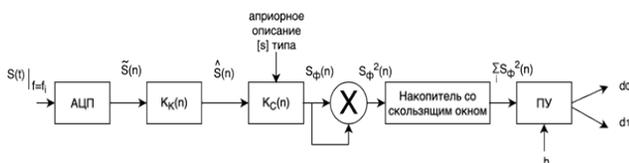


Fig. 4. Schematic of an autocorrelation detector of one receiving channel.

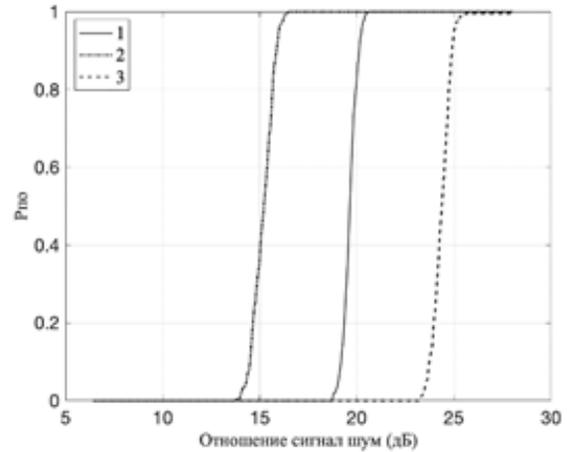


Fig. 5. Dependence of the probability of correct signal detection on the signal-to-noise ratio: 1 - using only a correction filter; 2 - use of a correction filter and a filter matched to the type of RSES; 3 - without using a corrective filter and a filter matched to the type of RSES.

The probabilistic characteristics of correct signal detection, presented in **Fig. 5**, are plotted for three cases:

1. Only a correction filter is used;
2. A correcting filter and a filter matched to the type of RSES are used;
3. Without the use of corrective and matched with the type of RSES filters.

It can be seen from the graphs that the combined use of correcting and type-matched RSES filters gives a gain in comparison with a receiver built without their use, in terms of the signal-to-noise ratio by 9 dB, which leads to an increase in the signal detection range by 2.8 times.

3. IMPROVING THE ACCURACY OF DETERMINING THE LOCATION OF IRI IN THE AVIATION ELECTRONIC SURVEILLANCE SYSTEM

3.1. CHARACTERISTICS OF DETERMINING THE LOCATION OF RSES IN THE AVIATION ELECTRONIC SURVEILLANCE SYSTEM

In the RESS aviation systems, the coordinates of the RSES and the errors of their estimation are most often determined from the set of measured bearings. For this, a method based on the Kalman filter, the method of least squares of angle corrections, etc. are used. The general

and necessary condition for the implementation of these methods is the assignment of an initial estimate of the coordinates of the RSES. The estimation result depends significantly on the accuracy of the initial estimate, especially with a small number of bearings. Due to the large error of the initial estimate of coordinates, in a number of cases, there is a "migration" of the estimate of the coordinates of the RSES during their further refinement. At the same time, the experience of using the RESS aviation systems has shown that a number of RSESs operate for a short time, which also leads to the need to determine their location based on a small number of bearings received.

The accuracy of the estimate is usually characterized [6] by the mean value of the squared error, which is equal to the sum of the variance of the estimate and the squared systematic error (bias) of the estimate. The expressions for calculating the variance D_λ of the estimate λ of the true coordinates of the RSES λ_0 are known. In modern aviation RESS systems, the systematic position determination error is not analyzed or eliminated, although the magnitude of this error, depending on the observation conditions, can be significant (up to several kilometers).

This part of the work is devoted to the analysis and elimination of the systematic error in determining the position of the radioactive sources in the AB RESS.

3.2. ANALYSIS OF THE SYSTEMATIC ERROR IN DETERMINING THE POSITION OF RADIOACTIVE SOURCES IN THE AB RESS

The systematic error in determining the location of the RSES can be defined as

$$\Delta = \mathbf{M}\{\lambda\} - \lambda_0,$$

where $\mathbf{M}\{\lambda\}$ is the mathematical expectation of the estimate, $\mathbf{M}\{\lambda\} = \int \lambda p(\lambda) d\lambda$; $p(\lambda)$ is the probability density of the estimate. For the case of a stationary local rectangular coordinate system OXY when the aircraft is flying along the abscissa axis $\lambda = [x, y]^T$, $\lambda_0 = [x_0, y_0]^T$, $\Delta = [\Delta x, \Delta y]^T$, where x, y – respectively, the abscissa

and ordinate of estimating the location of a stationary ground-based RSES, x_0, y_0 are its true abscissa and ordinate, respectively, $\Delta x, \Delta y$ is a systematic error in estimating the location of a stationary ground-based RSES on the abscissa and ordinate, respectively, t is the transposition sign. In the process of work, the expression for the probability density $p(x, y)$ of the initial estimate of coordinates was obtained [3] when it was formed from two bearings:

$$p(x, y) = \frac{|x_j - x_i|}{y_0^3} \exp \left\{ \frac{-[\arctg((x_0 - x_i) / y_0) - \alpha_i]^2 - [\arctg((x_0 - x_j) / y_0) - \alpha_j]^2}{2\sigma^2} \right\} \\ = \frac{2\pi\sigma^2 \left[1 + \left(\frac{x_0 - x_i}{y_0} \right)^2 \right] \left[1 + \left(\frac{x_0 - x_j}{y_0} \right)^2 \right]}{2\pi\sigma^2 \left[1 + \left(\frac{x_0 - x_i}{y_0} \right)^2 \right] \left[1 + \left(\frac{x_0 - x_j}{y_0} \right)^2 \right]},$$

where $(x_i, 0)$ and $(x_j, 0)$ are the coordinates of the aircraft, in which the initial and final bearings were measured, respectively; α_i and α_j – measurements of the initial and final bearings, respectively; σ – standard deviation (RMS) of bearing measurement.

Fig. 6 and 7 show, respectively, the projection of the normalized probability density $p(x, y)$ onto the OXY plane and the section of the probability density $p(x, y)$ by the plane perpendicular to the OX axis and passing through the point $(0, 0)$. In this case, the following are given: the true position of the RSES at the point $(0; 100)$ km; the bearing was measured when the aircraft was at the points $(-15; 0)$ km and $(15; 0)$ km, which corresponds to the bearing base $L = 30$ km; RMS of bearing

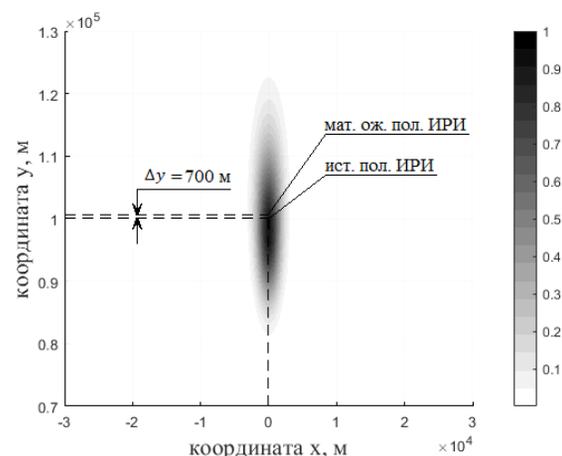


Fig. 6. Projection of the probability density $p(x, y)$ onto the OXY plane.

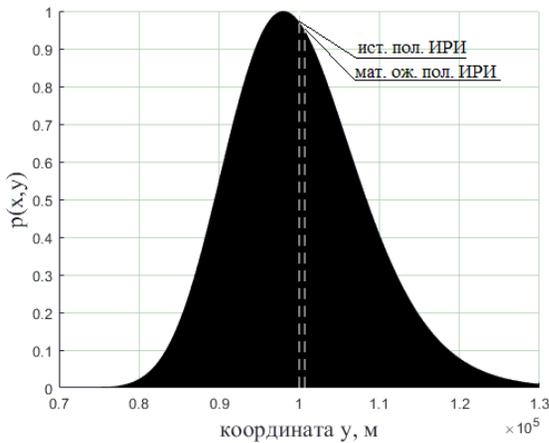


Fig. 7. Section of the probability density $p(x, y)$ by a plane perpendicular to the OX axis and passing through the point $(0,0)$.

measurement 1 degree. In Fig. 6, the value of the probability density is expressed in black and white gradient. From the analysis of the graphs it follows that the probability density distribution of the initial coordinates of the RSES $p(x, y)$ is unimodal and has a positive asymmetry. This type of probability density is confirmed experimentally when constructing a histogram. During the simulation, it was found that the probability density graph tends to a symmetric form when the angle between bearings tends to 109° .

Figs. 8, 9 show the graphs of the dependence of the systematic error Δy and Δx , respectively, of the initial estimate of the RSES on the distance D to the RSES. The range to

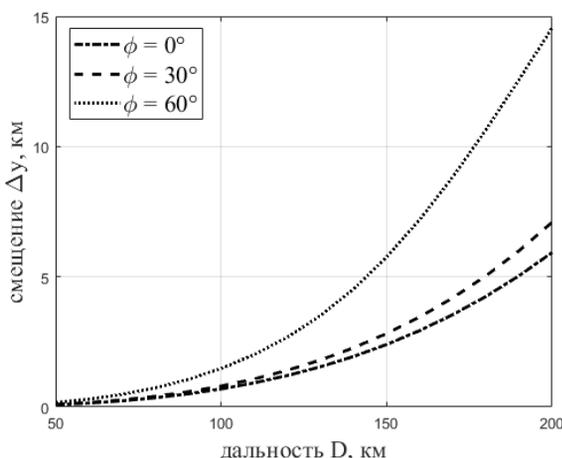


Fig. 8. Dependence of the displacement Δy on the range D at $\sigma = 1^\circ$ and $L = 30$ km.

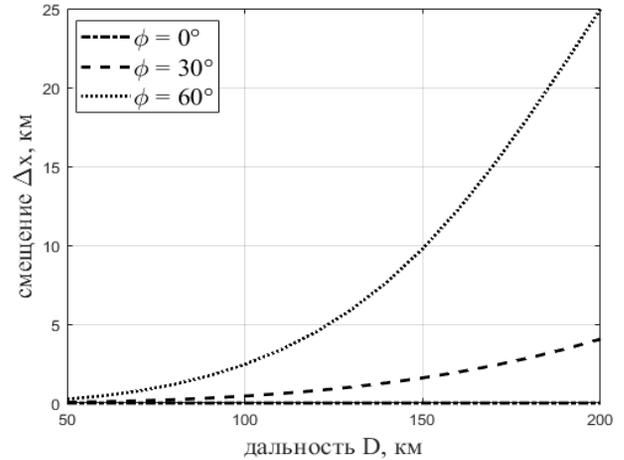


Fig. 9. Dependence of the displacement Δx on the range D at $\sigma = 1^\circ$ and $L = 30$ km.

the RSES is defined as the distance between the center of the DF base and the true location of the RSES. The plots were constructed with the RMS of the bearing measurement $\sigma = 1^\circ$ and the bearing base $L = 30$ km. The angle ϕ between the perpendicular to the center of the bearing base and the direction to the RSES changes in the values of $0^\circ, 30^\circ, 60^\circ$. The range D to RSES varies from 50 km to 200 km. From the analysis of the graphs it follows that the systematic error increases nonlinearly with an increase in the range to the RSES, and with the RMS of measuring the bearing of the order of a degree at long ranges, it can be several kilometers.

