ISSN 2500-316X (Online)

RUSSIAN TECHNOLOGICAL JOURNAL

РОССИЙСКИЙ ТЕХНОЛОГИЧЕСКИЙ ЖУРНАЛ

Information systems.
Computer sciences.
Issues of information security

Multiple robots (robotic centers) and systems.
Remote sensing and non-destructive testing

Modern radio engineering and telecommunication systems

Micro- and nanoelectronics.
Condensed matter physics

Analytical instrument engineering and technology

Mathematical modeling

Economics of knowledge-intensive and high-tech enterprises and industries.
Management in organizational systems

Product quality management. Standardization

Philosophical foundations of technology and society

12+
www.rtj-mirea.ru

МИРЭА
Российский технологический университет

11(5) 2023

ISSN 2500-316X (Online)

RUSSIAN TECHNOLOGICAL JOURNAL

РОССИЙСКИЙ ТЕХНОЛОГИЧЕСКИЙ ЖУРНАЛ

- Information systems. Computer sciences. Issues of information security
- Multiple robots (robotic centers) and systems. Remote sensing and non-destructive testing
- Modern radio engineering and telecommunication systems
- Micro- and nanoelectronics. Condensed matter physics
- Analytical instrument engineering and technology
- Mathematical modeling
- Economics of knowledge-intensive and high-tech enterprises and industries. Management in organizational systems
- Product quality management. Standardization
- Philosophical foundations of technology and society

- Информационные системы. Информатика. Проблемы информационной безопасности
- Роботизированные комплексы и системы. Технологии дистанционного зондирования и неразрушающего контроля
- Современные радиотехнические и телекоммуникационные системы
- Микро- и наноэлектроника. Физика конденсированного состояния
- Аналитическое приборостроение и технологии
- Математическое моделирование
- Экономика наукоемких и высокотехнологичных предприятий и производств. Управление в организационных системах
- Управление качеством продукции. Стандартизация
- Мировоззренческие основы технологии и общества

Russian Technological Journal
2023, Vol. 11, No. 5

Russian Technological Journal
2023, том 11, № 5

https://www.rtj-mirea.ru

https://doi.org/10.32362/2500-316X

ISSN 2500-316X (Online)

Russian Technological Journal
2023, Vol. 11, No. 5

Publication date September 30, 2023.

The peer-reviewed scientific and technical journal highlights the issues of complex development of radio engineering, telecommunication and information systems, electronics and informatics, as well as the results of fundamental and applied interdisciplinary researches, technological and economical developments aimed at the development and improvement of the modern technological base.

Periodicity: bimonthly.
The journal was founded in December 2013. The titles were «Herald of MSTU MIREA» until 2016 (ISSN 2313-5026) and «Rossiiskii tekhnologicheskii zhurnal» from January 2016 until July 2021 (ISSN 2500-316X).

Founder and Publisher:
Federal State Budget
Educational Institution of Higher Education
«MIREA – Russian Technological University»
78, Vernadskogo pr., Moscow, 119454 Russia.

The journal is included into the List of peer-reviewed science press of the State Commission for Academic Degrees and Titles of Russian Federation. The Journal is included in Russian State Library (RSL), Russian Science Citation Index, eLibrary, Socionet, Directory of Open Access Journals (DOAJ), Directory of Open Access Scholarly Resources (ROAD), Google Scholar, Ulrich's International Periodicals Directory.

The registration number ПИ № ФС 77 - 81733 was issued in August 19, 2021 by the Federal Service for Supervision of Communications, Information Technology, and Mass Media of Russia.

The subscription index of *Pressa Rossii*: **79641**.

Russian Technological Journal
2023, том 11, № 5

Дата опубликования 30 сентября 2023 г.

Научно-технический рецензируемый журнал освещает вопросы комплексного развития радиотехнических, телекоммуникационных и информационных систем, электроники и информатики, а также результаты фундаментальных и прикладных междисциплинарных исследований, технологических и организационно-экономических разработок, направленных на развитие и совершенствование современной технологической базы.

Периодичность: один раз в два месяца.
Журнал основан в декабре 2013 года. До 2016 г. издавался под названием «Вестник МГТУ МИРЭА» (ISSN 2313-5026), а с января 2016 г. по июль 2021 г. под названием «Российский технологический журнал» (ISSN 2500-316X).

Учредитель и издатель:
федеральное государственное бюджетное образовательное учреждение высшего образования
«МИРЭА – Российский технологический университет»
119454, РФ, г. Москва, пр-т Вернадского, д. 78.

Журнал входит в Перечень ведущих рецензируемых научных журналов ВАК РФ, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук и доктора наук, индексируется в РГБ, РИНЦ, eLibrary, Соционет, Directory of Open Access Journals (DOAJ), Directory of Open Access Scholarly Resources (ROAD), Google Scholar, Ulrich's International Periodicals Directory.

Регистрационный номер и дата принятия решения о регистрации СМИ ПИ № ФС 77 - 81733 от 19.08.2021 г. СМИ зарегистрировано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор).

Индекс по объединенному каталогу «Пресса России» **79641**.

https://www.rtj-mirea.ru

Editorial Board

Редакционная коллегия

Contents

Information systems. Computer sciences. Issues of information security

7 *Nikita G. Babak, Leonid Yu. Belorybkin, Shamil A. Otsokov, Alexey A. Terenin, Anastasia I. Shabrova*
Automatic depersonalization of confidential information

19 *Sergei V. Zuev*
Geometric properties of quantum entanglement and machine learning

Modern radio engineering and telecommunication systems

34 *Dmitry S. Vorunichev, Mikhail S. Kostin*
Investigation of the profilogram structure of microstrip microwave modules manufactured using additive 3D-printing technology

45 *Tatyana E. Gelfman, Alexey P. Pirkhavka*
Evaluation of the effectiveness of sliding redundancy of radioelectronic facilities

Micro- and nanoelectronics. Condensed matter physics

54 *Ekaterina S. Lepeshkina, Nikita D. Kustov, Vladislav Kh. Khanov*
Application of double-error correction codes to protect configuration programmable logic memory against space radiation

63 *Vladimir I. Musatov, Fedor A. Fedulov, Dmitrii V. Savelev, Ekaterina V. Bolotina, Leonid Y. Fetisov*
Nonlinear magnetoelectric effect in a ring composite heterostructure

71 *Alexander S. Sigov, Evgenij R. Lazarenko, Natalia B. Golovanova, Olga A. Minaeva, Sergei I. Anevsky, Roman V. Minaev, Pavel Yu. Pushkin*
Synchrotron radiation of a single electron application for optical spectroradiometry

Analytical instrument engineering and technology

81 *Mikhail Yu. Nikolshin, Alexander V. Frunze, Vladimir K. Bityukov*
DC/DC converter to power spectral lamps

Mathematical modeling

94 *Nikolay N. Karabutov*
On adaptive identification of systems having multiple nonlinearities

106 *Eduard M. Kartashov*
New energy effect in non-cylindrical domains with a thermally insulated moving boundary

Содержание

Информационные системы. Информатика. Проблемы информационной безопасности

7 *Н.Г. Бабак, Л.Ю. Белорыбкин, Ш.А. Оцоков, А.А. Теренин, А.И. Шаброва*
Автоматическое обезличивание конфиденциальной информации

19 *С.В. Зуев*
Геометрические свойства квантовой запутанности и машинное обучение

Современные радиотехнические и телекоммуникационные системы

34 *Д.С. Воруничев, М.С. Костин*
Исследование профилограммной структуры микрополосковых СВЧ-модулей, изготовленных по аддитивной технологии трехмерной печати

45 *Т.Э. Гельфман, А.П. Пирхавка*
Оценка эффективности скользящего резервирования радиоэлектронных средств

Микро- и наноэлектроника. Физика конденсированного состояния

54 *Е.С. Лепёшкина, Н.Д. Кустов, В.Х. Ханов*
Применение кодов с исправлением двух ошибок для защиты конфигурационной памяти программируемой логики от действия космической радиации

63 *В.И. Мусатов, Ф.А. Федулов, Д.В. Савельев, Е.В. Болотина, Л.Ю. Фетисов*
Нелинейный магнитоэлектрический эффект в кольцевой композитной гетероструктуре

71 *А.С. Сигов, Е.Р. Лазаренко, Н.Б. Голованова, О.А. Минаева, С.И. Аневский, Р.В. Минаев, П.Ю. Пушкин*
Использование синхротронного излучения отдельного электрона для спектрорадиометрии оптического диапазона

Аналитическое приборостроение и технологии

81 *М.Ю. Никольшин, А.В. Фрунзе, В.К. Битюков*
DC/DC-преобразователь для питания спектральных ламп

Математическое моделирование

94 *Н.Н. Карабутов*
Об адаптивной идентификации систем с несколькими нелинейностями

106 *Э.М. Карташов*
Новый энергетический эффект в областях нецилиндрического типа с термоизолированной движущейся границей

Information systems. Computer sciences. Issues of information security
Информационные системы. Информатика. Проблемы информационной безопасности

UDC 004.41, 004.89
https://doi.org/10.32362/2500-316X-2023-11-5-7-18

RESEARCH ARTICLE

Automatic depersonalization of confidential information

Nikita G. Babak [@, 1, 2], Leonid Yu. Belorybkin [2], Shamil A. Otsokov [3], Alexey A. Terenin [2], Anastasia I. Shabrova [2]

[1] *National Research University "Moscow Power Engineering Institute," Moscow, 111250 Russia*
[2] *Sberbank of Russia, Moscow, 117312 Russia*
[3] *MIREA – Russian Technological University, Moscow, 119454 Russia*
[@] *Corresponding author, e-mail: nikita.enrollee@gmail.com*

Abstract

Objectives. As the scope of personal data transmitted online continues to grow, national legislatures are increasingly regulating the storage and processing of digital information. This paper raises the problem of protecting personal data and other confidential information such as bank secrecy or medical confidentiality of individuals. One approach to the protection of confidential data is to depersonalize it, i.e., to transform it so that it becomes impossible to identify the specific subject to whom the data belongs. The aim of the work is to develop a method for the rapid and safe automation of the depersonalization process using machine learning technologies.

Methods. The authors propose the use of artificial intelligence models to implement a system for the automatic depersonalization of personal data without the use of human labor to preclude the possibility of recognizing confidential information even in unstructured data with sufficient accuracy. Rule-based algorithms for improving the precision of the depersonalization system are described.

Results. In order to solve this problem, a model of named entity recognition is trained on confidential data provided by the authors. In conjunction with rule-based algorithms, an F_1 score greater than 0.9 is achieved. For solving specific depersonalization problems, a choice between several implemented anonymization algorithm variants can be made.

Conclusions. The developed system solves the problem of automatic anonymization of confidential data. This opens an opportunity to ensure the secure processing and transmission of confidential information in many areas, such as banking, government administration, and advertising campaigns. The automation of the depersonalization process makes it possible to transfer confidential information in cases where it is necessary, but not currently possible due to legal restrictions. The distinctive feature of the developed solution is that both structured data and unstructured data are depersonalized, including the preservation of context.

Keywords: automated system, anonymization, information protection, cybersecurity, sensitive information, machine learning, neural networks, depersonalization, personal data, named entity recognition

• **Submitted:** 10.02.2023 • **Revised:** 14.06.2023 • **Accepted:** 13.07.2023

For citation: Babak N.G., Belorybkin L.Yu., Otsokov Sh.A., Terenin A.A., Shabrova A.I. Automatic depersonalization of confidential information. *Russ. Technol. J*. 2023;11(5):7–18. https://doi.org/10.32362/2500-316X-2023-11-5-7-18

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Автоматическое обезличивание конфиденциальной информации

Н.Г. Бабак [@, 1, 2], Л.Ю. Белорыбкин [2], Ш.А. Оцоков [3], А.А. Теренин [2], А.И. Шаброва [2]

[1] *Национальный исследовательский университет «МЭИ», Москва, 111250 Россия*
[2] *Публичное акционерное общество «Сбербанк России», Москва, 117312 Россия*
[3] *МИРЭА – Российский технологический университет, Москва, 119454 Россия*
[@] *Автор для переписки, e-mail: nikita.enrollee@gmail.com*

Резюме

Цели. В то время как объем персональных данных, передаваемых по сети, продолжает расти, законодательные органы все более жестко регулируют процессы хранения и обработки цифровой информации. В работе рассматривается проблема защиты персональных данных и другой конфиденциальной информации (КИ), например, банковской или врачебной тайны, физических лиц. Одним из способов защиты конфиденциальных данных является их обезличивание – преобразование, в результате которого становится невозможно установить принадлежность этих данных конкретному субъекту. Цель работы – построение автоматической системы, позволяющей быстро и безопасно обезличивать данные с помощью технологий машинного обучения.

Методы. Предлагается использовать модели искусственного интеллекта для реализации системы автоматического обезличивания КИ, т.к. это дает возможность распознавать КИ даже в неструктурированных данных с достаточно высокой точностью без привлечения человеческого труда. Для повышения точности всей системы обезличивания предлагается использовать алгоритмы на основе правил.

Результаты. На конфиденциальных данных, размеченных авторами для решения данной задачи, обучена модель распознавания именованных сущностей, которая в связке с алгоритмами на основе правил в результате имеет значение F_1-меры больше, чем 0.9. Реализовано несколько вариаций алгоритмов обезличивания, что позволяет выбирать между ними для каждой конкретной задачи.

Выводы. Разработанная система решает задачу автоматического обезличивания КИ. Это открывает возможность для безопасной обработки и передачи КИ во многих областях, например, в банковской деятельности, государственном управлении, рекламных кампаниях. Также автоматизация процесса обезличивания делает возможной передачу КИ в тех случаях, когда это необходимо, но невозможно в силу правовых ограничений. Отличительная особенность разработанного решения заключается в том, что обезличиваются как структурированные данные, так и неструктурированные, в т.ч. с сохранением контекста.

Ключевые слова: автоматизированная система, анонимизация, защита информации, кибербезопасность, конфиденциальная информация, машинное обучение, нейросети, обезличивание, персональные данные, распознавание именованных сущностей

• **Поступила:** 10.02.2023 • **Доработана:** 14.06.2023 • **Принята к опубликованию:** 13.07.2023

Для цитирования: Бабак Н.Г., Белорыбкин Л.Ю., Оцоков Ш.А., Теренин А.А., Шаброва А.И. Автоматическое обезличивание конфиденциальной информации. *Russ. Technol. J.* 2023;11(5):7–18. https://doi.org/10.32362/2500-316X-2023-11-5-7-18

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

In today's digitally intermediated world, the scope of stored and processed data is constantly growing, requiring increased reliability in terms of data protection. The issue of protecting personal data transmitted via computer networks and stored in information systems becomes particularly relevant. The list and procedure for processing personal data is outlined in Federal Law No. 152 "On Personal Data." Here, personal data is defined as any information pertaining to a directly or indirectly identified or identifiable individual[1].

Modern computer systems allow organizations to collect and process large amounts of data necessary for their effective functioning and development. The access to various kinds of data facilitated by the rapid development of information technology in turn increases the risk of information leakage [1]. The high risk of illegal access to confidential information (CI) makes the task of ensuring its protection particularly relevant.

One of the measures aimed at minimizing the risks of harm to an individual in the event of leakage of personal data from automated systems (AS) is depersonalization as required by law. The depersonalization of personal data comprises an action that makes it impossible to determine, without additional information, the specific subject to whom the personal data belongs. By means of such anonymization, legal data-processing requirements can be reduced, leading to lower costs for organizations when developing such systems. Thus, the depersonalization of personal data not only protects people from cyber threats, but also has a positive economic effect. This problem was considered in some works [2–5], but peculiarities of data processing in Russian, whose morphology has additional complexities, were not taken into account. Moreover, in these works, the detailing of recognizable CI entities was not adequately carried out, which reduces the quality of impersonal data.

[1] Federal Law No. 152-FZ dated July 27, 2006 "On Personal Data" (in Russ.). https://docs.cntd.ru/document/901990046. Accessed February 09, 2023.

1. TERMS OF REFERENCE

When carrying out data depersonalization, it is necessary to understand what data elements should be hidden. Therefore, we can say that the preliminary stage of depersonalization of CI (in particular personal data) is its separation from all other information. For this purpose, manual extraction of a certain type of information is not only more costly, but also subject to the risk of human error.

Based on above, there is a task of automatic recognition and subsequent CI depersonalization in the data processed and transmitted in the AS. Data can be transmitted in the form of exchanged files, various information flows, etc. In this regard, it is necessary to provide the ability to extract information from files of different extensions and byte representation.

2. CI RECOGNITION

There are several basic automated ways to recognize information, such as vocabulary search, regular expressions, and machine learning algorithms. While the recognition of any kind of information in structured data is quite often solved using rule-based systems, things are not so straightforward with unstructured data. Moreover, there is a large variety of data that directly or indirectly identify a person, such as name, first name and patronymic, passport series and number, phone number. For each type of data, large vocabularies will have to be compiled and constantly updated, along with the encoding of complex rules.

These problems can be solved by using machine learning algorithms to recognize personal data in structured and unstructured information. In particular, the task of personal data recognition is reduced to the task of Named Entity Recognition (NER) [6]. There are several basic ways to solve this problem:

- using statistical methods, for example, according to the number of certain characters;
- using rules based on vocabularies and regular expressions;
- using neural networks.

Statistical methods currently used to perform this task lack sufficient recognition quality, especially when dealing with unstructured data. Rule-based systems, although relatively fast, require more frequent updates and are prone to errors in more complex data, such as organization names, surnames, and first names. In addition, statistical and rule-based approaches do not take context into account. Neural networks can address these shortcomings. For the tasks of natural language processing and, in particular, NER, the most advanced are neural networks with transformer-type architecture [7]. Transformers transform natural language into a numerical vector representation called embedding, which in turn can be processed by machine. Such embeddings, unlike classical vectors, take into account the semantic proximity of token words.

When working with structured data, it is not always necessary to use neural networks to recognize some types of CI—simple rules and statistical methods suffice. Preliminary analysis and separation of data into structured and unstructured allows choosing a suitable recognition algorithm. For recognition of numerical data, regular expressions with check digits are more suitable, especially in structured data. It is also worth noting that some numeric personal data are well recognized by neural networks working with a sufficiently large set of unstructured data. In any case, in order to use machine learning algorithms, it is necessary to prepare a training sampling.

2.1. Data markup

Training sampling consists of data presented in a certain way and labeled with various attributes of CI. The text is divided into tokens, represented by words, which are assigned a tag (label) denoting belonging to a certain type of information.

Tags can be placed according to one of the following schemes:

- BIO/IOB, where B (Begin) is the beginning of the entity; I (Inside) is the continuation of the entity; O (Outside) is not related to the entity;
- BILUO/BILOU [8], where L (Last) is the end of an entity; U (Unit) is a single token entity; B, I, and O are decoded as in the BIO/IOB scheme.

Since the BIO scheme is more commonly-used, it is used in the present work.

Tagging of tokens may differ depending on the problem to be solved. In the Nested Named Entity Recognition (Nested NER) task [9, 10], two tags are assigned to each token: a summary tag and a nested tag. An example of markup is shown in Table 1. Tags manually applied by a qualified expert typically contain an abbreviated meaningful description of the information contained in the token being tagged. For example, the tag B-SNM is an abbreviation of Surname.

Table 1. Token tagging to recognize nested named entities

Token	Consolidated Tag	Nested Tag
Sidorov	B-PERSON	B-SNM
Ivan	I-PERSON	B-FNM
Petrovich	I-PERSON	B-PNM
has concluded	O	O
the contract	O	O
with	O	O
OOO	B-ORG	B-OPF
Romashka	I-ORG	B-ORG_NAME

When recognizing discontinuous named entities (Discontinuous NER) [11], tagging can be represented as a table where the number of columns depends on the maximum number of discontinuities for the discontinuous entity. Thus, the first word in a discontinuous entity is tagged with prefix B, while all subsequent tags are shifted by one column to the right at each discontinuity and prefixed with I (in the case of the BIO scheme).

Since the present work is aimed at solving the classical problem of named entities recognition, so the training sample is divided into words. Each word is matched with a label indicating that it belongs to one or another type of CI. The data set used contains various regulatory documents, memos and other documents involved in the production activities of the organization, which will later be depersonalized.

By tagging the data and training an artificial intelligence (AI) model, it is possible to recognize CI automatically, which in turn opens up the possibility for subsequent automatic depersonalization.

3. DEPERSONALIZATION

Once detected in the CI text, it can be depersonalized in a reversible or irreversible way. In most cases, depersonalization means irreversible implementation; if necessary, it is possible to save the substitution table in a protected loop to obtain reversible depersonalization.

The following depersonalization algorithms are possible in any implementation:

- setting to zero—deleting all or a significant part of the original value;
- replacement by constant—replacement of the original value by a non-zero constant;
- replacement with a value from the reference book—replacement of the original value with a random different value from the reference book, corresponding to the data type to be replaced;
- replacement by a character set—converting each character of the original value into a random character that matches the data type;
- shuffling—shuffling of individual values or groups of values of attributes of personal data in the array of personal data;

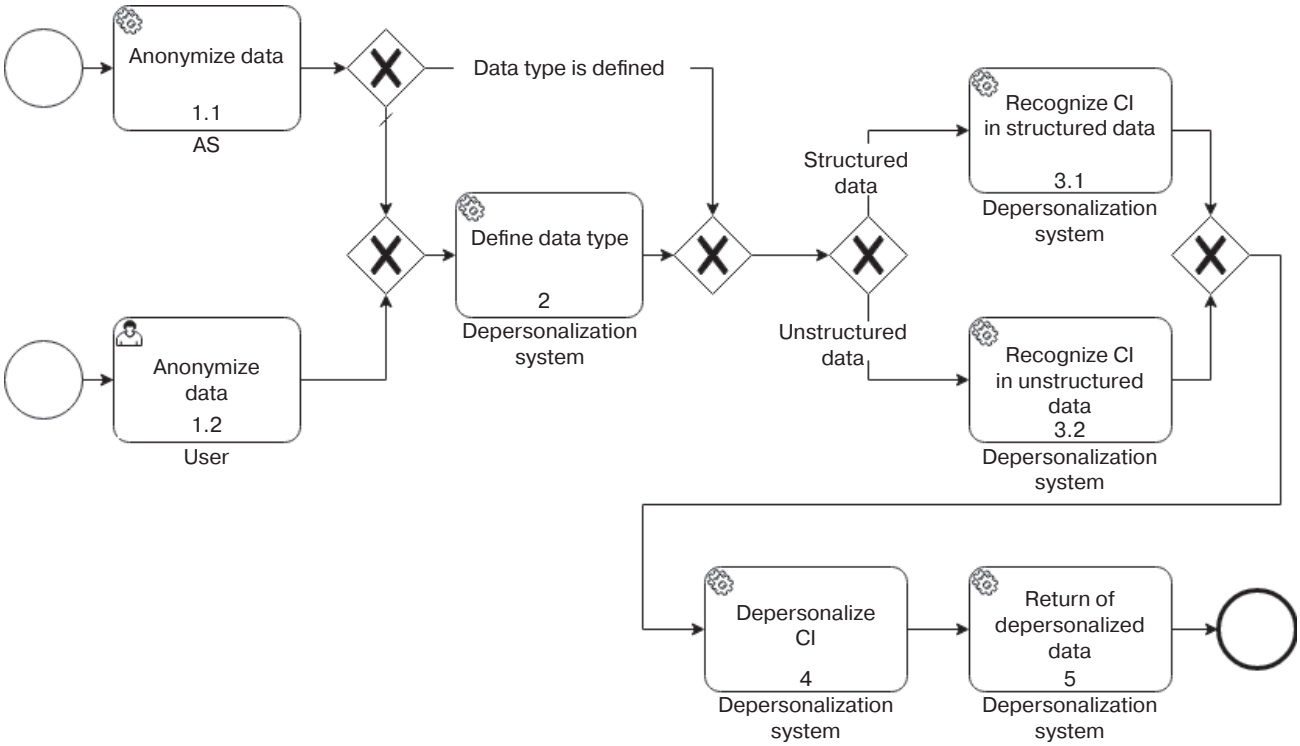


Fig. 1. CI depersonalization process

- blurring the sum and the date—replacing the original value by a random value close to the impersonal value;
- transformation based on a given expression—transformation of the initial value by an expression containing both constants and variables;
- masking—replacing part of the original value with a special character or a set of characters (mask);
- replacement by a random value—replacement of the original value with a randomly generated value;
- generation of pseudo-meaningful meanings—creation of text on the basis of language model or given expressions, allowing the correct text from the point of view of the basic linguistic norms and data parameters to be received. In addition to this method, we can refer generation of photos, taking into account gender and age of the person.

When choosing an approach to depersonalization of personal data, it is worth considering the guidance published by Roskomnadzor[2], according to which the main methods of depersonalization include: the method of introducing identifiers (replacing part of the data by identifiers and creating a table of matching identifiers with the original data); the method of changing the composition or semantics of personal data by replacing them with the results of statistical processing, transformation, summarization, or deleting parts of data; decomposition method (dividing the set of personal data).

[2] Order of Roskomnadzor dated September 05, 2013 No. 996 "On approval of requirements and methods for depersonalization of personal data" (in Russ.). https://rkn.gov.ru/docs/6_Trebovanija_i_metody_po_obezlichivaniju_personal6nykh_dannykh.docx. Accessed February 09, 2023.

Taking into account the recommendations of Roskomnadzor, the most suitable algorithms are pseudo-value generation and constant replacement. To implement reversible depersonalization, it is necessary to create a table of matching source data; here, it should be noted that the table itself should be stored separately from the depersonalized data, with only persons authorized to work with personal data in open form having access to it.

Depending on the problem to be solved, various algorithms may be used. For example, if it is necessary to unambiguously determine that an anonymization was performed and to understand what type of information was removed, a constant replacement algorithm is the best choice. If it is necessary to preserve the length of the value to be replaced at the same time as determining that an anonymization was performed, the partial masking algorithm will handle this task well. In the case where the depersonalized data needs to be used in almost the same way as the original data, for example, for training AI models, the best choice would be an algorithm for generating pseudo-meaningful values.

4. AUTOMATIC DEPERSONALIZATION SYSTEM

In order to work with CI as safely as possible, it is necessary to develop a system of automatic depersonalization. The process of automatic depersonalization by means of the system implemented by the authors of the present work consists of the following tasks (Fig. 1).

1.1 Request to the system for depersonalization according to API from a third-party AS.

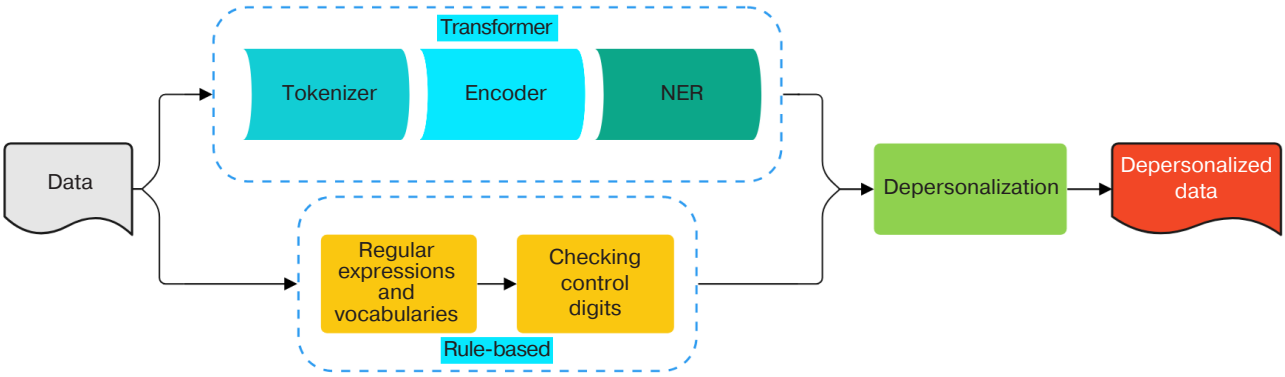


Fig. 2. Data processing by the depersonalization system

1.2 Request to the system for depersonalization through the interface from the user.
2 Data type definition and preprocessing.
3.1 CI recognition in structured data.
3.2 CI recognition in unstructured data.
4 Depersonalization of the recognized CI by the most suitable algorithm.
5 Return of a depersonalized document or data stream.

The need to separate the recognition of personal data in structured and unstructured information in the system arises due to the fact that different machine learning algorithms are used, in particular those using considering and not-considering syntactic features.

4.1. Data preparation

In total, about one million tokens, represented by individual words, were manually tagged by the authors for model training. For the markup we used service documents containing personal data, banking secrets and other CI. A BIO scheme was chosen as the markup scheme, where the first token within a confidential entity is prefixed with B, and all subsequent tokens are prefixed with I. This approach allows most pre-trained architectures to be compared and used, which simplifies the process of creating an AI model, at least in terms of reducing the time to train it.

Resulting set of marked data is divided into 3 parts, where 80% of the data is used to train the model, 10% is used to validate it, and 10% is used to calculate the metrics of the trained model. This is the ratio used, not 60/20/20, because some types of CI in the data set are not sufficient, and it would be irrational to further reduce their number in the training set.

When splitting text into tokens, it is necessary to save the indices of the splitting boundaries in order to anonymize it strictly within the specified boundaries following CI recognition.

4.2. Model training

Most advanced results in the tasks of named entities recognition are shown by neural networks based on transformer architecture. Transformers pre-trained on a large corpus of data are well reused in the tasks of natural language processing [12]. For this purpose, it is sufficient to fine-tune the model on its own data, thereby adjusting weights in order to better take into account the semantics of the input data.

The pre-trained rubert-base-cased model [13] is used as the basis, the use of other suitable architectures does not significantly affect the performance of the model. This is primarily due to the similarity of various transformers used to solve the NER problem, such as BERT [13], RoBERTa [14], and spaCy [15]. The pre-trained model comprises a Tokenizer and Encoder to which a NER classifier is added. In order to improve accuracy and reduce false positives, rule-based recognition algorithms are used, in particular, regular expressions and check digit checking [16]. The results of data processing by neural networks and rules are then summarized into a general assumption that the text belongs to one of the types of CI. A schematic representation of the data depersonalization process by the proposed system is shown in Fig. 2.

Due to the lack of context when processing structured data, preference is given to the rule-based recognition module.

A rule-based CI recognition model without neural networks and a PyTorch model based on a recurrent neural network (RNN) [17] are also implemented in the depersonalization system for the purposes of comparison.

Since most existing depersonalization systems are rule-based and have similar implementations, the rule-based model for CI recognition without neural networks serves to provide a baseline metric against which other solutions can be compared. Comparing an implemented depersonalization system with other implementations will knowingly present the proposed solution in a better light, since third-party implementations were designed for a different, most often structured, data set [18–20]. For example, some third-party systems work only with personal data and do not support bank secrecy depersonalization.

Since all personal data must be recognized and anonymized in the context of this task, the recall metric is important; however, precision is also important to ensure that the number of false positives does not undermine trust in the system. For this reason, the F_1-measure is used, which takes both of these metrics into account, and is calculated by the formula

$$F_1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}.$$

The metrics are calculated using the test part of the marked dataset described in Section 4.1. To begin with, a confusion matrix is constructed, in which the horizontal axis contains the true tags from the markup and the vertical axis contains the tags predicted by the AI model. Then the number of true CI attributes (TP, true positive), the number of true unrecognized attributes (TN, true negative), the number of false attributes recognized (FP, false positive) and the number of false unrecognized attributes (*FN*, false negative) are counted from the error matrix. After that, recall and precision are calculated using the formulas

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

and

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

and then their average harmonic F_1-measure is determined [21].

Table 2 shows the main attributes of CI and calculated weighted average F_1-measure by different models: rule-based model, recurrent neural network and BERT model. It is worth noting that the rule-based implementation works only on the basis of rules, while the other implementations use neural networks together with regular expressions and other rule-based algorithms.

Table 2. Main attributes of CI and calculated weighted average F_1-measure

CI attribute	F_1 (rule-based)	F_1 (RNN)	F_1 (BERT)
Surname	0.804	0.911	0.931
Name	0.819	0.876	0.929
Patronymic	0.874	0.883	0.943
Passport serial number	0.883	0.907	0.906
Authority that issued the passport	0.701	0.794	0.899
Phone number	0.959	0.969	0.967
E-mail	0.955	0.959	0.964
IP-address	0.929	0.932	0.930
Geolocation	0.904	0.919	0.922
Address	0.809	0.810	0.912
Date of birth	0.813	0.837	0.915
TIN	0.918	0.915	0.919
IIAN	0.931	0.935	0.934
OMI policy number	0.921	0.914	0.921
Bank account number	0.937	0.929	0.936
Bank card number	0.967	0.959	0.965
Military ID number	0.892	0.880	0.889
Primary State Registration Number of the Individual Entrepreneur	0.910	0.909	0.919
Job position	0.812	0.820	0.873
Organization name	0.817	0.899	0.951
Average weighted F_1-measure	0.878	0.898	0.926

The RoBERTa and spaCy models were additionally compared. These showed metrics similar to the BERT model with a scatter of F_1-measure values less than 0.01. In this regard, the BERT model was chosen because it is smaller than the RoBERTa model at the same time as having more flexible settings than the spaCy model; this becomes an important factor when implementing an industrial version of the model in a system.

As shown in Table 2, the rule-based solution is significantly inferior to machine learning models in terms of the values of the F_1-measure. The effect is especially noticeable in string data types, where context plays a significant role. Due to the heterogeneous set of documents, the recurrent neural network RNN also performs worse than BERT. Based on the values of the F_1-measure metric presented in Table 2, and the fact that transformer models have wide potential for development and reuse, the BERT model outperforms the other solutions by an average of 4%, for which reason it was selected in the final solution.

The main advantage of the depersonalization system using the BERT model over other solutions is the use of the self-attention mechanism, which allows better detection of CI through the analysis of context and importance of words in the text. The attention mechanism used in the model can be expressed by the formula

$$\text{attention} = \text{softmax}\left(\frac{\mathbf{QK}^T}{\sqrt{d_k}}\right)\mathbf{V},$$

where **Q** is the query vector; **K** is the key vector; **V** is the value vector; d_k is the dimensionality of vectors.

Vectors **Q**, **K**, and **V** are obtained by multiplying the token embedding by the corresponding matrices obtained by pre-training the rubert-base-cased model taken as the basis. Since in reality calculations performed over vector representations of several tokens **Q**, **K**, and **V** are matrices; therefore, before calculating the product of **Q** and **K**, the matrix **K** must be transposed [7]. In a practical implementation, the key vector and values are the same vector and serve to represent a token, while the query vector shows the significance of a given token with respect to other.

The Softmax function is expressed by the formula

$$\sigma(\mathbf{Z})_i = \frac{e^{z_i}}{\sum_{j=1}^{N} e^{z_j}},$$

where i and j are indices of a vector element in the range from 1 to N serves for normalization, i.e., it converts a vector **z** of dimension N to a vector **σ** of the same dimension, where all coordinates of the normalized vector $\boldsymbol{\sigma}_i$ are expressed by a number in the range from 0 to 1, and their sum is equal to one.

CI recognition in unstructured data, represented by images and audio recordings, is reduced to the processing of unstructured texts. For this purpose, the Optical Character Recognition (OCR) [22] and Automatic Speech Recognition (ASR) tasks [23] are solved in advance. [23].

4.3. Depersonalization

Having recognized a CI, the system depersonalizes it using one of the selected algorithms. The choice of algorithm is possible both at the level of the whole document or data set, as well as that of the separate type of CI. The system presented by the authors implements depersonalization algorithms based on the following methods:

- replacement with a constant (placeholder) of the form {Attribute_CI};
- masking to the * symbol;
- pseudo-meaningful value generation, including replacement with a value from a reference, conversion based on a given expression, and date blurring.

For example, having recognized by the AI model in the sentence “Alexander Sidorov (TIN 503199560259) received a transfer to the card 4561 2612 1234 5467” the CI represented by the surname, name, taxpayer identification number (TIN) and the bank card number, the system user can choose one of the depersonalization algorithms described above. When replaced with a placeholder, the sentence in question will take the following form: “{Name} {Surname} (TIN {TIN}) received a transfer to card {Bank Card Number},” where the CI is replaced with constants indicating what type of information was previously in the sentence. With partial masking, the CI is replaced by a mask, and the sentence in question will take the following form: “********* S****** (TIN 50**********) received transfer 4561 26** **** 5467,” where the parts of words that are not dangerous for the identification of the data subject, but which allow the indirect attributes, for example, the bank that issued the card, to be preserved. When substituted with pseudo meanings, the sentence in question will take the following form: “Vladislav Lazarev (TIN 503195234624) received a transfer to card 4561 2698 5513 5467.” The latter algorithm, unlike the previous two, works more slowly, as it generates pseudo meaningful data, but generates a fully meaningful text that can be used, for example, in machine learning tasks.

The selection of the depersonalization algorithm used depends on the task to be solved and is left to the user or the AS.

CONCLUSIONS

A total of about one million tokens are marked for training the AI model, so that a large number of data representation methods containing CI are covered. When the number of types of depersonalized documents is small, it is sufficient to partition a small set of data that includes all the necessary types of CI for model pre-training. Since, due to the use of transformer models, model retraining is not required in most cases, the developed system can be reused in different organizations "as is" or with adjustments on a small volume of data. The use of neural networks permits the removal of huge directories of surnames and names, as well as other data entities used to identify a person. Regular expressions, in turn, take into account structural features, such as existing series, codes and bank identification numbers, which makes it possible to detect even those data on which the model has not previously been trained.

The distinguishing advantage of the depersonalization system presented by the authors from the existing ones is the support of both structured and unstructured data. Moreover, in most known systems, depersonalization is performed in a destructive way, after which the data become unusable for many tasks, for example, for machine learning.

Average weighted F_1-measure of the implemented CI recognition model exceeded 0.9, indicating the high quality of the depersonalization system, which effectively eliminates the need for human labor for CI detection.

Implemented algorithms of depersonalization based on the method of constant replacement, masking and generation of pseudo-meaningful values cover all basic tasks of depersonalization: depersonalization with the possibility of unambiguous determination of the fact of masking, synonymous depersonalization, irreversible and reversible depersonalization, etc. The specified algorithms can also be used to automatically depersonalize the recognized confidential data. The practical value of the automatic depersonalization system developed by the authors lies in the fact that the depersonalized confidential data can be used similarly to the original data, but without the risk of violating cybersecurity requirements. Due to the automation of the process, the cost for depersonalization procedures can be practically reduced to zero.

Confidential data depersonalization system contains at least one processor and one memory connected to the processor, which contains machine-readable instructions. In addition, the depersonalization system may be run on a server, a programmable logic controller, or any other devices capable of executing a given sequence of instructions.