3.3. ALGORITHM FOR ELIMINATING THE SYSTEMATIC ERROR IN DETERMINING THE LOCATION OF THE RSES IN AB RESS

Elimination of the systematic error in determining the position of the RSES is possible only under the known observation conditions and its true position. In practice, this information is obtained only after performing the RESS and analyzing the data. Since there is no way to obtain the location of the RSES without preliminary collection of information [7], the elimination of the systematic error is also possible only after the end of data acquisition. In view of the fact that it is not always possible to obtain an unbiased estimate of the position of the RSES with the help of the RESS, a method is needed that allows, on the basis of the obtained observation

conditions for the estimation of the position of the RSES, to determine and eliminate the systematic estimation error.

As a result of the RESS, the following became known: the estimate $\hat{\lambda} = \{\hat{x}, \hat{y}\}$ of the coordinates of the RSES, hereinafter called the initial one; their estimation error \hat{R} ; coordinates $\lambda_i = \{x_i, y_i\}$ of the aircraft at the moments of bearing measurements; bearing values $\{a_i\}$, where $i = 1, 2, \dots, N$, N - the number of measured bearings; RMS deviation of bearing measurement σ . It is required to eliminate the systematic error Δ in determining the position of the RSES with an acceptable accuracy Δ_{add} .

In the developed algorithm, the systematic error is eliminated by the successive approximation method based on the RESS simulation. Simulation of the RESS consists in repeated tests, where the coordinates of the RSES are estimated under known initial conditions and the measurement of the bearing with an error introduced by the random number sensor, as well as in the estimation of the statistical parameters of the coordinates. The block diagram of the algorithm is shown in Fig. 10. The algorithm can be conditionally divided into three stages: 1) determination of the number of tests for simulation of RESS 2) determination of the straight line, on which the position of the RSES is refined by the method of successive approximation; 3) cyclic refinement of the location of the RSES.

At the first stage of the algorithm, the location estimate $\hat{\lambda}$ is taken as the true position λ_0^* , and statistical simulation of the RESS is performed with a limited number of tests N_0 . The standard deviation σ_{N_0} of the obtained estimate $\lambda_{N_0}^*$ and the number of tests N , for which the standard deviation σ_N of the coordinate estimates $\lambda_m^* = \{x_m^*, y_m^*\}$, determined in the cyclic refinement of the true location of λ_m^* RSES at the m -th step of the cycle, will be, with a confidence level P , an order of magnitude less than the permissible deviation Δ_{add} . The number of tests N is determined from the inequality [4]

$$N \geq \left[\frac{t(P)}{0.1\Delta_{perm}} \right]^2 \sigma_{N_0}^2,$$

where $t(P)$ is found from the equality $2\Phi(t) = P$, $\Phi(t)$ is the probability integral.

At the second stage of the algorithm, the statistical simulation of the RESS is repeated with the number of tests N and the estimate of coordinates λ^* , its systematic error Δ^* and the correlation matrix R^* are determined. A straight line is constructed passing through the points $\hat{\lambda}$ and λ^* . The direction of movement along a straight line is determined by the value of the angle γ between the extreme bearings to the point λ_0^* .

At the third stage of the algorithm, a cycle of refinement of the coordinates of the RSES is performed by selection on a straight line by the method of successive approximation of coordinates $\lambda_m = \lambda_{m-1} - k\Delta^*$ at the m -th step of the cycle, where $m = 1, 2, \dots, M$, M is the number steps required to exit the loop, k is the bias correction factor. At each step, statistical simulation of the RESS is performed with the number of tests N , and an estimate of the coordinates λ_m^* is determined. If at the coordinates λ_m the condition $|\lambda_m^* - \lambda^*| < \Delta_{доп}$ is satisfied, then the true position $\lambda_0^* = \lambda_m$ of the RSES is determined with sufficient accuracy, otherwise a new iteration of the correction of coordinates λ_m begins, while changing the sign

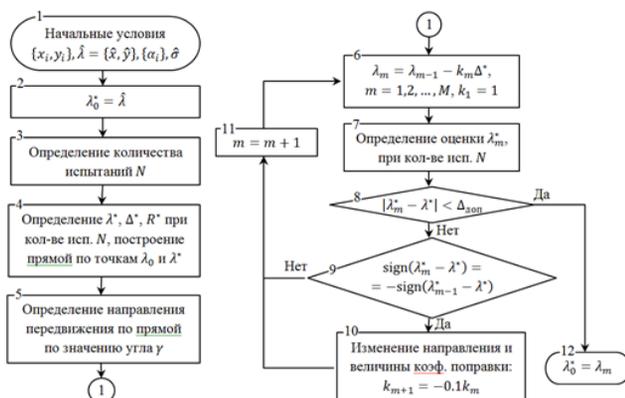


Fig. 10. Algorithm block diagram.

and magnitude of the correction in the case of "jumping over".

3.4. EVALUATION OF THE EFFICIENCY OF THE ALGORITHM FOR ELIMINATING THE SYSTEMATIC ERROR IN DETERMINING THE POSITION OF THE RADIOACTIVE SOURCE IN THE AVIATION ELECTRONIC SURVEILLANCE SYSTEM

The algorithm was tested on a PC using the MATLAB computing environment version R2019b. The performance of the PC was determined using the "bench (100)" function, which shows the averaged values of the execution time, among others, the "LU", "FFT" tests (operations with floating point numbers), respectively equal to 0.2174 s and 0, 1399 s.

The RESS simulation was carried out under the following initial conditions: the range D to the RSES varies from 50 to 200 km, the direction finding base is $L = 30$ km; 2 bearings measured symmetrically in one test; RMS of bearing measurement 1 degree; the number of tests for one iteration of the RSES coordinates correction is equal to $N = 10^7$. The execution time of one iteration was 160 s.

Fig. 11 shows the dynamics of changes in the deviation $\lambda_0^* - \lambda_0$ at $D = 200$ km, in Fig. 12 – the graph of the dependence of the deviation of the initial and corrected estimates from the true position of the RSES on the distance D .

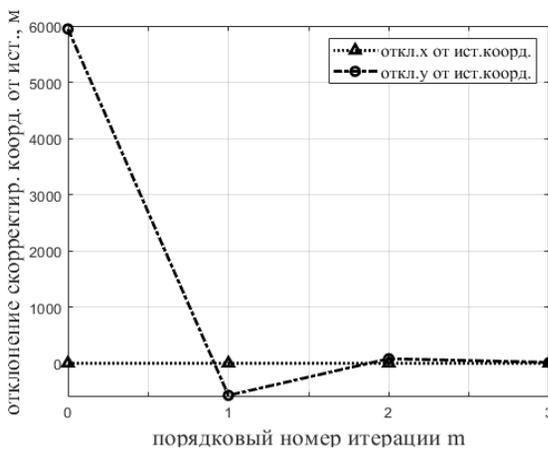


Fig. 11. Dynamics of the deviation of the corrected coordinates of the RSES from the true ones at $D = 200$ km.

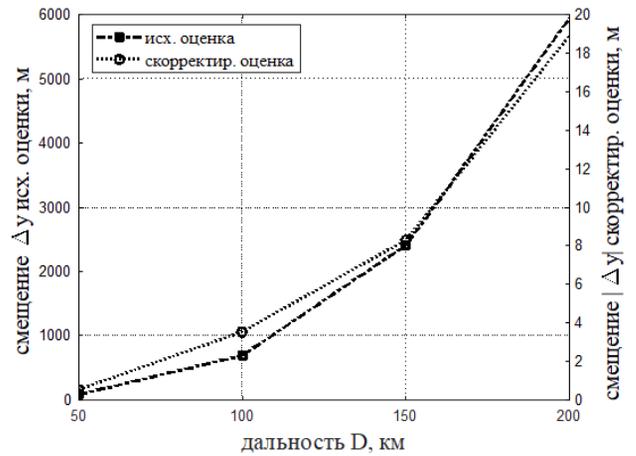


Fig. 12. A graph of the dependence of the deviation of the initial and corrected estimates from the true position of the RSES on the range D .

Analysis of the graphs shows that the systematic error decreases significantly already at the first iteration cycle of the algorithm, and further reduction is achieved after several subsequent iterations. At a distance of $D = 200$ km, the ordinate systematic error decreases from 6 km to 18 m; at smaller ranges, the result is significantly improved and amounts to a few meters. The systematic error along the abscissa in all cases does not exceed 1 m, which is insignificant within the framework of this problem. Thus, the developed algorithm provides effective elimination of the systematic error in estimating the position of the RSES.

4. IMPROVING THE EFFICIENCY OF RSES RECOGNITION

4.1 CHARACTERISTICS OF THE SOLUTION TO THE PROBLEM OF RECOGNITION OF RSES

For the case of stationary sources of radio emission, the distance between which is greater than the root-mean-square error in determining their coordinates, the problems of detection and estimation can be solved with high reliability. However, the analysis of modern RSES showed that most of them change their location, the distance between them may be less than the mean square error in determining their coordinates, widespread use in the region of the

same type of RSES, and the degree of overlap of possible values of parameters of different types is such that it will not allow solving the recognition problem definitely. At the same time, in the RESS aviation systems, as a rule, information about the RSES is received with significant frequency. All of the above factors significantly reduce the capabilities of the RESS aviation systems, not only in opening the RESS, and even more so in identifying and analyzing the changes that have occurred in the RESS area, accompanying the RESS and forming their radio technical portrait.

One of the ways to overcome these difficulties is to obtain from the received signals of the RSES, not only the values of radio technical parameters - identification signs already used in recognition, but also additional information. Additional information in the form of measured values, highlighted individual identifying signs, will increase the detail of recognition of radioactive sources up to a copy and, consequently, the effectiveness of their opening, tracking, identification and analysis of the changes that have occurred in the RESS area, clarification and addition of existing radio technical portraits of radioactive sources. This became possible due to the digital nature of information processing in modern aviation systems of the RESS, the intensive development of its algorithms and the corresponding element base.

However, the number of RSESs in the modern RESS area is so great that it is not possible to obtain additional information on all RSESs on a time scale close to real, due to the limited capabilities of the RESS aviation systems. A prerequisite for obtaining additional information about the RSES is the preservation of their received digitized signals for subsequent processing.

This part of the work is devoted to the preservation of the received RSES signal in a digitized form in the RESS aviation system.

4.2 ANALYSIS OF METHODS FOR STORING THE DIGITIZED SIGNAL OF RSES IN THE RESS SYSTEM

There are two known methods for storing the digitized signal of RSES in RESS systems.

The first method is the automatic saving of all digitized signals of the RSES. In view of the large number of RSESs and, accordingly, the signals received from them, this method places high demands on the RESS system in terms of the throughput of the channels for storing digitized signals, on the speed and volume of the memory device. In this regard, if a large number of digitized signals are transmitted over a radio channel, the secrecy of the operation and the noise immunity of the RESS system will decrease. At the same time, the experience of using RESS systems has shown that out of all the stored digitized signals, only a small part of them (less than 20%) can subsequently be claimed. The latter testifies to the inexpedient use of the computing resources of the RESS system in this method and the need to solve the problem of selecting the required digitized signals before storing them.

The second method is the automated saving of only the required digitized signals. This is achieved by changing the operator of the preset operating mode of the RESS system by adjusting it for frequency selection of the required signals and saving them in digital form. Due to the limited storage space of the RESS system, the recording time of a digitized signal may be shorter than its duration. As a result, the operations of recording the required signal by the operator are repeated many times, and in general a lot of time is spent on this. Therefore, this method is characterized by a low throughput for storing digitized signals.

The developed method implements the automatic storage of the required digitized signals from the RSES. It eliminates the disadvantages of the first method by selecting the required signals before storing them and

the disadvantages of the second method by automatically performing this procedure. In this method, the possibilities of selection are significantly expanded by including additional selection conditions in it: by the type of RSES; by radio technical parameters, including frequency; at the location of the RSES; on the reliability of information about the received signals.

4.3 ALGORITHM FOR AUTOMATIC SAVING OF THE REQUIRED DIGITIZED SIGNALS OF RSES

To automatically save the required digitized signals, tasks containing selection conditions are introduced into the RESS aviation system.

At the first stage of processing, after evaluating the radio technical parameters of the received signal, its belonging to the types of the radio emission source represented by the catalog of types can be determined. The assignment of the types of radiation sources, the digitized signals of which need to be stored, is considered as one of the ways of their selection in the space of radio technical parameters. Taking into account that not all interested RSESs or their signals, in particular new ones, are described in the catalog of types, another way of such selection is to determine whether the received signals belong to specified intervals of radio technical parameters.

At the second stage of processing, after assigning the received signal to one previously (at the previous step of observation) detected radio emission source, the coordinates of which have been determined, selection is carried out at the location of the RSES. For this, the areas of observation of the RSES are set, the signals of which must be stored in a digital form.

Since the information about the received signals significantly depends on the RESS conditions and has a probabilistic nature, in the process of its processing the parameters characterizing its reliability are estimated. To preserve reliable information in the developed algorithm, threshold values of these parameters are set: the reliability of the received signal, the probability of recognizing the type of RSES,

the size of the largest semi-axis of the ellipse of the error in determining the coordinates of the RSES. Exceeding the preset threshold values is an additional way of selecting the received signals for recording them in digital form.

The developed algorithm for automatic storage of the required digitized signals of RSES is presented in the form of a block diagram in **Fig. 13**.

The input data of the algorithm is presented in Blocks 2-4. Block 2 – Data on the received signal, received on the *i*-th observation interval: number of the type of RSES; the probability of recognizing the type of RSES; RTP of the received signal; the value of the RTP reliability parameter; the decision on the belonging of the received signal to the RSES, the coordinates of which are determined. Block 3 – Data on RSES on the *i*-1 observation interval: coordinates of RSES: value of the parameter of reliability of coordinates of RSES Block 4 – Task data for saving digitized files: number of the specified type; type recognition probability threshold; specified RTP intervals; the threshold of

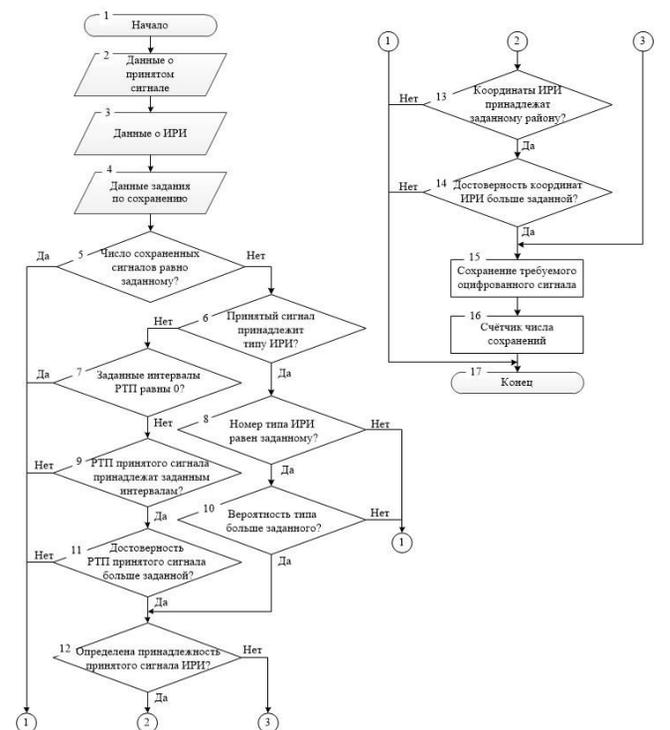


Fig. 13. Block diagram of the algorithm for automatic storage of the required digitized signals of RSES.

reliability of the parameters of the received signal; the given parameters of the observation area of the RSES; required number of stored digitized signals.

Checking the fulfillment of the selection conditions according to the task is represented by Blocks 5-14.

The completion of the task is achieved by executing, represented by Block 5, the equality of the number of previously stored digitized signals represented by Block 14 with a given number. It should be noted that when entering a task, the counter of the number of stored digitized signals is reset to zero.

Evaluation of the effectiveness of the developed algorithm showed that its implementation requires insignificant computing resources and will provide in modern aviation RESS systems automatic background storage of the received signals in digital form according to the task in a time scale close to real.

5. CONCLUSION

The analysis of solving the problems of detection, positioning and recognition of radioactive sources performed in this work substantiated the importance of the proposed methods for increasing their efficiency and determined the directions for the development of appropriate information processing algorithms in modern aviation RESS systems. The interdependence of the decisions made at each stage of information processing is used in the proposal for the use of correcting filters that are consistent with the type of RSES when detecting signals and saving them in digital form.

The performed modeling of the proposed methods and the developed algorithms for information processing confirmed their feasibility and increased efficiency in modern aviation RESS systems. The signal-to-noise ratio increases significantly in the receiving channels that use information about the types of radiation sources (in the example considered, the increase in the signal-to-noise ratio was 9 dB. With a

permissible accuracy equal to 0.1 MSE, the systematic error in determining the position of the RSES is eliminated. Provides a background automatic saving of the received signals in a digital form according to the task in a time scale close to real in order to obtain additional information about RSES, their individual recognition, tracking, identification and analysis of the changes that have occurred in the RESS area.

REFERENCES

1. Baylov VV, Plaksienko VS. Electronic surveillance equipment. Taganrog, TTI Publishing House, 2009, 104 p.
2. Biryukov ID, Silin SI. Proposals for the development of a quasi-optimal algorithm for detecting signals from radio sources in aircraft systems. *Proc. XIII All-Russian Scientific and Technical Conference "Radar location and radio communication"*, Moscow, 2019, pp. 133-137.
3. Buchuchan PV, Byldin AYu. Analysis of the error in determining the location of the source of radio emission by the direction finding method in the aviation electronic surveillance system. *Proc. XIII All-Russian Scientific and Technical Conference "Radar location and radio communication"*, Moscow, 2019, pp. 138-142.
4. Rumshiskiy LZ. *Mathematical processing of the experimental results*. Moscow, Nauka Publ., 1971, 192 p.
5. Timoshenko PI, Isakov VN. Effects of quantization of coefficients in non-recursive digital filters with a U-shaped amplitude-frequency characteristic. *Proc. International Scientific and Technical Conference "INTERMATIC-2018"*, Moscow, 2018, pp. 628-631.
6. Tikhonov VI. *Statistical radio engineering*. Moscow, Radio i svyaz' Publ., 1982, 624 p.
7. Erastov VE. *Metrology, standardization and certification*. Moscow, FORUM Publ., 2008, 208 p.

DOI: 10.17725/rensit.2020.12.529

Towards a New Paradigm

Gennady V. MishinskyJoint Institute for Nuclear Research, <http://www.jinr.ru/>

Dubna 141980, Moscow Region, Russian Federation

*E-mail: mysb@jinr.ru**Received May 31, 2020, peer-reviewed June 22, 2020, accepted June 29, 2020*

Abstract: The discoveries of new low-energy nuclear reactions and a new resonant interference exchange interaction explaining the course of these reactions give grounds to assert that a necessary and inevitable process of changing the paradigm is currently taking place.

Keywords: low-energy nuclear reactions, resonant interference exchange interaction, transatoms and transmolecules, history of science, resonant technology, noosphere, evolution, ecology

PACS: 01.70.+w; 03.75.Nt; 05.30.Jp; 24.30.-v; 25.60.Pj; 32.10.-f; 36.10.-k

For citation: Gennady V. Mishinsky. Towards a New Paradigm. *RENSIT*, 2020, 12(4):529-544. DOI: 10.17725/rensit.2020.12.529.

CONTENTS

1. INTRODUCTION (529)
 2. LOW-ENERGY NUCLEAR REACTIONS (530)
 - 2.1. COLD FUSION (530)
 - 2.2. LOW-ENERGY TRANSMUTATION OF ATOMIC NUCLEI OF CHEMICAL ELEMENTS (531)
 3. NEW PARADIGM (533)
 4. NEW STATE OF MATTER (534)
 5. NEW FUNDAMENTAL RIE X INTERACTION (534)
 6. PROPERTIES OF RIE X INTERACTION (536)
 7. TRANSMUTATION REACTIONS (539)
 8. EVOLUTION AND ECOLOGY (540)
 9. CONCLUSION (542)
- REFERENCES (542)

1. INTRODUCTION

In 1896, A. Becquerel discovered the natural radioactivity of uranium salts. Two years later, E. Rutherford and P. Villard showed that radioactive rays are composed of alpha, beta and gamma radiation. And in 1903, E. Rutherford and F. Soddy put forward a hypothesis about the transformation of chemical elements in the process of their radioactive decay. The nuclear nature of radioactivity was understood by Rutherford after he proposed a nuclear model of the atom in 1911 and came to the conclusion that radioactive radiation arises from nuclear processes that occur

inside the atomic nucleus. In 1934, this conclusion was confirmed by the discovery, by the spouses Irene and Frederic Joliot-Curie, of artificial radioactivity and the discovery, in 1938 by O. Hahn and F. Strassmann, of uranium fission under the action of neutrons. Since then, radioactivity and nuclear reactions with the transformation of some chemical elements into other elements have always gone “hand in hand”. And the scientific community came to a stable, firm opinion that nuclear reactions are always accompanied by radioactive radiation.

However, almost a century after the discovery of radioactivity, in 1989-1992, dramatic events took place in nuclear physics, marked by the unexpected discovery of “impossible”, radiationless and low-energy nuclear reactions.

Due to the development, at the end of the last century, of analytical instrumentation and computer technology, certified analytical laboratories for general use began to be created everywhere to perform the investigation of substances and materials. One of the goals of those studies was to determine the presence of chemical elements and their amount in selected materials and samples. Analytical laboratories typically include mass spectrometers with various types of ion sources, atom-emission- and X-ray-spectrometers, including electron microscopes,

which allow X-ray microprobe analysis of substances and materials. Researchers have now the opportunity, to obtain, with extreme sensitivity levels [1], independent, reliable information about the mass, elemental composition of both geological samples and materials obtained in various experiments.

It turned out that “extraneous” chemical elements were found in the medium, in many physical experiments related to electron impact on a condensed matter, after their completion, which were absent in that medium before the start of the impact. It is extremely important to note that the isotopes of the “extraneous” elements were stable, i.e. non-radioactive. The amount of “extraneous” elements obtained could not be explained by the impurities of chemical elements present in the reaction volumes. In some experiments, “extraneous” elements accounted for tens of percent of the total mass of the condensed matter. Later on, the production of “extraneous” elements in processes different from ordinary nuclear reactions was called low-energy transmutation of atomic nuclei of chemical elements. The transmutation is a transformation of some chemical elements into other chemical elements in weakly excited condensed matter. Still later, low-energy nuclear transmutation and Cold nuclear Fusion (CF) were combined under the general name Low Energy Nuclear Reactions (LENR) or Condensed Matter Nuclear Science (CMNS).

Low-energy nuclear reactions are not a special case of conventional nuclear reactions. They occur everywhere in the Universe and are the basis for the formation of a new paradigm. Just as the new paradigm does not include the old paradigm, so low-energy nuclear reactions do not include conventional, collisional nuclear reactions.

Let's identify this new paradigm.