The proposed automatic depersonalization system can be used to automatically recognize and depersonalize personal information at almost any cycle associated with its transfer and processing. Thanks to this, the risk of identity disclosure in case of data leakage is reduced. For example, automatic depersonalization can be used in banking, government services, the data science community [24], and other entities related to the processing of CI, in particular, personal data.

Authors' contribution. All authors equally contributed to the research work.

REFERENCES

1. Shabrova A.I., Terenin A.A., Babak N.G. Methodology for risk assessment from confidential information disclosure in data sources using data mining. *Sovremennye informacionnye tehnologii i IT-obrazovanie = Modern Information Technologies and IT-Education*. 2022;18(3):666–679 (in Russ.). https://doi.org/10.25559/SITITO.18.202203.666-679
2. Stolbov A.P. De-identification of personal data in health care. *Vrach i informacionnye tekhnologii = Medical Doctor and Information Technologies*. 2017;3:76–91 (in Russ.). Available from URL: https://elibrary.ru/zgyvot
3. Spevakov A.G., Kalutskiy I.V., Nikulin D.A., Shumailova V.A. Depersonalization of personal data during processing of information in automated systems. *Telekommunikatsii = Telecommunications*. 2016;10:16–20 (in Russ.). Available from URL: https://www.elibrary.ru/wwvxmt

СПИСОК ЛИТЕРАТУРЫ

1. Шаброва А.И., Теренин А.А, Бабак Н.Г. Методика оценки риска от разглашения конфиденциальной информации в источниках данных с использованием интеллектуального анализа данных. *Современные информационные технологии и ИТ-образование*. 2022;18(3):666–679. https://doi.org/10.25559/SITITO.18.202203.666-679
2. Столбов А.П. Обезличивание персональных данных в здравоохранении. *Врач и информационные технологии*. 2017;3:76–91. URL: https://elibrary.ru/zgyvot
3. Спеваков А.Г., Калуцкий И.В., Никулин Д.А., Шумайлова В.А. Обезличивание персональных данных при обработке в автоматизированных информационных системах. *Телекоммуникации*. 2016;10:16–20. URL: https://www.elibrary.ru/wwvxmt

4. Oleksy M., Ropiak N., Walkowiak T. Automated anonymization of text documents in Polish. *Procedia Computer Science*. 2021;192(1):1323–1333. https://doi.org/10.1016/j.procs.2021.08.136
5. Saluja B., Kumar G., Sedoc J., Callison-Burch C. Anonymization of Sensitive Information in Medical Health Records. In: *CEUR Workshop Proceedings*. 2019;2421:647–653. Available from URL: https://ceur-ws.org/Vol-2421/MEDDOCAN_paper_2.pdf
6. Roy A. *Recent Trends in Named Entity Recognition* (*NER*). arXiv. 2021. https://doi.org/10.48550/arxiv.2101.11420
7. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser L., Polosukhin I. Attention is all you need. In: *Advances in Neural Information Processing Systems*. 2017. https://doi.org/10.48550/arXiv.1706.03762
8. Ratinov L., Roth D. Design Challenges and Misconceptions in Named Entity Recognition. In: *Proceedings of the Thirteenth Conference on Computational Natural Language Learning* (*CoNLL 2009*). 2009. P. 147–155. Available from URL: https://aclanthology.org/W09-1119.pdf
9. Fisher J., Vlachos A. *Merge and label: A novel neural network architecture for nested NER*. arXiv. 2019. https://doi.org/10.48550/arXiv.1907.00464
10. Fu Y., Tan C., Chen M., Huang S., Huang F. Nested named entity recognition with partially-observed TreeCRFs. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. 2021;35(14):12839–12847. https://doi.org/10.1609/aaai.v35i14.17519
11. Dai X., Karimi S., Hachey B., Paris C. *An effective transition-based model for discontinuous NER*. arXiv. 2020. https://doi.org/10.48550/arXiv.2004.13454
12. Lothritz C., Allix K., Veiber L., Klein J., Bissyande T.F.D.A. Evaluating pretrained transformer-based models on the task of fine-grained named entity recognition. In: *Proceedings of the 28th International Conference on Computational Linguistics*. 2020. P. 3750–3760. http://doi.org/10.18653/v1/2020.coling-main.334
13. Kuratov Y., Arkhipov M. *Adaptation of deep bidirectional multilingual transformers for Russian language*. arXiv. 2019. https://doi.org/10.48550/arXiv.1905.07213
14. Conneau A., Khandelwal K., Goyal N., Chaudhary V., Wenzek G., Guzman F., Grave E., Ott M., Zettlemoyer L., Stoyanov V. *Unsupervised cross-lingual representation learning at scale*. arXiv. 2020. https://doi.org/10.48550/arXiv.1911.02116
15. Patel A.A., Arasanipalai A.U. *Applied Natural Language Processing in the Enterprise*. O'Reilly Media, Inc.; 2021. 336 p. ISBN 978-1-4920-6257-8. Available from URL: https://spacy.io/universe/project/applied-nlp-in-enterprise/
16. Singco V.Z., Trillo J., Abalorio C., Bustillo J.C., Bojocan J., Elape M. OCR-based Hybrid Image Text Summarizer using Luhn Algorithm with Finetune Transformer Models for Long Document. *Int. J. Emerging Technol. Adv. Eng*. 2023;13(02):47–56. http://doi.org/10.46338/ijetae0223_07
17. Soltau H., Shafran I., Wang M., Shafey L.E. *RNN Transducers for Nested Named Entity Recognition with constraints on alignment for long sequences*. arXiv. 2022. https://doi.org/10.48550/arXiv.2203.03543

4. Oleksy M., Ropiak N., Walkowiak T. Automated anonymization of text documents in Polish. *Procedia Computer Science*. 2021;192(1):1323–1333. https://doi.org/10.1016/j.procs.2021.08.136
5. Saluja B., Kumar G., Sedoc J., Callison-Burch C. Anonymization of Sensitive Information in Medical Health Records. In: *CEUR Workshop Proceedings*. 2019;2421:647–653. URL: https://ceur-ws.org/Vol-2421/MEDDOCAN_paper_2.pdf
6. Roy A. *Recent Trends in Named Entity Recognition* (*NER*). arXiv. 2021. https://doi.org/10.48550/arxiv.2101.11420
7. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser L., Polosukhin I. Attention is all you need. In: *Advances in Neural Information Processing Systems*. 2017. https://doi.org/10.48550/arXiv.1706.03762
8. Ratinov L., Roth D. Design Challenges and Misconceptions in Named Entity Recognition. In: *Proceedings of the Thirteenth Conference on Computational Natural Language Learning* (*CoNLL 2009*). 2009. P. 147–155. URL: https://aclanthology.org/W09-1119.pdf
9. Fisher J., Vlachos A. *Merge and label: A novel neural network architecture for nested NER*. arXiv. 2019. https://doi.org/10.48550/arXiv.1907.00464
10. Fu Y., Tan C., Chen M., Huang S., Huang F. Nested named entity recognition with partially-observed TreeCRFs. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. 2021;35(14):12839–12847. https://doi.org/10.1609/aaai.v35i14.17519
11. Dai X., Karimi S., Hachey B., Paris C. *An effective transition-based model for discontinuous NER*. arXiv. 2020. https://doi.org/10.48550/arXiv.2004.13454
12. Lothritz C., Allix K., Veiber L., Klein J., Bissyande T.F.D.A. Evaluating pretrained transformer-based models on the task of fine-grained named entity recognition. In: *Proceedings of the 28th International Conference on Computational Linguistics*. 2020. P. 3750–3760. http://doi.org/10.18653/v1/2020.coling-main.334
13. Kuratov Y., Arkhipov M. *Adaptation of deep bidirectional multilingual transformers for Russian language*. arXiv. 2019. https://doi.org/10.48550/arXiv.1905.07213
14. Conneau A., Khandelwal K., Goyal N., Chaudhary V., Wenzek G., Guzman F., Grave E., Ott M., Zettlemoyer L., Stoyanov V. *Unsupervised cross-lingual representation learning at scale*. arXiv. 2020. https://doi.org/10.48550/arXiv.1911.02116
15. Patel A.A., Arasanipalai A.U. *Applied Natural Language Processing in the Enterprise*. O'Reilly Media, Inc.; 2021. 336 p. ISBN 978-1-4920-6257-8. URL: https://spacy.io/universe/project/applied-nlp-in-enterprise/
16. Singco V.Z., Trillo J., Abalorio C., Bustillo J.C., Bojocan J., Elape M. OCR-based Hybrid Image Text Summarizer using Luhn Algorithm with Finetune Transformer Models for Long Document. *Int. J. Emerging Technol. Adv. Eng*. 2023;13(02):47–56. http://doi.org/10.46338/ijetae0223_07
17. Soltau H., Shafran I., Wang M., Shafey L.E. *RNN Transducers for Nested Named Entity Recognition with constraints on alignment for long sequences*. arXiv. 2022. https://doi.org/10.48550/arXiv.2203.03543

18. Abirkhaev E.A., Erokhin A.F., Pushkin P.Yu. Methods of depersonalizing data: overview and analysis. *Naukosfera*. 2021;6(2):57–31 (in Russ.). Available from URL: https://www.elibrary.ru/item.asp?id=46561812
19. Seryshev A.S., Krotov A.D., Efanova N.V. Development of an application for personal data depersonalization. In: *Digitalization of the Economy: Directions, Methods, Tools: Proceedings of the 3rd All-Russian Scientific and Practical Conference*. Krasnodar: Kuban State Agrarian University; 2021. P. 294–297 (in Russ.). ISBN 978-5-9074-3005-1. Available from URL: https://www.elibrary.ru/item.asp?id=44891383
20. Fot U.D., Korobova E.O. Depersonalization of personal data in the personnel management system of oil and gas sector enterprises. In: *The Role of the Oil and Gas Sector in the Technical and Economic Development of the Orenburg Region: Proceedings of the scientific-practical conference. Saratov: Amirit;* 2021. P. 161–168 (in Russ.). ISBN 978-5-0014-0888-8. Available from URL: https://www.elibrary.ru/item.asp?id=48392659
21. Williams C.K.I. The effect of class imbalance on Precision-Recall Curves. *Neural Computation*. 2021;33(4): 853–857. https://doi.org/10.1162/neco_a_01362
22. Du Y., Li C., Guo R., Yin X., Liu W., Zhou J., Bai Y., Yu Z., Yang Y., Dang Q., Wang H. *PP-OCR: A practical ultra lightweight OCR system*. arXiv. 2020. https://doi.org/10.48550/arXiv.2009.09941
23. Pan J., Shapiro J., Wohlwend J., Han K.J., Lei T., Ma T. *ASAPP-ASR: Multistream CNN and self-attentive SRU for SOTA speech recognition*. arXiv. 2020. https://doi.org/10.48550/arXiv.2005.10469
24. Ryffel T., Trask A., Dahl M., Wagner B., Mancuso J., Rueckert D., Passerat-Palmbach J. *A generic framework for privacy preserving deep learning*. arXiv. 2018. https://doi.org/10.48550/arXiv.1811.04017

18. Абирхаев Е.А., Ерохин А.Ф., Пушкин П.Ю. Методы обезличивальных данных: обзор и анализ. *Наукосфера*. 2021;6(2):57–31. URL: https://www.elibrary.ru/item.asp?id=46561812
19. Кротов А.Д., Серышев А.С., Ефанова Н.В. Разработка приложения для обезличивания персональных данных. В сб.: *Цифровизация экономики: направления, методы, инструменты*: *сб. материалов III всероссийской научно-практической конференции*. Краснодар: Кубанский государственный аграрный университет; 2021. С. 294–297. ISBN 978-5-9074-3005-1. URL: https://www.elibrary.ru/item.asp?id=44891383
20. Фот Ю.Д., Коробова Е.О. Обезличивание персональных данных в системе управления персоналом предприятий нефтегазового сектора. В сб.: *Роль нефтегазового сектора в технико-экономическом развитии Оренбуржья: сб. трудов научно-практической конференции*. Саратов: ООО «Амирит»; 2021. С. 161–168. ISBN 978-5-0014-0888-8. URL: https://www.elibrary.ru/item.asp?id=48392659
21. Williams C.K.I. The effect of class imbalance on Precision-Recall Curves. *Neural Computation*. 2021;33(4): 853–857. https://doi.org/10.1162/neco_a_01362
22. Du Y., Li C., Guo R., Yin X., Liu W., Zhou J., Bai Y., Yu Z., Yang Y., Dang Q., Wang H. *PP-OCR: A practical ultra lightweight OCR system*. arXiv. 2020. https://doi.org/10.48550/arXiv.2009.09941
23. Pan J., Shapiro J., Wohlwend J., Han K.J., Lei T., Ma T. *ASAPP-ASR: Multistream CNN and self-attentive SRU for SOTA speech recognition*. arXiv. 2020. https://doi.org/10.48550/arXiv.2005.10469
24. Ryffel T., Trask A., Dahl M., Wagner B., Mancuso J., Rueckert D., Passerat-Palmbach J. *A generic framework for privacy preserving deep learning*. arXiv. 2018. https://doi.org/10.48550/arXiv.1811.04017

About the authors

Nikita G. Babak, Postgraduate Student, Department of Computing Machines, Systems and Networks, Institute of Information Technologies and Computer Science, National Research University MPEI (14/1, Krasnokazarmennaya ul., Moscow, 111250 Russia); Chief Data Protection Officer, Cybersecurity Department, Sberbank of Russia (19, Vavilova ul., Moscow, 117312 Russia). E-mail: nikita.enrollee@gmail.com. ResearcherID HHY-9372-2022, RSCI SPIN-code 3687-6548, https://orcid.org/0000-0001-7129-1018

Leonid Yu. Belorybkin, Director of Data Protection Projects, Cybersecurity Department, Sberbank of Russia (19, Vavilova ul., Moscow, 117312 Russia). E-mail: lbelorybkin@gmail.com. https://orcid.org/0000-0002-8575-5773

Shamil A. Otsokov, Dr. Sci. (Eng.), Professor, Department of Intelligent Information Security Systems, Institute of Cybersecurity and Digital Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: shamil24@mail.ru. Scopus Author ID 57212622267, https://orcid.org/0000-0001-7451-5443

Alexey A. Terenin, Cand. Sci. (Eng.), Managing Director, Cybersecurity Department, Sberbank of Russia (19, Vavilova ul., Moscow, 117312 Russia). E-mail: aaterenin@yandex.ru. http://orcid.org/0000-0002-6242-6117

Anastasia I. Shabrova, Data Protection Architect, Cybersecurity Department, Sberbank of Russia (19, Vavilova ul., Moscow, 117312 Russia). E-mail: shabrova1113@gmail.com. https://orcid.org/0000-0002-4315-3061

Об авторах

Бабак Никита Григорьевич, аспирант, кафедра вычислительных машин, систем и сетей Института информационных и вычислительных технологий ФГБОУ ВО «Национальный исследовательский университет «МЭИ» (111250, Россия, Москва, Красноказарменная ул., д. 14, стр. 1); главный эксперт по защите данных, Департамент кибербезопасности ПАО «Сбербанк России» (117312, Россия, Москва, ул. Вавилова, д. 19). E-mail: nikita.enrollee@gmail.com. ResearcherID HHY-9372-2022, SPIN-код РИНЦ 3687-6548, https://orcid.org/0000-0001-7129-1018

Белорыбкин Леонид Юрьевич, директор проектов по защите данных, Департамент кибербезопасности ПАО «Сбербанк России» (117312, Россия, Москва, ул. Вавилова, д. 19). E-mail: lbelorybkin@gmail.com. https://orcid.org/0000-0002-8575-5773

Оцоков Шамиль Алиевич, д.т.н., профессор, кафедра КБ-4 «Интеллектуальные системы информационной безопасности» Института кибербезопасности и цифровых технологий ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: shamil24@mail.ru. Scopus Author ID 57212622267, https://orcid.org/0000-0001-7451-5443

Теренин Алексей Алексеевич, к.т.н., управляющий директор, Департамент кибербезопасности ПАО «Сбербанк России» (117312, Россия, Москва, ул. Вавилова, д. 19). E-mail: aaterenin@yandex.ru. http://orcid.org/0000-0002-6242-6117

Шаброва Анастасия Игоревна, архитектор по защите данных, Департамент кибербезопасности ПАО «Сбербанк России» (117312, Россия, Москва, ул. Вавилова, д. 19). E-mail: shabrova1113@gmail.com. https://orcid.org/0000-0002-4315-3061

Translated from Russian into English by Lyudmila O. Bychkova
Edited for English language and spelling by Thomas A. Beavitt

Information systems. Computer sciences. Issues of information security
Информационные системы. Информатика. Проблемы информационной безопасности

UDC 004.032.26
https://doi.org/10.32362/2500-316X-2023-11-5-19-33

RESEARCH ARTICLE

Geometric properties of quantum entanglement and machine learning

Sergei V. Zuev @

V.I. Vernadsky Crimean Federal University, Simferopol, 295007 Russia
@ *Corresponding author, e-mail: zuevsv@cfuv.ru*

Abstract

Objectives. Fast data analysis based on hidden patterns is one of the main issues for adaptive artificial intelligence systems development. This paper aims to propose and verify a method of such analysis based on the representation of data in the form of a quantum state, or, alternatively, in the form of a geometric object in a space allowing online machine learning.

Methods. This paper uses Feynman formalism to represent quantum states and operations on them, the representation of quantum computing in the form of quantum circuits, geometric transformations, topological classification, as well as methods of classical and quantum machine learning. The Python programming language is used as a development tool. Optimization tools for machine learning are taken from the SciPy module. The datasets for analysis are taken from open sources. Data preprocessing was performed by the method of mapping features into numerical vectors, then the method of bringing the data to the desired dimension was applied. The data was then displayed in a quantum state. A proprietary quantum computing emulator is used (it is in the public domain).

Results. The results of computational experiments revealed the ability of very simple quantum circuits to classify data without optimization. Comparative indicators of classification quality are obtained without the use of optimization, as well as with its use. Experiments were carried out with different datasets and for different values of the dimension of feature spaces. The efficiency of the models and methods of machine learning proposed in the work, as well as methods of combining them into network structures, is practically confirmed.

Conclusions. The proposed method of machine learning and the model of quantum neural networks can be used to create adaptive artificial intelligence systems as part of an online learning module. Free online optimization learning process allows it to be applied in data streaming, that is, adapting to changes in the environment. The developed software does not require quantum computers and can be used in the development of artificial intelligence systems in Python as imported modules.

Keywords: online learning, adaptive artificial intelligence, quantum machine learning, quantum entanglement

• **Submitted:** 13.02.2023 • **Revised:** 14.06.2023 • **Accepted:** 13.08.2023

For citation: Zuev S.V. Geometric properties of quantum entanglement and machine learning. *Russ. Technol. J.* 2023;11(5):19–33. https://doi.org/10.32362/2500-316X-2023-11-5-19-33

Financial disclosure: The author has no a financial or property interest in any material or method mentioned.

In addition to the noted affiliation, the author is an employee of the V.G. Shukhov Belgorod State Technological University (Belgorod, Russia). However, although this study is carried out independently of this organization, the author has an obligation to indicate his affiliation with it.

НАУЧНАЯ СТАТЬЯ

Геометрические свойства квантовой запутанности и машинное обучение

С.В. Зуев [@]

Крымский федеральный университет им. В.И. Вернадского, Симферополь, 295007 Россия
[@] *Автор для переписки, e-mail: zuevsv@cfuv.ru*

Резюме

Цели. Быстрая классификация данных на основе имеющихся в них закономерностей является одним из главных вопросов для построения систем адаптивного искусственного интеллекта. Цель работы – предложить и верифицировать метод такой классификации на основе представления данных в виде квантового состояния или (альтернативно) в виде геометрического объекта в пространстве, свойства которого позволяют производить машинное обучение «на лету» (онлайн-обучение).

Методы. В работе используется фейнмановский формализм для представления квантовых состояний и операций над ними, представление квантовых вычислений в виде квантовых схем, геометрические преобразования, топологическая классификация, а также методы классического и квантового машинного обучения. В качестве инструмента разработки использовался язык программирования Python, средства оптимизации для машинного обучения взяты из модуля SciPy. Размеченные данные для анализа взяты из открытых источников. Препроцессинг данных произведен методом отображения признаков в числовые векторы, затем применен метод приведения данных к нужной размерности и далее – отображение данных в квантовое состояние. Используется собственный эмулятор квантовых вычислений (находится в открытом доступе).

Результаты. Результаты вычислительных экспериментов выявили способность очень простых квантовых схем к классификации данных без оптимизации. Получены сравнительные показатели качества классификации без использования оптимизации, а также с ее использованием. Эксперименты проведены с различными датасетами и для различных значений размерности пространств признаков. Работоспособность предложенных в работе моделей и методов машинного обучения, а также методов их объединения в сетевые структуры, подтверждена практически.

Выводы. Предложенный метод машинного обучения и построения квантовых нейронных сетей может быть применен для создания систем адаптивного искусственного интеллекта в составе модуля онлайн-обучения. Отсутствие оптимизации в процессе онлайн-обучения позволяет применять его в потоке данных, т.е., адаптироваться к изменениям среды. Разработанное алгоритмическое обеспечение не требует наличия квантовых компьютеров и может быть применено при разработке программного обеспечения систем искусственного интеллекта на языке Python в качестве импортируемых модулей.

Ключевые слова: онлайн-обучение, адаптивный искусственный интеллект, квантовое машинное обучение, квантовая запутанность

• **Поступила:** 13.02.2023 • **Доработана:** 14.06.2023 • **Принята к опубликованию:** 13.08.2023

Для цитирования: Зуев С.В. Геометрические свойства квантовой запутанности и машинное обучение. *Russ. Technol. J.* 2023;11(5):19–33. https://doi.org/10.32362/2500-316X-2023-11-5-19-33

Прозрачность финансовой деятельности: Автор не имеет финансовой заинтересованности в представленных материалах или методах.

Кроме отмеченной аффилиации, автор является сотрудником БГТУ им. В.Г. Шухова (Белгород, Россия), но настоящая работа выполнена независимо от этой организации, хотя у автора имеется обязанность указать свою принадлежность к ней.

INTRODUCTION

Quantum algorithms are attracting more and more attention, since quantum computers are soon expected to be fully usable. On the other hand, quantum search and factorization algorithms are one of the main reasons for developing quantum computers. Several such computers are currently available around the world. However, their power is relatively low (the largest is the Chinese Jiuzhang with 76 qubits[1]), and they are still used for demonstration and research purposes.

The other pole of progress in information technology is artificial intelligence. Like most human knowledge, artificial intelligence is based on a natural phenomenon known as cognition, which still has no universally recognized quantitative theory. Quantum versions of such theories, inter alia, do not yet have proven clear advantages, although the results of this study can be seen as an indication of certain advantages of quantum methods in machine learning (ML).

The advantages of quantum computing and quantum computers in solving problems in the field of artificial intelligence are covered in review [1], which substantiates the relevance of studies in this area and indicates the main directions in one of which this paper is written. There is also a review of 2023 [2] containing references to all modern advances in this area.

The main possible advantage of quantum version of artificial intelligence is the exponential growth of computational capabilities. While the classical artificial neuron can process input data of N dimensions, quantum neuron can process 2^N-dimensional data. The application of the quantum version can significantly speed up execution of both learning and classification algorithms [3]. At the same time, one of the technical challenges in building a large-scale quantum computer is the need to ensure that there are "qubits that can be initialized with arbitrary values" [4]. This problem is relevant and is a significant obstacle to achieving quantum superiority.

In [5–9], prototypes of quantum neural networks based on constructing a quantum circuit with adjustable parameters were proposed. The present paper shows how this approach can be implemented in connection with the proposed neural network architecture and how such parameter settings can be dispensed with.

Quantum versions of the most popular ML algorithms have already been developed. The above-mentioned quantum neural networks work on a par with traditional ones. In [10], quantum support vector machines (SVMs) using the HHL algorithm [11] for inverting a matrix to generate a hyperplane were proposed. The image classification model presented in 2018 [12] is based on quantum k-nearest neighbors. The quantum linear regression using quantum data is proposed in [13]. A quantum analogue of the decision tree developed in [14] uses quantum accuracy and quantum entropy measurement, i.e., it develops the classical ID3 algorithm.

Several quantum ML methods have been developed for clustering in [15]. In particular, a quantum version of the k-means algorithm in different variants is presented in [16] and [17]. Another quantum clustering algorithm using Grover's algorithm to determine the cluster median is proposed in [18].

The quantum analogue of the quantum principal component analysis method [19] identifying eigenvectors related to the eigenvalues of an unknown state exponentially faster than any other solution has also been developed.

An area close to the topic of the paper is reinforcement learning, i.e., online learning taking into account the response of the environment. There are several quantum versions of reinforcement learning, such as [20] which uses a superposition of quantum states, and due to this parallelism is achieved and the learning speed increases.

Deep learning occupies a special place in ML. Deep learning methods require significant memory and time resources, thus making their development in the quantum area attractive. Among recent advances in this field is a series of works on quantum generative adversarial networks [21–23] with implementation in [24] using a superconducting quantum processor to generate and learn handwritten digital images by quantum generative Wasserstein adversarial networks [25]. It has also been shown that scalability and stability of quantum generative adversarial model learning improves on quantum Boltzmann machines [26, 27], quantum autocoders [28, 29], and quantum convolutional neural networks [7–9]. Among Russian works in this area, study [6] may be specified.

Improving optimization algorithms is also in the focus of research on quantum algorithms. Quantum enhanced optimization [30] as well as quantum gradient descent [31, 32] is used in quantum neural networks, e.g., in quantum Boltzmann machines [27].

Among recent works is experimental study [33] showing that SVMs outperform their classical counterparts by 3–4% on average, while quantum neural networks made on a quantum computer outperform quantum SVMs by 5% on average, and classical neural networks by 7%.

Quantum entanglement in connection with a model of learning was proposed in 2005. This is a model for the semantics of concept combinations created in a non-decomposition way. It deals with emergent properties/associations/inferences in connection with concept combinations[2]. In the paper, this idea is used

[1] https://en.wikipedia.org/wiki/Jiuzhang_(quantum_computer). Accessed January 01, 2023.

[2] Bruza P.D., Cole R.J. *Quantum Logic of Semantic Space: An Exploratory Investigation of Context Effects in Practical Reasoning*, 2006. https://arxiv.org/abs/quant-ph/0612178. Accessed January 01, 2023.

for a different purpose, i.e., to provide a way to separate labeled data. Although not directly related to learning, these ideas may allow a better way to be found to solve the following current problems of data analysis and artificial intelligence.

The first problem is online learning. It arises when the data environment changes and there is no time or resources for new learning in the system. The comprehensive theory of online learning is presented in the course by Massachusetts Institute of Technology, available online[3]. The main challenge in this problem is finding a compromise between quality and responsiveness. Quality-based learning is often time-consuming, while responsiveness-based learning may produce useless results. The better choice is to create a system that adjusts itself with allowance for the content of the data stream it receives. The paper proposes such a system based on entangled quantum states. Generalizing this idea, it is possible to approximate artificial intelligence systems to living intelligence in the sense of adapting to the environment.

The second problem is fast recognition, especially in the case of moving images. This problem is well described in various blogs and articles. For example, one of the current approaches to this issue is presented by Shao and Vitarsia in [34]. This research focuses on applying the BP neural network, i.e., an artificial neural network of forward propagation. The application of quantum algorithms for solving this problem has not been found in the literature. However, there is a software tool[4] designed to compare streaming video data which already works as a web service[5].

Any progress in solving these problems could lead to technological solutions in industries such as self-driving cars, unmanned aerial, and underwater vehicles, as well as video monitoring and other fields largely related to the detection of anomalies in a changing environment. It is not necessary to use a quantum computer to apply the results of this research, since the proposed algorithms can be implemented on emulators or reformulated in classical form. This would probably destroy quantum superiority, although the efficiency of low-dimensional data may be quite sufficient.

MATERIALS AND METHODS

This paper uses the quantum and classical data described below. The set $\{x^j\}$ of sets of n real numbers $\mathbf{x}^j = \{x_0^j, ..., x_{n-1}^j\}$ with label l^j defined for each set is the *classical* data. The set of quantum states

$$|q^j\rangle \equiv \sum_{k=0}^{2^n-1} a_k^j |k\rangle$$

is considered as *quantum* data. Components a_k^j of the quantum state vector are considered as given in a certain computational basis $|0\rangle, ..., |2^n-1\rangle$. These notations are commonly used in such books as [35]. Before the relationship between these data is established, we shall make a few preliminary remarks.

The type of data used to deliver information from system to system in nature is not obvious. However, human operations require classical information. It can be easily seen that 2^n-dimensional quantum system provides only n-dimensional classical data, although the quantum system operates in 2^n-dimensional state space during calculations. The problem of generating the initial quantum state arises from the fact that the source of quantum data, generally speaking, is unknown. It is certainly impossible to generate this data from classical ones. Thus, the only thing that can definitely be assumed is that the system has already had data in quantum form before the start of computation However, this means that all dependencies are already contained in quantum data, and the quantum intelligent system should use them. This is the basis for further consideration.

We shall first describe the state space structure of the system of n qubits. Proceeding from the way in which quantum data is represented, this space is embedded in $\mathbb{C}^N$, where $N = 2^n$ while $\mathbb{C}$ is the space of complex numbers. In addition, quantum states are described by vectors with an absolute value equal to 1, while vectors differing only by phase coefficient $e^{i\varphi}$ describe the same state. This suggests that the equivalence relation may be considered, as follows:

$$\frac{z^1}{z^0} \equiv w^1, ..., \frac{z^{N-1}}{z^0} \equiv w^{N-1}. \quad (1)$$

The space of such vectors $\mathbf{w}$ is called (*complex*) *projective space* $\mathbb{C}P^{N-1}$, which is a set of vectors with N complex coordinates $(z^1, \ldots, z^{N-1})$ connected by equivalence relation (1). Another condition which can be derived from (1) is the following:

$$\left|z^0\right|^2 + ... + \left|z^{N-1}\right|^2 = 1. \quad (2)$$

The phases of coordinates w^k are defined to the accuracy of the common multiplier $e^{-i\phi_0}$, where ϕ_0 is

[3] Rakhlin A. *Online Methods in Machine Learning. Theory and Applications*. TA: Arthur Flajolet. https://www.mit.edu/~rakhlin/6.883/. Accessed January 01, 2023.

[4] Biloushenko I.I., Zuev S.V. *Determining the degree of similarity of video fragments*, 2022; Certificate 2022685057 of 20.12.2022 issued by the Federal Service for Intellectual Property (in Russ.).

[5] https://ais.bstu.ru/services/1. Accessed July 05, 2023.

an arbitrary phase of coordinate z^0. Thus, space $\mathbb{C}P^{N-1}$ can be identified with the space of system states of n qubits while coordinates can be represented in the following form:

$$w^k = \frac{|z^k|}{|z^0|} e^{i(\phi_k - \phi_0)}.$$

We can assume without any restriction that $\phi_0 = 0$. According to the above procedure, the space $\mathbb{C}P^{N-1}$, is homeomorphic to surface (2) of dimension $2N-2$ since $z^0_{\text{im}} = 0$ and, therefore,

$$|z^0|^2 + ... + |z^{N-1}|^2 = \left(z^0_{\text{re}}\right)^2 + \left(z^1_{\text{re}}\right)^2 + \\ + \left(z^1_{\text{im}}\right)^2 + ... + \left(z^{N-1}_{\text{re}}\right)^2 + \left(z^{N-1}_{\text{im}}\right)^2 = 1.$$

This is sphere S^{N-2}, and each of its points z_j can be parameterized using the following generalized spherical coordinates:

$$\begin{aligned} z^0_j &= \cos\delta^0_j \cos\delta^1_j ... \cos\delta^{2^n-2}_j, \\ z^1_j &= \sin\delta^0_j \cos\delta^1_j ... \cos\delta^{2^n-2}_j e^{i\gamma^0_j}, \\ z^2_j &= \sin\delta^1_j \cos\delta^2_j ... \cos\delta^{2^n-2}_j e^{i\gamma^1_j}, \\ &... \\ z^{2^n-2}_j &= \sin\delta^{2^n-3}_j \cos\delta^{2^n-2}_j e^{i\gamma^{2^n-3}_j}, \\ z^{2^n-1}_j &= \sin\delta^{2^n-2}_j e^{i\gamma^{2^n-2}_j}, \end{aligned} \quad (3)$$

where

$$\delta^0_j = \frac{\pi}{2} d^0_j,\ \delta^1_j = \frac{\pi}{2} d^1_j,\ ...,\ \delta^{2^n-2}_j = \frac{\pi}{2} d^{2^n-2}_j, \\ \gamma^0_j = 2\pi d^{2^n-1}_j,\ ...,\ \gamma^{2^n-2}_j = 2\pi d^{2^{n+1}-3}_j$$

and d^i_j is the value of the ith feature in the jth data sample in the scaled data $(d^i_j \in [0,1))$.

Thus, it is possible to encode any training dataset into a quantum state using the following formula:

$$|q_j\rangle = \sum_{k=0}^{2^n-1} z^k_j(d)|k\rangle. \quad (4)$$

According to the postulates of quantum mechanics, if there are two systems with n_1 and n_2 qubits, respectively, then the states of the combined system have the following form:

$$|q\rangle = \sum_{k=0}^{2^{n_2}-1} \sum_{m=0}^{2^{n_1}-1} a^m b^k |m\rangle|k\rangle,$$

where a and b represent the state amplitudes of the first and the second system, respectively.

The state spaces for each system are $S^{2^{n_1}-2}$ and $S^{2^{n_2}-2}$. The set of states for the combined system is their direct product, i.e., $S^{2^{n_1}-2} \times S^{2^{n_2}-2}$. However, for topological reasons, this is definitely not $S^{2^{n_1+n_2}-2}$.

The part of the system of $(n_1 + n_2)$ qubits, which cannot be expressed as a product of subsystem states, forms a set of so-called *entangled states*. The main property of an entangled state is that in order to remove the system from the entangled state, it is necessary to perform a unitary transformation that significantly affects all its subsystems. Entangled states form a basis in the space of states, and further it is called *the entangled basis*.

If the state of a multi-cubit system is entangled, it is impossible to get out of it without affecting each cubit. At the same time, each state of the system can be written in the entangled basis. Thus each state component in this basis affects all qubits significantly. If the amplitudes of these components are measured, it can be seen how subsystems interact in this quantum system. If the state labels are given, then which basis vector corresponds to the label of interest needs to be defined. This can be determined from the statistics of measurement results for a given label. Moreover, if new states of the same system are measured in the same way, it can be predicted with a certain probability that they belong to the labeled class, corresponding to the measurement result that is most relevant to the labeled samples.

To a certain extent, this means that classification can be performed without optimization if the dependencies are already present in the data. The latter is an important addition, since classifying data without dependencies (e.g., when the data is a complete superposition of pure states) would fail. Hence, the dependencies resulting in a given class for classification need to be defined. This is essentially a quantum property related to entanglement. Certainly, this could be interpreted without resorting to quantum representations, but then it would be necessary to consider the topological properties of the set of states of the system, as well as the subsets of its entangled states, in order to build probabilistic models on them. At present, the interpretation in terms of quantum calculations appears simpler.

RESULTS

Quick online classification

We shall consider the marked data set $\{d^i_j, l_j\}$, where d^i_j is the value of the ith feature in the jth sample,

Table 1. The experiment with the heart disease dataset. F1 is the harmonic mean of method precision and recall

Method	Precision, %	Recall, %	F1, %	Learning time, ms	Operation time, ms
Classification by emulated quantum entangled basis	76	65	70	20.5	9
Classification by linear discriminant analysis	100	58	73	7.6	4.9
Classification by logistic regression	68	65	66	5.2	3.1

while l_j is the value of the label (class) for the jth sample. We shall separate all data into training and test samples denoting them by

$$\{d^i_{jt}, l_{jt}\} \text{ and } \{d^i_{jc}, l_{jc}\},$$

respectively. We assume that the values for all features are scaled, while the labels take values from 0 to $L-1$:

$$d^i_j \in [0,1),\ l_j \in \{0, ..., L-1\},$$

where $L = 2^l$.

The case of two qubits

We shall assume $n = 2$, $l = 1$, i.e., the number of features is 6 and the labels take on values 0 and 1. Then $i = 0, ..., 5$ and

$$\delta^0_j = \frac{\pi}{2} d^0_j,\ \ \delta^1_j = \frac{\pi}{2} d^1_j,\ \ \delta^2_j = \frac{\pi}{2} d^2_j,$$
$$\gamma^0_j = 2\pi d^3_j,\ \ \gamma^1_j = 2\pi d^4_j,\ \ \gamma^2_j = 2\pi d^5_j.$$

Data encoding into quantum states, according to (3) and (4), may be written as following:

$$z^0_j = \cos\delta^0_j \cos\delta^1_j \cos\delta^2_j,\ \ z^1_j = \sin\delta^0_j \cos\delta^1_j \cos\delta^2_j e^{i\gamma^0_j},$$
$$z^2_j = \sin\delta^1_j \cos\delta^2_j e^{i\gamma^1_j},\qquad z^3_j = \sin\delta^2_j e^{i\gamma^2_j}, \tag{5}$$

$$\left|q_j\right\rangle = \sum_{k=0}^{3} z^k_j \left|k\right\rangle. \tag{6}$$

We shall consider the following quantum circuit (Fig. 1). This is a well-known circuit for converting Bell states into vectors of computational basis. By using it, the probabilities of how the entangled basis vectors (Bell states) correspond to the vector given at the input is obtained.

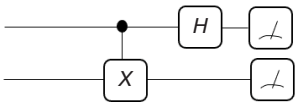


Fig. 1. Quantum circuit for converting Bell states into computational basis vectors. *H*—Hadamard gate, *X*—controlled X gate

The circuit shown in Fig. 1 can classify data containing six features. This is easily verified using the heart disease dataset taken from www.kaggle.com[6].

The original dataset contains 13 features and one label. The features are: age, sex, chest pain, pulse, cholesterol, and others. The label is the presence of heart disease. The dataset contains 303 data instances, of which 165 are labeled 1 while the rest are 0. The examples of analyzing this dataset by linear classifiers given on the website kaggle.com give values of the accuracy metric for predicting disease from 64% to 88%.

The following experiment is performed on this dataset. All feature values are translated into integer ranges from 0 to the feature-dependent limit value. The data is then reduced to six features without loss of information in the data, and parameterization in the form of quantum states of the two-particle system (5) is performed. The separation into training and test samples is done in a ratio of 65/35. The result obtained is compared with the result of the linear discriminant analysis (LDA) classifier[7]. The results are shown in Table 1.

The prepared dataset and program code are presented in the open-access archive[8]. The values shown in Table 1 are not high. However, they are obtained without optimization using the incoming data only, while the algorithm running time is spent mainly on emulation of quantum states and operations. Nevertheless, such classification procedure can work on almost any device in real time, since it does not require optimization and

[6] Akyildiz Ö. Heart disease data. https://www.kaggle.com/datasets/zgeakyldz/heart-desease-data. Accessed August 25, 2023.