2. LOW-ENERGY NUCLEAR REACTIONS

2.1. COLD FUSION

The cold fusion reaction (CF) was implemented by S. Pons and M. Fleischman in 1989 at the palladium (Pd) cathode electrolysis of a solution of deuterated lithium hydroxide in heavy water (0.1M

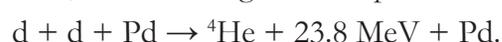
LiOD in a solution 99.5% D₂O + 0.5% H₂O) [2]. They reported that a significant amount of excess heat is released during electrolysis, which cannot be explained by chemical reactions. In addition, a weak neutron flux (n) and tritium (t) generation was detected in those experiments. Those results allowed the authors to draw a conclusion about the nuclear origin of excess heat and to assume that the following nuclear reactions with deuterons (d) take place in palladium electrode:



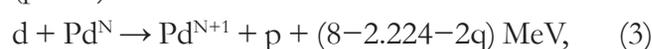
Similar thermonuclear reactions (1) and (2) begin to proceed at a temperature of ~ 100 million degrees (12.9 keV).

Some time later, other researchers [3] found that reactions with tritium yield were up to 10⁹ times more intense than reactions with neutron yield. Thermonuclear reactions (1) and (2) proceed with equal yield.

Further investigations have shown that the electrolysis gas resulting from electrolysis in CF reactions contains the helium isotope ⁴He, and its amount correlates with the energy released in the fusion reaction [4]. Thus, an energy of ~32 MeV is released per ⁴He atom. Since ⁴He is synthesized, then, the following reaction proceeds:



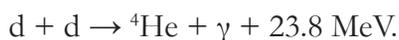
But the energy yielded in this reaction would be not enough for the existing correlation in the production of ⁴He atoms with the energy released. Therefore, other, additional CF reactions should proceed in parallel. For example, during the formation of a nuclear molecule (d-Pd^N), consisting of a deuteron and a palladium nucleus Pd^N, a reaction can occur with the capture of a neutron by palladium [5], which is a part of the deuteron nucleus, $d \equiv (p - n)$:



where *N* is the number of neutrons. In this reaction, an energy is released equal to the difference between the binding energy of a neutron in a palladium nucleus (on average ~8 MeV) and the binding energies of a 2.224 MeV

deuteron and a doubled nuclear molecular bond q ($d\text{-Pd}^N$).

Thermonuclear $d+d$ -reaction with the production of the helium isotope ${}^4\text{He}$ is well known, but it always proceeds with the emission of gamma quantum γ with a probability of 10^{-7} relative to reactions (1) and (2):



No gamma quanta were found in CF reactions.

As a result, we can conclude:

Cold fusion reactions differ from thermonuclear reactions in their properties, and they occur at "room" temperatures.

2.2. LOW-ENERGY TRANSMUTATION OF ATOMIC NUCLEI OF CHEMICAL ELEMENTS

S. Pons and M. Fleischmann used palladium as the cathode in their experiments with electrolysis. That was due to the fact that palladium dissolves hydrogen well. So, one volume of Pd dissolves, under normal conditions, 850 volumes of H_2 , or there are 7 hydrogen atoms per 10 palladium atoms. Therefore, in CF experiments carried out by different authors, other metals that dissolve hydrogen well were often used, for example: Ti, Fe, Co, Ni, Pt. Or other methods were used to saturate palladium with hydrogen. One of these methods is the method of cathode saturation by means of gas glow discharge.

The installation created by the authors of [6,7] for carrying out experiments with a gas glow discharge consisted of a chamber with a cathode and an anode filled with a working gas to a pressure of 300–1000 Pa. Hydrogen, deuterium, argon, xenon and their combinations were used as the working gas. The glow discharge was carried out at a current density of 10–50 mA/cm² and a discharge voltage of 500–1400 V. The experiments continued for up to 120 hours. The material for the cathodes was 100 μm thick foils made of palladium and other metals (Ti, Ag, Nb, etc.).

Samples of cathodes were analyzed for the detection of impurities of chemical elements in them before and after the experiments. They applied the following methods for the analysis of the samples: spark, secondary ion and secondary

neutral mass spectrometry, as well as the method of X-ray microprobe analysis. The content of elements in the cathodes was recorded in a near-surface layer 100 nm thick. The difference in the content of impurities of chemical elements before and after the experiments was interpreted as the production of "extraneous" nuclides. "Extraneous" elements were mainly contained in the outgrowths formed on the surface of palladium cathodes. The size of these formations reached 15 microns. The highest yield of "extraneous" nuclides was registered in a glow discharge in deuterium in a palladium cathode. The main nuclides (with a content of more than 1%) are ${}^7\text{Li}$, ${}^{12}\text{C}$, ${}^{15}\text{N}$, ${}^{20}\text{Ne}$, ${}^{29}\text{Si}$, ${}^{44,48}\text{Ca}$, ${}^{56,57}\text{Fe}$, ${}^{59}\text{Co}$, ${}^{64,66}\text{Zn}$, ${}^{75}\text{As}$, ${}^{107,109}\text{Ag}$, ${}^{110-112,114}\text{Cd}$, ${}^{115}\text{In}$. Fig. 1 shows the "extraneous" nuclides produced in the Pd cathode after its irradiation in a deuterium discharge for 22 hours at a discharge current of 50 mA. The absolute number of atoms of these nuclides amounts to 10^{17} .

For such elements as Li, B, C, Ca, Ti, Fe, Ni, Ga, Ge, etc., a change in the natural isotope ratio was recorded, for some elements by several tens of times. For example: depending on the location on the cathode, the ${}^{57}\text{Fe}/{}^{56}\text{Fe}$ ratio varies from 25 to 50 times, while the natural ratio is ${}^{57}\text{Fe}/{}^{56}\text{Fe} = 0.024$. At the same time, some basic isotopes are absent, for example: ${}^{58}\text{Ni}$, ${}^{70,73,74}\text{Ge}$, ${}^{113,116}\text{Cd}$. In addition, a change in the natural ratio of palladium isotopes is observed in Pd cathodes.

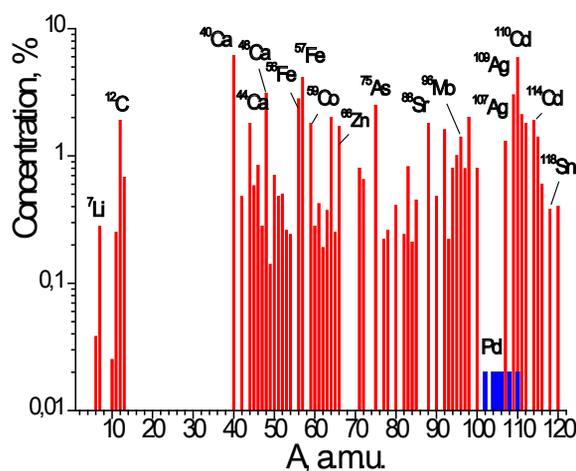


Fig. 1. "Extraneous" nuclides produced in a Pd cathode in a deuterium glow discharge. Palladium isotopes are highlighted with blue lines, without their relative abundance.

During the glow discharge and after it was turned off, gamma radiation was recorded, by means of a Ge(Li) detector, in the energy range 0.1–3.0 MeV. The analysis of gamma spectra showed that the emitters are neutron-rich nuclei with masses from $A = 16$ to $A = 136$, that yield β -radioactive decay chains. However, according to the authors' estimates, the number of stable isotopes formed as a result of transmutation is 10^9 - 10^{13} times greater than that of radioactive isotopes. In addition, the tracks of 3 MeV protons and 14 MeV α -particles with an intensity of 10 - $15 \text{ s}^{-1} \cdot \text{cm}^{-2}$ were recorded, with the help of CR-39 plastic detectors, in all experiments. The value of the energy of the registered protons allows, on the basis of the reaction equation (3), to estimate the energy of the nuclear molecular bond of the deuterium with the palladium nucleus ($d\text{-Pd}^N$): $3 \approx 8 - 2.224 - 2q$ (MeV). Therefore, $q \approx 1.4$ MeV.

The authors of [8,9] draw attention to the registration of unknown particles, which leave "strange" traces - tracks in X -ray and nuclear photographic emulsions. The size of the tracks varies from one to tens of millimeters. The form of these tracks is unusual and various; these are broken lines, rectilinear, curved and spiral lines that consist of separate spots. The spots, in turn, can be in the form of circles, ellipses, horseshoes. The authors underline the amazing ability of "strange" particles to penetrate into the metal and move around in it. Particles can escape from the metal, after they have changed its structure and composition and have left behind traces similar to those that remain on photographic emulsions.

Separately, the author [10, 11] investigated the emission of X -rays radiation from a palladium cathode in a high-current ~ 150 mA glow discharge of deuterium and hydrogen, as well as the production of excess thermal power.

In those experiments, X -ray radiation with an energy of 1.5-2 keV with an intensity of up to 100 roentgen/sec was registered and three various modes of X -ray emission were revealed when the parameters of the glow discharge were changed: diffuse X -ray radiation, radiation in the form of

narrowly directed X -ray microbeams, and super-powerful generation of X -ray radiation. The microbeam diameter at a distance of 200 mm from the cathode was estimated to be 10–20 μm , and the angular divergence was $\sim 10^{-4}$. The author notes the anomalously high penetrating power of X -ray microbeams in continuous metallic media. The stationary power of the super-powerful generation of X -ray radiation is estimated to be up to 10 W at a stationary electric discharge power of 50 W .

The excess power was measured with a water flow calorimeter. The measurement system made it possible to control the input electrical power and the thermal power removed by the cooling water with an accuracy of $\pm 0.5 \text{ W}$ with an absolute value of electrical power up to 120 W . In some experiments, the excess thermal power was several tens of W and amounted up to 50%.

The properties of reactions of low-energy transmutation of elements (hereinafter referred to as LTE or transmutation), revealed in experiments with a glow discharge, are characteristic of other experiments that have nothing to do with cold nuclear fusion. For example: in industrial, electronic, zone melting of zirconium ingots in a vacuum furnace [12]; in explosions of metal targets irradiated by a powerful pulse of electrons [13,14]; during explosions in liquid dielectric media of metal foils, through which a powerful pulse of electric current was passed [15,16]; when exposed to a pulsed current on a melt of lead with copper [17]; when applying electric current in water-mineral media [18]; with ultrasonic treatment of aqueous saline solutions [19]; when irradiated with braking quanta of condensed gases [20-22]; in growing biological structures [23-25] and in many others [18,26,27].

All above experiments were carried out by the authors dozens and hundreds of times. The results of experiments on the transmutation of chemical elements are guaranteed to be reproduced and therefore there are no doubts about them.

The main properties of low-energy transmutation reactions include:

- In all these experiments, new chemical elements appear that were absent in the starting material

before the beginning of the transmutation processes. This indicates that the atomic nuclei of some chemical elements are converted into atomic nuclei of other elements.

- In the products of transmutation, a ratio of isotopes of chemical elements is registered that differs from the natural ratio.
- As a rule, in most experiments, the products of transmutation are stable isotopes of elements. In special experiments, the conversion of radioactive isotopes into stable elements was carried out.
- Transmutation reactions are not accompanied by gamma and beta radioactive radiation.
- The yield of transmutation products in some experiments reaches tens of percent (10-25%) of the total mass of the condensed matter. This yield is incomparable with the yield of such products in conventional nuclear reactions.
- In the reactions of LTE, excess thermal, in some cases, electrical energy is released, the values of which cannot be explained by chemical reactions.
- The experimental and calculated values of the excess energy released in a separate transmutation reaction are small and range from tens of keV to several MeV.
- In some experiments, the authors note that the process of transmutation is accompanied by unknown radiation, which leaves its “strange” traces in photographic emulsions, on thin sections of metals, and which, when interacting with a substance, change its structure and chemical composition.
- The methods of the experiments on transmutation are extremely diverse and fundamentally different from the methods of nuclear physics.

As a result, two conclusions can be drawn:

1. **Nuclear reactions take place in transmutation reactions, just as in cold fusion reactions.**
2. **The properties of cold fusion reactions and transmutation reactions contradict to the properties of conventional nuclear reactions.**

3. NEW PARADIGM

The existing contradiction between conventional and low-energy nuclear reactions cannot be resolved otherwise than by way of a worldview jump, like those made by the community of scientists: in the transition from the concept of a flat earth to a spherical earth, from a geocentric system to a heliocentric system; when Newton discovered gravitational interaction and created classical mechanics; and in the transition from classical mechanics to quantum mechanics, special and general theory of relativity; at the discovery of electromagnetic, strong and weak interactions and at the discovery of the atom and atomic nucleus.

According to T. Kuhn [28]: “a scientific revolution occurs when scientists discover anomalies that cannot be explained using the current paradigm, within which scientific progress has taken place up to this point. Therefore, the new paradigm should be considered not just as a current theory, but as a change of a whole worldview, in which this paradigm exists along with all the conclusions made thanks to it”.

To make a real worldview transition, it is necessary to realize that in nature

- **In a condensed medium, in a strong magnetic field, nuclear reactions occur at low energies!** (In the reaction volume, excitation energy is < 1 eV/atom).

In a vacuum, nuclear reactions occur at high energies (> 10 keV/nucleus – thermonuclear fusion). Here, vacuum is understood as the residual gas pressure (less than 10^{-2} Pa) required to accelerate elementary particles or heavy ions to energies sufficient for their subsequent implementation of ordinary collisional nuclear reactions. Such a vacuum exists in interstellar space and in the vacuum chambers of accelerators. All the rest: stars and planets are condensed matter, moreover, excited condensed matter.

Modern nuclear physics studies nuclear reactions that take place in a vacuum. At the same time, the scientific community projects a part of the laws of nuclear physics operating in a vacuum onto condensed matter. In some cases, this is incorrect. Investigations of low-energy nuclear

reactions occurring in weakly excited condensed matter are necessary and inevitable.

4. NEW STATE OF MATTER

Low-energy transmutation reactions of interacting atomic nuclei cannot be explained within the framework of the old paradigm, within the framework of traditional physical concepts. There are three theoretical prohibitions for the phenomenon of transmutation [16,29]:

1. The impossibility for atomic nuclei to overcome, at their collision, the Coulomb barrier between them.
2. Extremely small probabilities of weak processes, which are responsible for the necessary conversion of neutrons into protons or vice versa in order to obtain stable isotopes in the output channel of transmutation reactions.
3. Low probabilities of multiatomic and, therefore, multinuclear reactions even in the absence of a Coulomb barrier. Multinuclear transmutation reactions must be introduced to explain the production of heavy chemical elements in many experiments in a medium consisting of light elements. Such heavy elements cannot be obtained in paired reactions that occur between light elements of the medium.

The above properties of transmutation reactions and prohibitions on their occurrence in the system analysis revealed the requirements that must be fulfilled for the implementation of low-energy transmutation reactions [30]:

1. The electronic structure of atoms and the nucleon structure of nuclei must change. Atoms must turn into transatoms, and nuclei must turn into transnuclei.
2. The electrons of the transatom must be close to the transnucleus. The electron wave functions should overlap significantly with nuclear wave functions.
3. Some of the electronic states of the transatom, and especially those, which are closest to the transnucleus, should not be occupied by electrons.
4. Transatoms must be attracted to each other.

5. During transmutation, the interaction of many transatoms and, accordingly, many transnuclei should occur simultaneously.
6. Transnuclei should be able to approach each other within the range of nuclear forces.
7. After transmutation in a condensed matter, transatoms and transnuclei should be transformed into conventional atoms and nuclei.

The requirements listed above for the course of transmutation reactions are, in fact, at the same time the properties of a new state of matter called a spin nuclide electron condensate [30,31]. A spin nuclide electron condensate is a transatom, in which electrons are paired into orthobosons $S = 1\hbar$. The paired electrons form a Bose-Einstein condensate. In the center of the transatom there is a transnucleus formed by the ultrastrong magnetic field of the electron Bose-Einstein condensate. The properties of the transnucleus differ from those of the conventional nucleus.

5. NEW FUNDAMENTAL INTERACTION

The most plausible scientific concept that satisfies all the requirements listed above is the theory of condensation of atomic electrons in the immediate vicinity of the nucleus due to their pairing into orthobosons with a spin equal to one $S=1\hbar$ [32], and the mechanism of automatic concentration of transnuclei and the implementation of transmutation reactions by them when transatoms combine their orthobosons.

The interaction responsible for low-energy nuclear reactions is the interaction, which is associated with both indistinguishability of identical objects: elementary particles, protons, neutrons, atomic nuclei, molecules, etc; and with the interaction of objects that are connected by resonant R-states. This interaction is called *resonant interference exchange* (RIEX) interaction.

RIEX interaction includes both the well-known exchange interaction between identical objects [33] and recently discovered exchange interaction between any objects A and B, which have resonant R-states, which belong to a composite system, which consists of objects A+B [5,34]. A composite system of A+B objects is not the result

of merging these objects. The resonant R-state of the composite system A+B is, in a certain sense, a certain “image” of objects A and B. This “image” is similar and “identical” to both object A and object B. The nature of the exchange interaction is associated with the overlap and interference of wave functions of identical objects or objects that have resonant R-states.

The principle of identity states that it is impossible, experimentally, to distinguish between identical objects or identical particles. So, if the places or states *a* and *b* of two identical particles 1 and 2 are interchanged: $\psi_a(1)\psi_b(2) \rightarrow \psi_a(2)\psi_b(1)$, then the result of the interaction between them will not change. Here $\psi_a(1) = [\psi_a(x_1, y_1, z_1)]S(1)$ and $\psi_b(2) = [\psi_b(x_2, y_2, z_2)]S(2)$ are the wave functions of particles, which are the products of their coordinate parts $[\psi_{a,b}(x, y, z)]$ by their spin parts *S*(1) and *S*(2), and $\psi_a(1)\psi_b(2)$ and $\psi_a(2)\psi_b(1)$ are wave functions of two particles.

The result of the interaction will not change if the wave function of particles is represented by a superposition of wave functions of two states - an eigenstate $\psi_a(1)\psi_b(2)$ and an identical state $\psi_a(2)\psi_b(1)$:

$$\psi^\pm(1,2) = \frac{1}{2} \{ \psi_a(1)\psi_b(2) \pm \psi_a(2)\psi_b(1) \}. \quad (4)$$

The plus sign in expression (4) describes bosons – particles with zero or integer spin, $s = 0, 1\hbar, 2\hbar, \dots$. The bosons obey Bose-Einstein statistics, in which the sign of the wave function does not change when the particles are rearranged. The minus sign describes fermions, i.e. particles with half-integer spin, $s = \hbar/2, 3\hbar/2, \dots$. Fermions obey the Fermi-Dirac statistics, in which the sign of the wave function $\psi^-(1,2)$ changes to the opposite when the particles are rearranged. Our Universe consists mainly of fermions: electrons, protons, neutrons, neutrinos.

By definition, the square of the wave function of particles is equal to the probability density of their being at a given point in space and at a given moment in time. If expression (4) is squared, then

$$|\psi^\pm(1,2)|^2 = \frac{1}{4} \{ (|\psi_a(1)\psi_b(2)|^2 + |\psi_a(2)\psi_b(1)|^2) \pm 2[\psi_a^*(1)\psi_b^*(2)\psi_a(2)\psi_b(1) + \psi_a(1)\psi_b(2)\psi_a^*(2)\psi_b^*(1)] \}.$$

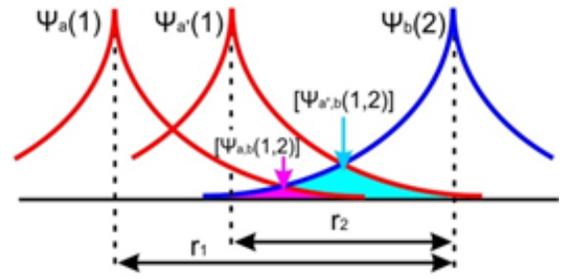


Fig. 2. Overlap of wave functions of identical particles.

The values in parentheses are the probability of particles being in their own and identical states. The value in square brackets is the probability of the particles being in an exchange state when each of the particles is simultaneously in two states *a* and *b*. An exchange state arises when the wave functions of identical particles overlap (Fig. 2). The more the wave functions of identical particles overlap, the greater value has the exchange state.

Fundamental exchange interactions. Since particles or other objects have masses, electric, baryonic, lepton charges, spins, spin magnetic moments, they participate in all fundamental interactions. The fundamental interactions that occur between objects in their eigenstates are the basic interactions between objects. Moreover, identical objects participate in additional, fundamental exchange interactions:

- Strong exchange interaction [F] – $\psi_a^*(1)\psi_b^*(2)[F]\psi_a(2)\psi_b(1)$;
- Electromagnetic exchange interaction [EM] – $\psi_a^*(1)\psi_b^*(2)[EM]\psi_a(2)\psi_b(1)$;
- Weak exchange interaction [W] – $\psi_a^*(1)\psi_b^*(2)[EM]\psi_a(2)\psi_b(1)$; and
- Inertial-gravitational exchange interaction [IG] – $\psi_a^*(1)\psi_b^*(2)[IG]\psi_a(2)\psi_b(1)$.

Fundamental resonant interference exchange interactions. If objects A and B have resonant R-states that belong to the composite system A+B, then these R-states are excited on the length of the wave functions of objects A and B. Each resonant R-state has its own wave function $\psi_R(R)$. The wave function $\psi_a(A)$ of object A is present in the R-state in proportion to the coefficient K_A $\psi_R(A) = K_A\psi_a(A)$ (Fig. 3). Accordingly, the wave function $\psi_b(B)$ of object B is present in the R-state in proportion to the

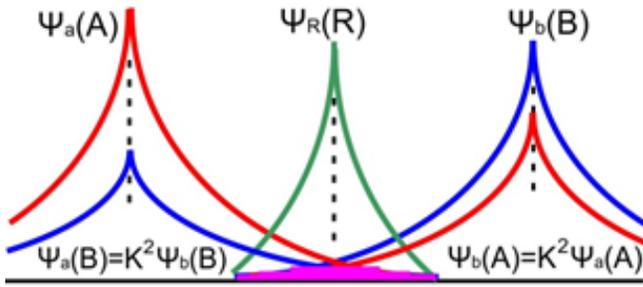


Fig. 3. *Overlap of wave functions of “identical” objects A and B during the formation of the R-state.*

coefficient K_B : $\psi_R(B) = K_B \psi_b(B)$. Coefficients K_A and K_B characterize the similarity of objects A and B to the resonant R-state. The coefficients K_A and K_B are individual for each R-state. Usually K_A and $K_B < 1$. And vice versa, the coefficients K_A and K_B characterize the similarity of the resonant R-state to objects A and B. Therefore, the wave function $\psi_a(A)$ of object A will be present in the *b*-state $\psi_b(A)$ with the coefficient $K_B K_A$: $\psi_b(A) = K_B \psi_R(A) = K_B K_A \psi_a(A)$ (Fig. 3). With the same coefficient $K_A K_B$, the wave function $\psi_b(B)$ of object B will be present in the *a*-state: $\psi_a(B) = K_A \psi_R(B) = K_A K_B \psi_b(B)$.