[7] https://scikit-learn.ru/1-2-linear-and-quadratic-discriminant-analysis (in Russ.). Accessed January 09, 2023.

[8] Program codes and datasets for the paper are archived at https://disk.yandex.ru/d/JK4dsbdGLP_ZaQ. Accessed January 09, 2023.

can work in parallel streams. The above experiment is carried out on a computer with Intel Core I5 processor and 8GB RAM (Intel, USA) in a single thread.

It may be doubted whether the proposed method is ML (due to the lack of optimization). However, the definition of ML [36] states that ML algorithms build a model based on sample data known as training data, in order to produce predictions or decisions without being explicitly programmed to do so. The method under consideration uses the data and the model contained in it and allows making predictions based on this model, i.e., it fully satisfies the ML definition.

The quality of classification can be improved using known techniques (bagging[9] and boosting[10]). However, it can also be improved by using ML in the traditional sense as a parametric transformation with optimization. The way to use it in the quantum case is shown, for example, in [37]. In the case under consideration, in the circuit shown in Fig. 1, two controlled gates (U_0, U_1) are embedded into the first register (controlled by 0 and by 1), then a simple gate V in the second register, and the second register is measured. The measurement result is related to the label value and the output state set to a clean state. The optimization parameters are the components of the gates. The schematic is shown in Fig. 2.

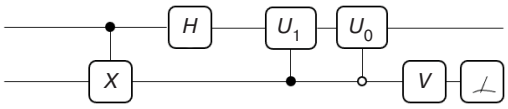


Fig. 2. The simplest quantum classifier with optimization

The two-particle state $|q\rangle = a_0|00\rangle + a_1|01\rangle + a_2|10\rangle + a_3|11\rangle$ is applied to the input and is converted further as follows:

$$|q\rangle \underset{CX}{\to} a_0|00\rangle + a_1|01\rangle + a_2|11\rangle + a_3|10\rangle \underset{H}{\to} \frac{a_0+a_3}{\sqrt{2}}|00\rangle +$$
$$+ \frac{a_1+a_2}{\sqrt{2}}|01\rangle + \frac{a_0-a_3}{\sqrt{2}}|10\rangle + \frac{a_1-a_2}{\sqrt{2}}|11\rangle \underset{CU_1,\,CU_0}{\to}$$
$$\underset{CU_1,\,CU_0}{\to} \frac{a_0+a_3}{\sqrt{2}}U_0|0\rangle|0\rangle + \frac{a_1+a_2}{\sqrt{2}}U_1|0\rangle|1\rangle +$$
$$+ \frac{a_0-a_3}{\sqrt{2}}U_0|1\rangle|0\rangle + \frac{a_1-a_2}{\sqrt{2}}U_1|1\rangle|1\rangle \underset{V}{\to}$$
$$\underset{V}{\to} \frac{a_0+a_3}{\sqrt{2}}U_0|0\rangle V|0\rangle + \frac{a_1+a_2}{\sqrt{2}}U_1|0\rangle V|1\rangle +$$
$$+ \frac{a_0-a_3}{\sqrt{2}}U_0|1\rangle V|0\rangle + \frac{a_1-a_2}{\sqrt{2}}U_1|1\rangle V|1\rangle.$$

[9] Bootstrap aggregating is a classification technique that uses compositions of algorithms each of which is trained independently. The result of classification is determined by voting.

[10] Boosting is a procedure of successive composition of machine learning algorithms, where each successive algorithm seeks to compensate for the compositional deficiencies of all previous algorithms.

If the parameters of gates U_0, U_1, and V are such that

$$U_0|0\rangle = \cos\alpha_0|0\rangle + \sin\alpha_0 e^{i\psi_0}|1\rangle,$$
$$U_0|1\rangle = -\sin\alpha_0 e^{i\phi_0}|0\rangle + \cos\alpha_0 e^{i(\phi_0+\psi_0)}|1\rangle,$$
$$U_1|0\rangle = \cos\alpha_1|0\rangle + \sin\alpha_1 e^{i\psi_1}|1\rangle,$$
$$U_1|1\rangle = -\sin\alpha_1 e^{i\phi_1}|0\rangle + \cos\alpha_1 e^{i(\phi_1+\psi_1)}|1\rangle,$$
$$V|0\rangle = \cos\beta|0\rangle + \sin\beta e^{i\nu}|1\rangle,$$
$$V|1\rangle = -\sin\beta e^{i\mu}|0\rangle + \cos\beta e^{i(\mu+\nu)}|1\rangle,$$

then the resulting state may be written in the following form:

$$A_{00}|00\rangle + A_{01}|01\rangle + A_{10}|10\rangle + A_{11}|11\rangle,$$

where

$$A_{00} = \frac{1}{\sqrt{2}}\left((a_0+a_3)\cos\alpha_0 - (a_0-a_3)\sin\alpha_0 e^{i\phi_0}\right)\cos\beta -$$
$$- \frac{1}{\sqrt{2}}\left((a_1+a_2)\cos\alpha_1 - (a_1-a_2)\sin\alpha_1 e^{i\phi_1}\right)\sin\beta e^{i\mu},$$

$$A_{01} = \frac{e^{i\nu}}{\sqrt{2}}\left((a_0+a_3)\cos\alpha_0 - (a_0-a_3)\sin\alpha_0\right)\sin\beta +$$
$$+ \frac{e^{i\nu}}{\sqrt{2}}\left((a_1+a_2)\cos\alpha_1 - (a_1-a_2)\sin\alpha_1 e^{i\phi_1}\right)\cos\beta e^{i\mu},$$

$$A_{10} = \frac{1}{\sqrt{2}}\left((a_0+a_3)\sin\alpha_0 e^{i\psi_0} + (a_0-a_3)\sin\alpha_0 e^{i\phi_0}\right)\cos\beta -$$
$$- \frac{e^{i(\mu+\psi_1)}}{\sqrt{2}}\left((a_1+a_2)\sin\alpha_1 + (a_1-a_2)\cos\alpha_1 e^{i\phi_1}\right)\sin\beta,$$

$$A_{11} = \frac{e^{i(\nu+\psi_0)}}{\sqrt{2}}\left((a_0+a_3)\sin\alpha_0 + (a_0-a_3)\cos\alpha_0 e^{i\phi_0}\right)\sin\beta +$$
$$+ \frac{e^{i(\mu+\nu+\psi_1)}}{\sqrt{2}}\left((a_1+a_2)\sin\alpha_1 + (a_1-a_2)\cos\alpha_1 e^{i\phi_1}\right)\cos\beta.$$

When labeled 0, this state should produce the result "0" with the highest probability when measured in the second register, which means the following:

$$A_{01} \to 0, A_{11} \to 0.$$

When labeled 1, on the contrary, it should be the following:

$$A_{00} \to 0, A_{10} \to 0.$$

The learning procedure is designed to find the best set of gate parameters (α_0, ϕ_0, ψ_0, α_1, ϕ_1, ψ_1, β, μ, and ν) which provides the best aspirations. This is an optimization problem.

The classifier shown in Fig. 2 is the smallest possible classifier of this type. It is an analogue of an artificial neuron with two-dimensional input: it receives 2-cubic quantum signal, while outputting 1-cubic quantum signal and 1-bit classical signal (unlike classical neuron, where only one signal is output). The presence of quantum signal at the output allows the further use of quantum information, i.e., creating a network. Training and operation of such a classifier is illustrated in the archive[11]. As the experiment with the dataset on heart disease shows, optimization increases the classification quality indicators but only insignificantly. This means that in the case under consideration, quantum machine learning (QML) based on quantum entanglement can be used, without any optimization. However, this is not a general statement; this may not be the case for higher-dimensional classifiers.

Quantum classifier training

The quantum circuit shown in Fig. 2 contains three gates (U_0, U_1, and V) with parameters that can be adjusted. For example, gate V can be written in the following form:

$$V = \cos\beta|0\rangle\langle 0| - \sin\beta e^{i\tau}|0\rangle\times$$
$$\times\langle 1| + \sin\beta e^{i\theta}|1\rangle\langle 0| + \cos\beta e^{i(\tau+\theta)}|1\rangle\langle 1|,$$

and the specified parameters are β, θ, and τ. However, changing parameter τ results in the multiplication of the resulting state by the phase multiplier only, i.e., it does not result in a significant effect. This parameter would play a role in systems of higher dimensionality while in the case of two qubits, only two parameters, β and θ, are varied for optimizing vent V.

The same is true for gates U_0 and U_1, which can be written as follows:

$$U_\varepsilon = \cos\alpha_\varepsilon|0\rangle\langle 0| - \sin\alpha_\varepsilon|0\rangle\langle 1| + \sin\alpha_\varepsilon e^{i\rho_\varepsilon}|1\rangle\times$$
$$\times\langle 0| + \cos\alpha_\varepsilon e^{i\rho_\varepsilon}|1\rangle\langle 1|,\ \varepsilon = 0, 1,$$

where the varying parameters are α_0, α_1, ρ_0, and ρ_1. In total, there are six varying parameters for the two-particle quantum circuit.

We shall construct the likelihood function as the sum of the moduli of the following scalar products:

$$\left(|\hat{a}_j|^2, |\hat{b}_j|^2\right) \text{ and } (1 - l_j, l_j),$$

where $\hat{a}_j$ and $\hat{b}_j$ are the amplitudes of the quantum state leaving the second register on the jth package while l_j is the label of the incoming quantum state. The following are the calculations for the initial state (6). Before the controlled gates:

$$|q_j\rangle \to z_j^0|00\rangle + z_j^1|01\rangle + z_j^2|11\rangle + z_j^3|10\rangle \to$$
$$\to \frac{1}{\sqrt{2}}\left(z_j^0 + z_j^3\right)|00\rangle + \frac{1}{\sqrt{2}}\left(z_j^1 + z_j^2\right)|01\rangle +$$
$$+ \frac{1}{\sqrt{2}}\left(z_j^0 - z_j^3\right)|10\rangle + \frac{1}{\sqrt{2}}\left(z_j^1 - z_j^2\right)|11\rangle$$

controlled valves U_1 and U_0 operate:

$$\frac{1}{\sqrt{2}}\left(\left(z_j^0 + z_j^3\right)\cos\alpha_0 - \left(z_j^0 - z_j^3\right)\sin\alpha_0\right)|00\rangle +$$
$$+\frac{1}{\sqrt{2}}\left(\left(z_j^1 + z_j^2\right)\cos\alpha_1 - \left(z_j^1 - z_j^2\right)\sin\alpha_1\right)|01\rangle +$$
$$+\frac{1}{\sqrt{2}}\left(\left(z_j^0 + z_j^3\right)\sin\alpha_0 e^{i\rho_0} + \left(z_j^0 - z_j^3\right)\cos\alpha_0 e^{i\rho_0}\right)|10\rangle +$$
$$+\frac{1}{\sqrt{2}}\left(\left(z_j^1 + z_j^2\right)\sin\alpha_1 e^{i\rho_1} + \left(z_j^1 - z_j^2\right)\cos\alpha_1 e^{i\rho_1}\right)|11\rangle.$$

Finally, valve V is active:

$$\frac{1}{\sqrt{2}}\left(\left(z_j^0 + z_j^3\right)\cos\alpha_0 - \left(z_j^0 - z_j^3\right)\sin\alpha_0\right)\times$$
$$\times\left(\cos\beta|00\rangle + \sin\beta e^{i\theta}|01\rangle\right) +$$
$$+\frac{1}{\sqrt{2}}\left(\left(z_j^0 + z_j^3\right)\sin\alpha_0 e^{i\rho_0} + \left(z_j^0 - z_j^3\right)\cos\alpha_0 e^{i\rho_0}\right)\times$$
$$\times\left(\cos\beta|10\rangle + \sin\beta e^{i\theta}|11\rangle\right) +$$
$$+\frac{1}{\sqrt{2}}\left(\left(z_j^1 + z_j^2\right)\cos\alpha_1 - \left(z_j^1 - z_j^2\right)\sin\alpha_1\right)\times$$
$$\times\left(-\sin\beta|00\rangle + \cos\beta e^{i\theta}|01\rangle\right) +$$
$$+\frac{1}{\sqrt{2}}\left(\left(z_j^1 + z_j^2\right)\sin\alpha_1 e^{i\rho_1} + \left(z_j^1 - z_j^2\right)\cos\alpha_1 e^{i\rho_1}\right)\times$$
$$\times\left(-\sin\beta|10\rangle + \cos\beta e^{i\theta}|11\rangle\right).$$

This is the state before the measurement and can be written in the following form:

$$B_{00}|00\rangle + B_{01}|01\rangle + B_{10}|10\rangle + B_{11}|11\rangle,$$

where

$$B_{00} = \frac{1}{\sqrt{2}}\left(\left(z_j^0 + z_j^3\right)\cos\alpha_0 - \left(z_j^0 - z_j^3\right)\sin\alpha_0\right)\cos\beta -$$
$$-\frac{1}{\sqrt{2}}\left(\left(z_j^1 + z_j^2\right)\cos\alpha_1 - \left(z_j^1 - z_j^2\right)\sin\alpha_1\right)\sin\beta,$$

[11] https://disk.yandex.ru/d/JK4dsbdGLP_ZaQ. Accessed January 09, 2023.

$$B_{01}=\frac{e^{i\theta}}{\sqrt{2}}\left(\left(z_j^0+z_j^3\right)\cos\alpha_0\sin\beta-\left(z_j^0-z_j^3\right)\sin\alpha_0\right)\sin\beta+$$
$$+\frac{e^{i\theta}}{\sqrt{2}}\left(\left(z_j^1+z_j^2\right)\cos\alpha_1\cos\beta-\left(z_j^1-z_j^2\right)\sin\alpha_1\right)\cos\beta,$$

$$B_{10}=\frac{e^{i\rho_0}}{\sqrt{2}}\left(\left(z_j^0+z_j^3\right)\sin\alpha_0+\left(z_j^0-z_j^3\right)\cos\alpha_0\right)\cos\beta-$$
$$-\frac{e^{i\rho_1}}{\sqrt{2}}\left(\left(z_j^1+z_j^2\right)\sin\alpha_1+\left(z_j^1-z_j^2\right)\cos\alpha_1\right)\sin\beta,$$

$$B_{11}=\frac{e^{i(\theta+\rho_0)}}{\sqrt{2}}\left(\left(z_j^0+z_j^3\right)\sin\alpha_0+\left(z_j^0-z_j^3\right)\cos\alpha_0\right)\sin\beta+$$
$$+\frac{e^{i(\theta+\rho_1)}}{\sqrt{2}}\left(\left(z_j^1+z_j^2\right)\sin\alpha_1+\left(z_j^1-z_j^2\right)\cos\alpha_1\right)\cos\beta.$$

The second register measurement would give 0 with probability $|\hat{a}_j|^2$ and 1 with probability $|\hat{b}_j|^2$, as follows:

$$|\hat{a}_j|^2=|B_{00}|^2+|B_{10}|^2,$$
$$|\hat{b}_j|^2=|B_{01}|^2+|B_{11}|^2.$$

It can be easily seen that the only remaining phase parameter is $\rho\equiv\rho_1-\rho_0$, and the likelihood function has the following form:

$$\Phi(\alpha_0,\alpha_0,\beta,\rho)=\sum_j|\hat{a}_j|^2(1-l_j)+\sum_j|\hat{b}_j|^2 l_j.$$

It should be maximized so that state $|1\rangle$ is expected in the second register, if $l_j = 1$. This is done using the COBYLA[12] method in the program code available in the archive[13].

The calculation of the likelihood function is given in order for it be used for calculations on a classical computer. In the case of a quantum computer, the likelihood function manually does not need to be calculated manually, since values $|\hat{a}_j|^2$ and $|\hat{b}_j|^2$ would be available as measurement results.

Arbitrary number of qubits

We shall generalize the constructed classifier to the case of an arbitrary number of particles starting with the version without optimization. For this, it is necessary to construct a multi-particle entangled basis. In the case of two particles, the entanglement appears in the form of a combination of the computational basis vector and its inverted vector. Such combinations for the case of three particles may be written as follows:

$$\frac{1}{\sqrt{2}}\left(|000\rangle+|111\rangle\right),\ \frac{1}{\sqrt{2}}\left(|001\rangle+|110\rangle\right),$$
$$\frac{1}{\sqrt{2}}\left(|010\rangle+|101\rangle\right),\ \frac{1}{\sqrt{2}}\left(|011\rangle+|100\rangle\right),$$
$$\frac{1}{\sqrt{2}}\left(|000\rangle-|111\rangle\right),\ \frac{1}{\sqrt{2}}\left(|001\rangle-|110\rangle\right),$$
$$\frac{1}{\sqrt{2}}\left(|010\rangle-|101\rangle\right),\ \frac{1}{\sqrt{2}}\left(|011\rangle-|100\rangle\right).$$

It can be seen easily that none of these states is the result of the tensor product of three one-particle or any two-particle and one-particle states. Thus, all of them are entangled states and form the basis which is also easy to check. This is the entangled basis for the 3-particle system. Similarly, the entangled basis for an arbitrary *n*-particle quantum system may be constructed, as follows:

$$\frac{1}{\sqrt{2}}\left(|0x\rangle+|1\bar{x}\rangle\right),\ \frac{1}{\sqrt{2}}\left(|0x\rangle-|1\bar{x}\rangle\right),$$

where x is a binary notation of a number from 0 to $2^{n-1}-1$ while the superscript denotes inversion.

The quantum circuit of the multi-particle classifier without optimization is shown in Fig. 3.

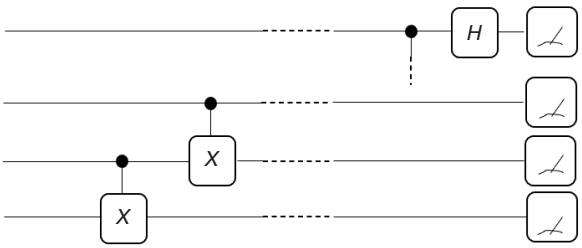


Fig. 3. Classifier without optimization

The generalization of the circuit shown in Fig. 2 is now obvious. It would be enough to set the controlled gates ($2n-2$ items, two for each control register) to the first register and one gate to the registers starting from the second. This is shown in Fig. 4.

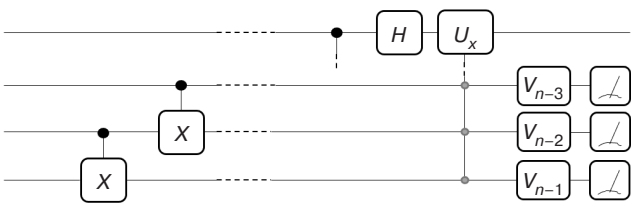


Fig. 4. Optimizable quantum classifier

[12] Constrained optimization by linear approximation (COBYLA).

[13] https://disk.yandex.ru/d/JK4dsbdGLP_ZaQ. Accessed January 09, 2023.

The gray dots in Fig. 4 indicate the 0 or 1 control on one of the registers. The circuit shown in Fig. 4 is a quantum neuron with an arbitrary number of inputs. The circuit shown in Fig. 3 can be viewed in the same way, if the first register is not measured since it would then contain a state which is a superposition of two vectors of the entangled basis. The vectors are determined by the measurement result. Learning without optimization is done, as before, on statistics and on the assumption that there are regularities in the data.

Anomaly detection on streaming data

When data is streaming, the last J incoming packages d^i_j, $j=\overline{0,J-1}$ may be taken into account. This data is not marked up a priori. The task is to detect an anomaly in the stream, i.e., a situation where an incoming series of data packages is significantly different from those in the stream before. The difference may not only be in the packages themselves but also in the order they are received.

The classifiers proposed above can be used to detect anomalies. The selected J packages are considered as basic for statistics; a series of measurements from 0 and 1 are obtained from them. For example, in the 2-particle case there are four possible measurement results (if the quantum state output is not considered). The frequencies of these measurement results form a vector of the system current state. If this vector remains approximately constant or takes values only in certain clusters whenever it is computed on J packages, then any new input from J packages on which the mentioned vector does not fall into any of the clusters would be an anomaly.

Two questions naturally arise. The first relates to the situation when there is a large number of features in streaming data, for example, in the case of a video stream. According to the consideration above, if there are F features, then $\log\left(F/2+1\right)$ qubits and $F/2+1$ possible states are required. This can be a quitelarge number. Therefore, a threshold on the frequency of the feature occurrences should be set, and this is a configurable parameter. In the experiment below with the student' dataset, the threshold frequency is set to 10%. If a feature appears in a given state in less than 10% of cases, then it is considered not to appear in that state at all. As this threshold increases, the number of data instances not classified by the network increases. When the threshold is decreased, the metrics deteriorates.

Another question concerns the J value. There are some recommendations for it: J should be less than the number of packages appearing in the maximum decision time and, at the same time, J should be relatively large in order for the statistics to be rich enough, i.e., for training to give effective prediction. If the anomaly is not detected, then the model is undertrained and J needs to be increased. Also if there are a lot of false positives, then the retraining takes place while J needs to be decreased.

An example code of the anomaly detector based on the classifier built on a tangled basis is given in the archive[14].

Artificial neural network from neuron quantum analogs

The quantum neuron (q-neuron) shown in Fig. 4 or 3 can be used to build a quantum neural network (QNN). In general terms, such a network, receiving I quantum states (qubits) at the input, returns $Q < I$ quantum states together with $C = I - Q$ classical bits of information. The QNN can be trained on the basis of the classical information received. Training procedure in the case of q-neurons with optimization coincides with that of the classical case. The parameters of quantum gates act as weights.

Copying of quantum states is not allowed. Therefore, QNNs do not contain branching. However, q-neuron can contain more than one quantum output that can be used for creating networks of different architecture.

Let there be Q_1 q-neurons in the first layer. Clearly, that $Q_1 \le \frac{I}{2}$ since any q-neuron should have at least two qubits at the input. Accordingly, there would be $C_1 \le I - Q_1$ classical bits of information at the layer output. The next layer may have Q_2 q-neurons, with $Q_2 \le \frac{I-C_1}{2}$. This is shown schematically in Fig. 5, where all q-neurons have one quantum output each while $I = 20$, $Q_1 = 8$, and $Q_2 = 3$. If no second layer is added, the network shown in Fig. 5 would produce 8 qubits and 12 bits at the output. If one more q-neuron is added after the second layer (as a third layer), there would be 1 qubit and 19 bits at the output.

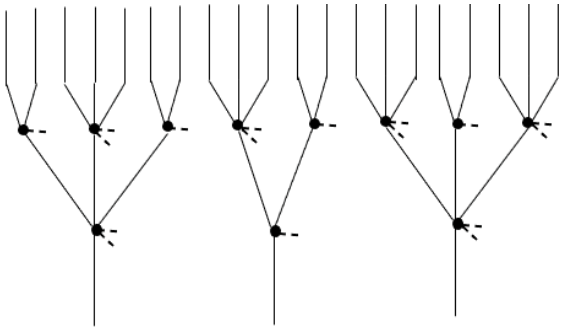


Fig. 5. QNN on 11 q-neurons (black dots): 20 qubits (lines) at the input, 3 qubits at the output, and 17 bits (dashed lines) of classical information

[14] https://disk.yandex.ru/d/JK4dsbdGLP_ZaQ. Accessed January 09, 2023.

The QNN shown in Fig. 5 is similar to the fractal neural network discussed in [38] in architecture. Although classical neural networks with ordinary neurons are considered in [38], the properties noted there such as high learnability and the ability to work with high dimensionality of input vectors remain the same for QNN.

Quantum neural network training

Training a single q-neuron has already been discussed above. We shall generalize this procedure to training the network. For simplicity, only the case of two layers is focused on. We shall take for example the network shown in Fig. 5 in the center containing five qubits at the input, two q-neurons and three bits of classical information on the first layer, and one q-neuron and one bit of classical information on the second layer.

By performing a forward pass on the training sample, the statistics of matching labels and sets of classical bits is obtained. We shall match them with each other. Let the labels take only two possible values in this case. One q-neuron output allows 2 classical bits to be attained when both registers are measured, since there is no task of quantum state propagation further and so it can be measured. Therefore, initially, the most frequent outputs (it could be either 00, 01, 10, or 11) can be allocated to label 0 approximately in proportion to the share of this label in the training sample, while the rest allocated to label 1. Let label ε correspond to e_0, e_1, e_2, and e_3 of all n_0, n_1, n_2, and n_3 outputs of species 00, 01, 10, and 11, respectively. This label has a fraction f_ε in the training sample. We shall select ε with the highest fraction, sort $\frac{e_0}{n_0},\frac{e_1}{n_1},\frac{e_2}{n_2},\frac{e_3}{n_3}$ in descending order and choose the sum of the first elements of the resulting list, with the closest to f_ε fraction of the sum of the corresponding n_i in the total number of packages. The corresponding outputs are allocated to label ε.

Example. A sample of 5,000 batches is separated by labels of 4000 (0) and 1000 (1). For label 0 at the network output, e_0 = 1500, e_1 = 500, e_2 = 1400, and e_3 = 600 is obtained while total numbers of outputs are n_0 = 1600, n_1 = 700, n_2 = 2000, and n_3 = 700. We have:

$$\frac{e_0}{n_0},\frac{e_1}{n_1},\frac{e_2}{n_2},\frac{e_3}{n_3}=\frac{1500}{1600},\frac{500}{700},\frac{1400}{2000},\frac{600}{700}.$$

The order of n_i is as follows: n_0, n_3, n_1, and n_2. This label has a fraction of 0.8, and the closest matching sum is $n_0 + n_3 + n_1$ = 3000 having a fraction of 0.6. Hence, outputs 00, 01, and 11 should be allocated to label 0, while output 10 to label 1. Although, of course, with so many batches in the sample, it would be better to increase the number of features and, consequently, the number of qubits at the input (the quantum network cannot expand due to the prohibition of copying).

Unlike a conventional neural network and a QNN with optimization, QNN without optimization can be trained in the forward direction instead of backward, i.e., simultaneously with its operation. This can be performed according to the scenario described above, since the same logic may be used to calculate statistics for the first layer: separating outputs by labels and keeping on doing so further in the next layers. The classification quality would improve anyway from layer to layer.

DISCUSSION

Many works are devoted to QMLs and QNNs. In the Russian-language literature, work [1] as already mentioned, in particular, points out the importance of "searching for a QNN model which is optimal in terms of utilizing all the advantages presented by quantum computing and neural networks, as well as ML algorithms". One of the most recent reviews in these areas [39], contains the following text in the Outlook section: "The first quantum advantages in QML will likely arise from hidden parameter extraction from quantum data. This can be for quantum sensing or quantum state classification/regression. Fundamentally, we know from the theory of optimal measurement that non-local quantum measurements can extract hidden parameters using fewer samples. Using QML, one can form and search over a parameterization of hypotheses for such measurements." This paper presents one possible path for classifying quantum states.

The QML considered in the paper is of QC or QQ class, i.e., it uses quantum data on classical (emulating) or quantum devices. The proposed q-neuron is ideologically close to known concepts (described in [6] and [39], among others). However, at the same time it has a new essential feature that is exploiting quantum entanglement. In particular, q-neuron can operate without learning in the usual sense; optimization and error back propagation are not needed.

This can be exemplified by the experiment with the analysis of the dataset of student states[15]. For this experiment, the dataset is used without preprocessing. The entire preprocessing is captured in the analysis script available in the archive[16]. The dataset is a labeled one with 34 feature columns and 4424 data instances. The binary label used in the experiment is Dropout = 1, with other values (Enrolled, Graduate) = 0. The task

[15] https://www.kaggle.com/datasets/thedevastator/higher-education-predictors-of-student-retention. Accessed January 09, 2023.

[16] https://disk.yandex.ru/d/JK4dsbdGLP_ZaQ. Accessed January 09, 2023.

Table 2. Experiment with students' dataset

Method	Precision, %	Recall, %	F1, %	Unclassified states, %	Learning time, ms	Operation time, ms
Classification by emulated quantum entangled basis	39	100	57	40	2968	1901
Classification by linear discriminant analysis	100	29	45	0	138	23

is to predict dropout. The features are mapped to integer intervals, the feature space is transformed to dimension 62 without loss of information. This is done for the use of QNN of two q-neurons in the first layer and one q-neuron in the second layer. Training is carried out without optimization. The network configuration implies the first q-neuron with three inputs and the second one with two inputs, in the first layer. The second layer contains a single q-neuron with two inputs. Some of the data instances could not be classified by the neural network, since they are approximately equally close to both 1 and 0 label values. The results can be compared with the classical LDA algorithm taken from the scikit-learn package of the Python language. The results are shown in Table 2.

It can be seen from the experiment that QNN does not ascribe label 0 to the student if he/she drops out: false negatives are equal to 0. At the same time, the share of false positives is quite high: more than half of the dropout labels (1) are false. The QNN considered 40% of the data unsuitable for classification. Certainly, these results are quite different from those of the LDA classifier, and it is unclear in which direction (better or worse). Nevertheless, it can be seen that LDA training time exceeds the running time by 6 times, while QNN has only 1.5 times.

A rigorous examination of the QNN performance of different architectures is planned in the following studies.

CONCLUSIONS

A new direction of QML development applying quantum entanglement significantly is proposed. It allows for the building of intelligent systems working on streaming data and learning online, taking into account changes in the data environment but not reduced to reinforcement learning. The proposed learning method could be called "reinforcement learning in reverse". In reinforcement learning, the agent calculates the classification quality while the environment remains an external factor. However, in the proposed approach, the environment is the carrier of classification patterns and they are recovered directly from it.

Such systems can be used in control systems of unmanned vehicles of any kind, as well as in security systems and intelligent business-assistants. In this case, the use of quantum computers is not mandatory.

Acknowledgments

This study is supported partially by the Russian Foundation for Basic Research (project No. 19-29-09056).

REFERENCES

1. Sigov A.S., Andrianova E.G., Zhukov D.O., Zykov S.V., Tarasov I.E. Quantum informatics: overview of the main achievements. *Russian Technological Journal*. 2019;7(1):5–37 (in Russ.). https://doi.org/10.32362/2500-316X-2019-7-1-5-37
2. Zeguendry A., Jarir Z., Quafafou M. Quantum Machine Learning: A Review and Case Studies. *Entropy*. 2023;25(2):287. https://doi.org/10.3390/e25020287
3. Tacchino F., Macchiavello C., Gerace D., et al. An artificial neuron implemented on an actual quantum processor. *npj Quantum Inf*. 2019;5(1):26. https://doi.org/10.1038/s41534-019-0140-4
4. DiVincenzo D.P. The Physical Implementation of Quantum Computation. *Fortschritte der Physik*. 2000;48(9–11);771–783. https://doi.org/10.1002/1521-3978(200009)48:9/11%3C771::AID-PROP771%3E3.0.CO;2-E
5. Menneer T., Narayanan A. Quantum-inspired neural networks. In: *Proceedings of the Neural Information Processing Systems 95*. Denver, CO, USA, 27–30 November 1995. Available from URL: https://www.researchgate.net/publication/2267350_Quantum-inspired_Neural_Networks
6. Gushanskiy S.M., Buglov V.E. Quantum deep learning of convolutional neural network using variational quantum circuit. *Izvestiya YuFU. Tekhnicheskie nauki = Izvestiya SFedU. Engineering Sciences*. 2021;7(224):167–174 (in Russ.).
7. Cong I., Choi S., Lukin M.D. Quantum convolutional neural networks. *Nat. Phys*. 2019;15(12):1273–1278. https://doi.org/10.1038/s41567-019-0648-8
8. Kerenidis I., Landman J., Prakash A. *Quantum algorithms for deep convolutional neural networks*. arXiv:1911.01117. https://doi.org/10.48550/arXiv.1911.01117
9. Henderson M., Shakya S., Pradhan S., Cook T. Quanvolutional neural networks: Powering image recognition with quantum circuits. *Quantum Mach. Intell*. 2020;2(1):2. https://doi.org/10.1007/s42484-020-00012-y
10. Rebentrost P., Mohseni M., Lloyd S. Quantum Support Vector Machine for Big Data Classification. *Phys. Rev. Lett*. 2014;113(13):130503. https://doi.org/10.1103/physrevlett.113.130503
11. Harrow A.W., Hassidim A., Lloyd S. Quantum Algorithm for Linear Systems of Equations. *Phys. Rev. Lett*. 2009;103(15):150502. https://doi.org/10.1103/physrevlett.103.150502
12. Dang Y., Jiang N., Hu H., Ji Z., Zhang W. Image classification based on quantum K-Nearest-Neighbor algorithm. *Quantum Inf. Process*. 2018;17(9):239. https://doi.org/10.1007/s11128-018-2004-9
13. Schuld M., Sinayskiy I., Petruccione F. Prediction by linear regression on a quantum computer. *Phys. Rev. A*. 2016;94(2):022342. https://doi.org/10.1103/PhysRevA.94.022342
14. Lu S., Braunstein S.L. Quantum decision tree classifier. *Quantum Inf. Process*. 2014;13(3):757–770. https://doi.org/10.1007/s11128-013-0687-5
15. Lloyd S., Mohseni M., Rebentrost P. *Quantum algorithms for supervised and unsupervised machine learning*. arXiv:1307.0411. https://doi.org/10.48550/arXiv.1307.0411
16. Lloyd S. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* 1982;28(2):129–137. https://doi.org/10.1109/TIT.1982.1056489
17. Kerenidis I., Landman J., Luongo A., Prakash A. *q-means: A quantum algorithm for unsupervised machine learning*. arXiv: 1812.03584. https://doi.org/10.48550/arXiv.1812.03584
18. Aïmeur E., Brassard G., Gambs S. Quantum speed-up for unsupervised learning. *Mach. Learn.* 2013;90(2):261–287. https://doi.org/10.1007/s10994-012-5316-5
19. Lloyd S., Mohseni M., Rebentrost P. Quantum principal component analysis. *Nat. Phys.* 2014;10(9):631–633. https://doi.org/10.1038/nphys3029
20. Dong D., Chen C., Li H., Tarn T.-J. Quantum Reinforcement Learning. *IEEE Trans. Syst. Man Cybern. Part B* (*Cybern*.) 2008;38(5):1207–1220. https://doi.org/10.1109/TSMCB.2008.925743
21. Lloyd S., Weedbrook C. Quantum Generative Adversarial Learning. *Phys. Rev. Lett*. 2018;121(4):040502. https://doi.org/10.1103/PhysRevLett.121.040502
22. Dallaire-Demers P.-L., Killoran N. Quantum generative adversarial networks. *Phys. Rev. A*. 2018;98(1):012324. https://doi.org/10.1103/PhysRevA.98.012324
23. Situ H., He Z., Wang Y., Li L., Zheng S. Quantum generative adversarial network for generating discrete distribution. *Inf. Sci*. 2020;538:193–208. https://doi.org/10.1016/j.ins.2020.05.127
24. Huang H.-L., Du Y., Gong M., Zhao Y., et al. Experimental Quantum Generative Adversarial Networks for Image Generation. *Phys. Rev. Appl*. 2021;16(2):024051. https://doi.org/10.1103/PhysRevApplied.16.024051
25. Chakrabarti S., Yiming H., Li T., Feizi S., Wu X. *Quantum Wasserstein generative adversarial networks*. arXiv:1911.00111. https://doi.org/10.48550/arXiv.1911.00111
26. Kieferová M., Wiebe N. Tomography and generative training with quantum Boltzmann machines. *Phys. Rev. A*. 2017;96(6):062327. https://doi.org/10.1103/PhysRevA.96.062327
27. Amin M.H., Andriyash E., Rolfe J., Kulchytskyy B., Melko R. Quantum Boltzmann Machine. *Phys. Rev. X*. 2018;892:021050. https://doi.org/10.1103/PhysRevX.8.021050
28. Romero J., Olson J.P., Aspuru-Guzik A. Quantum autoencoders for efficient compression of quantum data. *Quantum Sci. Technol*. 2017;2(4):045001. https://doi.org/10.1088/2058-9565/aa8072
29. Khoshaman A., Vinci W., Denis B., Andriyash E., Sadeghi H., Amin M.H. Quantum variational autoencoder. *Quantum Sci. Technol*. 2018;4(1):014001. https://doi.org/10.1088/2058-9565/aada1f
30. Jiang Z., Rieffel E.G., Wang Z. Near-optimal quantum circuit for Grover's unstructured search using a transverse field. *Phys. Rev. A*. 2017;95(6):062317. https://doi.org/10.1103/PhysRevA.95.062317
31. Farhi E., Goldstone J., Gutmann S. *A quantum approximate optimization algorithm*. arXiv:1411.4028. https://doi.org/10.48550/arXiv.1411.4028
32. Kerenidis I., Prakash A. Quantum gradient descent for linear systems and least squares. *Phys. Rev. A*. 2020;101(2):022316. https://doi.org/10.1103/PhysRevA.101.022316

33. Simões R.D.M., Huber P., Meier N., Smailov N., Füchslin R.M., Stockinger K. Experimental Evaluation of Quantum Machine Learning Algorithms. *IEEE Access*. 2023;11:6197–6208. https://doi.org/10.1109/ACCESS.2023.3236409
34. Shao Yu., Witarsyah D. Fast recognition method of moving video images based on BP neural networks. *Open Physics*. 2018;16(1):1024–1032. https://doi.org/10.1515/phys-2018-0123
35. Nielsen M.A., Chuang I.L. *Quantum Computation and Quantum Information*. 2nd ed. Cambridge, UK: Cambridge University Press; 2010. 866 p. ISBN 978-1-107-00217-3
36. Samuel A.L. Some Studies in Machine Learning Using the Game of Checkers. *IBM J. Res. Develop*. 1959;44(1–2):207–226. (Reprinted from *IBM J. Res. Develop*. 1959;3(3):210–229. https://doi.org/10.1147/rd.33.0210)
37. Maheshwari D., Sierra-Sosa D., Garcia-Zapirain B., Variational Quantum Classifier for Binary Classification: Real vs Synthetic Dataset. *IEEE Access*. 2022;10:3705–3715. https://doi.org/10.1109/ACCESS.2021.3139323
38. Zuev S., Kabalyants P., Polyakov V., Chernikov S. Fractal Neural Networks. In: *2021 International Conference on Electrical, Computer and Energy Technologies* (*ICECET*). 2021. https://doi.org/10.1109/ICECET52533.2021.9698649
39. Cerezo M., Verdon G., Huang H.Y., et al. Challenges and opportunities in quantum machine learning. *Nat. Comput. Sci*. 2022;2(9):567–576. https://doi.org/10.1038/s43588-022-00311-3

About the author

Sergei V. Zuev, Cand. Sci. (Phys.-Math.), Associate Professor, Department of Computer Engineering and Modelling, Institute of Physics and Technologies, V.I. Vernadsky Crimean Federal University (4, Vernadskogo pr., Simferopol, 295007 Republic of Crimea, Russia). E-mail: zuevsv@cfuv.ru. Scopus Author ID 57292501000, ResearcherID U-1055-2017, https://orcid.org/0000-0003-4237-0491

Об авторе

Зуев Сергей Валентинович, к.ф.-м.н., доцент, кафедра компьютерной инженерии и моделирования Физико-технического института ФГАОУ ВО «Крымский федеральный университет им. В.И. Вернадского» (295007, Россия, Республика Крым, Симферополь, пр-т Вернадского, д. 4). E-mail: zuevsv@cfuv.ru. Scopus Author ID 57292501000, ResearcherID U-1055-2017, https://orcid.org/0000-0003-4237-0491

Translated from Russian into English by Kirill V. Nazarov
Edited for English language and spelling by Dr. David Mossop

Modern radio engineering and telecommunication systems

Современные радиотехнические и телекоммуникационные системы

UDC 658.5.012.14 + 621.37
https://doi.org/10.32362/2500-316X-2023-11-5-34-44

RESEARCH ARTICLE

Investigation of the profilogram structure of microstrip microwave modules manufactured using additive 3D-printing technology

Dmitry S. Vorunichev [@], **Mikhail S. Kostin**

MIREA – Russian Technological University, Moscow, 119454 Russia
[@] *Corresponding author, e-mail: vorunichev@mirea.ru*

Abstract

Objectives. The aim of the work is to study the surface roughness of the current-carrying topology and dielectric of the upper (Top Layer) and lower (Bottom Layer) sides of microwave modules manufactured using additive three-dimensional printing technology when prototyping prototypes of microwave modules on a 3D printer of DragonFly 2020 LDM multilayer printed circuit boards.