Thus, we can say that objects A and B are “identical” to each other with a generalized similarity coefficient $K^2 \equiv K_A K_B$. The spatial area of the RIEX interaction for all fundamental potentials: F, EM, W, and IG, is determined by the reduced lengths of the wave functions of “identical” objects A and B with the coefficient K^2 : $K^2 \cdot \lambda_{A,B}$ (Fig. 3).

When the reduced lengths of the wave functions of objects A and B overlap, then they are simultaneously in two states: eigenstate $\psi_a(A)\psi_b(B)$ and identical $\psi_b(A)\psi_a(B)$. And their total wave function is equal to $\psi^\pm(A, B) = \psi_a(A)\psi_b(B) \pm \psi_a(B)\psi_b(A)$.

The eigenpart $\psi_a(A)$ interacts, in its place *a*, with the identical part $\psi_a(B)$, and the eigenpart $\psi_b(B)$ interacts, in its place *b*, with $\psi_b(A)$. Thus, due to the resonant interference exchange interaction, the short-range strong F and local weak W interactions become “long-range”.

As already mentioned, “identical” objects A and B, which have resonant R-states, participate in additional, fundamental RIEX interactions:

in strong RIEX interactions [F] – $\psi_a^*(A)\psi_b^*(B)[F]\psi_a(B)\psi_b(A)$;
 in electromagnetic RIEX interaction [EM] – $\psi_a^*(A)\psi_b^*(B)[EM]\psi_a(B)\psi_b(A)$;
 in weak RIEX interaction [W] – $\psi_a^*(A)\psi_b^*(B)[W]\psi_a(B)\psi_b(A)$ and
 in inertial-gravitational RIEX interaction [IG] – $\psi_a^*(A)\psi_b^*(B)[IG]\psi_a(B)\psi_b(A)$.

The energy of the RIEX interaction is an additional contribution to the total energy of interacting objects. This additional energy is a resonance energy.

6. PROPERTIES OF RIEX INTERACTION

The main properties of fundamental resonant interference exchange interactions include:

1. Fundamental RIEX interactions between “identical” objects occur the more intensively, the more their wave functions overlap.
2. At the reduced lengths of the wave functions of “identical” objects, the short-range strong F and local weak W interactions become, due to the RIEX interaction, “long-range” interactions.
3. The energy of the RIEX interaction E_c is an additional contribution to the total energy of the E system of interacting objects.
4. The energy of the RIEX interaction E_c can be positive or negative, depending on the type of fundamental interaction.
5. The sign of the contribution of the exchange energy to the total energy of the system can be different: plus or minus $\pm E_c$, depending on whether the coordinate part of the total wave function of objects is symmetric or antisymmetric. Therefore, the exchange energy E_c can reduce, and in some cases, completely compensate for the main part of the energy of the system C, $E^C \equiv C - E_c$.

One of the consequences of the exchange Coulomb interaction is that it allows, in an atom in a strong magnetic field $B > 30$ T, electrons with parallel spins to pair into orthobosons with $S = 1h$.

As indicated above, most transmutation experiments are induced through the electronic action on a condensed matter by powerful pulse

of electrons or powerful currents. The directed motion of electrons creates a magnetic field both due to the transfer of the electric charges of electrons e^- and due to the transfer of their magnetic moments μ_e . The magnetic moments of the electron flow, due to the property of helicity, are directed in one direction, towards their momenta. The spins of electrons and neutrinos (e^- and ν_e) are directed against the momentum - they have left helicity, and the spins of positrons and antineutrinos (e^+ and $\bar{\nu}_e$) are directed along the momentum - they have the right helicity of particles. The electron's magnetic moment is directed against the spin. The \mathbf{B}_s magnetic field created by the magnetic moments is described by the Landau equation [35]:

$$\mathbf{B}_s = \mu_0 \sum_i \frac{3\mathbf{n}_i(\boldsymbol{\mu}_e \cdot \mathbf{n}_i) - \boldsymbol{\mu}_e}{r_i^3}, \quad (5)$$

where $\mu_0 = 1.26 \cdot 10^{-6}$ H/m is magnetic constant; $\mu_e = 9.29 \cdot 10^{-24}$ J/T = $5.79 \cdot 10^{-5}$ eV/T, r is the distance from the electron to the point at which the \mathbf{B}_s field is calculated; \mathbf{n} is a unit vector in the r direction, i is the number of electrons with parallel spins. According to formula (5), the magnetic moment of an electron $\boldsymbol{\mu}_e$ creates a magnetic field equal to 30 T at a distance of 0.092 nanometers along its axis (the diameter of a hydrogen atom is 0.106 nm). The same magnetic field of 30 T is created in the center of an electron lattice cell with unidirectional magnetic moments of electrons and with a side of $1.6 \cdot 10^{-10}$ m, which is comparable with the size of atoms. So an orthohelium atom, whose electron spins are parallel, has magnetic fields in the nuclear region of ~ 410 T and ~ 70 T at its diameter of $\sim 1.75 \cdot 10^{-10}$ m. The strong magnetic field > 30 T that arises in a condensed matter gives rise to pairing of atomic electrons into orthobosons with spins $S = 1\hbar$, thus forming a “magnetic nesting doll” with the creation of a spin nuclide electron condensate [31].

The paper [36] shows that, in a strong magnetic field, pairing of atomic electrons with parallel spins occurs due to:

first, the exchange interaction of electrons, which has the character of attraction and,

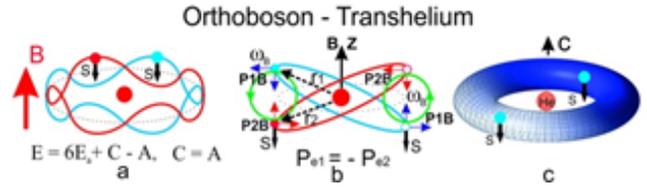


Fig. 4. *a* - electron oscillations; *b* – electron oscillations in the longitudinal and in the transverse magnetic field \mathbf{B} directions; *c* – orthoboson - transhelium.

secondly, the appearance of oscillations of electrons around their orbitals (Fig. 4a,b).

The $\boldsymbol{\ell} + \mathbf{s} \parallel \mathbf{j} + \mathbf{j}$ bonds of all atomic electrons are broken, and their orbital moments $\boldsymbol{\ell}$ are “frozen” into the field in a strong magnetic field \mathbf{B} . The interaction between electrons makes them oscillate around the orbitals. These oscillations are quantized by introducing a new, oscillatory quantum number n_b . The exchange interaction between two electrons and their oscillations with quantum numbers $\pm n_b$ allow electrons to create an orthoboson with $S = 1\hbar$. The quantum numbers of oscillations of paired electrons are equal to each other in absolute value, but opposite in sign $n_b^1 = -n_b^2$, $n_b = 1, 2, 3, \dots$. Therefore, the Pauli principle is fulfilled for them. Due to the exchange interaction, correlated oscillations appear in two electrons (Fig. 4a,b). – The sum of the momenta of two electrons in a pair is equal to zero, i.e. electrons in a pair have momenta $\mathbf{P}_{1e} = -\mathbf{P}_{2e}$, which are equal in magnitude and opposite in direction. – Electrons in a pair oscillate both along and across the magnetic field \mathbf{B} (Fig. 4b). Since electrons in a pair oscillate in antiphase $\mathbf{P}_{1e} = -\mathbf{P}_{2e}$ ($n_b^1 = -n_b^2$), this motion allows two electrons in the same energy states to be in non-intersecting spatial regions (Fig. 4b). The trajectories of electron motion can be represented as closed spirals nested into each other, located on the surface of the toroid (Fig. 4c). The two electron spirals are similar to the double helix of a DNA molecule.

The trajectories of several orthobosons in a multielectron transatom create a **toroidal spin electron magnetic twist** - a torsem-twist of spirals nested into each other, which resembles the DNA code of the main character of the film “The Fifth Element” - Leeloo [37]. The torsem-twist is located on the surface of the toroid. The

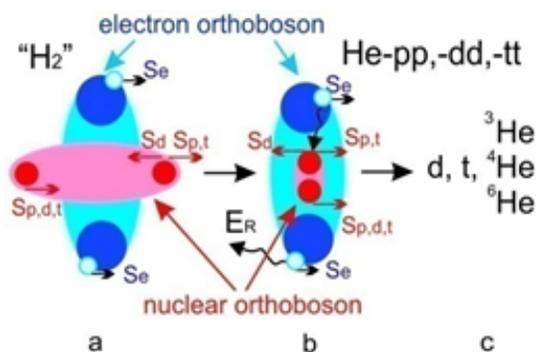


Fig. 5. Electronic and nuclear orthobosons in the hydrogen transmolecule and in the “helium” transmolecule.

multielectron atom forms a Transatom, which has an electron Bose-Einstein condensate (Fig. 6). The atomic electrons inevitably pair, in a strong magnetic field, into orthobosons, ordinary atoms inevitably transform into transatoms.

Since the coordinate part of the total wave function of electrons in the orthoboson is antisymmetric $n_b^1 = -n_b^2$, the total Coulomb energy of electrons is $E = 6E_a + C - Ec$ (Fig. 4a), where $6E_a$ is the interaction energy of two paired electrons with a nucleus, E_a is the binding energy of a single electron with a nucleus, C is the basic, ordinary Coulomb energy of repulsion of two electrons, and Ec is their exchange Coulomb energy. The energies C and Ec are positive. And since both electrons are in equal energy states, then $C = Ec$ and $E^C \equiv C - Ec = 0$. The exchange Coulomb attraction of two electrons completely compensates their Coulomb repulsion. A quantum paradox arises: “The Waves extinguish the Wind”.

Two hydrogen atoms, which interact in a strong magnetic field, combine, due to the occurrence of electron oscillations ω_p , to form a hydrogen transmolecule “H₂”, which electrons are paired into an orthoboson [36] (Fig. 5a). This orthoboson creates an electromagnetic potential well in the hydrogen transmolecule “H₂” with a magnetic induction vector at the center of $\sim 10^4$ T. The protons ($S_p = \hbar/2$) in such a ultrastrong and inhomogeneous magnetic field will have parallel spins $\uparrow\uparrow$. The protons in a hydrogen transmolecule form, like electrons, a bound state – a nuclear orthoboson $S = 1\hbar$ due to their own exchange interaction and their own correlated oscillations. Just like electrons, the exchange

Coulomb interaction of protons completely compensates for their Coulomb repulsion. This will lead to the approaching of protons to each other to nuclear distances and the formation of a “helium-pp” (“He-pp”) transmolecule [36] (Fig. 5b). The protons in the “He-pp” transmolecule can be replaced with deuterons d or tritons t . Then transmolecules “He-dd” and “He-tt” are formed.

One of the perturbing potentials between electrons and between protons in orthobosons is the above Coulomb interaction: $V_C = k \cdot e^2 / r_{12}$, where r_{12} is the distance between electrons or protons, $k = 1/4\pi\epsilon_0 = 8.99 \cdot 10^9 \text{ N} \cdot \text{m}^2 / \text{C}^2$ ($\epsilon_0 = 8.85 \cdot 10^{-12} \text{ F/m}$ is the electrical constant), $e = 1.6022 \cdot 10^{-19} \text{ C}$ is the electron and proton charges. Another disturbing potential between electrons and between protons in orthobosons is the gravitational interaction: $V_G = G \cdot m^2 / r_{12}$, where $G = 6.67 \cdot 10^{-11} \text{ m}^3 \cdot \text{s}^{-2} \cdot \text{kg}^{-1}$ is the gravitational constant, m is the electron mass $9.11 \cdot 10^{-31} \cdot \text{kg}$ or proton $1.67 \cdot 10^{-27} \cdot \text{kg}$. The gravitational energy of electrons or protons, or other identical nuclei bound into orthobosons, is equal to $E^G = \mathfrak{D} - E_G$, where \mathfrak{D} is the usual gravitational energy of attraction of two electrons or two protons, or two other identical nuclei to each other, and E_G is their exchange gravitational energy. Gravitational energies \mathfrak{D} and E_G of electrons, protons and other objects have negative values. And, as in the case of Coulomb interaction, the exchange gravitational repulsion of electrons, protons, neutrons and other identical nuclei, which form an orthoboson, fully compensates for their gravitational attraction $E^G = 0$.