Methods. Methods of metallographic analysis in bright and dark fields, surface roughness profiling, and computer modeling were used.

Results. Experimental samples of microstrip microwave elements of modules of multilayer boards of a given configuration, telemetry sensors, printed circuit board (PCB) antennas were obtained. The topological and radiophysical features of the additively formed upper and lower surface layers of experimental samples of boards of strip modules were studied. Optical profilogram measurements of the roughness of the outer sides of the board were carried out at 10 points, amounting to 2 μm for the upper layer of the topology and 0.3 μm for the lower layer; the average grain size of the dielectric base was determined at 0.007 mm^2. The roughness of the conductive topology and upper side dielectric was shown to correspond to an accuracy class of 6–7, while the roughness of the microstrip conductive topology and the dielectric of the lower side of the board corresponds to an accuracy class of 10–12.

Conclusions. It is established that an uneven formation of the lower and upper strip layers of the printed module can affect the inhomogeneity of the distribution of radiophysical parameters (dielectric permittivity, surface conductivity, etc.), as well as the instability of the structural (adhesion ability, thermal conductivity, etc.) characteristics of the strip module, which must be taken into account when prototyping devices using inkjet 3D printing technology, including when adapting Gerber projects of PCB modules created for classical board production technology.

Keywords: 3D printing, microwave module, prototyping, additive technology, nanoink, microstrip sensors, microwave elements, multilayer printed circuit boards, radiophysical parameters, structural heterogeneity

• **Submitted:** 08.02.2023 • **Revised:** 20.04.2023 • **Accepted:** 10.07.2023

For citation: Vorunichev D.S., Kostin M.S. Investigation of the profilogram structure of microstrip microwave modules manufactured using additive 3D-printing technology. *Russ. Technol. J*. 2023;11(5):34–44. https://doi.org/10.32362/2500-316X-2023-11-5-34-44

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Исследование профилограммной структуры микрополосковых СВЧ-модулей, изготовленных по аддитивной технологии трехмерной печати

Д.С. Воруничев [@], М.С. Костин

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[@] Автор для переписки, e-mail: vorunichev@mirea.ru

Резюме

Цели. Целью работы является исследование шероховатости поверхности токонесущей топологии и диэлектрика верхней (Top Layer) и нижней (Bottom Layer) сторон СВЧ-модулей, изготовленных по аддитивной технологии трехмерной печати при прототипировании опытных образцов СВЧ-модулей на 3D-принтере многослойных печатных плат DragonFly 2020 LDM.

Методы. Использованы методы металлографического анализа в светлом и темном поле, профилографирование шероховатости поверхностей, компьютерное моделирование.

Результаты. Получены экспериментальные образцы микрополосковых СВЧ-элементов модулей многослойных плат заданной конфигурации, датчиков телеметрии, PCB-антенн (антенн на печатных платах). Исследованы топологические и радиофизические особенности аддитивно сформированных верхнего и нижнего поверхностных слоев экспериментальных образцов плат полосковых модулей. Проведены оптические профилограммные измерения шероховатости наружных сторон платы по 10 точкам, которые составили для верхнего слоя топологии – 2 мкм, для нижнего – 0.3 мкм, а также определен средний размер зерна диэлектрической основы – 0.007 мм^2. Показано, что шероховатость токопроводящей топологии и диэлектрика верхней стороны соответствует 6–7 классам точности. При этом шероховатость микрополосковой токопроводящей топологии и диэлектрика нижней стороны платы соответствует 10–12 классам точности.

Выводы. Установлено, что неравномерное формирование нижнего и верхнего полосковых слоев печатного модуля способно оказывать влияние на неоднородность распределения радиофизических параметров (диэлектрическую проницаемость, поверхностную проводимость и т.д.), а также на нестабильность конструктивных характеристик (адгезионной способности, теплопроводности и т.д.) полоскового модуля, что необходимо учитывать при прототипировании устройств по технологии струйной 3D-печати, в т.ч. при адаптации Gerber-проектов PCB-модулей, созданных под технологию классического производства плат.

Ключевые слова: 3D-печать, СВЧ-модуль, прототипирование, аддитивная технология, наночернила, микрополосковые датчики, СВЧ-элементы, многослойные печатные платы, радиофизические параметры, структурная неоднородность

• **Поступила:** 08.02.2023 • **Доработана:** 20.04.2023 • **Принята к опубликованию:** 10.07.2023

Для цитирования: Воруничев Д.С., Костин М.С. Исследование профилограммной структуры микрополосковых СВЧ-модулей, изготовленных по аддитивной технологии трехмерной печати. *Russ. Technol. J.* 2023;11(5):34–44. https://doi.org/10.32362/2500-316X-2023-11-5-34-44

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

Additive 3D-printing technologies using homogeneous materials are constructively applied in various branches of science and technology to reduce the time involved in the pre-production prototyping of industrial products. High-precision printing techniques using polymers or powder metals for the prototyping of enclosures, parts, and assembly units used in radioelectronic products have been in operation for over 15 years. Having by now been sufficiently researched and systematically worked out, such 3D printing technologies are widely used by various companies of technological equipment and composites presented on the industrial market [1]. However, in the radioelectronic sector, industrial technologies for printed electronics appeared only in 2019 due to the need to solve technological problems related to the development of special materials collectively known as nanoink comprising a solution of nanoparticles in a dispersed medium [2]. Additive two-component printing solution requires the combined use of nanoink (conductive and dielectric) offering compatible radiophysical properties to provide printing with the required resolution and uniformity in terms of the distribution of the specified parameters in the multilayer structure of the printed module. The successful development and creation of suitable nanoink with low sintering temperature characteristics has opened a new direction of additive printed electronics in the radioelectronic industry [2–6]. Today, the rapid development of printed electronics used to solve various prototyping problems including in the field of microwave technology comprises a new layer of technological solutions used in rapid production processes. The possibility of printing device prototypes on multilayer printed circuit boards directly in design centers bypasses the need for third-party contract manufacturers, reducing the time needed to prototype and debug new designs. Modified versions produced as part of product debugging cycles require the cycle to be repeated, often several times, until the final circuitry and device layout are released. However, printing prototype production files directly in the design center on a 3D printer takes only a few hours, after which the prototype can be assembled and tested for performance, followed by making all necessary changes to repeat the cycle. Due to ongoing research and innovation in the field, the possibilities of printed electronics are constantly expanding. In particular, printed microwave electronics technology offers more flexibility to designers, allowing printed elements such as capacitors, transformers, antennas, radio frequency identification (RFID) tags to be printed on various substrates including flexible at an accuracy of a few microns on the board [7, 8]. Additive 3D technology is additionally used to protect the original product design and intellectual property from reverse engineering prior to its serial production at the production-technological level of development protection[1], [2] [9].

The aim of the present work is an experimental study of the standard modes of additive technology of three-dimensional printing of microwave electronics products and elements using a DragonFly (Nano Dimension, Israel) 3D printer to achieve optimal results in terms of the resultant uniformity of conducting and dielectric layers, as well as surface roughness topology of multilayer printed circuit boards microwave modules, which is important for ensuring the stability of radio physical parameters.

1. PRINTED RADIOELECTRONICS USING THE DRAGONFLY 3D PRINTER

Research was performed using DragonFly LDM 2020 3D-printer operating in the Radioelectronic Technologies megalaboratory cluster of the Institute of Radio Electronics and Informatics of RTU MIREA. This equipment represents a “minifab” or minifactory (Fig. 1), comprising an automated full-cycle system for manufacturing electronic devices by the additive two-component ink-jet printing method. Printing is performed by two ink-jet heads (conductive and dielectric) with subsequent layer-by-layer curing: infrared for conductive ink and UV system for dielectric ink. Although the setup can be used to print up to 55 circuit board layers for testing in laboratory conditions, the working area of printing inside the machine room of the system represents the actual technological limitation. The main characteristics and technological standards of radio electronics printing on 3D printer DragonFly LDM 2020 are shown in Table 1, as well as those relating to the conductive and dielectric ink.

Prototype samples of the printed boards with microwave element topology for studying the surface structure of the outer sides (layers) carried out according to the technology of printed radioelectronics are shown in Fig. 2 and in Fig. 3.

[1] Khesin S. The DragonFly 3D printer is a revolution in the manufacture of multilayer printed circuit boards. *Vektor Vysokikh Tekhnologii = The Hi-Tech Vector Research and Practice Journal*. 2018;4(39):38–41 (in Russ.). https://ostec-group.ru/upload/iblock/3fd/hesin.pdf. Available February 08, 2023.

[2] DragonFly LDM. Inks User Guide NanoDimension. Ness Ziona: Nano Dimension Technologies Ltd. 2020. 52 p. https://www.nano-di.com/ame-dragonfly-ldm-2-0. Available February 08, 2023.

Table 1. Technological capabilities of the DragonFly LDM 2020 3D printer

Process parameter	Parameter value/characteristic
Conductor/gap	100 μm
Minimum layer thickness	17 μm
Dielectric droplet diameter	3 μm
Droplet diameter of a current-conducting ink	0.4 μm
Number of layers	Up to 55
Minimum diameter of the metallized through holes	400 μm
Minimum diameter of the non-metallized through holes	400 μm
Minimum diameter of the transition holes	<200 μm
Maximum board dimensions	160 × 160 × 3 mm
Board printing time	From 3 to 20 h
Maximum soldering temperature	165°C
Conductive ink	AgCite 90072 Silver Nanoparticle Conductive Ink (Nano Dimension, Israel)
Dielectric ink	Dielectric Ink 1092 – Dielectric UV Curable Acrylates Ink (Nano Dimension, Israel)

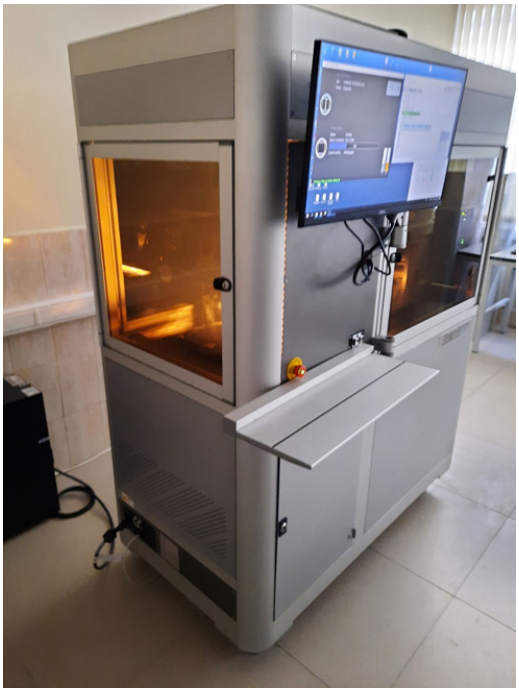


Fig. 1. DragonFly LDM 2020 3D printer operating in the Radioelectronic Technologies megalaboratory cluster at the Institute of Radio Electronics and Informatics of the RTU MIREA

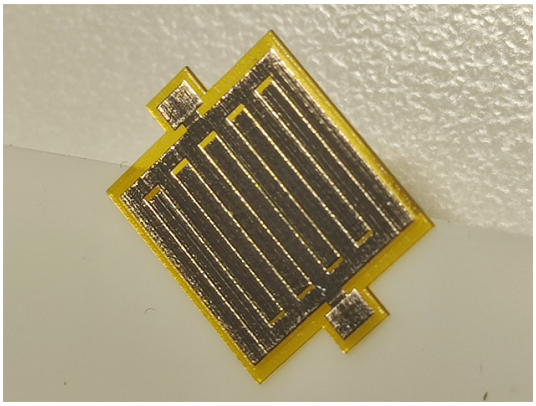


Fig. 2. Printed microwave sensor produced by RTU MIREA in the course of design research for Technopark of St. Petersburg

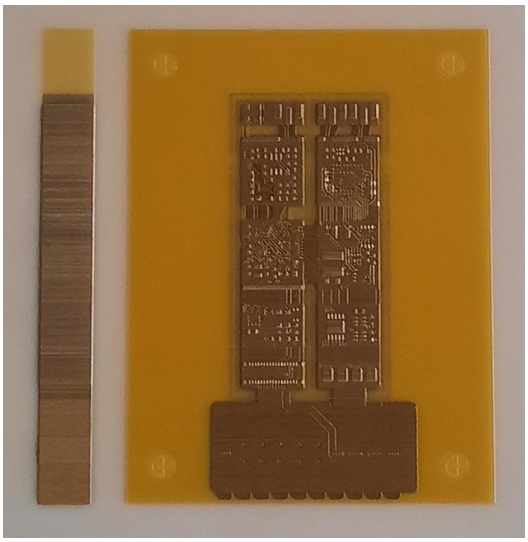


Fig. 3. Printed microwave module of a four-layer board for a global navigation satellite system manufactured by the RTU MIREA in the course of design work at the Technological Center Research and Production Complex (total thickness of the printed module – 0.123 mm)

As part of the technological cycle of additive 3D printing of a multilayer microstrip line board of a microwave module, its top layer (Top Layer side) and bottom layer (Bottom Layer side) are shown to be formed under different environmental conditions (contact/noncontact position of the sample topology formation). Unlike conventional multilayer printed circuit board technology having the same structure of copper traces and dielectric on both outer layers, the resultant surface roughness of the sides varies by an order of magnitude. This circumstance may affect design characteristics of the board (adhesion of the conductive layer to the dielectric, gas adsorption, residual mechanical stress, thermal properties, miscibility of solder on the adhesive conductive layer, etc.) and lead to the uneven distribution of its radiophysical parameters (surface resistivity, dielectric permittivity, impedance heterogeneity, signal delay, etc.), which must be taken into account when prototyping microwave modules implemented by technology of additive printed electronics.

Surface roughness, which determines the homogeneity of the structure of the conducting layer, is typically estimated by profile irregularities (Fig. 4), which are obtained by comparing the actual surface with a plane [5].

From the estimated parameters of the profilogram characteristic of the board roughness the following parameters are determined:

1. Average deviation of the profile R_{av}

$$R_{av} = \frac{1}{l}\int_0^l |\, y(x)\,|\, dx, \qquad (1)$$

where l is the base length; y is the profile deviation; x is the horizontal coordinate.

With the discrete method of profile processing, the parameter R_{av} is calculated by the formula:

$$R_{av} = \frac{1}{n}\sum_{i=1}^{n} |\, y_i\,|, \qquad (2)$$

where y_i are the measured profile deviations in discrete points; n is the number of measured discrete deviations on the base length.

2. Height of profile irregularities by ten points R_z.

$$R_z = \frac{\sum_{i=1}^{5} |\, y_{pi}\,| + \sum_{i=1}^{5} |\, y_{vi}\,|}{5}, \qquad (3)$$

where y_{pi} is the height of the ith highest protrusion of the profile; y_{vi} is the depth of the ith lowest hollow of the profile.

3. The greatest height of profile curvature R_{max}.
4. The average pitch of the profile curvature S_m.
5. Average pitch of the local protrusions S.

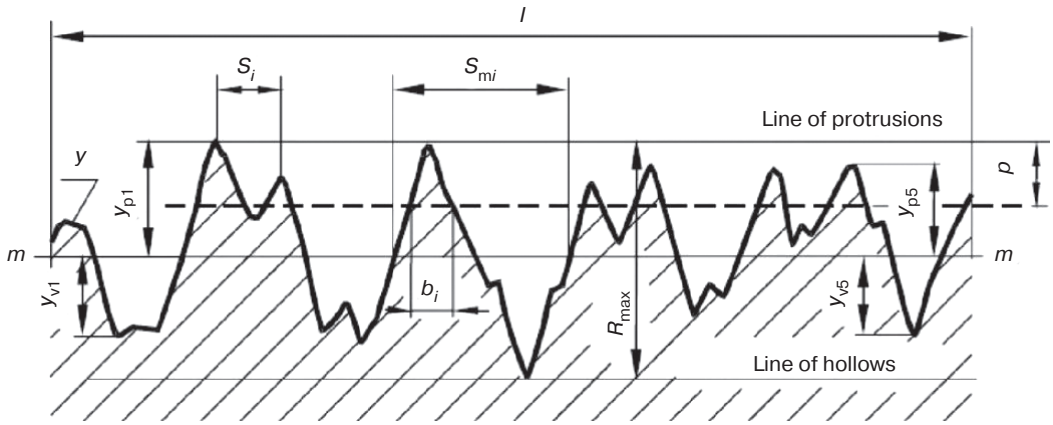


Fig. 4. Profilogram characteristics of the board roughness

6. Relative reference length of the profile t_p (where p are the values of the section level of the profile), determined by the formula:

$$t_p = \frac{1}{l}\left(\sum_{i=1}^{n} b_i\right)\cdot 100\%, \quad (4)$$

where $\sum_{i=1}^{n} b_i$ is the reference length of the profile; b_i is the components of the reference length.

When combined, the calculated formulas (1)–(4) allow the profilogram characteristics of the microstrip line structure of the microwave module to be further estimated.

2. RESEARCH OF THE PROFILOGRAM STRUCTURE OF THE OUTER LAYERS OF A MULTILAYER PRINTED CIRCUIT BOARD OF MICROWAVE MODULES

The top side (Top Layer) of a multilayer printed circuit board of a microwave module is formed by additive printing on the previous layer. The outer layer is printed with nanoink, which solidifies when coming into contact with the air environment. In this process mode, the top layer has a much higher surface roughness and granularity than the bottom layer. Figure 5 shows a two-dimensional microvision profilogram of the top layer of a prototype printed microwave module board (Fig. 3) with measurement of the grain size of the dielectric layer.

The surface structure study and measurements were carried out using a dark-field technique on a μVizo-MET-221 microimager (LOMO, Russia). The significant dielectric part grain area, which, as shown on the microvision image of the fragment, is on average 0.007 mm^2 or more, is caused by printing the layer with nanoparticles with their further baking by the curing system. The surface roughness was measured by changing the sharpness between the protrusion and the depression according to the two-dimensional profilogram obtained by the optical method from the microscope; here the focal difference between the sharpness is the depth of the measured groove. The measurements were carried out at ten points of the fragment of the upper side of the multilayer printed circuit board presented in the profilogram (Fig. 6). The measured roughness of the upper side is 2 μm.

The bottom layer is additively formed on a smooth single-use mylar substrate placed on a vacuum stage. Although other layers will be added to the underside during printing, only the bottom layer is directly physically printed on the smooth substrate, reducing the roughness by an order of magnitude compared to the top layer. Measurements were made at ten points on the roughness of the underside using a similar technique, and a profilogram is shown in Fig. 7. The measured roughness of the underside is 0.3 μm.

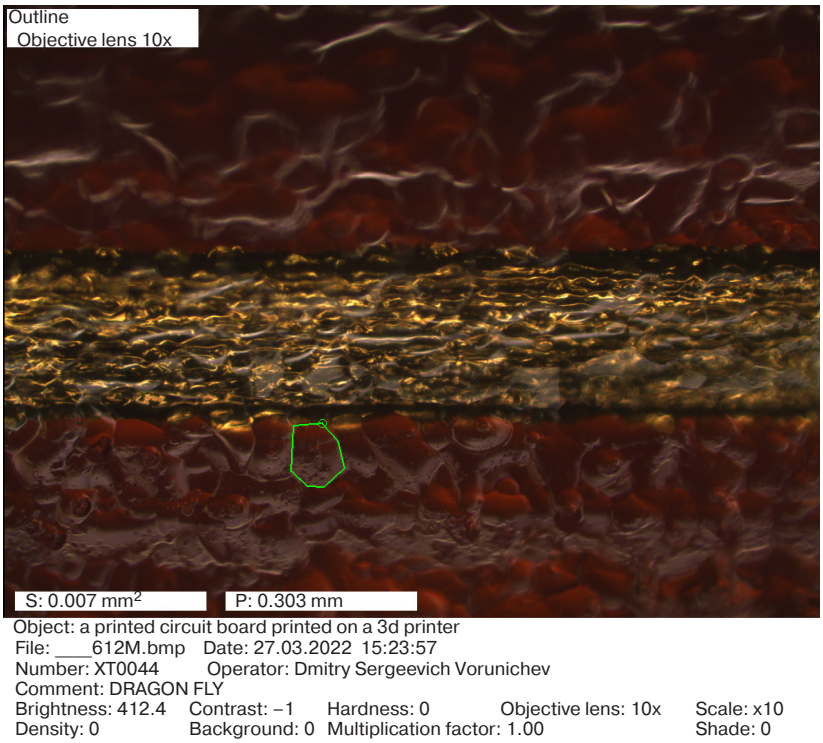


Fig. 5. Fragment of a printed circuit board of a microwave module prototype with superimposed measurement of dielectric grain size (in the center—conductor printed with conductive ink; the remainder—dielectric part printed with dielectric ink)

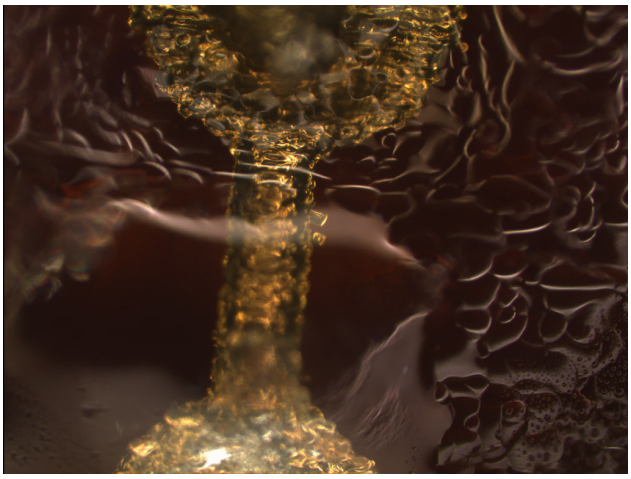


Fig. 6. Profilogram of the top side of the printed microwave module sample board

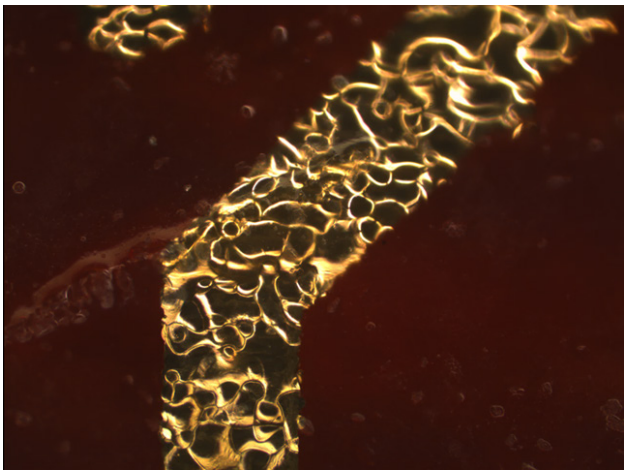


Fig. 7. Profilogram of the underside of the printed microwave module sample board

Using a software profilograph (Mitutoyo, Japan), we built profilograms of the top and undersides of the multilayer board of the microwave module sample for the conductive and dielectric structural layers. The results of fragmentary measurements of the microstrip line profilograms of the conductive topology and dielectric base of the prototype samples are shown in the Table 2.

The results of the profilograph measurements correspond to the optical measurements by the microvisor and are as follows:

- Top Layer: roughness of the microstrip line element (conductive layer of the printing topology) R_{av} = 1.6–2.1 (6–7 grade), dielectric roughness R_{av} = 1.7–1.8 (6–7 grades);
- Bottom Layer: roughness of the microstrip line element (conductive layer of the printing topology) R_{av} = 0.04–0.12 (10–12 grade), dielectric roughness R_{av} = 0.03–0.08 (11–12 grades).

When using an additive technology, the irregular profilogram formation of the lower and upper strip layers of a microwave module determines the regularity of the internal material structure, which can impact on the non-uniformity of the distribution of the radiophysical parameters (dielectric permittivity, surface conductivity, microstrip line impedance characteristic, signal delay, etc.)[3] [9–14], as well as on the stability of the structural characteristics of the strip module (adhesion of the conductive layer with the dielectric, gas adsorption, residual mechanical stress, thermal properties, solder miscibility on the adhesive conductive layer, etc.). This should be taken into account when prototyping devices

[3] DragonFly LDM. Inks User Guide NanoDimension. Ness Ziona: Nano Dimension Technologies Ltd. 2020. 52 p. https://www.nano-di.com/ame-dragonfly-ldm-2-0. Available February 08, 2023.

Table 2. Profilogram characteristics of the printing topology elements (fragments) of multilayer printed circuit board of the microwave module prototype

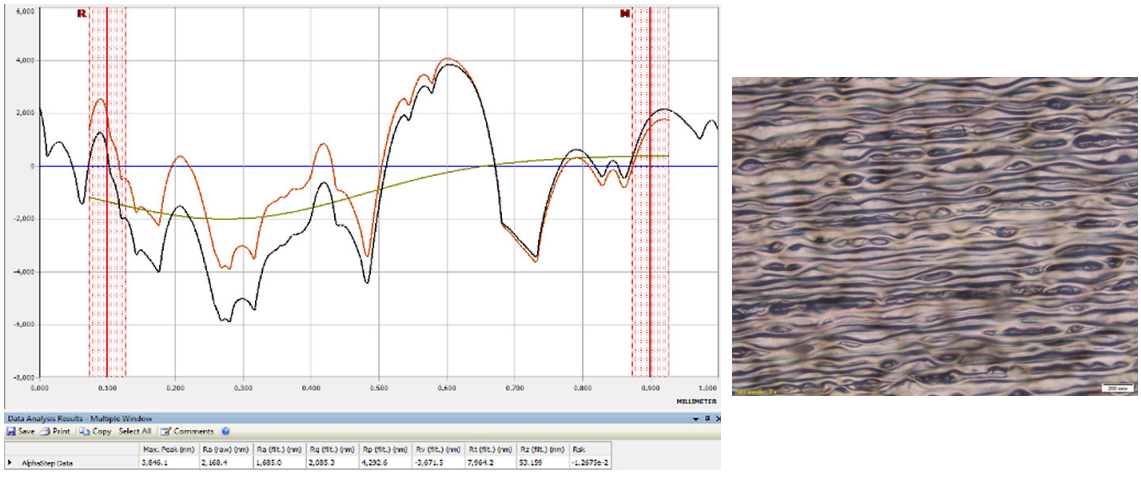
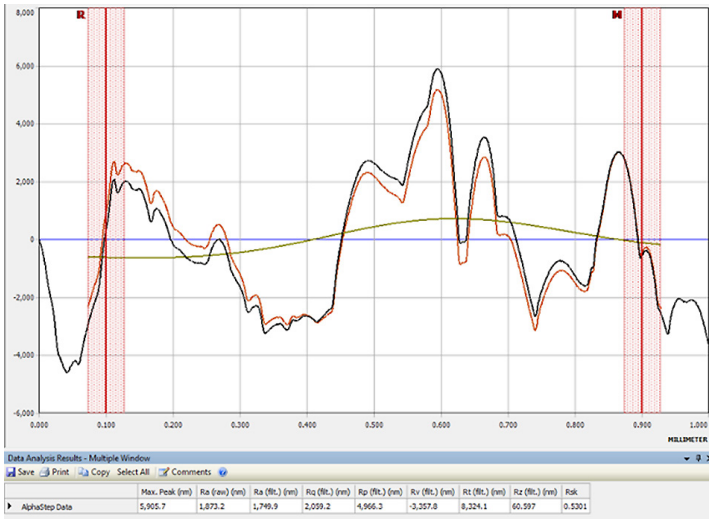
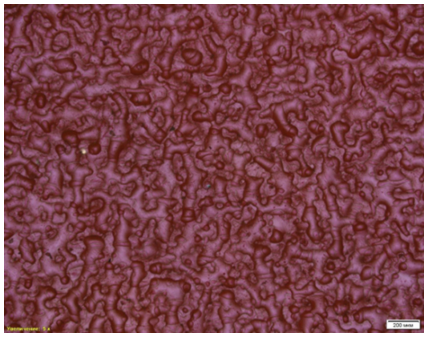
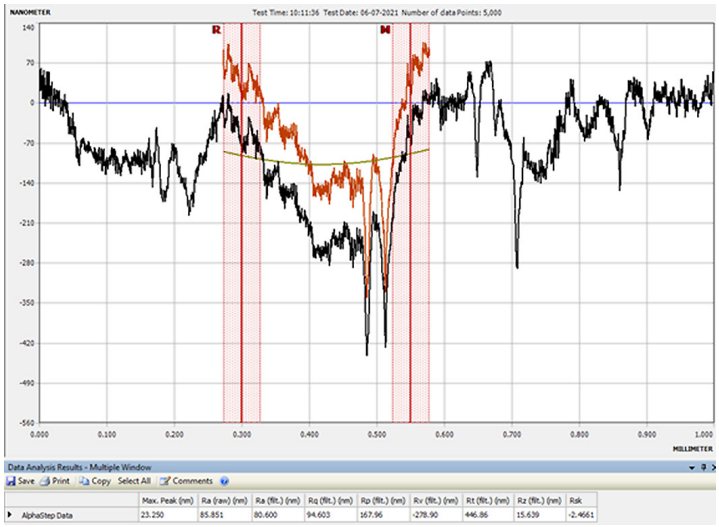
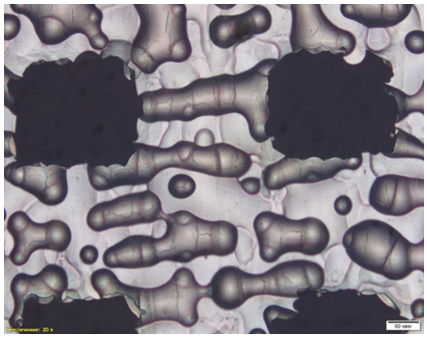
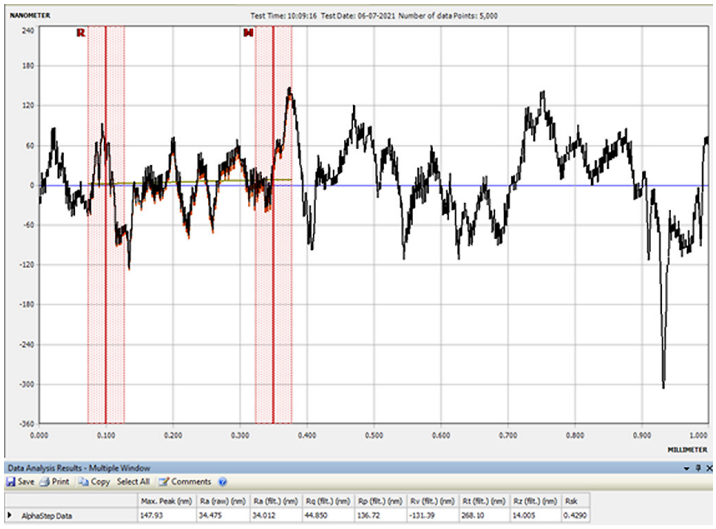
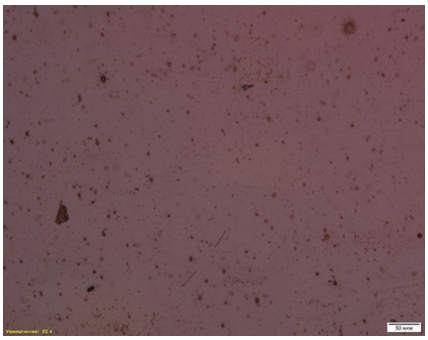
Upper side (Top Layer)	
Microstrip line element, multiple magnification ×5	

Table 2. Continued

Upper side (Top Layer)	
Dielectric, multiple magnification ×5	 
Underside (Bottom Layer)	
Microstrip line element, multiple magnification ×20	 
Underside (Bottom Layer)	
Dielectric, multiple magnification ×20	 

using additive 3D printing technology, including the adaptation of the Gerber-projects[4] of PCB-modules created for classic board production technology.

CONCLUSIONS

Topological and radiophysical features of additively formed upper and lower surface layers of experimental samples of strip-module boards were investigated using the DragonFly LDM 2020 3D printer. Optical profilogram measurements of the roughness of the outer sides of the board in the dark field were caried out on a metallographic microscope μVizo-MET-221 along with a profilograph used to construct the corresponding characteristics.

The conducted research of the prototypes showed a significant difference in surface roughness: in microvision measurements, upper side—2 μm, lower side—0.3 μm; in profilograph measurements, upper side—1.8 μm, lower side—0.1 μm. This parameter can influence the structural, technological, and electrical characteristics of the product. The average grain size of the dielectric part—0.007 mm^2—was determined along with the reasons for the difference in the roughness of the sides.

The results of the implemented research into side roughness and profilograms of the multilayer printed circuit boards can be used in the design of microstrip line products manufactured by printed electronics, as well as to develop solutions for improving the technology in order to compare the characteristics of the vector analysis of structural heterogeneity with the topology of prototypes of microwave modules.

Authors' contribution. All authors equally contributed to the research work.

[4] Gerber-format files for PCB modules production.

REFERENCES

1. Kharalgin S.V., Kulikov G.V., Kotelnikov A.B., Snastin M.V., Dobychina E.M. Prototyping of microwave devices with specified electrodynamic characteristics using additive 3D printing technology. *Russ. Technol. J.* 2019;7(1):80–101 (in Russ.). https://doi.org/10.32362/2500-316X-2019-7-1-80-101
2. Sokol D., Yamada M., Nulman J. Design and Performance of Additively Manufactured In-Circuit Board Planar Capacitors. *IEEE Transactions on Electron Devices*. 2021;68(11): 5747–5752. https://doi.org/10.1109/TED.2021.3117934
3. Li M., Yang Y., Nulman J., Yamada M., Iacopi F. Unique multi-level metal layer electronics solutions offered by advanced 3D printing. In: *2022 6th IEEE Electron Devices Technology & Manufacturing Conference* (*EDTM*). 2022. P. 144–144. https://doi.org/10.1109/EDTM53872.2022.9798133
4. de Marzo G., Mastronardi V.M., Algieri L., et al. Sustainable, Flexible, and Biocompatible Enhanced Piezoelectric Chitosan Thin Film for Compliant Piezosensors for Human Health. *Advanced Electronic Materials*. 2022. https://doi.org/10.1002/aelm.202200069
5. Wu S.-Y., Yang C., Hsu W., Lin L. 3D-printed microelectronics for integrated circuitry and passive wireless sensors. *Microsyst. Nanoeng*. 2015;1(1):15013. https://doi.org/10.1038/micronano.2015.13
6. Li M., Yang Y., Iacopi F., Yamada M., Nulman J., et al. Compact multilayer bandpass filter using low-temperature additively manufacturing solution. *IEEE Trans. Electron Devices*. 2021;68(7):3163–3169. https://doi.org/10.1109/TED.2021.3072926
7. Li Y., Ge L., Wang J., et al. A *Ka*-band 3-D-printed wideband stepped waveguide-fed magnetoelectric dipole antenna array. *IEEE Trans. Antennas Propag*. 2020;68(4): 2724–2735. https://doi.org/10.1109/TAP.2019.2950868

СПИСОК ЛИТЕРАТУРЫ

1. Харалгин С.В., Куликов Г.В., Котельников А.Б., Снастин М.В., Добычина Е.М. Прототипирование СВЧ-устройств с заданными электродинамическими характеристиками по технологии аддитивной 3D-печати. *Russ. Technol. J*. 2019;7(1):80–101. https://doi.org/10.32362/2500-316X-2019-7-1-80-101
2. Sokol D., Yamada M., Nulman J. Design and Performance of Additively Manufactured In-Circuit Board Planar Capacitors. *IEEE Transactions on Electron Devices*. 2021;68(11): 5747–5752. https://doi.org/10.1109/TED.2021.3117934
3. Li M., Yang Y., Nulman J., Yamada M., Iacopi F. Unique multi-level metal layer electronics solutions offered by advanced 3D printing. In: *2022 6th IEEE Electron Devices Technology & Manufacturing Conference* (*EDTM*). 2022. P. 144–144. https://doi.org/10.1109/EDTM53872.2022.9798133
4. de Marzo G., Mastronardi V.M., Algieri L., et al. Sustainable, Flexible, and Biocompatible Enhanced Piezoelectric Chitosan Thin Film for Compliant Piezosensors for Human Health. *Advanced Electronic Materials*. 2022. https://doi.org/10.1002/aelm.202200069
5. Wu S.-Y., Yang C., Hsu W., Lin L. 3D-printed microelectronics for integrated circuitry and passive wireless sensors. *Microsyst. Nanoeng*. 2015;1(1):15013. https://doi.org/10.1038/micronano.2015.13
6. Li M., Yang Y., Iacopi F., Yamada M., Nulman J., et al. Compact multilayer bandpass filter using low-temperature additively manufacturing solution. *IEEE Trans. Electron Devices*. 2021;68(7):3163–3169. https://doi.org/10.1109/TED.2021.3072926
7. Li Y., Ge L., Wang J., et al. A *Ka*-band 3-D-printed wideband stepped waveguide fed magnetoelectric dipole antenna array. *IEEE Trans. Antennas Propag*. 2020;68(4): 2724–2735. https://doi.org/10.1109/TAP.2019.2950868

8. Kornilov D.Yu., Tkachev S.V., Zaitsev E.V., Kim V.P., Kushnir A.E. Printer technologies in electronics. Materials and Devices for Printing – the First Russian Seminar (Moscow, December 15, 2017). *RENSIT*. 2017;9(2):181–204. https://doi.org/10.17725/rensit.2017.09.181
9. Vorunichev D.S., Vorunicheva K.Yu. Current capabilities of prototyping technologies for multilayer printed circuit boards on a 3D printer. *Russ. Technol. J*. 2021;9(4):28–37 (in Russ.). https://doi.org/10.32362/2500-316X-2021-9-4-28-37
10. Huang G.-L., Han C.-Z., Xu W., Yuan T., Zhang X. A compact 16-way high-power combiner implemented via 3-D metal printing technique for advanced radio-frequency electronics system applications. *IEEE Trans. Ind. Electron*. 2019;66(6):4767–4776. https://doi.org/10.1109/TIE.2018.2863219
11. Vorunichev D.S. Application of Additive Technology for 3D Printing of Electronic Devices as a Way to Reduce Prototyping Time. In: *Proceedings of the 2021 IEEE International Conference on Quality Management, Transport and Information Security, Information Technologies* (*IT&QM&IS*). 2021. P. 480–483. https://doi.org/10.1109/ITQMIS53292.2021.9642806
12. Vorunichev D.S. Prototyping Electronic Devices on a Dragonfly 3D Printer as a Means of Preserving Original Developments inside the Design Center. In: *Proceedings of the 2021 IEEE International Conference on Quality Management, Transport and Information Security, Information Technologies* (*IT&QM&IS*). 2021. P. 472–475. https://doi.org/10.1109/ITQMIS53292.2021.9642911
13. Kostin M.S., Yarlykov A.D. Radiowave technology of resonant gas-sensor microwave telemetry. *Russ. Technol. J*. 2021;9(1):18–28 (in Russ.). https://doi.org/10.32362/2500-316X-2021-9-1-18-28
14. Li M., Yang Y., Iacopi F., Nulman J., Chappel-Ram S. 3D-printed low-profile single-substrate multi-metal layer antennas and array with bandwidth enhancement. *IEEE Access*. 2020;8:217370–217379. https://doi.org/10.1109/ACCESS.2020.3041232

8. Корнилов Д.Ю., Ткачев С.В., Зайцев Е.В., Ким В.П., Кушнир А.Е. Принтерные технологии в электронике. Материалы и устройства для печати – первый российский семинар (Москва, 15.12.2017). *РЭНСИТ*. 2017;9(2):181–204.
9. Воруничев Д.С., Воруничева К.Ю. Текущие возможности технологии прототипирования многослойных печатных плат на 3D-принтере. *Russ. Technol. J*. 2021;9(4):28–37. https://doi.org/10.32362/2500-316X-2021-9-4-28-37
10. Huang G.-L., Han C.-Z., Xu W., Yuan T., Zhang X. A compact 16-way high-power combiner implemented via 3-D metal printing technique for advanced radio-frequency electronics system applications. *IEEE Trans. Ind. Electron*. 2019;66(6):4767–4776. https://doi.org/10.1109/TIE.2018.2863219
11. Vorunichev D.S. Application of Additive Technology for 3D Printing of Electronic Devices as a Way to Reduce Prototyping Time. In: *Proceedings of the 2021 IEEE International Conference on Quality Management, Transport and Information Security, Information Technologies* (*IT&QM&IS*). 2021. P. 480–483. https://doi.org/10.1109/ITQMIS53292.2021.9642806
12. Vorunichev D.S. Prototyping Electronic Devices on a Dragonfly 3D Printer as a Means of Preserving Original Developments inside the Design Center. In: *Proceedings of the 2021 IEEE International Conference on Quality Management, Transport and Information Security, Information Technologies* (*IT&QM&IS*). 2021. P. 472–475. https://doi.org/10.1109/ITQMIS53292.2021.9642911
13. Костин М.С., Ярлыков А.Д. Радиоволновая технология резонансной газосенсорной СВЧ-телеметрии. *Russ. Technol. J*. 2021;9(1):18–28. https://doi.org/10.32362/2500-316X-2021-9-1-18-28
14. Li M., Yang Y., Iacopi F., Nulman J., Chappel-Ram S. 3D-printed low-profile single-substrate multi-metal layer antennas and array with bandwidth enhancement. *IEEE Access*. 2020;8:217370–217379. https://doi.org/10.1109/ACCESS.2020.3041232

About the authors

Dmitry S. Vorunichev, Deputy Director of the Institute of Radio Electronics and Informatics; Lecturer of the Department of Design and Production of Radioelectronic Means, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: vorunichev@mirea.ru. Scopus Author ID 57204939440, RSCI SPIN-code 5559-6656, http://orcid.org/0000-0003-0512-7572

Mikhail S. Kostin, Dr. Sci. (Eng.), Head of the Department of Radio Wave Processes and Technologies, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: kostin_m@mirea.ru. Scopus Author ID 57208434671, RSCI SPIN-code 5819-2178, http://orcid.org/0000-0002-5232-5478

Об авторах

Воруничев Дмитрий Сергеевич, заместитель директора Института радиоэлектроники и информатики; преподаватель, кафедра конструирования и производства радиоэлектронных средств Института радиоэлектроники и информатики ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: vorunichev@mirea.ru. Scopus Author ID 57204939440, SPIN-код РИНЦ 5559-6656, http://orcid.org/0000-0003-0512-7572

Костин Михаил Сергеевич, д.т.н., доцент, заведующий кафедрой радиоволновых процессов и технологий Института радиоэлектроники и информатики ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: kostin_m@mirea.ru. Scopus Author ID 57208434671, SPIN-код РИНЦ 5819-2178, http://orcid.org/0000-0002-5232-5478

Translated from Russian into English by Lyudmila O. Bychkova
Edited for English language and spelling by Thomas A. Beavitt

Modern radio engineering and telecommunication systems

Современные радиотехнические и телекоммуникационные системы

UDC 621.396.49
https://doi.org/10.32362/2500-316X-2023-11-5-45-53

RESEARCH ARTICLE

Evaluation of the effectiveness of sliding redundancy of radioelectronic facilities

Tatyana E. Gelfman [@], **Alexey P. Pirkhavka**

MIREA – Russian Technological University, Moscow, 119454 Russia
[@] *Corresponding author, e-mail: gelfman@mirea.ru*

Abstract

Objectives. The increased reliability of radioelectronic facilities can be achieved by the application of structural and load redundancy. Structural redundancy is achieved taking into account multiplicity of redundancy and the intensity of failures of elements of radioelectronic facilities, while load redundancy involves an easing of electrical, thermal, and mechanical operating modes of the elements. The choice of a redundancy method is determined according to reliability indicator requirements, which may often be contradictory. Therefore, the problem of how to effectively combine structural redundancy and load redundancy methods is very topical. In long-life radioelectronic facilities, for example, in satellite communication repeater systems, sliding redundancy is applied when limiting mass-dimensional parameters and consequently consumed energy. The aim of the work is to evaluate the efficiency of sliding redundancy according to various reliability indicators when altering redundancy multiplicity, reserve operating mode, element failure intensity, and switching device type.

Methods. To describe the structure of a complex sliding redundancy system, a logical-probabilistic method is used, in which the dependence of the system reliability indicators on the reliability indicators of the elements is formulated as a logical function of operability. Graph-analytical methods are used to compare different variants of reliability logic schemes.

Results. Mathematical models have been obtained to evaluate the effectiveness of sliding reservation. A comparative analysis of the efficiency of sliding redundancy with a loaded and unloaded reserve was carried out in terms failure-free operation probability, as well as gamma-percentage resource, failure rate when changing the fractional multiplicity of the redundancy, and element failure rate. The influence of the reliability of the switching device on the efficiency of the sliding redundancy is considered.

Conclusions. Practical recommendations on the selection of the redundancy mode are presented according to different reliability indices and constructed mathematical models of the sliding redundancy efficiency coefficients. The correlation between the reliability indices of elements and the switching device whose reliability can be discounted, is determined. To increase the efficiency of sliding redundancy of radioelectronic facilities, it is necessary to combine multiplicity of redundancy and the operating mode of the reserve with approaches aimed at reducing the intensity of failure of elements.

Keywords: reliability, radioelectronic facilities, probability of failure-free operation, gamma-percentage resource, failure rate, sliding redundancy

• **Submitted:** 23.03.2023 • **Revised:** 27.04.2023 • **Accepted:** 06.07.2023

For citation: Gelfman T.E., Pirkhavka A.P. Evaluation of the effectiveness of sliding redundancy of radioelectronic facilities. *Russ. Technol. J*. 2023;11(5):45–53. https://doi.org/10.32362/2500-316X-2023-11-5-45-53

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Оценка эффективности скользящего резервирования радиоэлектронных средств

Т.Э. Гельфман [@], **А.П. Пирхавка**

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[@] *Автор для переписки, e-mail: gelfman@mirea.ru*

Резюме

Цели. Повышение надежности радиоэлектронных средств достигается применением структурного и нагрузочного резервирования. Эффективность структурного резервирования зависит от кратности резервирования и от интенсивности отказов элементов радиоэлектронных средств. При нагрузочном резервировании путем облегчения электрических, тепловых и механических режимов работы элементов можно снизить их интенсивности отказов. Выбор способа резервирования определяется требованиями к показателям безотказности, которые часто находятся в противоречии. Поэтому весьма актуальной является проблема эффективного сочетания методов структурного резервирования и способов нагрузочного резервирования. В радиоэлектронных средствах длительного срока действия, например, в ретрансляторах спутниковых систем связи, при ограничении на массогабаритные параметры и потребляемую энергию применяется скользящее резервирование. Цель работы – оценка эффективности скользящего резервирования по различным показателям надежности при изменении кратности резервирования, режима работы резерва, интенсивности отказов элементов и переключающего устройства.

Методы. Для описания структуры сложной системы скользящего резервирования используется логико-вероятностный метод, в котором зависимость показателей надежности системы от показателей надежности элементов формулируется в виде логической функции работоспособности. Для сравнения различных вариантов логических схем надежности применяются графоаналитические методы.

Результаты. Получены математические модели для оценки эффективности скользящего резервирования. Проведен сравнительный анализ эффективности скользящего резервирования с нагруженным и ненагруженным резервом по вероятности безотказной работы, по гамма-процентному ресурсу, по интенсивности отказов при изменении дробной кратности резервирования и интенсивности отказов элементов. Исследовано влияние надежности переключающего устройства на эффективность скользящего резервирования.

Выводы. Построенные математические модели коэффициентов эффективности скользящего резервирования по разным показателям надежности позволяют дать практические рекомендации по выбору режима резерва. Определено соотношение показателей безотказности элементов и переключающего устройства, при котором надежностью переключающего устройства можно пренебречь. Для повышения эффективности скользящего резервирования радиоэлектронных средств необходимо сочетать кратность резервирования, режим работы резерва и способы уменьшения интенсивности отказов элементов.

Ключевые слова: надежность, радиоэлектронное средство, вероятность безотказной работы, гамма-процентный ресурс, интенсивность отказов, скользящее резервирование

• **Поступила:** 23.03.2023 • **Доработана:** 27.04.2023 • **Принята к опубликованию:** 06.07.2023

Для цитирования: Гельфман Т.Э., Пирхавка А.П. Оценка эффективности скользящего резервирования радиоэлектронных средств. *Russ. Technol. J.* 2023;11(5):45–53. https://doi.org/10.32362/2500-316X-2023-11-5-45-53

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

When ensuring the reliability of radioelectronic facilities (REF), a contradiction arises between the need to combine increased system complexity with limited increases in the reliability of the element base, which justifies the search for new solutions. Such indicators as efficiency, durability, availability, survivability, and safety depend on the reliability of complex REFs. Thus, in order to achieve high indicators of REF reliability, requirements according to many indicators must be should satisfied: probability of no-failure operation, average operating time between failure, gamma-percentage resource, intensity of failures, etc. For example, to achieve a given REF efficiency it is necessary to guarantee a certain value of probability of no-failure operation, and to provide durability according to a certain gamma-percentage resource value [1, 2].

REF efficiency and reliability indicators have much in common. Many problems related to justification of reliability requirements and the development of reliability assurance programs, as well as the selection of maintenance, control and operation systems, can be considered in terms of efficiency research problems. Examples of such problems include the following:

- determination of reasonable development time;
- selecting the optimal range of REF;
- selecting the best combination of REF design parameters;
- selecting from several logical reliability diagrams the diagram which provides operability at the highest number of failures of any of its elements;
- comparison of redundancy methods and construction topologies.

Assessment of efficiency of redundant REF is typically carried out by comparing redundancy types for different reliability indicators. For this purpose, a coefficient of efficiency [1] is introduced, which shows which type of redundancy is more effective according to the investigated reliability index.

All reliability indicators or criteria are connected by single-valued mathematical models, since the desire to satisfy several criteria at once often leads to contradictory requirements, including excessive redundancy, depending on the redundancy multiplicity. For this reason, the choice of a reliability assurance method essentially depends on the criterion used. Therefore, structural redundancy is used to provide probability of no-failure operation [2–6], while load redundancy is used to provide mean time between failures of the long-lasting system [1, 3, 7]. In practice, it is often necessary to combine structural and load redundancy the methods.

In complex systems, various kinds of failures occur [1, 8], all having a random character. In calculations, the independence of these failures is assumed. Their influence can be estimated by different mathematical models: for example, for sudden failures of elements at a constant failure rate, the exponential model is used, while the normal distribution is used to analyze the influence of gradual failures. However, gradual failures due to ageing change the probability of occurrence of sudden failures and failures, complicating the analysis of REF reliability [9, 10].

The high efficiency of satellite communication networks [11–13] is due to the creation of repeaters when using them in sliding redundancy, significantly increasing reliability along with a relatively small gain in weight, size, and energy consumption [3]. One, two, or more redundant elements can be used to ensure the redundancy of individual elements, each of which can be connected in place of any of the main elements. For example, in the sliding redundancy of onboard transmitters or their power amplifiers, the reserve can be loaded and unloaded [14, 15]. For modern geostationary communication satellites, the multiplicity of sliding redundancy of transmitters (ratio of the number of working and redundant transmitters) can vary from 1/5 to 1/2[1] [4].

When using loaded or hot redundancy, there is no need to activate or allow the redundant device to operate in active loaded standby mode, as would be the case unloaded or cold redundancy. For this reason, sliding redundancy with loaded redundancy increases system availability and responsiveness, but also increases operating costs

[1] Dinges S.I., Ivanyushkin R.Yu., Kozyrev V.B., et al. *Radio transmitting devices.* Textbook for universities. Ivanyushkin R.Yu. (Ed.). Moscow: Goryachaya liniya – Telekom; 2021. 1150 p. (in Russ.). ISBN 978-5-9912-0774-4.

because such a structure requires additional power and a more complex switching device. It is noted in [4, 5, 15] that in non-loaded reserve substitution redundancy, the automatic control and switching machine reduces the probability of fault-free operation of the redundant system, but no recommendations are given for the selection of the fault-free performance of the switching device. Therefore, the issue of increasing the reliability and efficiency of sliding redundancy, closely related to the choice of redundancy multiplicity and reserve operation mode, and with the provision of faultlessness indicators of the switching device, which do not reduce the reliability of the system as a whole, is relevant.

FAILURE PROBABILITY AND GAMMA-PERCENT LIFETIME OF SLIDING REDUNDANCY DIAGRAMS

The structural diagram of sliding redundancy reliability with multiplicity m/n is shown in Fig. 1.

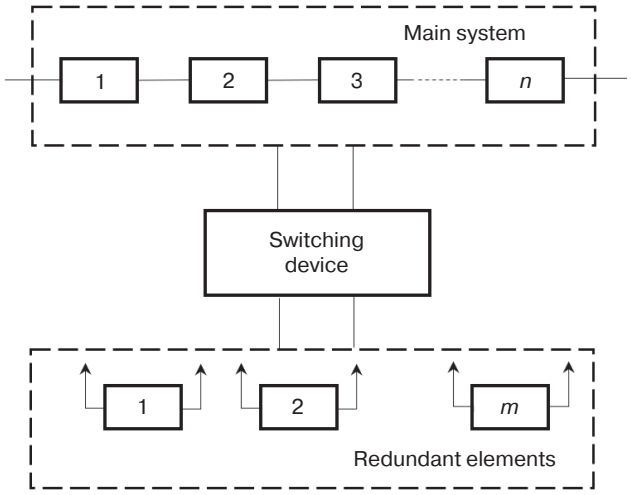


Fig. 1. Structural diagram of sliding redundancy reliability with multiplicity m/n

For a diagram with unloaded redundancy and a redundancy factor of $1/n$, the dependence of probability of no-failure operation on time is described by the expression[2]:

$$P_1(t) = \left(n\frac{\lambda}{\lambda_k} + 1 \right) e^{-n\lambda t} - n\frac{\lambda}{\lambda_k} e^{-(\lambda_k + n\lambda)t}, \quad (1)$$

where λ is the failure rate of the main and redundant elements; n is the number of elements of the main system; λ_k is the failure rate of the switching device.

When the redundancy multiplicity is $1/n$, the expression for the time dependence of the probability of faultless operation of a sliding redundancy diagram with loaded redundancy can be obtained without taking into account the reliability of the switching device [1]:

$$P_2(t) = e^{-n\lambda t}\left[1 + n\left(1 - e^{-\lambda t}\right)\right]. \quad (2)$$

The expression for the probability of failure-free operation of a sliding redundancy diagram with a redundancy factor of 2/3 can also be obtained without taking into account the reliability of the switching device [1]:

- for an unloaded redundancy diagram

$$P_3(t) = e^{-3\lambda t}\left(1 + 3\lambda t + 4.5\lambda^2 t^2\right), \quad (3)$$

- for a loaded redundancy diagram

$$P_4(t) = e^{-3\lambda t}\left(10 + 6e^{-2\lambda t} - 15e^{\lambda t}\right). \quad (4)$$

Figure 2 illustrates dependencies (1) and (2) for $n = 2$ and $n = 4$ at different failure intensities of the main and redundant elements. For a diagram with unloaded reserve, a switching device failure rate $\lambda_k = 10^{-8}$ h^{-1} is assumed when constructing the dependencies of Fig. 2.

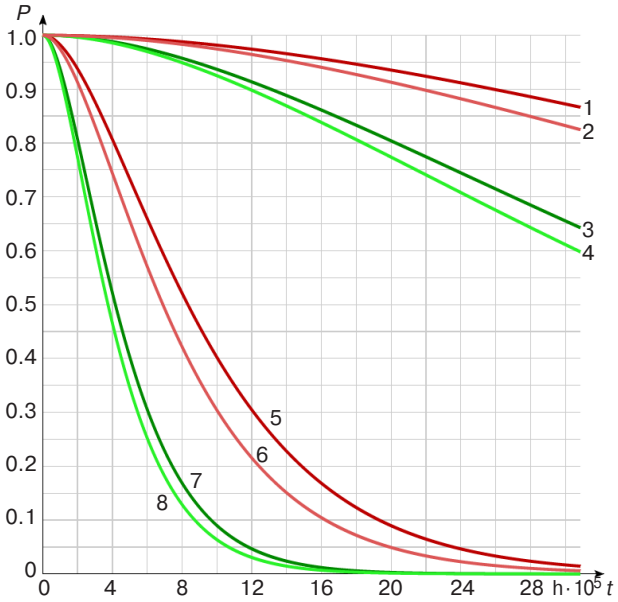


Fig. 2. Time dependencies of the probability of failure-free operation of the sliding redundancy diagram: 1, 2, 3, 4: $\lambda = 10^{-7}$ h^{-1}; 5, 6, 7, 8: $\lambda = 10^{-6}$ h^{-1}; 1, 3, 5, 7: unloaded redundancy; 2, 4, 6, 8: loaded redundancy; 1, 2, 5, 6: redundancy multiplicity 1/2; 3, 4, 7, 8: redundancy multiplicity 1/4

According to the graphs depicted in Fig. 2, unloaded redundancy provides a higher probability of no-failure operation, even when taking into account the reliability of the switching device. For example, for the unloaded reserve with $\lambda = 10^{-6}$ h^{-1} and $\lambda_k = 10^{-8}$ h^{-1} at redundancy

[2] Yanshin A.A. *Theoretical foundations of EVA design, technology and reliability*: textbook for universities. Moscow: Radio i svyaz'; 1983. 128 p. (in Russ.).

multiplicities 1/4 and 1/2, the probability of non-failure operation for time $4 \cdot 10^5$ h is respectively 0.52 and 0.8. In the case of application of the loaded redundancy, the corresponding probability values of no-failure operation are 0.46 and 0.74, respectively.

From the graphs given in Fig. 2, we can determine the index of durability-gamma percent resource [2] of the redundancy diagrams. At use of elements with intensity of failures $\lambda = 10^{-7}$ h^{-1}, the gamma-percent resource of the scheme with unloaded redundancy is $1.3 \cdot 10^6$ h at multiplicity of redundancy 1/4 and $2.6 \cdot 10^6$ h at multiplicity of redundancy 1/2. The corresponding values of gamma-percent resource of the diagram with loaded redundancy are $1.2 \cdot 10^6$ and $2.2 \cdot 10^6$ h, respectively. In other words, the diagram with unloaded redundancy is more durable.

It also follows from the dependencies of Fig. 2 that the gamma-percent resource of any of the diagrams under consideration deteriorates by an order of magnitude when the failure rate of the main and redundant elements increases to $\lambda = 10^{-6}$ h^{-1}.

Figure 3 shows the graphs of time dependencies of no-failure operation probability of the diagram with unloaded redundancy having multiplicity of redundancy 1/4 and 1/2, $\lambda = 10^{-7}$ h^{-1} at different failure rates of the switching device constructed in accordance with expression (2).

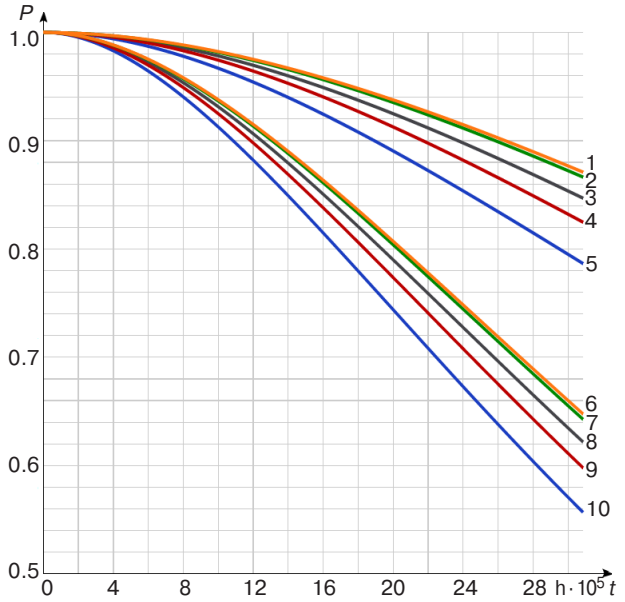


Fig. 3. Time dependencies of the probability of failure-free operation of the diagram with unloaded redundancy: 1, 2, 3, 4, 5: redundancy multiplicity 1/2; 6, 7, 8, 9, 10: redundancy multiplicity 1/4; 1, 6: $\lambda_k = 10^{-9}$ h^{-1}; 2, 7: $\lambda_k = 10^{-8}$ h^{-1}; 3, 8: $\lambda_k = 5 \cdot 10^{-8}$ h^{-1}; 4, 9: $\lambda_k = 10^{-7}$ h^{-1}; 5, 10: $\lambda_k = 2 \cdot 10^{-7}$ h^{-1}

The effect of switching device reliability on the durability of a sliding redundancy diagram with unloaded redundancy can be estimated on the basis of the graphs given in Fig. 3. For example, when a switching device with failure rates of $2 \cdot 10^{-7}$ h^{-1}, $5 \cdot 10^{-8}$ h^{-1}, and 10^{-8} h^{-1}, respectively, is used in a scheme with 1/4 redundancy multiplicity, the gamma percent resource is $1.1 \cdot 10^6$, $1.24 \cdot 10^6$, and $1.3 \cdot 10^6$ h, respectively. The corresponding values of gamma percent resource for a scheme with 1/2 redundancy multiplicity are $1.9 \cdot 10^6$, $2.4 \cdot 10^6$, and $2.6 \cdot 10^6$ h, respectively. It follows from the closeness of curves 1 and 2, as well as 6 and 7, that the reliability of the switching device has almost no effect on the durability of the unloaded sliding redundancy diagram at a ratio $\lambda/\lambda_k > 10$.

REDUNDANCY EFFICIENCY COEFFICIENT BY PROBABILITY OF FAILURE-FREE OPERATION

Let us analyze the effect of the state of redundancy in sliding redundancy on the efficiency of redundancy using the coefficient of redundancy efficiency K_p on the probability of the failure-free operation:

$$K_p(t) = \frac{P_1(t)}{P_2(t)}.$$

For a redundancy multiplicity of 1/2 according to (1) and (2), the efficiency factor is defined by the expression:

$$K_{p\,1/2}(t) = \frac{2\frac{\lambda}{\lambda_k}\left(1 - e^{-\lambda_k t}\right) + 1}{3 - 2e^{-\lambda t}}.$$

For a redundancy multiplicity of 1/4, the formula for the efficiency factor is as follows:

$$K_{p\,1/4}(t) = \frac{4\frac{\lambda}{\lambda_k}\left(1 - e^{-\lambda_k t}\right) + 1}{5 - 4e^{-\lambda t}}.$$

For a redundancy multiplicity of 2/3 without taking into account the reliability of the switching device, the formula for the coefficient of effectiveness by the probability of failure-free operation according to (3) and (4) is as follows:

$$K_{p\,2/3}(t) = \frac{1 + 3\lambda t + 4.5\lambda^2 t^2}{10 + 6e^{-2\lambda t} - 15e^{-\lambda t}}.$$

Figure 4 illustrates the time dependencies of the redundancy effectiveness coefficient for a switching device failure rate 10^{-8} h^{-1}, redundancy multiplicities 1/2 and 1/4, and different element failure rates. Figure 5 depicts graphs of the dependence on time

of the redundancy effectiveness coefficient for a redundancy multiplicity of 2/3 at different element failure intensities.

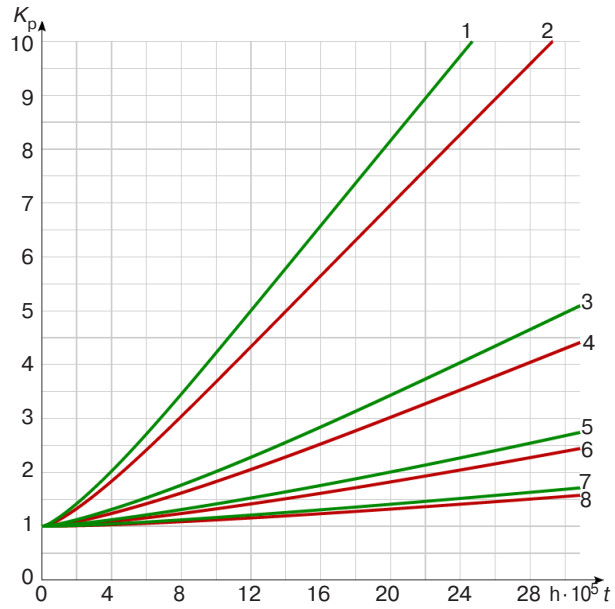


Fig. 4. Time dependencies of the redundancy efficiency coefficient on the probability of failure-free operation: 1, 2: $\lambda = 5 \cdot 10^{-6}$ h^{-1}; 3, 4: $\lambda = 2 \cdot 10^{-6}$ h^{-1}; 5, 6: $\lambda = 10^{-6}$ h^{-1}; 7, 8: $\lambda = 5 \cdot 10^{-7}$ h^{-1}; 1, 3, 5, 7: redundancy multiplicity 1/4; 2, 4, 6, 8: redundancy multiplicity 1/2

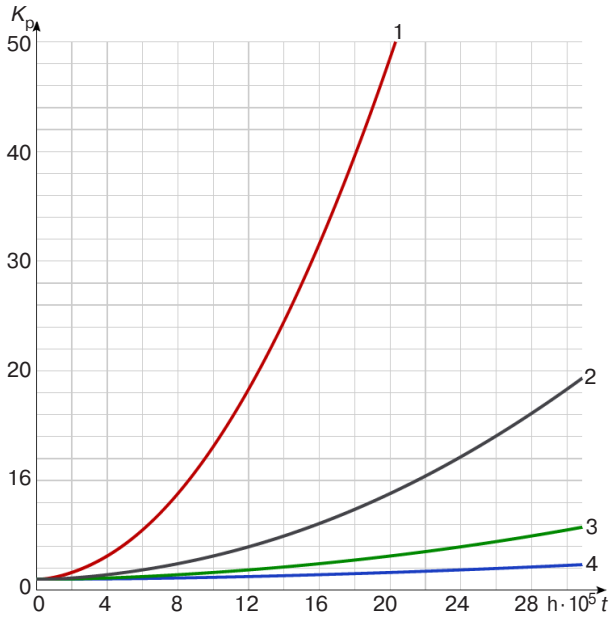


Fig. 5. Time dependencies of the redundancy effectiveness coefficient on the probability of failure-free operation for the redundancy multiplicity of 2/3. 1: $\lambda = 5 \cdot 10^{-6}$ h^{-1}; 2: $\lambda = 2 \cdot 10^{-6}$ h^{-1}; 3: $\lambda = 10^{-6}$ h^{-1}; 4: $\lambda = 5 \cdot 10^{-7}$ h^{-1}

Analysis of the graphs of Fig. 4 and Fig. 5 shows that the coefficient of efficiency of the probability of the failure-free operation increases with time, as well as with the increase of λ. Thus, for example, for the time of 10^6 h and $1.4 \cdot 10^6$ h at a redundancy multiplicity of 1/4 $\lambda = 5 \cdot 10^{-6}$ h^{-1} and $\lambda_k = 10^{-8}$ h^{-1} the coefficient of efficiency is equal to 4.2 and 5.8, respectively. When the redundancy multiplicity is 1/2, the similar values of K_p are 3.7 and 5.0. In case of the decrease of the failures rate down to $\lambda = 10^{-6}$ h^{-1} for the time of 10^6 h and the redundancy multiplicities 1/4 and 1/2, the efficiency coefficient is equal to 2.0 and 1.8, respectively.

In the case of application of the redundancy multiplicity 2/3 at $\lambda = 5 \cdot 10^{-6}$ h^{-1} for the time 10^6 h and $1.4 \cdot 10^6$ h, the redundancy efficiency coefficient is equal to 13.3 and 24.5, respectively (Fig. 5). At decrease of intensity of failures down to $\lambda = 10^{-6}$ h^{-1}, corresponding values of K_p are equal to 1.7 and 2.2, respectively.

FAILURE RATE OF SLIDING REDUNDANCY DIAGRAMS

One of the main indicators of the reliability of REF is the failure rate. Let's investigate how the failure rate of sliding redundancy diagrams with loaded and unloaded redundancy changes for different values of λ and λ_k when the redundancy multiplicity is being changed.

If the redundancy multiplicity is equal to 1/n, the failure rate of the scheme with unloaded redundancy is determined by the expression:

$$\lambda_1(t) = -\frac{dP_1(t)}{dt} \Big/ P_1(t) = \frac{n\lambda\left(n\frac{\lambda}{\lambda_k}+1\right)\left(1-e^{-\lambda_k t}\right)}{n\frac{\lambda}{\lambda_k}\left(1-e^{-\lambda_k t}\right)+1}, \quad (5)$$

and for a diagram with a loaded redundancy the failure rate is determined by the expression:

$$\lambda_2(t) = -\frac{dP_2(t)}{dt} \Big/ P_2(t) = \frac{n\lambda(n+1)\left(1-e^{-\lambda t}\right)}{n\left(1-e^{-\lambda t}\right)+1}. \quad (6)$$

In the case of a redundancy multiplicity of 2/3, the equations for the failure rate are as follows:

- for the diagram with unloaded redundancy

$$\lambda_3(t) = -\frac{dP_3(t)}{dt} \Big/ P_3(t) = \frac{27\lambda^3 t^2}{2+6\lambda t+9\lambda^2 t^2}, \quad (7)$$

- for the diagram with loaded redundancy

$$\lambda_4(t) = -\frac{dP_4(t)}{dt} \Big/ P_4(t) = \frac{30\lambda\left(1+e^{-2\lambda t}-2e^{-\lambda t}\right)}{10+6e^{-2\lambda t}-15e^{-\lambda t}}. \quad (8)$$

From the equations (5)–(8) it follows that for all considered sliding redundancy diagrams at $t = 0$ the failure rate is equal to zero, while when the failure rate is equal to $n\lambda$, the failure rate is the same as for a non-redundant system. Such systems are referred to as ageing.

The dependencies of failure rate on the time of sliding redundancy schemes with redundancy multiplicity of 1/2, 1/4, and 2/3 at $\lambda = 10^{-6}$ h^{-1} are given in Fig. 6. When constructing the dependencies, the value of the switching device failure rate $\lambda_k = 10^{-8}$ h^{-1} was taken for a diagram with unloaded redundancy for multiplicity of redundancy 1/2 and 1/4.

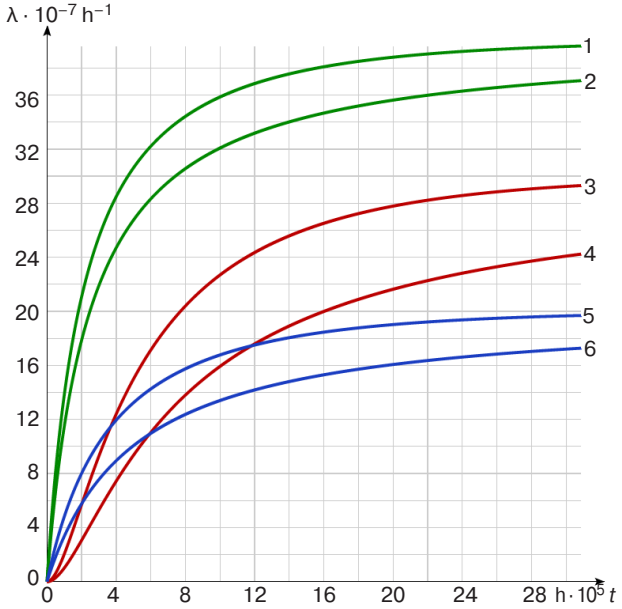


Fig. 6. Dependencies of failure rate on time of sliding redundancy diagrams:
1, 2: multiplicity of redundancy 1/4;
3, 4: multiplicity of redundancy 2/3;
5, 6: multiplicity of redundancy 1/2;
1, 3, 5: loaded redundancy;
2, 4, 6: unloaded redundancy

For comparing the failure intensities of systems with various redundant multiplicities, we will introduce an ageing rate indicator—time of achievement by intensity of failures of level 0.9 from the maximum time, i.e., when redundancy already practically does not influence intensity of failures. The table shows the values of this time determined by equations (5)–(8) and graphs of Fig. 6 values of this time in hours for some values of λ at different variants of redundancy.

Increasing the reliability of REF, in particular satellite communication systems, is associated with an increase in the active life and is determined by functional and structural solutions, allowing to postpone the processes of ageing and degradation of elements and systems as a whole. From the equations (5)–(8), as well as the graphs presented in Fig. 6 graphs and Table data it follows that

- it is possible to slow down the ageing process of the systems of sliding redundancy with unloaded and loaded redundancy at different multiplies of redundancy by application of load redundancy;
- the ageing rate of the considered schemes with sliding redundancy depends on the redundancy multiplicity and is smaller for the multiplicity of 2/3.

CONCLUSIONS

The considered methodology for determining the efficiency of sliding redundancy using mathematical models of the probability of failure-free operation and intensity of failures allows us to draw the following conclusions:

1. Diagram of sliding redundancy with unloaded redundancy is more effective than the diagram with loaded redundancy even when taking into account the intensity of failures of the switching device. Moreover, the efficiency of sliding redundancy with unloaded redundancy increases with time, as well as with the increase of the failure rate of the main and redundant elements.
2. If the ratio of element failure rate to switchgear failure rate is greater than 10, the reliability of the switchgear has almost no effect on the durability of the unloaded sliding redundancy diagram.
3. Combination of load redundancy to reduce the failure rate of elements and structural redundancy makes it possible to achieve a decrease in the ageing rate of systems with sliding redundancy, i.e., to extend the duration of redundancy with

Table. Ageing rate of a redundant system

λ, h^{-1}	Redundancy multiplicity					
	1/2		1/4		2/3	
	Loaded redundancy	Unloaded redundancy	Loaded redundancy	Unloaded redundancy	Loaded redundancy	Unloaded redundancy
10^{-6}	$1.38 \cdot 10^6$	$4.36 \cdot 10^6$	$1.03 \cdot 10^6$	$2.22 \cdot 10^6$	$1.75 \cdot 10^6$	$6.29 \cdot 10^6$
$5 \cdot 10^{-7}$	$2.79 \cdot 10^6$	$8.93 \cdot 10^6$	$2.02 \cdot 10^6$	$4.17 \cdot 10^6$	$3.51 \cdot 10^6$	$12.7 \cdot 10^6$

unloaded and loaded redundancy. The best results are obtained by the diagram with fractional multiplicity of 2/3.

Thus, to improve the efficiency of the long-life REF sliding redundancy, it is necessary to combine the multiplicity of redundancy with the reserve operating mode and ways to reduce the intensity of failure of elements.

Authors' contributions

T.E. Gelfman—research idea, scientific editing.

A.P. Pirkhavka—conducting research, writing the text of the article.