Consequently, long-range Coulomb and gravitational interactions are absent in transmolecules for pairs of electronic and nuclear orthobosons. But in these transmolecules, strong and weak interactions act at the lengths of the wave functions of nuclear orthobosons. And weak interaction takes place in electronic orthobosons and between them and nuclear orthobosons.

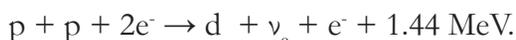
The lengths of the wave functions λ of the orthobosons in the ground state with $n = 1$ are equal to the radii of the orthobosons R : $\lambda = R = \hbar/P$, where n is the principal quantum number, P is the momentum of the particles that make up

the orthoboson. Therefore, the wave functions of electrons in a torsem-twist overlap significantly with nuclear wave functions.

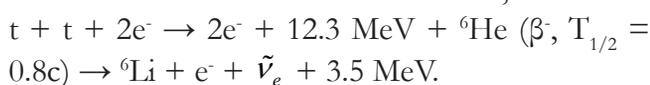
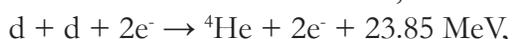
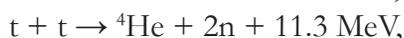
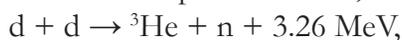
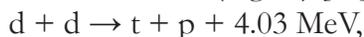
Thus, the resonant interference exchange interaction is a new, Fifth fundamental interaction. It is a Universal Interaction as it always includes all other four fundamental interactions. Moreover, RIEX interaction changes the action of components of its four interactions and controls them, like "The Fifth Element" from the film of the same name, which controls the four elements [37].

7. TRANSMUTATION REACTIONS

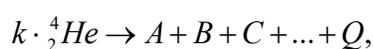
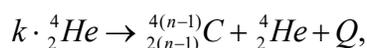
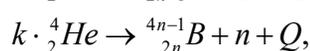
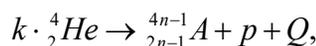
Since the protons in the "He-pp" transmolecule are at a nuclear distance, this will lead to a nuclear transmutation reaction with the participation of electronic orthoboson. This will synthesize a deuteron:



With the production of deuterium and tritium, transmolecules "He-dd" and "He-tt" will be formed, which are also nuclear orthobosons. They enter into nuclear transmutation reactions without a Coulomb barrier, including reactions with the participation of electronic orthoboson, $2e^-$, with the formation of protons, neutrons, tritons, ^3He , ^4He , $^6\text{He} \rightarrow ^6\text{Li}$ (Fig. 5c) [32]:



At the same time, orthohelium atoms will form multinuclear transmolecules $k \cdot ^4_2\text{He}$ with helium Bose-Einstein condensate. The creation of such transmolecules leads to multinuclear reactions, with the emission of protons, neutrons, alpha particles and heavy chemical elements: A, B, C ..., with a nuclear charge $Z \geq 6$ [38]:



where Q is the energy released as a result of the reaction.

Since the spins \mathbf{s} and the magnetic moments of the electrons μ_e in the Bose-condensate are directed in the same direction, and they generate, in and around transatoms, an ultrastrong directed inhomogeneous and anisotropic magnetic field up to $\mathbf{B}_s \sim 10^5\text{--}10^{10} \text{ T}$ (5) [30]. In this case, the inhomogeneity and anisotropy of the magnetic field ΔB_s exist at the dimensions of the interacting transnuclei. This leads to the uncertainty of the energies of nucleons with magnetic moments μ_N in transnuclei - $\Delta B_s \mu_N$, which is equivalent to the inhomogeneity of time. Thus, we can say that the RIEX interaction changes the space-time structure and properties. Consequently, during the interaction of transnuclei moving in a transmolecule, the integrals of motion are not preserved: the law of conservation of momentum, the law of conservation of angular momentum (spin) and the law of conservation of energy are violated. Thus, the conclusion supposes that the study of the physical vacuum is impossible otherwise than through the study of the condensed state of matter in extreme conditions.

The internal ultrastrong magnetic field B_s^0 , when it interacts with the magnetic spin and magnetic orbital moments of nucleons in the nucleus, changes the structure of the nucleus, and turns it into a Transnucleus.

External ultrastrong magnetic fields B_s^R of transatoms attract them to each other ($B_s^R * \mu_e > 10^5 * 5.8 \cdot 10^{-5} \text{ eV}$). Electronic Bose-condensates of two transatoms are combined into a common condensate. A double nuclear transmolecule is formed from the transnuclei. Other transnuclei can join it. A multinuclear transmolecule is formed, in which multinuclear reactions take place, including those that involve electronic orthobosons. Thus, nuclear-electron or strong-weak reactions occur, the products of which are non-radioactive.

These reactions are carried out due to the resonant interference exchange interaction. **Fig. 6** shows the formation of a sodium transmolecule $^{23}_{11}\text{Na}^M$ from the transatoms of boron $^{11}_5\text{B}^T$ and

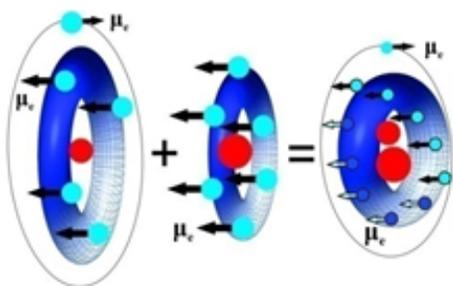


Fig. 6. The formation of sodium transmolecule from boron and carbon transatoms.

carbon ${}^{12}_6\text{C}^T$. The transnuclei ${}^{11}_5\text{B}^T$ and ${}^{12}_6\text{C}^T$ in the transmolecule ${}^{23}_{11}\text{Na}^M$ cannot merge because of the Coulomb barrier between them. But strong and electroweak forces act due to the existence of resonant R-states in the transmolecule between the transnuclei and between the transnuclei and their orthobosons.

The sodium transmolecule ${}^{23}_{11}\text{Na}^M$ can be composed of other transatoms, for example: transhelium ${}^4_2\text{He}^T$ and transfluorine ${}^{19}_9\text{F}^T$. From the same number of nucleons, regardless of whether they are protons or neutrons, other transmolecules can be composed, for example: magnesium ${}^{23}_{12}\text{Mg}^M$ from transhelium ${}^3_2\text{He}^T$ and transneon ${}^{20}_{10}\text{Ne}^T$. All these transmolecules have their own resonant R-states, but they have different nuclear and nuclear molecular binding energies. The energetically closer to each other are the binding energies of these transmolecules, the more their resonant R-states overlap and the more likely the transmutation reaction occurs. For this reason, the energy released in transmutation reactions is insignificant, if compared to conventional nuclear reactions: from tens of keV to several MeV. The **Fig. 7** shows all possible computer-calculated transmutation reactions. It is not excluded that the

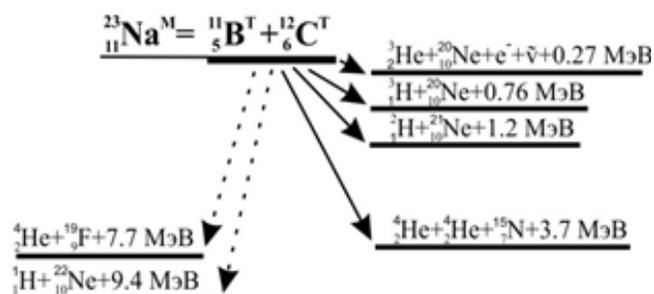


Fig. 7. Transmutation reactions on the example of sodium transmolecule transformation.

transformation of the sodium transmolecule into the conventional sodium atomic nucleus may occur ${}^{23}_{11}\text{Na}: {}^{23}_{11}\text{Na}^M \rightarrow {}^{23}_{11}\text{Na} + 13.36 \text{ MeV}$. The energy released in the 13.36 MeV reaction is transferred to the electron orthobosons that surround the nucleus. Thus, transmutation reactions can be represented as reactions of nucleon and multinucleon transfers between transnuclei with a possible conversion of protons into neutrons and vice versa, as well as reactions of radiationless fusion and fission of transnuclei.

The atomic nuclei scatter after the implementation of low-energy nuclear reactions. And, if they are not in a strong magnetic field, then the reaction products form conventional nuclei and conventional atoms.

It is remarkable that the fission of the uranium-235 nucleus by a thermal neutron is associated with the RIEX interaction [5,34]. The process of a thermal neutron capture by a uranium-235 nucleus has a pronounced resonance character when the neutron energy is close to one of the values corresponding to the R-level of the composite system: a neutron plus a uranium-235 nucleus. After RIEX capture by the uranium-235 nucleus of a neutron to the R-level, it carries out an electromagnetic transition to the highly excited level of their common nucleus, uranium-236. The energy of this excited state $\sim 7\text{-}8$ MeV is greater than the energy of the 6 MeV Coulomb fission barrier. Therefore, the uranium-236 nucleus with a greater degree of probability, determined by nuclear interaction, will split into two fragments than it finds itself in the ground state, the transition to which is determined by electromagnetic interaction. Thus, nuclear power engineering, unlike all other types of power engineering, is a resonant technology.

8. EVOLUTION AND ECOLOGY

In the last century, the evolutionary development of living nature, the biosphere passed into a new state, in the noosphere. In 1944, V.I. Vernadsky wrote [39]: “In the twentieth century, for the first time in the history of the Earth, the man recognized and embraced the entire biosphere. The Humankind, taken as a whole, has become

a powerful geological force, ever growing force. And the Humankind, its thinking and work face the question of restructuring the biosphere in the interests of free-thinking humanity as an integral whole. This new state of the biosphere is the noosphere”.

In the 21st century, ecology, which studies, among other things, the influence of human activity on the environment and wildlife, came to the conclusion that at this stage of evolution, the noosphere is characterized by a contradiction between its irresistible development and the necessary preservation. The main problem in the preservation of the noosphere is environmental pollution caused by the ever increasing production activity of mankind. The pollution of the environment and wildlife has reached geological proportions. The scale of pollution is so huge that the Earth's biosphere is no longer able to dispose and neutralize them. For this reason, the humankind faces the fundamental question of the existence of the Living ecosystem, in general, and the existence of humanity itself, in particular.

The main sources of pollution are the combustion of hydrocarbons for the needs of the energy sector and the processing of mineral materials. The combustion of hydrocarbons is accompanied by an increase in the level of carbon monoxide in the atmosphere, which leads to a greenhouse effect that catastrophically changes the Earth's climate. The processing of minerals entails the appearance of waste in the form of new chemical compounds, foreign and toxic materials, harmful to living organisms and not capable of being disposed in the biosphere.

Despite all the efforts of the humankind, including the creation of alternative energy sources, it is becoming more and more obvious that the modern technology is not able to eliminate the existing contradiction between the development and preservation of the noosphere.

With the discovery of low-energy nuclear reactions and resonant interference exchange interactions, this contradiction can be overcome. This conclusion is based on the idea of creating

resonant technologies as a way to obtain maximum results at minimum costs.