REFERENCES

1. Polovko A.M., Gurov S.V. *Osnovy teorii nadezhnosti* (*Fundamentals of Reliability Theory*). St. Petersburg: BKhV-Peterburg; 2008. 704 p. (in Russ.). ISBN 978-5-94157-541-1
2. Gelfman T.E., Pirkhavka A.P., Skripachev V.O. Analysis of the effectiveness of methods for ensuring the reliability of a communication satellite transponder. *Russ. Technol. J.* 2023;11(1):51–59. https://doi.org/10.32362/2500-316X-2023-11-1-51-59
3. Sevastyanov N.N., Andreev A.I. *Osnovy upravleniya nadezhnost'yu kosmicheskikh apparatov s dlitel'nymi srokami ekspluatatsii* (*Fundamentals of Reliability Management of Spacecraft with Long Service Life*). Tomsk: TGU; 2015. 266 p. (in Russ.). ISBN 978-5-94621-460-5
4. Saleh J.H., Hassan R., Torres-Padilla J.P., Hastings D.E., Newman D.J. Impact of subsystem reliability on satellite revenue generation and present value. *Journal of Spacecraft and Rockets*, 2005;42(6):1122–1129. https://doi.org/10.2514/1.13137
5. Gelfman T.E., Kalmykov M.N., Serditov A.A., Chuev E.A., Bogachev V.N., Kharitonov A.Yu. Reliability of space communication systems. In: *Fundamental Problems of Radioengineering and Device Construction. Proceedings of the International Scientific and Technical Conference "INTERMATIC 2012"*. 2012;12(6):157–161 (in Russ.).
6. Kislaev A.G., Khropov A.N. Optimal redundancy as a method for improving the reliability of space communication systems. *Fundamental Problems of Radioengineering and Device Construction. Proceedings of the International Scientific and Technical Conference "INTERMATIC 2010"*. 2010;10(3):109–112 (in Russ.).
7. Zhadnov V.V. Calculation assessment of durability indicators of spacecrafts and systems electronic means. *Nadezhnost' i kachestvo slozhnykh system = Reliability and Quality of Complex Systems.* 2013;2:65–73 (in Russ.). Available from URL: https://nikas.pnzgu.ru/files/nikas.pnzgu.ru/zhadnov_65_73.pdf
8. Yampurin N.P., Baranova A.V. *Osnovy nadezhnosti elektronnykh sredstv* (*Fundamentals of Reliability of Electronic Means*). Moscow: Akademiya; 2010. 240 p. (in Russ.). ISBN 978-5-7695-5908-2
9. Alyoshin V.F., Kolobov A.Yu., Petrov J.A. Challenging issues for predicting and validating reliability of sustained spacecraft operation. *Nauka i obrazovanie: nauchnoe izdanie MGTU im. N.E. Baumana = Science and Education of Bauman MSTU.* 2015;6:31–41 (in Russ.). Available from URL: https://www.elibrary.ru/item.asp?id=23852897

СПИСОК ЛИТЕРАТУРЫ

1. Половко А.М., Гуров С.В. *Основы теории надежности*. СПб.: БХВ-Петербург; 2008. 704 с. ISBN 978-5-94157-541-1
2. Гельфман Т.Э., Пирхавка А.П., Скрипачев В.О. Анализ эффективности методов обеспечения надежности ретранслятора спутника связи. *Russ. Technol. J.* 2023;11(1):51–59. https://doi.org/10.32362/2500-316X-2023-11-1-51-59
3. Севастьянов Н.Н., Андреев А.И. *Основы управления надежностью космических аппаратов с длительными сроками эксплуатации*. Томск: ТГУ; 2015. 266 с. ISBN 978-5-94621-460-5
4. Saleh J.H., Hassan R., Torres-Padilla J.P., Hastings D.E., Newman D.J. Impact of subsystem reliability on satellite revenue generation and present value. *Journal of Spacecraft and Rockets*, 2005;42(6):1122–1129. https://doi.org/10.2514/1.13137
5. Гельфман Т.Э., Калмыков М.Н., Сердитов А.А., Чуев Е.А., Богачев В.Н., Харитонов А.Ю. Надежность систем космической связи. В сб.: *Фундаментальные проблемы радиоэлектронного приборостроения. Материалы Международной научно-технической конференции «INTERMATIC – 2012»*. 2012;12(6):157–161.
6. Кислаев А.Г., Хропов А.Н. Оптимальное резервирование, как метод повышения надежности систем космической связи. *Фундаментальные проблемы радиоэлектронного приборостроения. Материалы Международной научно-технической конференции «INTERMATIC – 2010»*. 2010;10(1–3):109–112.
7. Жаднов В.В. Расчетная оценка показателей долговечности электронных средств космических аппаратов и систем. *Надежность и качество сложных систем.* 2013;2:65–73. URL: https://nikas.pnzgu.ru/files/nikas.pnzgu.ru/zhadnov_65_73.pdf
8. Ямпурин Н.П., Баранова А.В. *Основы надежности электронных средств*. М.: Академия; 2010. 240 с. ISBN 978-5-7695-5908-2
9. Алешин В.Ф., Колобов А.Ю., Петров Ю.А. Проблемные вопросы прогнозирования и подтверждения надежности космических аппаратов длительного функционирования. *Наука и образование: научное издание МГТУ им. Н.Э. Баумана*. 2015;6:31–41. https://www.elibrary.ru/item.asp?id=23852897
10. Kofanov Y.N., Lvov B.G., Meleh N.A., Sotnikova S.Y. Reliability of space infocommunications equipment. In: *Proceedings of the 2018 International Conference*

10. Kofanov Y.N., Lvov B.G., Meleh N.A., Sotnikova S.Y. Reliability of space infocommunications equipment. In: *Proceedings of the 2018 International Conference "Quality Management, Transport and Information Security, Information Technologies"* (*IT&QM&IS*). IEEE. 2018. P. 354–357. https://doi.org/10.1109/ITMQIS.2018.8525088
11. Jung S., Choi J.P. End-to-end reliability of satellite communication network systems. *IEEE Systems Journal*. 2021;15(1):791–801. https://doi.org/10.1109/JSYST.2020.2980760
12. Jung S., Choi J.P. Reliability of small satellite networks with software-defined radio and enhanced multiple access protocol. *IEEE Transactions on Aerospace and Electronic Systems.* 2021;57(3):1891–1902. https://doi.org/10.1109/TAES.2021.3050652
13. Gelfman T.E., Pirkhavka A.P. The operational readiness factor of satellite communication networks. *Russ. Technol. J.* 2022;10(1):35−40 (in Russ.). https://doi.org/10.32362/2500-316X-2022-10-1-35-40
14. Askinazi G.B., Bykov V.L., Dyachkova M.N., et al. *Sputnikovaya svyaz' i veshchaniye:* spravochnik (*Satellite Communication and Broadcasting.* Handbook). Kantor L.Ya. (Ed.). Moscow: Radio i svyaz'; 1988. 344 p. (in Russ.).
15. Mehmet N., Selman D., Hasan H.E., Cenk S. Reliability and cost focused optimization approach for a communication satellite payload redundancy allocation problem. *International Journal of Electrical, Electronic and Communication Sciences*, 2018;11.0(5). https://doi.org/10.5281/zenodo.1316576

"Quality Management, Transport and Information Security, Information Technologies" (*IT&QM&IS*). IEEE. 2018. P. 354–357. https://doi.org/10.1109/ITMQIS.2018.8525088
11. Jung S., Choi J.P. End-to-end reliability of satellite communication network systems. *IEEE Systems Journal*. 2021;15(1):791–801. https://doi.org/10.1109/JSYST.2020.2980760
12. Jung S., Choi J.P. Reliability of small satellite networks with software-defined radio and enhanced multiple access protocol. *IEEE Transactions on Aerospace and Electronic Systems.* 2021;57(3):1891–1902. https://doi.org/10.1109/TAES.2021.3050652
13. Гельфман Т.Э., Пирхавка А.П. Коэффициент оперативной готовности спутниковых сетей связи. *Russ. Technol. J.* 2022;10(1):35−40. https://doi.org/10.32362/2500-316X-2022-10-1-35-40
14. Аскинази Г.Б., Быков В.Л., Дьячкова М.Н. и др. *Спутниковая связь и вещание:* справочник; ред. Л.Я. Кантор. М.: Радио и связь; 1988. 344 с.
15. Mehmet N., Selman D., Hasan H.E., Cenk S. Reliability and cost focused optimization approach for a communication satellite payload redundancy allocation problem. *Int. J. Electrical, Electronic and Communication Sciences*. 2018;11.0(5). https://doi.org/10.5281/zenodo.1316576

About the authors

Tatyana E. Gelfman, Associate Professor, Department of Radio Electronic Systems and Complexes, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: gelfman@mirea.ru. https://orcid.org/0000-0002-8302-6999

Alexey P. Pirkhavka, Cand. Sci. (Eng.), Associate Professor, Department of Radio Electronic Systems and Complexes, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: pirkhavka@mirea.ru. https://orcid.org/0000-0002-2460-7507

Об авторах

Гельфман Татьяна Элевна, доцент, кафедра радиоэлектронных систем и комплексов Института радиоэлектроники и информатики ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: gelfman@mirea.ru. https://orcid.org/0000-0002-8302-6999

Пирхавка Алексей Петрович, к.т.н., доцент, кафедра радиоэлектронных систем и комплексов Института радиоэлектроники и информатики ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: pirkhavka@mirea.ru. https://orcid.org/0000-0002-2460-7507

Translated from Russian into English by Lyudmila O. Bychkova
Edited for English language and spelling by Thomas A. Beavitt

Micro- and nanoelectronics. Condensed matter physics

Микро- и наноэлектроника. Физика конденсированного состояния

UDC 004.052.2
https://doi.org/10.32362/2500-316X-2023-11-5-54-62

RESEARCH ARTICLE

Application of double-error correction codes to protect configuration programmable logic memory against space radiation

Ekaterina S. Lepeshkina [@], Nikita D. Kustov, Vladislav Kh. Khanov

Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, 660037 Russia
[@] Corresponding author, e-mail: klepka1111.93@mail.ru

Abstract

Objectives. Programmable logic integrated circuits of the field programmable gate array (FPGA) type based on static configuration memory are widely used in the electronics of onboard spacecraft systems. Under the influence of space radiation, errors may occur in the FPGA configuration memory. The main methods of protection against such errors involve various options for reservation triggers, as well as the use of error-correcting codes in special error detection and correction circuits. The purpose of the present work is to determine which error-correcting codes are best suited to the implementation of internal scrubbing of the FPGA configuration memory taking redundancy into account.

Methods. The paper analyses various methods for scrubbing FPGA configuration memory, which are used to correct errors caused by the action of space radiation. It is proposed to increase the efficiency of internal scrubbing of the FPGA configuration memory using codes that correct both single- and double-adjacent SEC-DED-DAEC errors. In this case, the need to perform external scrubbing of the configuration memory is reduced by overwriting it with a reference configuration from non-volatile radiation-resistant memory; in this way, FPGA downtime caused by the external scrubbing procedure is reduced. Due to the known SEC-DED-DAEC codes having a non-zero probability of erroneous detection and subsequent erroneous correction of a double non-adjacent error, as well as various redundancy and implementation complexities, a study was made of the most efficient code for internal scrubbing.

Results. The results showed that the Datta, Neale and Hoyoon–Yongsurk codes are optimal from the indicated positions. Recommendations are given for selecting a specific code depending on the specific requirements for a particular planned space mission.

Conclusions. The study confirms the effectiveness of protecting the memory of programmable logic by using two-error-correcting codes.

Keywords: programmable logic integrated circuits, faults in configuration memory, methods for clearing configuration memory from faults, double-adjacent error-correcting codes

• **Submitted:** 30.01.2023 • **Revised:** 17.05.2023 • **Accepted:** 07.07.2023

For citation: Lepeshkina E.S., Kustov N.D., Khanov V.Kh. Application of double-error correction codes to protect configuration programmable logic memory against space radiation. *Russ. Technol. J*. 2023;11(5):54–62. https://doi.org/10.32362/2500-316X-2023-11-5-54-62

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Применение кодов с исправлением двух ошибок для защиты конфигурационной памяти программируемой логики от действия космической радиации

Е.С. Лепёшкина [@], Н.Д. Кустов, В.Х. Ханов

Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева, Красноярск, 660037 Россия
[@] *Автор для переписки, e-mail: klepka1111.93@mail.ru*

Резюме

Цели. Программируемая логика типа field programmable gate array (FPGA) на основе статической конфигурационной памяти широко применяется в электронике бортовых систем космических аппаратов. Под воздействием космической радиации в конфигурационной памяти FPGA могут возникать ошибки. Основными методами защиты от них являются различные варианты резервирования триггеров, а также применение помехоустойчивых кодов в специальных схемах детектирования и исправления ошибок. Цель работы – определение из группы помехоустойчивых кодов тех, которые с учетом их избыточности наилучшим образом подходят для реализации внутреннего скраббинга конфигурационной памяти программируемых логических интегральных схем.

Методы. В работе рассмотрены методы скраббинга конфигурационной памяти FPGA, которые применяются для ее очистки от ошибок, вызванных действием космической радиации. Предлагается для повышения эффективности внутреннего скраббинга конфигурационной памяти FPGA использовать коды, исправляющие как однократные, так и двукратные смежные ошибки SEC-DED-DAEC. В этом случае уменьшается необходимость выполнения внешнего скраббинга конфигурационной памяти путем ее перезаписи эталонной конфигурацией из энергонезависимой радиационно-стойкой памяти. Таким образом, снижается время неработоспособного состояния FPGA, вызванное процедурой внешнего скраббинга. В связи с тем, что известные коды SEC-DED-DAEC имеют ненулевую вероятность ошибочного детектирования, а затем – ошибочного исправления двойной несмежной ошибки, а также обладают разной избыточностью и сложностью реализации, было приведено исследование наиболее эффективного кода для внутреннего скраббинга.

Результаты. Результаты исследования показали, что наилучшими с указанных позиций являются коды Датта, Нила и Хоюна – Йонгсурка. Приведены результаты сравнения кодов по выбранным критериям. Даны рекомендации для выбора конкретного кода в зависимости от возможных требований к планируемой космической миссии.

Выводы. Проведенное исследование показало эффективность защиты памяти программируемой логики с помощью применения кодов с исправлением двух ошибок.

Ключевые слова: программируемая логическая интегральная схема, сбой в конфигурационной памяти, методы очистки конфигурационной памяти от сбоев, коды с исправлением двух смежных ошибок

• **Поступила:** 30.01.2023 • **Доработана:** 17.05.2023 • **Принята к опубликованию:** 07.07.2023

Для цитирования: Лепёшкина Е.С., Кустов Н.Д., Ханов В.Х. Применение кодов с исправлением двух ошибок для защиты конфигурационной памяти программируемой логики от действия космической радиации. *Russ. Technol. J.* 2023;11(5):54–62. https://doi.org/10.32362/2500-316X-2023-11-5-54-62

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

Currently, there is a significant progress in the development of space electronics, which has significantly expanded the functionality of modern spacecraft, while reducing the size, mass and power consumption of on-board systems. At the same time, the problem of protecting hardware from radiation failures in electronic components remains urgent, especially since cheap components from the commercial electronic component base (ECB), which lacks failure protection mechanisms typical of radiation-resistant space ECB, have begun to be widely used on small spacecraft. In hardware based on commercial ECBs, only single event upset (SEU) protection methods, which set out to correct single events of charged radiation particles hitting electronic components, are for the most part capable of application [1]. Such failures cause logic errors without destroying the component, i.e., they do not cause irreversible processes in the semiconductor structure of the component [2]. However, SEU failures tend to occur in response to triggers, with which any electronic equipment can be easily saturated. The main drawback of such an approach is the relatively high probability of such failures.

The main methods of protection against SEU failures are various options of redundancy, including triple modular redundancy (TMR) [3, 4] of the triggers, as well as the use of error correction codes (ECC) in special error detection and correction circuits (EDAC) [5, 6], which typically correct a single error (one-bit error) in the structure of several triggers, for example, in the SRAM word (static random access memory).

One of the widely used components in space instrumentation consists in the SRAM FPGA (static random access memory field programmable gate array) type of programmable logic integrated circuits. Such programmable logic integrated circuits (FPGAs), which mainly belong to the commercial category of microcircuits, are characterized by a large number of programmable elements, low power consumption, and a high speed of operation of the circuits implanted in them, but low fault tolerance to SEU events [7]. SEU errors can occur not only in the triggers of the SRAM-implemented FPGA circuit, but also in the SRAM configuration memory. The latter type of error is the most dangerous due to the change to the structure of the implemented circuit, resulting in constant periodic failures whenever signals pass through the damaged part of the circuit [8].

The main method of SRAM FPGA protection from SEU is scrubbing, which consists in overwriting the faulty contents of the configuration memory by the reference configuration whenever an error is detected in the implemented circuit [9, 10].

In this paper, various scrubbing variants are considered, and the possibilities of applying ECCs for scrubbing are analyzed. In this connection, the efficiency of the known SEC-DED-DAEC codes for the purpose of scrubbing SRAM FPGA configuration memory is evaluated assuming the application of SEC-DED-DAEC (single error correction, double error detection and double-adjacent error correction) type codes.

RELATED WORKS

Since the content of FPGA SRAM configuration memory can be corrupted by radiation exposure, configuration memory errors must be detected and corrected quickly in order to maintain correct operation in a radiation-exposed environment such as outer space. Configuration memory errors are eliminated by scrubbing, comprising a method for overwriting memory with a reference configuration.

There are two main overwriting approaches: full and partial. Full scrubbing consists in overwriting the entire contents of the configuration memory, while partial scrubbing comprises a block-by-block overwriting of the configuration memory [11]. In the second case, the system inoperability time due to loading the full configuration is reduced.

Scrubbing can also be “blind” and “sighted” [12]. “Blind” scrubbing consists in periodic overwriting of full or sequentially block configuration memory. Here,

the overwriting period is selected from the anticipated error rate, for example, depending on the orbital altitude of the spacecraft. "Sight" scrubbing is performed following the detection of an error, possibly when the configuration memory block has been localized, errors have accumulated, and it is impossible to correct them by any other additional method.

There are direct and indirect approaches for detecting errors in the configuration memory. The indirect method consists in identifying errors in the internal FPGA system configured with the contents of the configuration memory. If there is a failure in the configuration memory, it will appear in the system in the form of an error, which can be used to detect (and rectify) TMR blocks distributed in the system. In case of a repeated (3–5 times consecutively) detected error, the scrubbing procedure should be started [11].

The direct method consists in periodic block-by-block comparison of the contents of the configuration memory with the corresponding block of the reference configuration, which is located in non-volatile and radiation-resistant memory. Here, ECCs can be used to accelerate the work. In this case, each block of the reference configuration in the non-volatile memory also stores its checksum. When the corresponding block is read back from the FPGA configuration memory, its checksum is calculated and compared with the corresponding sum from the non-volatile memory. If the sums do not match, it means that a failure has occurred in this block of the configuration memory and that this block must consequently be overwritten [8].

Modern Xilinx[1] FPGAs have built-in mechanisms based on the application of EDAC and single error correction and double error detection (SEC-DED) codes for block-by-block scanning of the configuration memory and automatic correction of single errors in it, as well as double errors detection in the scanned block [8]. This approach, further referred to in the paper as internal scrubbing, allows to reduce the number of runs of external partial scrubbing of configuration memory to correct double and larger multiplicity errors, and thus to reduce downtime of the FPGA-implemented system associated with the execution of external scrubbing.

PROBLEM STATEMENT

SEC-DED codes are known to fix one fault in a memory word. The SEC-DED code used must be low redundancy, fast, and easy to implement (for FPGA, it comprises a small number of logic elements for implementing a particular code). These criteria are best met by the Hsiao code (39, 32) [13] (where 39 is the codeword size in bits, 32 is informational), which belongs to the group of modified Hamming codes [14].

[1] https://www.xilinx.com/. Accessed February 16, 2023.

With the development of new ECB creation technologies and the transition to finer component production processes, the probability of multibit failures—primarily two-bit failures—increases [15]. One of the approaches to solving this problem consists in the use of SEC-DED-DAEC codes [16]. Codes belonging to this group correct 2 adjacent errors (adjacent means located in two adjacent bits of one memory word). Such codes also belong to the group of modified Hamming codes, meaning that they are fast and have low redundancy. Any SEC-DED-DAEC code copes well with the tasks of correcting single- and double adjacent errors. However, such codes have a specific disadvantage, consisting in the probability of incorrect correction of non-adjacent double errors; it can happen that one code from this group has a higher probability, while for another code, the probability is lower.

It should be noted that the probability of a contiguous double error in a memory word is significantly higher than the probability of a double non-contiguous error. A double contiguous error ensues from a single SEU event spanning two adjacent bits of the same word, while a double non-adjacent error comprises an accumulation of errors. First there is a failure that causes an error in one bit of a memory word, then, as time passes, another failure occurs causing an error in another bit of the same memory word, resulting in a double non-contiguous error. Clearly, although this process is unlikely, it cannot be ruled out. Therefore, it would be wrong to exclude from the SEC-DED-DAEC code analysis the estimation of the probability of incorrectly correcting unrelated double errors.

SEC-DED-DAEC codes are proposed as a means to improve the efficiency of internal scrubbing of the FPGA configuration memory. In this case, both single and double contiguous errors are corrected in the configuration memory to permit less recourse to external scrubbing of the configuration memory by overwriting it with a reference configuration from the non-volatile radiation-resistant memory.

The research task then appears as follows. Due to the fact that the number of SEC-DED-DAEC codes comprises some set, each of them has different parameters; therefore, it is necessary to conduct research in order to obtain an estimate of the probability of false detection (and subsequent false correction) of a twofold unrelated error along with the required resources for the implementation of the encoder/decoder, thus determining which of them are best fitted to implement internal FPGA configuration memory scrubbing given their additional redundancy.

RESEARCH PROVISION

As it was noted earlier, for SEC-DED-DAEC codes there is a probability of erroneous detection of double non-contiguous error as double contiguous, which

subsequently leads to erroneous correction of code word bits. The determination of such probability for this or that code is carried out by means of functional modeling.

For this purpose, a functional model was developed, presented as a program in C++, developed in the *Microsoft Visual Studio*[2] environment. The main logical parts of the program are described below:

1) variable initialization block (generating **G**-matrix and validating **H**-matrix);
2) information word generation block (one-dimensional Boolean array) using a pseudorandom function;
3) information word coding block by enumerating the columns of **G**-matrix. When one appears, an XOR operation is performed on the current bit of the codeword and the corresponding bit of the information word;
4) a block for introducing a double non-adjacent error using a pseudorandom function to determine random non-adjacent positions and subsequent inversion of the codeword bits;
5) error syndrome detection block by enumerating columns of the **H**-matrix. When one appears, an XOR operation is performed on the current bit of the syndrome and the corresponding bit of the codeword;
6) block of double adjacent error detection by comparing the resulting syndrome with the syndromes of double adjacent errors resulting from the XOR operation on two neighboring columns of the **H**-matrix;
7) block for detecting a double non-adjacent error. A double non-adjacent error is detected if a double adjacent error is not detected.

The algorithm goes through a large number of iterations. The probability of error-free detection of a double non-adjacent error is determined by the ratio of the number of outcomes of double non-adjacent error detection to the total number of code words with introduced double non-adjacent error that have passed through the algorithm.

In order to determine the resources required to implement the encoder and decoder codes (the number of logical elements), a simulation was performed. For this purpose, a program was developed using the VHDL language in the *Quartus*[3] development environment. Functional debugging was performed in the *ModelSim*[4] environment.

The logical parts of the encoder are organized as follows:

1) initialization block for variables (CLK clock signal, input data word and output codeword);
2) coding block. When the CLK signal changes, a cycle of encoding the checksum bits according to the **G**-matrix is started using the XOR operation;
3) codeword output block (initial information word and control sum).

Logical parts of a decoder:

1) initialization block for variables (CLK clock signal, input codeword, output corrected word);
2) error syndrome detection block. When the CLK signal changes, the loop for the codeword bits is started; the error syndrome is calculated according to the **H**-matrix;
3) decoding block. Firstly, the case without an error is checked (if the syndrome is zero). If the syndrome is non-zero, the resulting syndrome is compared to single error syndromes and error correction is performed. If there are no coincidences with single error syndromes, comparison with double adjacent error syndromes and error correction takes place. If there are no matches with syndromes, a non-adjacent error is detected;
4) block of corrected word output.

Simulation was performed for FPGA Cyclone IV E EP4CE6E22A7 (Intel, USA). As a result of the synthesis, the number of used encoder and decoder logic elements was counted.

Functional and simulation models for several codes were developed in accordance with the presented description. The main criterion for the selection of the codes under study is the explicit description in the scientific and technical literature of the H-matrices of the codes. In this case, the probability of incorrect generation of the check matrix is excluded, the research process is simplified, and the probability of an error occurring in the simulated results is minimized. Thus, the following codes with 32-bit information word were chosen as the codes under study: Dutta (39, 32) [17], Datta (42, 32) [18], Neale (42, 32) [19], Reviriego (39, 32) [20], Cha–Yoon (39, 32) [21], Hoyoon–Yongsurk (41, 32) [22].

In addition to the SEC-DED-DAEC codes, we present for comparison data for one SEC-DED code, namely, the Hsiao code [13], which is widely used in the implementation of the EDAC mechanism for memory.

RESULTS

The results of the studies are presented in the table. The given data show that the leading positions are occupied by the Datta, Neale and Hoyoon–Yongsurk codes. While the Hsiao code shows the average values

[2] https://visualstudio.microsoft.com/ru/ (in Russ.). Accessed February 16, 2023.

[3] http://altera.ru/soft_quartus.html. Accessed February 16, 2023.

[4] https://altera.co.uk/products/software/quartus-ii/modelsim/qts-modelsim-index.html. Accessed February 16, 2023.

for the selected comparison criteria, it should be kept in mind that it does not correct two adjacent errors. Comparing them with each other in relation to the use of FPGA configuration SRAM memory scanning, the following recommendations can be made.

Table. Comparison of SEC-DED-DAEC codes according to the criteria of error-free detection of a double non-adjacent error and implementation complexity

Code	Probability of error-free detection of a double non-adjacent error	Number of encoder/decoder logical elements
Hsiao (39, 32)	63.4	164/332
Dutta (39, 32)	43.5	170/407
Datta (40, 32)	78.9	164/354
Neale (42, 32)	84.4	57/391
Reviriego (39, 32)	38.4	175/373
Cha–Yoon (39, 32)	60.7	123/272
Hoyoon–Yongsurk (41, 32)	95.7	178/384

For the Datta code, the redundancy of the codeword is aligned to the byte dimension, which can simplify the construction of the FPGA configuration memory. While the probability of error detection of double non-adjacent error is the worst of the considered three codes, it is slightly inferior to the Neale code. The complexities of the coder and decoder code are low; the complexity of the coder is average.

For the Neale code, the redundancy of the codeword is not aligned with the byte boundary. The probability of erroneous detection of double non-adjacent error is average, but at a rather high level. Encoder complexity is very low; decoder complexity is the highest.

The redundancy for the Hoyoon–Yongsurk code is also not aligned with the byte boundary. Here, while the probability of error detection of double non-adjacent error is the best and encoder complexity is the highest, decoder complexity is average.

The selection of one or another code choice for implementation should take into account the initial data for the planned space mission. If simplicity and speed of implementation are critical with an acceptable probability of false detection of double non-adjacent error under conditions of low intensity of radiation failures, the Datta configuration memory code can be recommended for use in the EDAC mechanism. The Neale code can also qualify for this position, but it has the highest redundancy, and is not aligned on the byte boundary. If the planned space mission is long-term and will be carried out under conditions of a high failure rate, it may be better to use the Hoyoon–Yongsurk code, despite its relative implementation complexity and low performance.

In addition, we note that the complexity of the analyzed triplet codes, expressed in the number of required logical elements for the implementation of the encoder/decoder, on average corresponds to the complexity of the commonly used Hsiao code. Therefore, problems do not arise in terms of the use of FPGA resources with their implementation for the purpose of configuration memory scanning when scrubbing.

CONCLUSIONS

The present work proposes that the algorithm of the internal scanning method be changed for the internal scrubbing of the FPGA SRAM configuration memory: instead of one of the SEC-DED codes, such as the Hsiao code, it is proposed to use one of the SEC-DED-DAEC codes. In this case, both a single error and two adjacent errors will be corrected in the configuration SRAM, which will reduce the number of external partial configuration memory scrubbing runs to correct double and larger multiplicity errors, thus reducing the downtime of the FPGA-implemented system associated with scrubbing implementation.

SEC-DED-DAEC codes have one negative property: non-zero probability of erroneous detection (and then erroneous correction) of double non-adjacent error. In addition, they have different redundancies for storing the checksum in SRAM memory, as well as various complexities of implementation, which can be estimated by the required number of logical elements to implement the encoder/decoder. A study was carried out to determine the most efficient code according to these criteria among SEC-DED-DAEC codes with known check matrices. The results of the study showed that the Datta, Neale, and Hoyoon–Yongsurk codes are optimal from these positions. While the Hoyun–Yongsurk code has almost zero probability of error detection of double adjacent error, at the same time, it has the greatest complexity. The Dutta code is the easiest to implement, but the probability of detecting a double non-adjacent error in error is about 20%. The Neale code occupies an intermediate position. When selecting error correction approaches, the choice of a particular code should be determined by the requirements of the planned space mission.

Acknowledgments

The reported study was supported by the Russian Foundation for Basic Research, project No. 19-38-90052.

Authors' contribution. All authors equally contributed to the research work.

REFERENCES

1. Maksimenko S.L., Melekhin V.F., Filippov A.S. Analysis of the problem of radiation-tolerant information and control-systems implementation. *Informatsionno-upravlyayushchie sistemy = Information and Control Systems.* 2012;2(57):18–25 (in Russ.). Available from URL: http://www.i-us.ru/index.php/ius/article/view/13788
2. Gaillard R. Single Event Effects: Mechanisms and Classification. In: Nicolaidis M. (Ed.). *Soft Errors in Modern Electronic Systems. Frontiers in Electronic Testing.* Boston, MA: Springer; 2011. V. 41. P. 27–54. https://doi.org/10.1007/978-1-4419-6993-4_2
3. Kastensmidt F.L., Sterpone L., Carro L., Reorda M.S. On the Optimal Design of Triple Modular Redundancy Logic for SRAM-based FPGAs. In: *Proceedings of the Design, Automation and Test in Europe Conference and Exhibition*. 2005. P. 1290–1295. https://doi.org/10.1109/DATE.2005.229
4. Cheng D., Qi D., Chen M. Radiation-hardened Test Design for Aerospace SoC. In: *2020 IEEE 5th International Conference on Integrated Circuits and Microsystems* (*ICICM*). 2020. P. 213–217. https://doi.org/10.1109/ICICM50929.2020.9292308
5. Mang I., Mang E., Popescu C. VHDL implementation of an error detection and correction module based on Hamming code. *J. Comput. Sci. Control Syst.* 2011;4(2):43–46.
6. Baviera E., Schettino G.M., Tuniz E., Vatta F. Software Implementation of Error Detection and Correction Against Single-Event Upsets. In: *2020 International Conference on Software, Telecommunications and Computer Networks* (*SoftCOM*). 2020. https://doi.org/10.23919/SoftCOM50211.2020.9238173
7. Maksfild K. *Proektirovanie na PLIS. Arkhitektura, sredstva i metody: kurs molodogo boitsa* (*The Design Warrior's Guide to FPGAs: Devices, Tools and Flows*.): transl. from Engl. Moscow: Dodeka-XXI; 2007. 408 p. (in Russ.). [Maxfield C. *The Design Warrior's Guide to FPGAs: Devices, Tools and Flows*. Oxford, UK: Jordan Hill; 2004. 560 p.]
8. Virtlin M., Khardint A. Hybrid Clear Configuration for Xilinx FPGAs. In: *PLIS i parallel'nye struktury v aerokosmicheskoi oblasti. Programmnye oshibki i otkazoustoichivoe proektirovanie* (*FPGAs and Parallel Architectures for Aerospace Applications. Soft Errors and Fault-Tolerant Design*). Kastensmidt F., Rech P. (Eds.). Moscow: Tekhnosfera; 2018. 326 p. (in Russ.).
9. Sokolov A. Software and hardware methods of radiation hardening of microprocessors based control systems. *Sovremennaya Elektronika*. 2014;6:30–33 (in Russ.).
10. Zhang R., Xiao L., Li J., Cao X., Li L. An Adjustable and Fast Error Repair Scrubbing Method Based on Xilinx Essential Bits Technology for SRAM-Based FPGA. *IEEE Transactions on Reliability*. 2020;69(2):430–439. https://doi.org/10.1109/TR.2019.2896897
11. Tsetin E., Dizzel' O., Li T., et al. Review and study of methods for detecting and eliminating single failures for heterogeneous systems based on FPGAs. In: *PLIS i parallel'nye struktury v aerokosmicheskoi oblasti. Programmnye oshibki i otkazoustoichivoe proektirovanie* (*FPGAs and Parallel Architectures for Aerospace Applications. Soft Errors and Fault-Tolerant Design*).

СПИСОК ЛИТЕРАТУРЫ

1. Максименко С.Л., Мелехин В.Ф., Филиппов А.С. Анализ проблемы построения радиационно-стойких информационно-управляющих систем. *Информационно-управляющие системы*. 2012;2(57):18–25. URL: http://www.i-us.ru/index.php/ius/article/view/13788
2. Gaillard R. Single Event Effects: Mechanisms and Classification. In: Nicolaidis M. (Ed.). *Soft Errors in Modern Electronic Systems. Frontiers in Electronic Testing.* Boston, MA: Springer; 2011. V. 41. P. 27–54. https://doi.org/10.1007/978-1-4419-6993-4_2
3. Kastensmidt F.L., Sterpone L., Carro L., Reorda M.S. On the Optimal Design of Triple Modular Redundancy Logic for SRAM-based FPGAs. In: *Proceedings of the Design, Automation and Test in Europe Conference and Exhibition*. 2005. P. 1290–1295. https://doi.org/10.1109/DATE.2005.229
4. Cheng D., Qi D., Chen M. Radiation-hardened Test Design for Aerospace SoC. In: *2020 IEEE 5th International Conference on Integrated Circuits and Microsystems* (*ICICM*). 2020. P. 213–217. https://doi.org/10.1109/ICICM50929.2020.9292308
5. Mang I., Mang E., Popescu C. VHDL implementation of an error detection and correction module based on Hamming code. *J. Comput. Sci. Control Syst.* 2011;4(2):43–46.
6. Baviera E., Schettino G.M., Tuniz E., Vatta F. Software Implementation of Error Detection and Correction Against Single-Event Upsets. In: *2020 International Conference on Software, Telecommunications and Computer Networks* (*SoftCOM*). 2020. https://doi.org/10.23919/SoftCOM50211.2020.9238173
7. Максфилд К. *Проектирование на ПЛИС. Архитектура, средства и методы: курс молодого бойца*: пер. с англ. М.: Додэка-XXI; 2007. 408 с.
8. Виртлин М., Хардинт А. Гибридная очистка конфигурации для ПЛИС компании Xilinx. В кн.: *ПЛИС и параллельные структуры в аэрокосмической области. Программные ошибки и отказоустойчивое проектирование*; под ред. Ф. Канстеншмидт, П. Реха. М.: Техносфера; 2018. 326 с.
9. Соколов А. Программно-аппаратные методы повышения радиационной стойкости микросхем SRAM ПЛИС. *Современная электроника*. 2014;6:30–33.
10. Zhang R., Xiao L., Li J., Cao X., Li L. An Adjustable and Fast Error Repair Scrubbing Method Based on Xilinx Essential Bits Technology for SRAM-Based FPGA. *IEEE Transactions on Reliability*. 2020;69(2):430–439. https://doi.org/10.1109/TR.2019.2896897
11. Цетин Э., Диззель О., Ли Т. и др. Обзор и исследование методов обнаружения и устранения одиночных сбоев для гетерогенных систем на основе ПЛИС. В кн.: *ПЛИС и параллельные структуры в аэрокосмической области. Программные ошибки и отказоустойчивое проектирование*; под ред. Ф. Канстеншмидт, П. Реха. М.: Техносфера; 2018. 326 с.
12. Флеминг Ш.Т., Томас Д.В., Винтерстейн Ф. Энергосберегающая адаптивная платформа FDIR, применяющая модули гетерогенных систем на кристалле. В кн.: *ПЛИС и параллельные структуры в аэрокосмической*

Kastensmidt F., Rech P. (Eds.). Moscow: Tekhnosfera; 2018. 326 p. (in Russ.).
12. Fleming Sh., Tomas D., Vintersteyn F. Energy-saving adaptive FDIR platform using heterogeneous system-on-chip modules. In book: *PLIS i parallel'nye struktury v aerokosmicheskoi oblasti. Programmnye oshibki i otkazoustoichivoe proektirovanie* (*FPGAs and Parallel Architectures for Aerospace Applications. Soft Errors and Fault-Tolerant Design*). Kastensmidt F., Rech P. (Eds.). Moscow: Tekhnosfera; 2018. 326 p. (in Russ.).
13. Hsiao M.Y. A class of optimal minimum odd-weight-column SEC-DED codes. *IBM J. Res. Develop*. 1970;14(4):395–401. https://doi.org/10.1147/rd.144.0395
14. Hamming R.W. Error detecting and error correcting codes. *Bell System technical journal*. 1950;29(2): 147–160. Available from URL: https://www.sci-hub.ru/10.1002/j.1538-7305.1950.tb00463.x
15. Chumakov A.I., Sogoyan A.V., Boruzdina A.B., et al. Mechanisms of multiple cell upsets in memory. *Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh sistem* (*MES*) = *Problems of Advanced Micro- and Nanoelectronic Systems Development.* 2016;4:145–152 (in Russ.).
16. Krasnyuk A.A., Petrov K.A. Application features of the error correction coding in sub-100-nm memory microcircuits for cosmic systems. *Russ. Microelectron.* 2013;42(1): 53–58. https://doi.org/10.1134/S1063739712040087 [Original Russian Text: Krasnyuk A.A., Petrov K.A. Application features of the error correction coding in sub-100-nm memory microcircuits for cosmic systems. *Mikroelektronika.* 2012;41(6):450–456 (in Russ.).]
17. Dutta A., Touba N.A. Multiple Bit Upset Tolerant Memory Using a Selective Cycle Avoidance Based SEC-DED-DAEC Code. In: *25th IEEE VLSI Test Symposium* (*VTS'07*). 2007. P. 349–354. https://doi.org/10.1109/VTS.2007.40
18. Datta R., Touba N.A. Exploiting unused spare columns to improve memory ECC. In: *27th IEEE VLSI Test Symposium*. 2009. P. 47–52. https://doi.org/10.1109/VTS.2009.52
19. Neale A., Sachdev M. A new SEC-DED error correction code subclass for adjacent MBU tolerance in embedded memory. *IEEE Transactions on Device and Materials Reliability.* 2013;13(1):223–230. https://doi.org/10.1109/TDMR.2012.2232671
20. Reviriego P., Liu S.S., Sánchez-Macián A., Xiao L., Maestro J.A. Unequal error protection codes derived from SEC-DED codes. *Electron. Lett*. 2016;52(8):619–620. https://doi.org/10.1049/el.2016.0077
21. Cha S., Yoon H. Efficient implementation of single error correction and double error detection code with check bit pre-computation for memories. *JSTS: J. Semiconductor Technol. Sci.* 2018;12(4):418–425. https://doi.org/10.5573/JSTS.2012.12.4.418
22. Hoyoon Jun, Yongsurk Lee. Protection of On-chip Memory Systems against Multiple Cell Upsets Using Double-adjacent Error Correction Codes. *Int. J. Computer Inform. Technol*. 2014;3(6):1316–1320. Available from URL: https://www.ijcit.com/archives/volume3/issue6/Paper030621.pdf

области. Программные ошибки и отказоустойчивое проектирование; под ред. Ф. Канстеншмидт, П. Реха. М.: Техносфера; 2018. 326 с.
13. Hsiao M.Y. A class of optimal minimum odd-weight-column SEC-DED codes. *IBM J. Res. Develop*. 1970;14(4):395–401. https://doi.org/10.1147/rd.144.0395
14. Hamming R.W. Error detecting and error correcting codes. *Bell System technical journal*. 1950;29(2):147–160. URL: https://www.sci-hub.ru/10.1002/j.1538-7305.1950.tb00463.x
15. Чумаков А.И., Согоян А.В., Боруздина А.Б. и др. Механизмы многократных сбоев в микросхемах памяти. *Проблемы разработки перспективных микро- и наноэлектронных систем* (*МЭС*). 2016;4:145–152. URL: http://www.mes-conference.ru/data/year2016/pdf/D188.pdf
16. Краснюк А.А., Петров К.А. Особенности применения методов помехоустойчивого кодирования в суб-100-нм микросхемах памяти для космических систем. *Микроэлектроника*. 2012;41(6):450–456.
17. Dutta A., Touba N.A. Multiple Bit Upset Tolerant Memory Using a Selective Cycle Avoidance Based SEC-DED-DAEC Code. In: *25th IEEE VLSI Test Symposium* (*VTS'07*). 2007. P. 349–354. https://doi.org/10.1109/VTS.2007.40
18. Datta R., Touba N.A. Exploiting unused spare columns to improve memory ECC. In: *27th IEEE VLSI Test Symposium*. 2009. P. 47–52. https://doi.org/10.1109/VTS.2009.52
19. Neale A., Sachdev M. A new SEC-DED error correction code subclass for adjacent MBU tolerance in embedded memory. *IEEE Transactions on Device and Materials Reliability.* 2013;13(1):223–230. https://doi.org/10.1109/TDMR.2012.2232671
20. Reviriego P., Liu S.S., Sánchez-Macián A., Xiao L., Maestro J.A. Unequal error protection codes derived from SEC-DED codes. *Electron. Lett*. 2016;52(8):619–620. https://doi.org/10.1049/el.2016.0077
21. Cha S., Yoon H. Efficient implementation of single error correction and double error detection code with check bit pre-computation for memories. *JSTS: J. Semiconductor Technol. Sci.* 2018;12(4):418–425. https://doi.org/10.5573/JSTS.2012.12.4.418
22. Hoyoon Jun, Yongsurk Lee. Protection of On-chip Memory Systems against Multiple Cell Upsets Using Double-adjacent Error Correction Codes. *Int. J. Computer Inform. Technol*. 2014;3(6):1316–1320. URL: https://www.ijcit.com/archives/volume3/issue6/Paper030621.pdf

About the authors

Ekaterina S. Lepeshkina, Assistant, Information Technology Security Department; Engineer, "Small Satellite" Laboratory, Reshetnev Siberian State University of Science and Technology (Reshetnev University). (31, Krasnoyarsky Rabochy pr., Krasnoyarsk, 660037 Russia). E-mail: klepka1111.93@mail.ru. Scopus Author ID 57218577296, RSCI SPIN-code 8746-6555, https://orcid.org/0000-0001-5116-6260

Nikita D. Kustov, Assistant, Information Technology Security Department; Engineer, "Small Satellite" Laboratory, Reshetnev Siberian State University of Science and Technology (Reshetnev University). (31, Krasnoyarsky Rabochy pr., Krasnoyarsk, 660037 Russia). E-mail: kustovnd@yandex.ru. Scopus Author ID 57218577358, RSCI SPIN-code 9031-4516, https://orcid.org/0000-0002-3362-3971

Vladislav Kh. Khanov, Cand. Sci. (Eng.), Associate Professor, Information Technology Security Department; Head of the "Small Satellite" Laboratory, Reshetnev Siberian State University of Science and Technology (Reshetnev University). (31, Krasnoyarsky Rabochy pr., Krasnoyarsk, 660037 Russia). E-mail: khvkh@mail.ru. Scopus Author ID 56491191500, RSCI SPIN-code 5197-5699, https://orcid.org/0000-0001-6720-9405

Об авторах

Лепёшкина Екатерина Сергеевна, ассистент, кафедра безопасности информационных технологий; инженер, лаборатория «Малые космические аппараты», Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева (660037, Россия, Красноярск, просп. им. газеты «Красноярский рабочий», д. 31). E-mail: klepka1111.93@mail.ru. Scopus Author ID 57218577296, SPIN-код РИНЦ 8746-6555, https://orcid.org/0000-0001-5116-6260

Кустов Никита Дмитриевич, ассистент, кафедра безопасности информационных технологий; инженер, лаборатория «Малые космические аппараты», Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева (660037, Россия, Красноярск, просп. им. газеты «Красноярский рабочий», д. 31). E-mail: kustovnd@yandex.ru. Scopus Author ID 57218577358, SPIN-код РИНЦ 9031-4516, https://orcid.org/0000-0002-3362-3971

Ханов Владислав Ханифович, к.т.н., доцент, кафедра безопасности информационных технологий; заведующий лабораторией «Малые космические аппараты», Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева (660037, Россия, Красноярск, просп. им. газеты «Красноярский рабочий», д. 31). E-mail: khvkh@mail.ru. Scopus Author ID 56491191500, SPIN-код РИНЦ 5197-5699, https://orcid.org/0000-0001-6720-9405

Translated from Russian into English by Lyudmila O. Bychkova
Edited for English language and spelling by Thomas A. Beavitt

Micro- and nanoelectronics. Condensed matter physics

Микро- и наноэлектроника. Физика конденсированного состояния

UDC 537.63: 538.955: 538.956
https://doi.org/10.32362/2500-316X-2023-11-5-63-70

RESEARCH ARTICLE

Nonlinear magnetoelectric effect in a ring composite heterostructure

Vladimir I. Musatov [@], Fedor A. Fedulov, Dmitrii V. Savelev, Ekaterina V. Bolotina, Leonid Y. Fetisov

MIREA – Russian Technological University, Moscow, 119454 Russia
[@] Corresponding author, e-mail: musatov.v.i@mail.ru

Abstract

Objectives. The relevance of the study of magnetoelectric (ME) effect in ring ferromagnetic–piezoelectric heterostructures is due to the possibility of creating various ME devices having improved characteristics. A detailed investigation of the nonlinear ME effect in a ring composite heterostructure based on lead zirconate titanate (PZT) piezoceramics and Metglas® amorphous ferromagnetic (FM) alloy under circular magnetization is presented.

Methods. The ME effect was measured by the low-frequency magnetic field modulation method. Excitation alternating- and constant magnetic bias fields were created using toroidal coils wound on a ring heterostructure for circular magnetization of the FM layer.

Results. When excited with circular magnetic fields in a non-resonant mode, the ME ring heterostructure generates a nonlinear ME voltage of higher harmonics. The field and amplitude dependencies of the first three ME voltage harmonics were investigated. ME coefficients were obtained for the linear ME effect $\alpha^{(1)}$ = 5.2 mV/(Oe·cm), the nonlinear ME effect $\alpha^{(2)}$ = 6 mV/(Oe^2·cm), and $\alpha^{(3)}$ = 0.15 mV/(Oe^3·cm) at an excitation magnetic field frequency f = 1 kHz. The maximum amplitudes of the 1st and 3rd harmonics were observed at a constant bias magnetic field H ~ 7 Oe, which is almost two times smaller than in planar PZT–Metglas® heterostructures.

Conclusions. A nonlinear ME effect was observed and investigated in a ring heterostructure based on PZT piezoceramics and Metglas® amorphous FM alloy. Due to the absence of demagnetization during circular magnetization of the closed FM layer, nonlinear ME effects are detected at significantly lower amplitudes of the exciting alternating and constant bias magnetic fields as compared to planar heterostructures. The investigated ring heterostructures are of potential use in the creation of frequency multipliers.

Keywords: nonlinear magnetoelectric effect, composite heterostructure, magnetostriction, ferromagnet, piezoelectric effect

• **Submitted:** 23.03.2023 • **Revised:** 27.04.2023 • **Accepted:** 11.07.2023

For citation: Musatov V.I., Fedulov F.A., Savelev D.V., Bolotina E.V., Fetisov L.Y. Nonlinear magnetoelectric effect in a ring composite heterostructure. *Russ. Technol. J*. 2023;11(5):63–70. https://doi.org/10.32362/2500-316X-2023-11-5-63-70

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Нелинейный магнитоэлектрический эффект в кольцевой композитной гетероструктуре

В.И. Мусатов [@],
Ф.А. Федулов,
Д.В. Савельев,
Е.В. Болотина,
Л.Ю. Фетисов

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[@] Автор для переписки, e-mail: musatov.v.i@mail.ru

Резюме

Цели. Актуальность исследования магнитоэлектрических (МЭ) характеристик кольцевых гетероструктур «ферромагнетик-пьезоэлектрик» обусловлена созданием на их основе МЭ-устройств с улучшенными характеристиками. Целью настоящей работы является детальное исследование нелинейного МЭ-эффекта в кольцевой композитной гетероструктуре на основе пьезокерамики цирконата-титаната свинца (ЦТС) и аморфного ферромагнитного (ФМ) сплава Metglas® при ее циркулярном намагничивании.

Методы. МЭ-эффект исследован методом низкочастотной модуляции магнитного поля. Возбуждающее переменное и постоянное магнитные поля смещения были созданы при помощи тороидальной катушки, намотанной на гетероструктуру, для циркулярного намагничивания ферромагнитного слоя.

Результаты. Обнаружен нелинейный МЭ-эффект, заключающийся в генерации высших гармоник МЭ-напряжения при возбуждении структуры циркулярными магнитными полями в нерезонансном режиме. Исследованы полевые и амплитудные зависимости первых трех гармоник МЭ-напряжения. Получены МЭ-коэффициенты для линейного МЭ-эффекта $\alpha^{(1)}$ = 5.2 мВ/(Э·см) и для нелинейного МЭ-эффекта $\alpha^{(2)}$ = 6 мВ/(Э^2·см) и $\alpha^{(3)}$ = 0.15 мВ/(Э^3·см) при частоте переменного магнитного поля f = 1 кГц. Максимумы амплитуд 1-й и 3-й гармоник наблюдались при постоянном магнитном поле H ~ 7 Э, что почти в два раза меньше, чем в планарных гетероструктурах ЦТС–Metglas®.

Выводы. Обнаружен и исследован нелинейный МЭ-эффект в кольцевой структуре на основе пьезокерамики ЦТС и аморфного ФМ-сплава Metglas®. Вследствие отсутствия размагничивания при циркулярном намагничивании замкнутого ФМ-слоя нелинейные МЭ-эффекты проявляются при значительно меньших амплитудах возбуждающего переменного и управляющего постоянного магнитных полей по сравнению с планарными гетероструктурами. Исследуемые кольцевые структуры могут быть использованы для создания на их основе умножителей частоты.

Ключевые слова: нелинейный магнитоэлектрический эффект, композитная гетероструктура, магнитострикция, ферромагнетик, пьезоэлектрический эффект

• **Поступила:** 23.03.2023 • **Доработана:** 27.04.2023 • **Принята к опубликованию:** 11.07.2023

Для цитирования: Мусатов В.И., Федулов Ф.А., Савельев Д.В., Болотина Е.В., Фетисов Л.Ю. Нелинейный магнитоэлектрический эффект в кольцевой композитной гетероструктуре. *Russ. Technol. J.* 2023;11(5):63–70. https://doi.org/10.32362/2500-316X-2023-11-5-63-70

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

Magnetoelectric (ME) effects in composite multiferroic heterostructures containing mechanically coupled piezoelectric (PE) and ferromagnetic (FM) layers manifest themselves in the appearance of sample polarization when it is placed in an external magnetic field due to a combination of magnetostriction of the FM layer and the piezoelectric effect of the PE layer [1]. Such structures are of great interest due to their practical application for the creation of highly sensitive sensors of alternating and permanent magnetic fields, actuators, radio signal processing devices, transformers, etc. [2–5].

There is currently increasing interest in the study of ring composite heterostructures consisting of PE and FM layers. The characteristics of a linear ME effect observed in such heterostructures based on lead zirconate titanate (PZT) piezoceramics and Ni-based FM layers and amorphous Metglas® FM alloy when magnetized by an external field in the plane or along the ring axis have been studied [6–8]. Tunable transformers [9] and inductors [10], in which the inductance is tunable by 1000% under the action of constant magnetic and electric fields, were fabricated based on the ME ring composite heterostructures. A number of current sensor designs based on ring-ME structures have been proposed [11].

There is a wide interest in the study of nonlinear effects in ME heterostructures, such as the generation of harmonics and combinational frequencies, bistability, hysteresis suppression [12], forming the basis the design of devices for frequency multiplication and alternating magnetic field spectrum analysis. Due to the significant influence of demagnetization effects in the FM layer, the study of such effects in planar heterostructures requires constant and alternating magnetic fields of the order of units to tens of Oe.

At the same time, circular magnetization of ring heterostructures seems a promising avenue of enquiry [13]. By such means, it may be possible to achieve a significant decrease in demagnetization effects at the same time as increasing the efficiency of ME conversion at reduced bias magnetic fields due to the closed magnetic flux in the ring FM layer. In connection with the above, the study of nonlinear ME effects in ring heterostructures during circular magnetization is of great interest. To the best of the present authors' knowledge, the present work represents the first study into nonlinear ME effects in ring heterostructures excited by circular magnetic fields. Here, the nonlinear ME effect of voltage harmonics generation in a two-layer ring heterostructure consisting of PZT-19 piezoceramics and Metglas® tape of amorphous magnetic alloy under circular magnetization by alternating and constant magnetic field was observed and studied.

SAMPLES AND MEASURING METHODOLOGY

A schematic representation and an appearance of the investigated two-layer heterostructure is depicted in Fig. 1. The investigated heterostructure comprises a two-layer ring with an outer layer made of PZT-19 piezoceramic (NII ELPA, Russia), bounded with an inner FM layer made of Metglas® 2605SA1 amorphous magnetic alloy (Metglas® Inc., USA). The ends of the FM layer have been overlapped. The layers were joined using a cyanoacrylate adhesive at a thickness of ~3 μm (Weiss, CA-500.200, Germany). The overlap of the opposite ends of the amorphous tape was less than 1 mm. The FM layer had a length L_m = 50.2 mm, thickness $a_m \approx 27$ μm, width w_m = 5 mm, saturation magnetostriction $\lambda_S \approx 25 \cdot 10^{-6}$, and magnetic permeability $\mu \approx 10^4$. Radially polarized piezoceramic ring of $Pb(Zr_xTi_{1-x})O_3$ (PZT) composition had an inner diameter of 16 mm, thickness a_p = 1 mm and width w_p = 5 mm. Ag-electrodes were deposited to the outer and inner surfaces of the PZT ring. PZT-19 piezoceramics is characterized by piezoelectric modulus values d_{31} = 175 pC/N and relative dielectric permeability ε = 1750.

Two toroidal coils of copper wire of 0.2 mm diameter having a number of evenly distributed windings N = 90 are wound on the ring. One coil generates a circular bias constant magnetic field H = 0–115 Oe by passing through it a current of I_{dc} = 0–5 A from an AKTAKOM APS-7305 supply source (AKTAKOM, Russia). A circular alternating magnetic field $h\cos(2\pi ft)$ having an amplitude up to h = 3.45 Oe and frequency f = 0–100 kHz was created by the second coil connected to an Agilent 33210A waveform generator (Agilent Technologies Inc., USA). The amplitude of the magnetic field was determined analytically using the formula $H = \frac{NI}{2\pi r}$, where r is the middle line of the toroidal coil,

N is the number of windings, and I is the current passing through the coil. Frequency spectra of the ME voltage were recorded using a Tektronix TDS3032B digital oscilloscope (Tektronix Inc., USA).

During the course of the study, voltage (u) dependencies and Fourier spectra were obtained at different excitation amplitudes (h) and constant magnetic fields (H), from which the field and amplitude dependencies of the 1st, 2nd, and 3rd harmonics of output ME voltage were obtained.

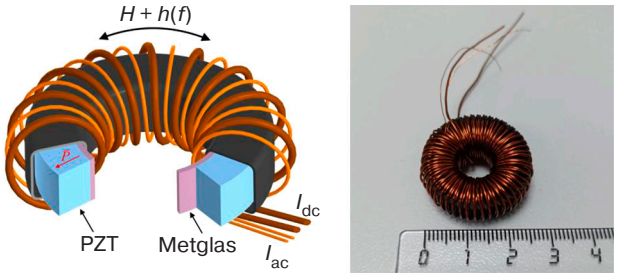


Fig. 1. Schematic representation (a) and appearance (b) of a PZT–Metglas® ring heterostructure with toroidal coils

RESULTS

At excitation of the structure by an alternating magnetic field of amplitude h = 0–3.45 Oe with a frequency f_1 = 1 kHz in a circular constant field H = 0–115 Oe, a nonlinear effect of ME voltage harmonics excitation was detected. The measurements were carried out away from the resonant frequency of the structure $f_0 \approx 54.2$ kHz. In the spectrum measured at H = 0.45 Oe and h = 3.45 Oe (Fig. 2), sixteen peaks are observed corresponding to the ME voltage harmonics with frequencies $f_n = f_1 \cdot n$, where n is an integer. The amplitudes of the first three peaks were $u^{(1)}$ = 1.8 mV, $u^{(2)}$ = 7.15 mV, and $u^{(3)}$ = 0.6 mV, respectively.

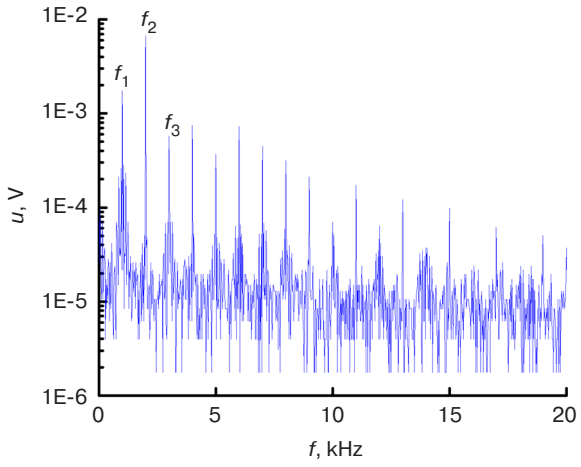


Fig. 2. Fourier spectrum of ring heterostructure ME voltage at h = 3.45 Oe with a frequency f_1 = 1 kHz and H = 0.45 Oe

Values of ME coefficients of the 1st, 2nd, and 3rd harmonics can be estimated as $\alpha^{(n)} = u^{(n)}/(h^n a_p)$. Hence, $\alpha^{(1)}$ = 5.2 mV/(Oe·cm), $\alpha^{(2)}$ = 6 mV/(Oe^2·cm), and $\alpha^{(3)}$ = 0.15 mV/(Oe^3·cm), respectively. In the absence of a constant magnetic field, only even harmonics were observed. These harmonics can be attributed to the nonlinear dependence of the magnetostriction $\lambda(H)$ of the FM layer on the constant magnetic field H [13].

Figure 3 shows the dependencies of the amplitudes of the 1st, 2nd, and 3rd harmonics on the constant magnetic field H at h = 3.45 Oe based on the obtained Fourier spectra of the ME voltage. The shape of curve $u^{(1)}(H)$ for the 1st harmonic (Fig. 3a) is typical for the linear ME effect. The amplitude maximum $u^{(1)}_{max} = 13.4$ mV achieved in the field $H^{(1)}_m \approx 6.8$ Oe followed by a monotonic decrease as the constant magnetic field increases up to H = 115 Oe. The maximum ME voltage corresponds to the highest value of the piezomagnetic coefficient of the FM layer $\lambda^{(1)}(H) = \partial\lambda / \partial H\big|_H$, where $\lambda(H)$ is the field dependence of the magnetostriction of the FM layer.

Figure 3b shows the dependencies of the amplitudes of the 2nd $u^{(2)}(H)$ and 3rd $u^{(3)}(H)$ harmonics on the constant magnetic field H. The amplitude of the second harmonic $u^{(2)}_{max} = 7.15$ mV is highest at $H^{(2)}_m = 0$, and then monotonically decreases with increasing field to reach a minimum $u^{(2)}_{min} = 0.1$ mV at $H = H^{(2)}_m \approx 5.6$ Oe. Note that the graph $u^{(2)}(H)$ does not show a local minimum or subsequent local maximum typical for the field dependence of the 2nd harmonic in planar heterostructures [14, 15].

The amplitude of the 3rd harmonic at H = 0 has a value $u^{(3)} \approx 0.4$ mV. When the magnetic field increases, two local maxima $u^{(3)}_{1max} = 1.2$ mV and $u^{(3)}_{2max} = 1.4$ mV are observed in the fields $H^{(3)}_{1m} \approx 0.9$ Oe and $H^{(3)}_{2m} \approx 1.4$ Oe, respectively. Subsequently, the amplitude of the 3rd harmonic monotonically decreases to zero at $H \sim 5.6$ Oe.

Figure 4 shows the dependencies of the ME voltage amplitudes of the 1st, 2nd, and 3rd harmonics on the amplitude of the excitation alternating magnetic field h. Measurements were performed in the optimal magnetic fields H_m obtained from the curves shown in Fig. 3: $H^{(1)}_m \approx 6.8$ Oe, $H^{(2)}_m = 0$ Oe, and $H^{(3)}_m = 1.4$ Oe, respectively.

The amplitude of the 1st harmonic can be seen to linearly depend on the value of h throughout the range of amplitudes of alternating magnetic fields h = 0–3.45 Oe. The amplitude of the 2nd harmonic increased in proportion to h^2 across the entire range while the amplitude of the 3rd harmonic is proportional to h^3. This type of dependence corresponds to theoretical calculations of the ME voltage $u^{(n)} \sim h^n$ [14], where n is the sequence number of the harmonic; h is the amplitude of the alternating magnetic field.

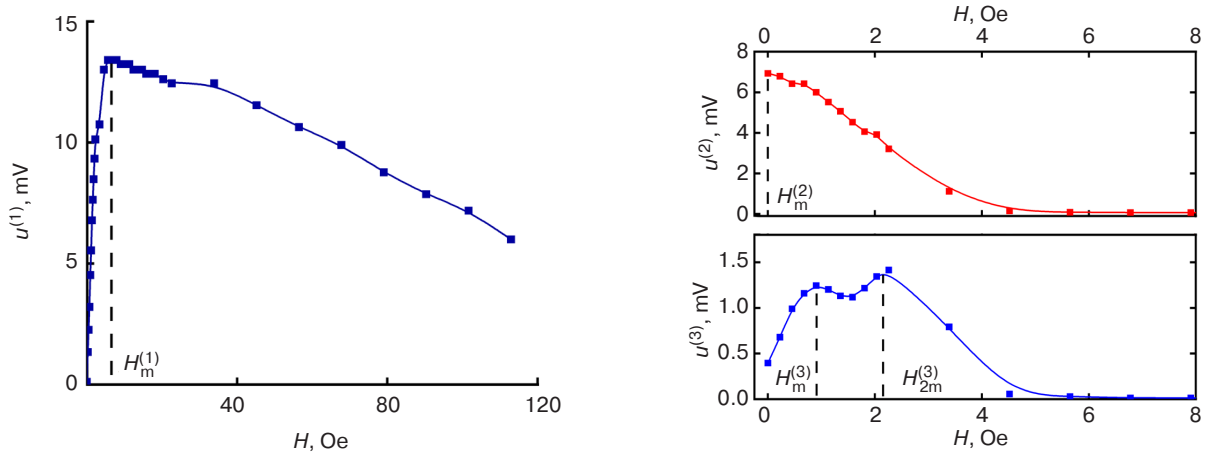


Fig. 3. Dependencies of ME voltage amplitudes of the 1st (a), 2nd, and 3rd (b) harmonics of the ring heterostructure ME voltage on the constant magnetic field *H*

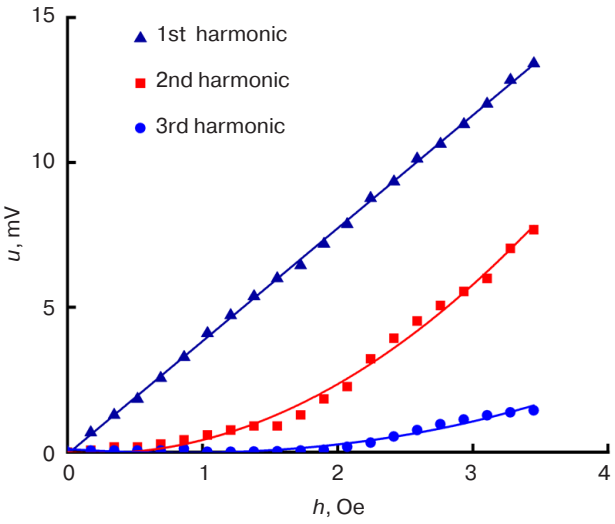


Fig. 4. Dependencies of the ME voltage amplitudes of the 1st, 2nd, and 3rd harmonics of the ME voltage on the field amplitude *h* at $H_m^{(1)} = 6.8$ Oe, $H_m^{(2)} = 0$ Oe, and $H_m^{(3)} = 1.4$ Oe, respectively. The solid lines are the power-law approximation

DISCUSSION

The appearance of the ME voltage harmonics is due to the nonlinear dependence of the magnetostriction of the FM layer on the magnetic field. By decomposing the function describing the magnetostriction $\lambda(H)$ into a Taylor series in the vicinity of the field H at $H << h$, we obtain the equation:

$$\lambda(H) = \lambda(H_0) + qh + \frac{ph^2}{2} + \frac{mh^3}{4}..., \qquad (1)$$

where q, p, and m are the 1st, 2nd, and 3rd derivatives of magnetostriction by magnetic field, respectively.

The expression for the ME voltage can be written in the following form

$$u(H) = Ad\lambda(H), \qquad (2)$$

where A is the coefficient that depends on the parameters of the PE and FM layers and the method of its constraint; d is the piezoelectric modulus module of the PE layer.

Substituting Eq. (1) in Eq. (2) and taking into account that the alternating magnetic field is given as $h = h_0 \cos(2\pi ft)$, we obtain:

$$u(H) = u_0(H_0) + u_1 \cos(2\pi ft) + \\ + u_2 \cos(4\pi ft) + u_3 \cos(6\pi ft)\ldots. \qquad (3)$$

The first term in (3) denotes the constant component of the ME voltage, the second term denotes the generated ME voltage at the excitation field frequency, and the third and fourth terms describe the generation of the 2nd and 3rd harmonics of the ME voltage, respectively [16].

To the authors' knowledge, there are currently no published works in which multiple harmonics of higher orders have been observed in a planar heterostructure of similar composition. Although ~100 harmonics were previously observed in the langatate–Metglas® heterostructure [17], these occurred at a larger pumping amplitude h = 20–25 Oe of the alternating field.

Some peculiarities can be noted in the results of the study of the nonlinear MPE effect in the FM–PE ring structure. In contrast to ring structures, the magnetic field inside the FM layer H_{in} in planar structures is connected with the external field H by the relation $H_{in} = \frac{H}{1 + N\mu}$, where N is the demagnetization factor and μ is the magnetic permeability of the FM layer [18].

Hence, it follows that demagnetization reduces both the constant and the excitation alternating field inside the FM layer in a planar heterostructure. In a ring heterostructure with circular magnetization, the magnetic flux in the FM layer is closed ($N \approx 0$). Therefore, inside the FM layer $H_{in} \approx H$ and $h_{in} \approx h$, which leads to a change in the shape of the magnetostriction dependence on the external field $\lambda(H)$, in particular, to a decrease in the magnetostriction saturation field.

In the studied structure (Fig. 3), the maximum of the 1st harmonic was observed in the field $H_m^{(1)} \approx 6.8$ Oe. Amplitude $u^{(1)}$ reached 90% of the maximum in the field ~4 Oe, which is several times less than in planar PZT–Metglas® heterostructures [19, 20].

Due to the absence of demagnetization in the ring heterostructure, the nonlinear ME effects appear at lower amplitudes of the external excitation field h. Apparently, this is associated with a change in the shape of the $\lambda(H)$ dependence in the ring structure. In the dependence plot of the amplitude of the 2nd harmonic, the local maximum has disappeared, while in the dependence plot of the 3rd harmonic, it remains [17].

Values of ME coefficients of the 2nd and 3rd harmonics can be estimated as $\alpha^{(2)} = u^{(2)}/(h^2 a_p) = 6$ mV/(Oe^2·cm) and $\alpha^{(3)} = u^{(3)}/(h^3 a_p) = 0.34$ mV/(Oe^3·cm), respectively. The obtained ME coefficient values are comparable to the ME coefficients in the PZT–Metglas® planar structure of similar composition $\alpha^{(2)} = 9.6$ mV/(Oe^2·cm) and $\alpha^{(3)} = 0.4$ mV/(Oe^3·cm) measured at the same excitation field frequency [16].

CONCLUSIONS

In the present work, a nonlinear ME effect of the generation of higher harmonics of the ME voltage was observed when a ring two-layered PZT–Metglas® heterostructure is excited by circular magnetic fields in non-resonant mode. In the course of the study, we observed sixteen harmonics of the ME voltage in nonlinear mode when the structure is excited by an alternating magnetic field with a frequency of 1 kHz. Nonlinear ME effects in the ring heterostructure were detected in smaller magnetic fields compared to planar structures; this is associated with almost zero demagnetization of the FM layer. In particular, the maximum amplitudes of the 1st and 3rd harmonics were observed at $H \sim 7$ Oe, which is almost half as much as in planar PZT–Metglas® heterostructures.

The obtained results based on the investigated ring ME heterostructures demonstrate the possibility of creating efficient solid-state frequency multipliers excited by alternating magnetic fields of the order of oersted units.

Acknowledgments

This work was financially supported by the Russian Science Foundation, grant No. 19-79-10128-P. The samples were made with the financial support of grant "For young scientists 2022–2023 RTU MIREA" (NICh-57). Measurements of Fourier spectra were performed using equipment of the RTU MIREA Center for Collective Use.

Authors' contributions

V.I. Musatov—preparing test samples and experimental setup, conducting research, data curation, preparing graphic information, and writing the text of the article.

F.A. Fedulov—preparing test samples and experimental setup, data checking, and writing the text of the article.

D.V. Savelev—conducting research, data curation, description of the research methodology.

E.V. Bolotina—conducting research, preparing graphic information.

L.Y. Fetisov—data checking and writing the text of the article.

REFERENCES

1. Kopyl S., Surmenev R., Surmeneva M., Fetisov Y., Kholkin A. Magnetoelectric effect: principles and applications in biology and medicine – a review. *Materials Today Bio.* 2021;12:100149. https://doi.org/10.1016/j.mtbio.2021.100149
2. Azam T., Bukhari S.H., Liaqat U., Miran W. Emerging Methods in Biosensing of Immunoglobin G – A Review. *Sensors.* 2023;23(2):676. https://doi.org/10.3390/s23020676
3. Herrera D.L., Kruk R., Leistner K., Sort J. Magnetoelectric materials, phenomena, and devices. *APL Materials.* 2021;9(5):050401. https://doi.org/10.1063/5.0053631
4. Fetisov Y.K., Chashin D.V. Magnetoelectric coil-free voltage transformer based on monolithic ferrite-piezoelectric heterostructure. *Sensors and Actuators A: Physical.* 2022;344:113737. https://doi.org/10.1016/j.sna.2022.113737
5. Tu C., Chu Z.Q., Spetzler B., Hayes P., et al. Mechanical-resonance-enhanced thin-film magnetoelectric heterostructures for magnetometers, mechanical antennas, tunable RF inductors, and filters. *Materials.* 2019;12(14):2259. https://doi.org/10.3390/ma12142259
6. Xu L., Yan Y., Qiao L., Wang J., Pan D., Yang S., Volinsky A.A. Layer thickness and sequence effects on resonant magnetoelectric coupling in Ni/Pb(Zr,Ti)O_3 cylindrical composites. *Materials Lett.* 2016;185:13–16. https://doi.org/10.1016/j.matlet.2016.08.042
7. Yakubov V., Xu L., Volinsky A.A., Qiao L., Pan D. Edge geometry effects on resonance response of electroplated cylindrical Ni/PZT/Ni magnetoelectric composites. *AIP Advances*. 2017;7(8):085305. https://doi.org/10.1063/1.4998947

8. Giang D.T.H., Tam H.A., Khanh V.T.N., Vinh N.T., Tuan P.A., Van Tuan N.V., et al. Magnetoelectric vortex magnetic field sensors based on the metglas/PZT laminates. *Sensors*. 2020;20(10):2810. https://doi.org/10.3390/s20102810
9. Zhang S., Leung C.M., Kuang W., Or S.W., Ho S.L. Concurrent operational modes and enhanced current sensitivity in heterostructure of magnetoelectric ring and piezoelectric transformer. *J. Appl. Phys*. 2013;113(17):17C733. https://doi.org/10.1063/1.4801390
10. Fetisov L.Y., Saveliev D.V., Chashin D.V., Gladyshev I.V., Fetisov Y.K. Circular Magnetoelectric Heterostructure Based Inductor Tuned with Magnetic and Electric Fields. *J. Commun. Technol. Electron.* 2021;66(12):1402–1412. https://doi.org/10.1134/S1064226922020036
11. Ge B., Zhang J., Zhang Q., Filippov D.A., Wu J., Tao J., et al. Ultra-low anisotropy magnetoelectric sensor in Ferrite/piezoelectric toroidal composites. *J. Magn. Magn. Mater.* 2022;564(Part 2):170115. https://doi.org/10.1016/j.jmmm.2022.170115
12. Burdin D.A., Chashin D.V., Ekonomov N.A., Fetisov Y.K., Stashkevich A.A. High-sensitivity dc field magnetometer using nonlinear resonance magnetoelectric effect. *J. Magn. Magn. Mater.* 2016;405(52):244–248. https://doi.org/10.1016/j.jmmm.2015.12.079
13. Wu G., Zhang R. Giant circumferential magnetoelectric effect in Pb(Zr,Ti)O_3/Mn-Zn-ferrite cylindrical composite. *Sensors and Actuators A: Physical*. 2021;330(14):112845. https://doi.org/10.1016/j.sna.2021.112845
14. Fetisov L.Y., Baraban I.A., Fetisov Y.K., Burdin D.A., Vopson M.M. Nonlinear magnetoelectric effects in flexible composite ferromagnetic–Piezopolymer structures. *J. Magn. Magn. Mater.* 2017;441:628–634. https://doi.org/10.1016/j.jmmm.2017.06.013
15. Burdin D., Chashin D., Ekonomov N., Gordeev S., Fetisov Y. Nonlinear magnetoelectric effect in a layered ferromagnetic-piezoelectric heterostructure excited by transverse magnetic field. *Appl. Phys. Lett.* 2020;116(7):072901. https://doi.org/10.1063/1.5136088
16. Burdin D., Chashin D., Ekonomov N., Fetisov L., Fetisov Y., Shamonin M. DC magnetic field sensing based on the nonlinear magnetoelectric effect in magnetic heterostructures. *J. Phys. D: Appl. Phys.* 2016;49(37):375002. https://doi.org/10.1088/0022-3727/49/37/375002
17. Fetisov L.Y., Burdin D.A., Ekonomov N.A., Chashin D.V., Zhang J., Srinivasan G., et al. Nonlinear magnetoelectric effects at high magnetic field amplitudes in composite multiferroics. *J. Phys. D: Appl. Phys.* 2018;51(15):154003. http://doi.org/10.1088/1361-6463/aab384
18. Joseph R.I., Schlömann E. Demagnetizing field in nonellipsoidal bodies. *J. Appl. Phys*. 1965;36(5):1579–1593. https://doi.org/10.1063/1.1703091
19. Fang F., Zhao C.P., Yang W. Thickness effects on magnetoelectric coupling for Metglas/PZT/Metglas laminates. *Science China: Physics, Mechanics and Astronomy*. 2011;54(4):581–585. https://doi.org/10.1007/s11433-011-4268-2
20. Li M., Wang Y., Hasanyan D., Li J., Viehland D. Giant Converse magnetoelectric effect in multi-push-pull mode Metglas/Pb(Zr,Ti)O_3/metglas laminates. *Appl. Phys. Lett.* 2012;100(13):132904. https://doi.org/10.1063/1.3698114

About the authors

Vladimir I. Musatov, Postgraduate Student, Department of Nanoelectronics, Institute for Advanced Technologies and Industrial Programming, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: musatov.v.i@mail.ru. Scopus Author ID 57416814900, https://orcid.org/0000-0002-2995-8824

Fedor A. Fedulov, Cand. Sci. (Eng.), Researcher, Scientific and Educational Center "Magnetoelectric materials and devices," MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: ostsilograf@yandex.ru. Scopus Author ID 57194284263, https://orcid.org/0000-0003-2188-0011

Dmitrii V. Savelev, Research Engineer, Scientific and Educational Center "Magnetoelectric materials and devices," MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: dimsav94@gmail.com. Scopus Author ID 57196479660, ResearcherID D-8952-2019, RSCI SPIN-code 3273-4160, https://orcid.org/0000-0001-7762-9198

Ekaterina V. Bolotina, Student, Department of Nanoelectronics, Institute for Advanced Technologies and Industrial Programming, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: ekaterina.bolotina1@mail.ru. https://orcid.org/0000-0003-1004-2821

Leonid Y. Fetisov, Dr. Sci., Professor, Department of Nanoelectronics, Institute for Advanced Technologies and Industrial Programming, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: fetisovl@yandex.ru. Scopus Author ID 26431336600, ResearcherID D-1163-2013, RSCI SPIN-code 9788-0680, https://orcid.org/0000-0002-3699-4321

Об авторах

Мусатов Владимир Иванович, аспирант, кафедра наноэлектроники Института перспективных технологий и индустриального программирования, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: musatov.v.i@mail.ru. Scopus Author ID 57416814900, https://orcid.org/0000-0002-2995-8824

Федулов Федор Александрович, к.т.н., научный сотрудник, Научно-образовательный центр «Магнитоэлектрические материалы и устройства», ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: ostsilograf@yandex.ru. Scopus Author ID 57194284263, https://orcid.org/0000-0003-2188-0011

Савельев Дмитрий Владимирович, инженер-исследователь, Научно-образовательный центр «Магнитоэлектрические материалы и устройства», ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: dimsav94@gmail.com. Scopus Author ID 57196479660, ResearcherID D-8952-2019, SPIN-код РИНЦ 3273-4160, https://orcid.org/0000-0001-7762-9198

Болотина Екатерина Витальевна, студент, кафедра наноэлектроники Института перспективных технологий и индустриального программирования, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: ekaterina.bolotina1@mail.ru. https://orcid.org/0000-0003-1004-2821

Фетисов Леонид Юрьевич, д.ф.-м.н., доцент, профессор кафедры наноэлектроники Института перспективных технологий и индустриального программирования, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: fetisovl@yandex.ru. Scopus Author ID 26431336600, ResearcherID D-1163-2013, SPIN-код РИНЦ 9788-0680, https://orcid.org/0000-0002-3699-4321

Translated from Russian into English by Lyudmila O. Bychkova
Edited for English language and spelling by Thomas A. Beavitt

Micro- and nanoelectronics. Condensed matter physics
Микро- и наноэлектроника. Физика конденсированного состояния

UDC 681.2.0821.083:535.2:543.063
https://doi.org/10.32362/2500-316X-2023-11-5-71-80

RESEARCH ARTICLE

Synchrotron radiation of a single electron application for optical spectroradiometry

Alexander S. Sigov [1], Evgenij R. Lazarenko [2], Natalia B. Golovanova [1], Olga A. Minaeva [1, @], Sergei I. Anevsky [1], Roman V. Minaev [3], Pavel Yu. Pushkin [1]

[1] *MIREA – Russian Technological University, Moscow, 119454 Russia*
[2] *Federal Agency for Technical Regulation and Metrology (Rosstandart), Moscow, 125039 Russia*
[3] *Elektrosteklo, Moscow, 119571 Russia*
[@] *Corresponding author, e-mail: minaeva_o@mirea.ru*

Abstract

Objectives. The investigations of optical radiation sources and metrological detector characteristics in the infrared (IR), visible, and air ultraviolet (UV) spectral regions are partially based on the unique metrological properties of synchrotron radiation. The aim of this work is to develop a high-precision method for determining the storage ring accelerated electron number with synchrotron radiation of a single electron to establish spectroradiometry and photometry units.