First of all, this concerns the development of new energy sources based on low-energy nuclear reactions: cold nuclear fusion reactions and nuclear transmutation reactions. Due to the lack of carbon dioxide production, harmful emissions and effluents, due to the absence of radiation and radioactive waste, such resonant energy sources (RES) are environment friendly energy sources. In future, they will replace energy sources using fossil fuels. As a result, the need for the extraction, transportation and processing of fossil fuels for energy purposes will disappear. Powerful and compact, RES can be deployed in centralized or dispersed configurations. RES are cheap and practically inexhaustible sources of energy. For these reasons, they can be widely used in the elimination of waste products harmful to the biosphere. Waste disposal facilities can use resonant technologies in the same way as industrial production itself should use such technologies.

The composition of many modern materials that are used in the manufacture of industrial products include rare chemical elements and elements scattered in nature. The concentration of such elements in mineral deposits can be tenths of a gram per ton. Obviously, the production of such elements in industrial quantities requires the processing of hundreds of millions of tons of ore. Since some chemical elements are converted into others in transmutation reactions, these resonance reactions can be used in the production of rare elements and their isotopes from cheap and widespread chemical elements. This will save huge energy, material and human resources.

Based on the properties of low-energy nuclear reactions and the properties of transatoms, it is possible to create other, both obvious and non-trivial, disruptive innovation [29]. Such resonant technologies will radically change the technological structure of the noosphere. And, as a consequence, they will inevitably affect the evolution of the noosphere and save it at the same time.

9. CONCLUSION

Low-energy nuclear reactions in condensed matter occur due to resonant interference exchange interaction. The new paradigm is based on new nuclear reactions, on a new state of matter: a spin nuclide electron condensate, and, first of all, on a new resonant interference exchange interaction. The RIEEX interaction is a universal interaction not only because it includes and controls all the other four fundamental interactions, but also because its actions extend to the Whole of Nature, from elementary particles to complex biological and social systems.

The new paradigm gave a start to the formation of new technological ways and a new civilizational paradigm.

REFERENCES

1. Strashnov I, Fernando R, Izosimov I. Trace analysis of radioisotopes by laser spectroscopy and mass spectrometry. *J. Radioanalytical and Nuclear Chemistry*, 2019, 322:1437-1445.
2. Fleishmann M, Pons S, Hawkins M. Electrochemically induced nuclear fusion of deuterium. *J. Electroanal. Chem.*, 1989, 261:301-308.
3. Iyengar PK. Cold Fusion Results in BARC Experiments. *Proc. 5 Intern. Conference On Emerging Nuclear Energy Systems (ICENES)*, Karlsruhe, FRG, Singapore, 1989; Iyengar PK. and Srinivasan M. Overview of BARC Studies in Cold Fusion. *Proc. 5 Intern. Conference on Cold Fusion (ICCF-5)*, 1990, Utah, Sald Lake City.
4. Miles MH. Heat and helium production in cold fusion experiments. *Proc. 2nd Int. Conf. on Cold Fusion "The Science of Cold Fusion"*, Como, Italy, Societa Italiana di Fisica, Bologna, Italy, 1990. Miles MH, Hollins RA, Bush BF, Lagowski JJ, Miles RE. *J. Electroanal. Chem.*, 1993, 346:99; Miles MH. *Proc. ICCF-10*, 2003, Cambridge.
5. Mishinsky GV. Resonant interference exchange interaction. *Radioelectronics. Nanosystems. Information Technologies (RENSIT)*, 2019, 11(3):261-278. DOI: 10.17725/rensit.2019.11.261.
6. Karabut AB, Kucherov YaR, Savvatimova IB. Nuclear product ratio for glow discharge in deuterium. *Phys. Letters A*, 1992, 170:265-272.
7. Savvatimova IB, Karabut AB. The products of nuclear reactions recorded at the cathode after experiments in a glow discharge in deuterium. *Povernoc't' [Surface Investigation: X-Ray, Synchrotron and Neutron Techniques]*, 1996, 1:63-75-81.
8. Savvatimova IB. Transmutation Effects in the Cathode Exposed Glow Discharge. Nuclear Phenomena Or Ion Irradiation Results. *Proc. 7th Int. Conf. on Cold Fusion (ICCF)*, Canada, 1998: 342-350; Reproducibility of Experiments in Glow Discharge and Processes Accompanying Deuterium ions Bombardment. *Proc. 8th ICCF*, Italy, 2000:277-283.
9. Nesterovich AV, Rodionov BU, Savvatimova IB. Formation of tracks during cold transmutation of atomic nuclei. *Proc. of 8th Russian Conference on Cold Transmutation of Nuclei of Chemical Elements (RCCTNCE)*, Moscow, 2001:211-215.
10. Karabut AB. Analiz rezultatov registratsii izbytochnoy teplovoy moshchnosti, primesnykh nuklidov s izmenennym sootnosheniem izotopov i pronikayushchego izlucheniya v eksperimentakh s silnotochnym tleyushchim razryadom [Analysis of the results of recording excess heat power, impurity nuclides with a modified natural ratio of isotopes and penetrating radiation in experiments with a high-current glow discharge]. *RCCTNCE-7*, Moscow, 2000, p. 27-35 (in Russ.).
11. Karabut AB. Zavisimost' proizvodstva izbytochnoy teplovoy moshchnosti, produktov yadernykh reaktsiy i emissii rentgenovskogo izlucheniya ot eksperimentalnykh parametrov silnotochnogo tleyushchego razryada [Dependence of the production of excessive thermal power, products of nuclear reactions and emission of X-ray radiation on the experimental parameters of a high-current glow discharge. *RCCTNCE-9*, Moscow, 2002, p. 86-98 (in Russ.).
12. Solin MI. Experimental facts of spontaneous appearance of condensate of soliton charges with the formation of nuclear fusion products

- in liquid zirconium. *Physical Idea in Russia*, 2001, 1:43-58 (in Russ.).
13. Adamenko SV, Selleri F, A. van der Merwe (eds.). Controlled Nucleosynthesis Breakthroughs in Experiment and Theory. *Series: Fundamental theories in Physics*, Springer, 2007, 156:780, <http://www.springer.com/physics/elementary/book/978-1-4020-5873-8>.
 14. Adamenko SV. The concept of artificially initiated collapse of matter and the main results of the first stage of its experimental implementation. *Preprint*, 2004, Kiev. (in Russ.), http://proton-21.com.ua/publ/Preprint_ru.pdf.
 15. Urutskoev LI, Liksonov VI, Tsinoev VG. Observation of transformation of chemical elements during electric discharge. *Annales de la Fondation Louis de Broglie*, 2002, 27(4):701-726.
 16. Kuznetsov VD, Mishinsky GV, Penkov FM, Arbuzov VI, Zhemenuk VI. Low energy transmutation of atomic nuclei of chemical elements. *Annales de la Fondation Louis de Broglie*, 2003, 28(2):173-214.
 17. Krivitsky VA. Paradoxes of transmutation and the development of the Earth. Moscow, NITs "Academica" Publ., 2016, 239 p.
 18. Balakirev VF, Krymsky VV, Bolotov BV, Vasilieva NV, Vachaev AV, Ivanov NI, Kazbanov VI, Pavlova GA, Solin MI, Trofimov VI, Urutskoev LI. *Vzaimoprevrashcheniya khimicheskikh elementov* [Mutual transformation of chemical elements]. Ekaterinburg, Urals Branch of RAS Publ., 2003, 96 p.
 19. Kladov AF. *Cavitation destruction of matter* (in Russ.); <http://roslo.narod.ru/rao/rao1.htm>.
 20. Didyk AYU, Wiśniewski R. Nuclear reactions, induced by gamma quanta, in palladium saturated with deuterium surrounded by dense deuterium gas. *Europhys. Lett*, 2012, 99(2):22001.
 21. Wiśniewski R, Mishinsky GV, Gulbekian GG, Wilczyńska-Kitowska T, Semin VA. Sintez khimicheskikh elementov i tverdotelnykh struktur pri obluchenii γ -kvantami kondensirovannykh gazov [Synthesis of chemical elements and solid-state structures under irradiation by γ -quanta in condensed gases]. *International Journal of Unconventional Science (IJUS)*, 2017, 17-18(5):6-15 (in Russ.), <http://www.unconv-science.org/>.
 22. Didyk AYU, Wiśniewski R, Wilczyńska-Kitowska T, Mishinsky GV, Semin VA. Synthesis of chemical elements under irradiation by bremsstrahlung gamma-rays of palladium in condensed gases. *RENSIT*, 2019, 11(2):143-160, doi: 10.17725/rensit.2019.11.143.
 23. Vysotskii VI, Kornilova AA. *Nuclear fusion and transmutation of isotopes in biological systems*. Moscow, Mir Publ., 2003, 304 p.
 24. Vysotskii VI, Kornilova AA. Transmutation of stable isotopes and deactivation of radioactive waste in growing biological systems. *Annals of Nuclear Energy*, 2013, 62:626-633.
 25. Kornilova AA, Vysotskii VI. Synthesis and transmutation of stable and radioactive isotopes in biological systems. *RENSIT*, 2017, 9(1):52-64. DOI: 10.17725/rensit.2017.09.052.
 26. Proceedings of the 1-26th Russian Conferences on Cold Transmutation of Nuclei of Chemical Elements, 1993-2018.
 27. Proceedings of the 1-22th International Conferences on Condensed Matter Nuclear Science (Cold Fusion).
 28. Kuhn Thomas S. *The Structure of Scientific Revolutions*. University of Chicago Press, 1962, 172 p.
 29. Mishinsky GV, Kuznetsov VD, Penkov FM. Nizkoenergeticheskaya transmutatsiya atomnykh yader khimicheskikh elementov. Raspredelenie po elementam v produktakh transmutatsii. nukleosintez [Low energy transmutation of atomic nuclei of chemical elements. Element distribution in the products of low energy transmutation. Nucleosynthesis]. *IJUS*, 2017, 17-18(5):61-81 (in Russ.).
 30. Mishinsky GV. Magnitnye polya transatomov. Spinovy-nuklidny-elektronny kondensat [Magnetic fields of transatoms. Spin-nuclide-electronic condensate]. *IJUS*, 2017, 15-16(5):6-25 (in Russ.).
 31. Mishinsky GV. Spin electron condensate. Spin nuclide electron condensate. *RENSIT*,

- 2018, 10(3):411-424. DOI: 10.17725/rensit.2018.10.411.
32. Mishinsky GV. Atom in a strong magnetic field. Transformation of atoms to transatoms. *RENSIT*, 2017, 9(2):147-160. DOI: 10.17725/rensit.2017.09.147.
33. Heisenberg W. Über die Spektra von Atomsystemen mit zwei Elektronen. *Z. Phys.*, 1926, 39(7):499-518.
34. Mishinsky GV. Theory of cold fusion reactions. *RENSIT*, 2019, 11(2):125-142. DOI: 10.17725/rensit.2019.11.125.
35. Landau LD, Lifshitz EM. *The Classical theory of Fields*. Oxford, Butterworth-Heinemann, 1975, 402 p.
36. Mishinsky GV. Non-Coulomb nuclear reactions of transatoms. Stellar energy and nucleosynthesis. *RENSIT*, 2018, 10(1):35-52. DOI: 10.17725/rensit.2018.10.035.
37. Besson L, Kamen RM. *The Fifth Element*, film, 1997.
38. Mishinsky GV. Multinuclear reactions in condensed helium. *RENSIT*, 2017, 9(1):94-105. DOI: 10.17725/rensit.2017.09.094.
39. Vernadsky VI. *Khimicheskoe stroenie biosfery Zemli i ee okruzheniya* [The chemical structure of the Earth's biosphere and its environment]. Moscow, Nauka Publ., 1965, 374 p.