Methods. By determining the number of accelerated electrons, any storage ring can be used to calculate the synchrotron radiation characteristics at wavelengths of many large then the critical wavelength in the visible, air UV, and IR regions of the spectrum. This makes it possible to determine the main metrological characteristics normalized to the number of electrons, such as luminous intensity, luminance, illuminance, radiant power, radiance, irradiance, etc., regardless of the energy of the electrons.

Results. When applying the method for determining the number of accelerated electrons at low currents of the electronic storage ring, a total standard deviation of the number of accelerated electrons is less than 0.01% for an exposure range of the CCD matrix from 10^{-2} to $3 \cdot 10^3$ s in a wide dynamic range of $1–10^{10}$ electrons per orbit.

Conclusions. The use of a CCD-based radiometer-comparator calibrated by responsivity on a synchrotron radiation source is particularly relevant in monitoring luminance contrast thresholds and spatial distribution of object and background brightness, as well as determining metrological characteristics of optoelectronic measuring instruments, including CCD cameras, radiometers, spectroradiometers and photometers.

Keywords: synchrotron radiation, responsivity threshold, luminance contrast, luminance spatial distribution, measuring instruments

• **Submitted:** 05.06.2023 • **Revised:** 23.06.2023 • **Accepted:** 11.07.2023

For citation: Sigov A.S., Lazarenko E.R., Golovanova N.B., Minaeva O.A., Anevsky S.I., Minaev R.V., Pushkin P.Yu. Synchrotron radiation of a single electron application for optical spectroradiometry. *Russ. Technol. J*. 2023;11(5):71–80. https://doi.org/10.32362/2500-316X-2023-11-5-71-80

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Использование синхротронного излучения отдельного электрона для спектрорадиометрии оптического диапазона

А.С. Сигов [1], Е.Р. Лазаренко [2], Н.Б. Голованова [1], О.А. Минаева [1, @], С.И. Аневский [1], Р.В. Минаев [3], П.Ю. Пушкин [1]

[1] *МИРЭА – Российский технологический университет, Москва, 119454 Россия*
[2] *Федеральное агентство по техническому регулированию и метрологии (Росстандарт), Москва, 125039 Россия*
[3] *ООО «Электростекло», Москва, 119571 Россия*
[@] *Автор для переписки, e-mail: minaeva_o@mirea.ru*

Резюме

Цели. Исследование метрологических характеристик источников и приемников оптического излучения в инфракрасной (ИК), видимой и ближней ультрафиолетовой (УФ) областях спектра в значительной мере основано на использовании уникальных метрологических свойств синхротронного излучения. Целью работы является развитие высокоточного метода определения числа ускоренных электронов накопительного кольца, основанного на использовании синхротронного излучения отдельного электрона для воспроизведения единиц величин спектрорадиометрии и фотометрии.

Методы. Определение числа ускоренных электронов позволяет для любого накопительного кольца рассчитать характеристики синхротронного излучения на длинах волн, намного бóльших критической длины волны, т.е. в видимой, ближней УФ- и ИК-областях спектра. Это обеспечивает возможность, вне зависимости от энергии электронов, определить нормированные на число электронов основные метрологические характеристики, такие как сила света, яркость, освещенность, сила излучения, энергетическая освещенность, энергетическая яркость и другие.

Результаты. Применение метода определения числа ускоренных электронов при малых токах электронного накопительного кольца позволяет обеспечить в широком динамическом диапазоне $1–10^{10}$ электронов на орбите значение суммарного среднеквадратического отклонения не более 0.01% для диапазона экспозиций приборов с зарядовой связью (ПЗС-матрицы) от 10^{-2} до $3 \cdot 10^3$ с.

Выводы. Применение радиометра-компаратора на основе телескопа с ПЗС-матрицей, откалиброванного по чувствительности на источнике синхротронного излучения, особенно актуально при контроле пороговых значений яркостного контраста и пространственного распределения яркости объекта и фона, а также определения метрологических характеристик оптико-электронных средств измерений, включая ПЗС-камеры, радиометры, спектрорадиометры и фотометры.

Ключевые слова: синхротронное излучение, порог чувствительности, яркостный контраст, пространственное распределение яркости, средства измерений

• **Поступила:** 05.06.2023 • **Доработана:** 23.06.2023 • **Принята к опубликованию:** 11.07.2023

Для цитирования: Сигов А.С., Лазаренко Е.Р., Голованова Н.Б., Минаева О.А., Аневский С.И., Минаев Р.В., Пушкин П.Ю. Использование синхротронного излучения отдельного электрона для спектрорадиометрии оптического диапазона. *Russ. Technol. J.* 2023;11(5):71–80. https://doi.org/10.32362/2500-316X-2023-11-5-71-80

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

The solution of topical spectroradiometry and photometry problems is based on the use of reference sources and receivers of radiation in the infrared (IR), visible and ultraviolet (UV) regions of the spectrum [1–3]. Spectroradiometry methods play an important role in various fields of science and technology, including plasma diagnostics, photochemistry and photobiology, as well as astronavigation, Earth remote sensing, solar activity diagnostics, the study of fluid properties in the terahertz range, localization of remote objects, and nanoelectronics [4, 5].

Despite the success of national metrology institutes (NMIs) in creating spectroradiometry standards, significant difficulties are faced when assessing the quality of radiometers and photometers used in scientific research, as well as when evaluating efficiency and risks involved in UV radiation and photometric characteristics of emitters in areas of production, transport, labor protection, sanitary and epidemiological surveillance, etc. [6–8]. International key comparisons of K2c absolute spectral sensitivity of reference detectors UV radiation conducted by the International Bureau of Weights and Measures (BIPM) showed that, out of 14 participants, NMIs Physikalisch-Technische Bundesanstalt[1] (Germany), VNIIOFI[2] (Russia), and NIST[3] (USA) meet accuracy requirements in the spectral range of 200–400 nm [9, 10].

Blackbody models and synchrotron radiation sources are used as primary standards for photometry and spectroradiometry units in the infrared, visible, and air-ultraviolet regions of the spectrum [11, 12]. However, due to the limitation of the radiance temperature of the blackbody model to 3500 K, it is not possible to extend the working spectral range into the short-wave UV region of the spectrum, while the radiance temperature of synchrotron radiation is regulated by changing the energy of the electrons from thousands to tens of millions of degrees Kelvin.

The determination of photometry and spectroradiometry units by primary standards are associated with a number of problems on which leading NMIs have been working for many years. For example, the blackbody model requires a determination of the radiation temperature, accurate registration of weak fluxes of UV radiation against the background of powerful IR radiation, and ensuring the equal luminance of the emitting area. When using a synchrotron as a reference emitter, it is necessary to ensure the accuracy of electron beam diagnostics.

When using cryogenic radiometers as primary reference receivers, the main errors are related to low responsivity and heat exchange between the receiving cavity, housing and superconducting elements, as well as the need to compare radiation fluxes across a wide dynamic range [13]. The high intensity of synchrotron radiation, absence of lines in the spectrum, and high radiance temperature allow the use, along with a cryogenic radiometer, of an ionization chamber and a Golay detector [14]. In addition, synchrotron radiation can be used to conduct metrological studies of the characteristics of multilayer surface coatings and calculate the spectral responsivity of secondary reference radiation detectors using the dependence of detectors signals on the energy of accelerated electrons [15, 16].

Thus, in order to solve the problems arising in the investigations of sources and detectors metrological characteristics in the infrared, visible, and air-ultraviolet regions, the development of techniques based on the use

[1] https://www.ptb.de/cms/en.html. Accessed June 05, 2023.

[2] All-Russian Research Institute of Optical and Physical Measurements. https://www.vniiofi.ru/ (in Russ.). Accessed June 05, 2023.

[3] National Institute of Standards and Technology. https://www.nist.gov/. Accessed June 05, 2023.

of the unique metrological properties of synchrotron radiation is of particular interest.

SYNCHROTRON RADIATION RADIOMETRY

The spectral characteristics of synchrotron radiation are calculated with high-accuracy measurements of the electron orbital radius, energy, and number of accelerated electrons [17]. The distribution of the spectral radiance of synchrotron radiation over the emitting area is described by the following expression:

$$L(x,y)=\frac{27Ne^2c}{32\pi^3R^3D\sigma_{x'}\sigma_{y'}}\left(\lambda_{\tilde{n}}/\lambda\right)^4\gamma^8\left[1+(\gamma\Psi)^2\right]^2\times$$
$$\times\left\{K^2{}_{2/3}(\xi)+K^2{}_{1/3}(\xi)(\gamma\Psi)^2/\left[1+(\gamma\Psi)^2\right]\right\}\times \quad (1)$$
$$\times\exp\left[-\frac{\left(x'-x'_0\right)^2}{2\sigma_{x'}{}^2}-\frac{\left(y'-y'_0\right)^2}{2\sigma_{y'}{}^2}\right],$$

where L is the spectral radiance of the synchrotron radiation (W/m^3·sr); x' and y' are coordinates of two-dimensional emitting region in the orbital and perpendicular planes; x'_0 and y'_0 are coordinates of maximum of synchrotron radiation spectral radiance distribution; N is the number of accelerated electrons; $\gamma = E/E_0$ is the relativistic factor; E is the energy of the accelerated electron, E_0 = 0.511 MeV is the rest energy of the electron; e is the charge of the electron; c is the speed of light; R is the radius of the electron orbit in the radiation point; D is the integral of the two-dimensional Gaussian distribution; $\sigma_{x'}$ and $\sigma_{y'}$ is the standard deviation of the spatial distribution of the spectral radiance of the electron clot in the orbit plane and perpendicular plan; $\lambda_c = (4/3)\pi R\gamma^{-3}$ is the critical wavelength; λ is the wavelength; Ψ is the deflection angle from the orbital plane; $K_{1/3}$ and $K_{2/3}$ are modified Bessel functions (McDonald functions); $\xi = [\lambda_c/(2\lambda)][1 + (\gamma\Psi)^2]^{3/2}$ is the argument.

The two-dimensional Gaussian distribution integral D is calculated according to the formula:

$$D=\int\int\frac{1}{2\pi\sigma_{x'}\sigma_{y'}}\exp\left[-\frac{\left(x'-x'_0\right)^2}{2\sigma_{x'}{}^2}-\frac{\left(y'-y'_0\right)^2}{2\sigma_{y'}{}^2}\right]dx'dy'.$$

Figure 1 shows the dependence of the modified Bessel functions (McDonald functions) $K_{1/3}$, $K_{2/3}$ on the wavelength normalized to the critical wavelength λ_c of the synchrotron radiation spectrum [17].

Equation (1) describes the spectral and angular distribution of the spectral radiance in the polarization σ- and π-components of synchrotron radiation. The polarization vector of the σ-component, which lies in the plane of the electron orbit perpendicular to the induction vector of the deflecting magnetic field, is described by the first term of Eq. (1), while the polarization vector of the π-component lying in the plane parallel to the magnetic field induction vector is described by the second term.

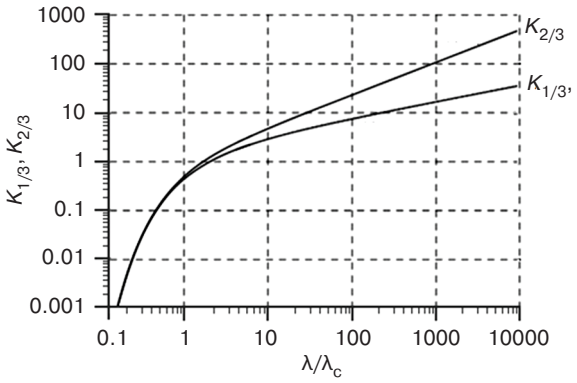


Fig. 1. Dependence of modified Bessel functions (McDonald functions) $K_{1/3}$, $K_{2/3}$ on wavelength

A universal function $f(\lambda/\lambda_0)$ of the spectral distribution of the synchrotron radiation flux is obtained by integrating Eq. (1) over the emitting region of the electron clot and the deflection angle from the electron orbital plane [18].

The universal function shown in Fig. 2 is used to calculate the characteristics of synchrotron radiation by specifying the critical wavelength λ_c of the electron storage rings.

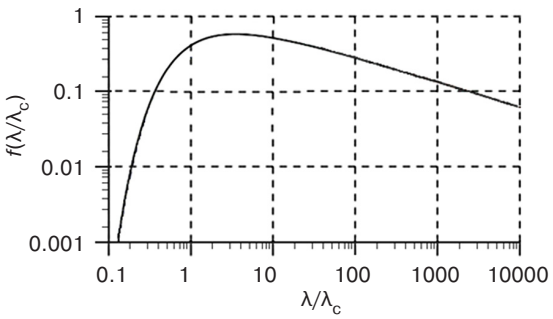


Fig. 2. Universal spectral distribution function of the synchrotron radiation flux

The energy of the electrons in the electron storage ring is determined by absolute measurements of the magnetic field induction at the emitting point of the electron orbit, by the wavelength of the Compton backscattering of photons on accelerated electrons, or by relative spectral measurements of the synchrotron radiation flux. The orbital radius of the electron storage ring is determined by the frequency of the accelerating high-frequency field.

The most important and challenging task of synchrotron radiation spectroradiometry is determining the number of accelerated electrons with high accuracy. At synchrotron radiation sources, the number of

accelerated NMI electrons is measured by comparing the spectral radiance of the blackbody model and the synchrotron in the visible region of the spectrum; i.e., for spectroradiometry, the relative spectral distribution of synchrotron radiation spectral radiance with absolutization in the visible range is used according to the black body model. Over a number of years, an accurate method for measuring the number of electrons with an error not exceeding a hundredth of a percent was developed using synchrotron radiation sources for spectroradiometry and photometry. The method of particle number measurements is based on Fourier-transformation of the signal of a telescope with a charge-coupled device (CCD) matrix proportional to the spectral radiance of synchrotron radiation; here, isolation of a single electron in a relativistic orbit of an electron storage ring is used with filtering of high-frequency spatial harmonics to ensure a wide linear range of responsivity of the detector.

RADIOMETER–COMPARATOR

A radiometer–comparator (made in Russia) includes an achromatic refractor telescope with a focal distance of 6 m and an aperture of 150 mm, containing a cooled CCD matrix, a set of interference filters for the UV, visible, and IR spectrum ranges, as well as filters for spectral responsivity correction according to the relative spectral luminous efficiency of monochromatic radiation [19]. The use of a cooled CCD matrix for dark signal subtraction ensures the possibility of measurements at decreasing beam current in a wide range of exposures from 0.1 to 4000 s. A general view of the comparator radiometer on the synchrotron radiation channel is shown in Fig. 3.

Fig. 3. General view of the radiometer comparator on the synchrotron radiation channel

The cooled CCD matrix comprises 3326 × × 2504 pixels of 5.4 × 5.4 μm with 16-bit sampling of signal values. The comparator radiometer is mounted on the white channel of the electron storage ring without monochromatization of synchrotron radiation at a fixed distance from the emitting point of the orbit.

MEASUREMENT OF THE NUMBER OF ELECTRONS IN THE ORBIT

Synchrotron radiation flux, which is characterized by a uniform angular distribution in the horizontal plane of the electron orbit, has a complex angular dependence of the intensity of polarization components and the degree of polarization in the vertical plane [18]. The influence of axial oscillations of the clot electrons further complicates the angular dependence of the intensity of polarization components in the vertical direction, as determined by the convolution of the angular distribution of the synchrotron radiation of an individual electron and the distribution of clot electrons by angles of deviation from the orbital plane. The choice of telescope aperture is determined by the need to integrate the flux of synchrotron radiation in the vertical plane over the receiving surface of the CCD matrix.

To exclude distortions in the distribution of the energy illumination recorded during the integration of the CCD matrix pixel signals, the pixel responsivity is equalized to eliminate the zone inhomogeneity of the telescope transmission coefficient. Figure 4 shows the distribution of CCD matrix pixel signals corresponding to the Gaussian distribution of spectral radiance over the emitting region of the synchrotron.

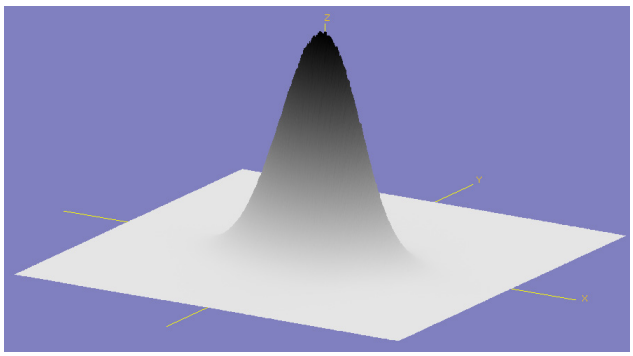


Fig. 4. Distribution of the CCD matrix pixel signals

The image on the CCD matrix is formed by multiple passage of electrons through a relativistic orbit. To reduce the influence of the CCD matrix noise when recording small fluxes of synchrotron radiation, forward and reverse Fourier transforms with filtration of high-frequency spatial harmonics are used. Calculation of energetic characteristics of synchrotron radiation of a single electron allows estimating the noise level and responsivity threshold for the comparator radiometer which does not exceed 10^{-10} W/(nm·m^2·sr). An important metrological property of synchrotron radiation is the perfect Gaussian distribution of the spectral radiance

over the radiating region of the electron storage ring due to the axial, radial, and phase oscillations of the clot electrons. As the number of electrons decreases, the integral signal decreases proportionally. In this case, the relative distribution of the spectral radiance over the emitting region does not change, allowing the Gaussian distribution to be used to obtain a uniform responsivity of the comparator radiometer over the receiving surface of the CCD matrix.

Adjustment of the electron beam current of the storage ring in a wide dynamic range is carried out while controlling the relative number of electrons by the integral flux of synchrotron radiation keeping the spectral and angular distribution constant. For the first time, registration of a single electron in orbit was carried out on a first-generation electron storage ring using a photomultiplier tube. When a single electron is singled out in orbit for high-precision reproduction of spectroradiometry and photometry units the spectral radiance of synchrotron radiation has Gaussian distribution due to high frequency of electron circulation, which allows higher spatial harmonics to be excluded during signal registration, radically reducing noise level and increasing accuracy when integrating pixel signals in a wide exposure range from milliseconds to one hour using direct and inverse Fourier transform with high-frequency filtering.

First, it is necessary to provide a responsivity linearity range of CCD matrix in ten orders of magnitude taking into account random and systematic errors of shutter response time, signal reading noise and possible saturation of pixel charges. The CCD matrix signal is measured when the exposure time changes, but at a constant value of the synchrotron radiation flux. The high stability of energy characteristics of synchrotron radiation at fixed electron beam current allows CCD matrix signals to be compared even at maximum exposures.

At the beginning of the cycle of absolute electron number measurements using the synchrotron radiation flux of a single electron, the choice of the minimum exposure of the CCD matrix corresponds to the maximum beam current of the electron storage ring. The sequential decrease in the electron-beam current along with the decrease in the signal proportional to the synchrotron radiation flux is compensated by increasing the exposure time. The difference in the signals obtained before the current decrease and following the exposure time increase indicates the nonlinear dependence of the CCD matrix responsivity on the exposure time, thus requiring the introduction of correction coefficients. To reduce the electron-beam current, a scriber is used, which allows electrons to be removed from the orbit of the storage ring in steps, so that the last electron remains on the orbit, which can be held for several hours.

Figure 5 shows the step change of the CCD matrix signals when the electrons are removed from the orbit of the electron storage ring.

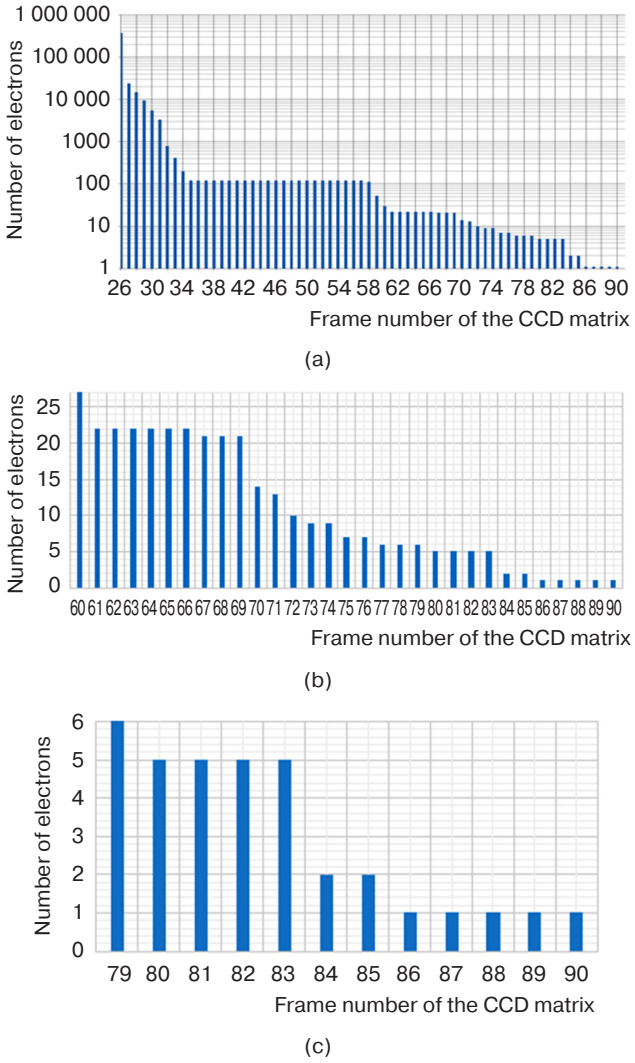


Fig. 5. Diagram of the step decrease of the CCD matrix signals when electrons are removed from the orbit: (a) diagram from 400 000 to 1 electron; (b) fragment of the diagram from 27 to 1 electron; (c) fragment of the diagram from 6 to 1 electron

Figure 5a shows the step change in the CCD matrix signal proportional to the number of electrons in the orbit on a logarithmic scale, starting at frame 26, with a signal of 400 000 relative units. According to the scriber, the signal decreased rapidly from frame 26 to frame 35 to 110 units. The stability of the synchrotron radiation flux was checked for several hours to estimate the standard deviation of the CCD matrix signals from frame 35 to frame 57. During this time, the electron clot did not lose a single electron; the signal standard deviation did not exceed 0.01%. Frames 58 and 60 illustrate a signal decrease by a factor of 5; from frame 61, synchrotron flux stability was repeated with estimation of the signal standard deviation.

Figure 5b shows on a linear scale the smallest reduction in signals at frames 67, 72, 78, and 80, which comprises a multiple of the step changes in signals corresponding to the synchrotron emission flux of a single electron.

Figure 5c portrays the results of comparing the signals of the last frames on a linear scale; in frame 80, one electron was removed and five electrons remained in the orbit. After checking the stability of the signals on frame 84, three electrons are removed using the scriber; the signal corresponds to the remaining two electrons. After frame 85, another electron is removed from the orbit, while in frame 86, the CCD matrix signal corresponds to one electron which is kept in the orbit for a long time. This allows the responsivity of the CCD matrix to be connected with the synchrotron radiation flux of a single electron. By determining the ratio of signals in the last frames shown in Fig. 5c, it becomes possible to precisely specify the number of electrons for each captured frame and select in the graphs the scale at which the relative unit of the signal corresponds to the synchrotron radiation flux of a single electron.

CONCLUSIONS

The method for determining the number of accelerated electrons of the storage ring based on the use of synchrotron radiation of a single electron is designed to diagnose an electron clot for establishing the spectroradiometry and photometry units based on the fundamental physical constants: electron charge and the speed of light in a vacuum. When determining the number of electrons from 1 to 10^{10} in the exposure range from 10^{-2} to $3 \cdot 10^{3}$ s, the total standard deviation does not exceed 0.01%.

According to the developed method, the spectral radiance of synchrotron radiation can be calculated for any electron storage ring at wavelengths significantly longer than the critical wavelength, i.e., in the visible, air UV, and IR regions of the spectrum. By this means, basic metrological characteristics can be determined normalized to the number of electrons regardless of electron energy, given in terms of luminance intensity, luminance, illuminance, radiant power, radiance, irradiance, etc. Thus, the spectroradiometric and photometric characteristics of synchrotron radiation in the visible, air UV, and IR regions of the spectrum, normalized to the number of accelerated electrons serve as constants for each electron storage ring.

The use of a comparator radiometer based on a telescope with a CCD matrix, which is calibrated by sensitivity at a synchrotron radiation source, is especially relevant for controlling the threshold values of luminance contrast and spatial luminance distribution of the object and background, as well as for determining the metrological characteristics of optoelectronic measuring instruments, including CCD cameras, radiometers, spectroradiometers and photometers.

Authors' contribution. All authors equally contributed to the research work.

REFERENCES

1. El'kin G.I., Salamatov V.Yu., Krutikov V.N., Novikov N.Yu. The state and main directions of work in the field of ensuring the uniformity of measurements in the Russian Federation. *Zakonodatel'naya i prikladnaya metrologiya = Legislative and Applied Metrology*. 2010;3:5–10 (in Russ.). Available from URL: https://metrob.ru/html/Stati/metrolob/napravlenia.html
2. Anevskii S.I., Minaeva O.A., Krutikov V.N., Minaev R.V., et al. *Metrological Support of Nanotechnologies and Nanoindustry Products*. Moscow: Logos; 2011. 592 p. (in Russ.).
3. Richter M., Ulm G. Metrology with Synchrotron Radiation. In: Jaeschke E., Khan S., Schneider J., Hastings J. (Eds.). *Synchrotron Light Sources and Free-Electron Lasers*. Springer; 2020. P. 1–35. https://doi.org/10.1007/978-3-319-04507-8_63-1
4. Shevelko A.P. Spectral instruments for x-ray and VUV plasma diagnostics. *J. Phys.: Conf. Ser*. 2018;1115(2):022042. http://doi.org/10.1088/1742-6596/1115/2/022042

СПИСОК ЛИТЕРАТУРЫ

1. Элькин Г.И., Саламатов В.Ю., Крутиков В.Н., Новиков Н.Ю. Состояние и основные направления работ в области обеспечения единства измерений в Российской Федерации. *Законодательная и прикладная метрология*. 2010;3:5–10. URL: https://metrob.ru/html/Stati/metrolob/napravlenia.html
2. Аневский С.И., Минаева О.А., Крутиков В.Н., Минаев Р.В. и др. *Метрологическое обеспечение нанотехнологий и продукции наноиндустрии.* М.: Издательство «Логос»; 2011. 592 с.
3. Richter M., Ulm G. Metrology with Synchrotron Radiation. In: Jaeschke E., Khan S., Schneider J., Hastings J. (Eds.). *Synchrotron Light Sources and Free-Electron Lasers*. Springer; 2020. P. 1–35. https://doi.org/10.1007/978-3-319-04507-8_63-1
4. Shevelko A.P. Spectral instruments for x-ray and VUV plasma diagnostics. *J. Phys.: Conf. Ser*. 2018;1115(2):022042. http://doi.org/10.1088/1742-6596/1115/2/022042

5. Torretti F., Liu F., Bayraktar M., Scheers J., Bouza Z., Ubachs W., Hoekstra R., Versolato O. Spectral characterization of an industrial EUV light source for nanolithography. *J. Phys. D: Appl. Phys.* 2019;53(5):055204. http://doi.org/10.1088/1361-6463/ab56d4
6. Berni L.Â., Vilela W.A., Beloto A.F., de Sena F.O. System for measuring the angular response of radiometers. In: *Proc. 8th Iberoamerican Optics Meeting and 11th Latin American Meeting on Optics, Lasers, and Applications*. 2013. V. 8785. https://doi.org/10.1117/12.2019888
7. Fryc I. Spectral correction of detector used in illuminance measurements. In: *Proc. 11th Slovak-Czech-Polish Optical Conference on Wave and Quantum Aspects of Contemporary Optics* (*SPIE 3820*). 1999. V. 3820. P. 343–348. https://doi.org/10.1117/12.353083
8. Han L. The principle and characteristics of photoelectric sensors. *Science and Technology Innovation and Application*. 2020;10:77–78.
9. Lutz W. The CCPR K2.c key comparison of spectral responsivity from 200 nm to 400 nm. *Metrologia.* 2014;51(6):S336. http://doi.org/10.1088/0026-1394/51/6/S336
10. Anevskii S.I., Zolotarevskii Y.M., Ivanov V.S., et al. Spectroradiometry of ultraviolet radiation. *Meas. Tech.* 2016;28(11):1216–1222. http://doi.org/10.1007/s11018-016-0873-9
[Original Russian Text: Anevskii S.I., Zolotarevskii Yu.M., Ivanov V.S., Krutikov V.N., Minaeva O.A., Minaev R.V. Spectroradiometry of ultraviolet radiation. *Izmeritel"naya Tekhnika*. 2015;11:26–30 (in Russ.).]
11. König S., Gutschwager B., Taubert R.D., Hollandt J. Metrological Characterization and Calibration of Thermographic Cameras for Quantitative temperature measurement. *Sens. Sens. Syst*. 2020;9(2):425–442. https://doi.org/10.5194/jsss-9-425-2020
12. Gottwald A., Krumrey M., Scholze F., et al. Metrology with synchrotron radiation at PTB. *Eur. Phys. J. Plus.* 2022;137(11):1238. https://doi.org/10.1140/epjp/s13360-022-03417-9
13. Woods S., Neira J., Proctor J., Rice J., Tomlin N., White M., Stephens M., Lehman J. Generalized Electrical Substitution Methods and Detectors for Absolute Optical Power Measurements. *Metrologia.* 2022;59(4):044002. https://doi.org/10.1088/1681-7575/ac72dc
14. Anevsky S., Ivanov V., Kuznetsov V., Minaeva O., et al. Primary UV-radiation detector standards. *Metrologia*. 2003;40(1):S25. http://doi.org/10.1088/0026-1394/40/1/307
15. Sigov A.S., Minaeva O.A., Anevsky S.I., Lebedev A.M., Minaev R.V. Metrological studies of the characteristics of multilayer surface coatings using synchrotron radiation. *Russ. Technol. J.* 2021;9(1):38–47 (in Russ.). https://doi.org/10.32362/2500-316X-2021-9-1-38-47
16. Sigov A.S., Golovanova N.B., Minaeva O.A., Anevsky S.I., Shamin R.V., Ostanina O.I. Solution of topical specroradiometric problems using synchrotron radiation. *Russ. Technol. J.* 2022;10(3):34–44 (in Russ.). https://doi.org/10.32362/2500-316X-2022-10-3-34-44

5. Torretti F., Liu F., Bayraktar M., Scheers J., Bouza Z., Ubachs W., Hoekstra R., Versolato O. Spectral characterization of an industrial EUV light source for nanolithography. *J. Phys. D: Appl. Phys.* 2019;53(5):055204. http://doi.org/10.1088/1361-6463/ab56d4
6. Berni L.Â., Vilela W.A., Beloto A.F., de Sena F.O. System for measuring the angular response of radiometers. In: *Proc. 8th Iberoamerican Optics Meeting and 11th Latin American Meeting on Optics, Lasers, and Applications*. 2013. V. 8785. https://doi.org/10.1117/12.2019888
7. Fryc I. Spectral correction of detector used in illuminance measurements. In: *Proc. 11th Slovak-Czech-Polish Optical Conference on Wave and Quantum Aspects of Contemporary Optics* (*SPIE 3820*). 1999. V. 3820. P. 343–348. https://doi.org/10.1117/12.353083
8. Han L. The principle and characteristics of photoelectric sensors. *Science and Technology Innovation and Application*. 2020;10:77–78.
9. Lutz W. The CCPR K2.c key comparison of spectral responsivity from 200 nm to 400 nm. *Metrologia.* 2014;51(6):S336. http://doi.org/10.1088/0026-1394/51/6/S336
10. Аневский С.И., Золотаревский Ю.М., Иванов В.С., Крутиков В.Н., Минаева О.А., Минаев Р.В. Спектрорадиометрия ультрафиолетового излучения. *Измерительная техника*. 2015;11:26–30.
11. König S., Gutschwager B., Taubert R.D., Hollandt J. Metrological Characterization and Calibration of Thermographic Cameras for Quantitative temperature measurement. *Sens. Sens. Syst*. 2020;9(2):425–442. https://doi.org/10.5194/jsss-9-425-2020
12. Gottwald A., Krumrey M., Scholze F., et al. Metrology with synchrotron radiation at PTB. *Eur. Phys. J. Plus.* 2022;137(11):1238. https://doi.org/10.1140/epjp/s13360-022-03417-9
13. Woods S., Neira J., Proctor J., Rice J., Tomlin N., White M., Stephens M., Lehman J. Generalized Electrical Substitution Methods and Detectors for Absolute Optical Power Measurements. *Metrologia.* 2022;59(4):044002. https://doi.org/10.1088/1681-7575/ac72dc
14. Anevsky S., Ivanov V., Kuznetsov V., Minaeva O., et al. Primary UV-radiation detector standards. *Metrologia*. 2003;40(1):S25. http://doi.org/10.1088/0026-1394/40/1/307
15. Сигов А.С., Минаева О.А., Лебедев А.М., Минаев Р.В. Метрологические исследования характеристик многослойных поверхностных покрытий с использованием синхротронного излучения. *Russ. Technol. J*. 2021;9(2):38–46. https://doi.org/10.32362/2500-316X-2021-9-1-38-47
16. Сигов А.С., Голованова Н.Б., Минаева О.А., Аневский С.И., Шамин Р.В., Останина О.И. Решение актуальных задач спектрорадиометрии с использованием синхротронного излучения. *Russ. Technol. J*. 2022;10(3):34–44. https://doi.org/10.32362/2500-316X-2022-10-3-34-44
17. Schwinger J. On the Classical Radiarion of Accelerated Electrons. *Phys. Rev*. 1949;75(12):1912. https://doi.org/10.1103/PhysRev.75.1912
18. Wiedemann H. *Particle Accelerator Physics*. Springer Science & Business Media; 2007. P. 815–894.

17. Schwinger J. On the Classical Radiarion of Accelerated Electrons. *Phys. Rev*. 1949;75(12):1912. https://doi.org/10.1103/PhysRev.75.1912
18. Wiedemann H. *Particle Accelerator Physics*. Springer Science & Business Media; 2007. P. 815–894.
19. Anevskii S.I., Zolotarevskii Y.M., Krutikov V.N., et al. The Use of a Standard Source of Synchrotron Radiation for Calibration of the Sensitivity of a Telescope with CCD Array and High Angular Resolution. *Meas. Tech.* 2015;58(5):520–525. https://doi.org/10.1007/s11018-015-0747-6
[Original Russian Text: Anevskii S.I., Zolotarevskii Yu.M., Krutikov V.N., Lebedev A.M., Minaev R.V., Senin D.S., Stankevich V.G. The Use of a Standard Source of Synchrotron Radiation for Calibration of the Sensitivity of a Telescope with CCD Array and High Angular Resolution *Izmeritel'naya Tekhnika*. 2015;5:33–36 (in Russ.).]
19. Аневский С.И., Золотаревский Ю.М., Крутиков В.Н., Лебедев А.М., Минаев Р.В., Сенин Д.С., Станкевич В.Г. Использование эталонного источника синхротронного излучения для калибровки чувствительности телескопа с ПЗС-матрицей с высоким угловым разрешением. *Измерительная техника*. 2015;5:33–36

About the authors

Alexander S. Sigov, Academician of the RAS, Dr. Sci. (Phys.-Math.), Professor, President of the MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: sigov@mirea.ru. ResearcherID L-4103-2017, Scopus Author ID 35557510600, RSCI SPIN-code 2869-5663, https://www.researchgate.net/profile/A_Sigov

Evgenij R. Lazarenko, Deputy Head of the Federal Agency for Technical Regulation and Metrology (Rosstandart) (10/2, Presnenskaya nab., Moscow, 125039 Russia). E-mail: elazarenko@rst.gov.ru.

Natalia B. Golovanova, Dr. Sci. (Econ.), Deputy First Vice-Rrector, MIREA – Russian Technological University, (78, Vernadskogo pr., Moscow, 119454 Russia) E-mail: golovanova@mirea.ru. Scopus Author ID 57191447039, RSCI SPIN-code 7197-9948, https://orcid.org/0000-0002-9901-8897

Olga A. Minaeva, Dr. Sci. (Eng.), Head of the Department of Metrology and Standardization, Institute of Advanced Technologies and Industrial Programming, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: minaeva_o@mirea.ru. Scopus Author ID 6603019847, RSCI SPIN-code 5081-1976, https://orcid.org/0000-0001-9465-3210

Sergei I. Anevsky, Dr. Sci. (Eng.), Professor, Department of Metrology and Standardization, Institute of Advanced Technologies and Industrial Programming, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: anevskij@mirea.ru. https://orcid.org/0000-0001-9197-0034

Roman V. Minaev, Cand. Sci. (Eng.), Head of the Research Department, Elektrosteklo (5, 26 Bakinskikh Komissarov ul., Moscow, 119571 Russia). E-mail: minaev@eltktrosteklo.ru. Scopus Author ID 22235214600, https://orcid.org/0009-0009-6335-5531

Pavel Yu. Pushkin, Cand. Sci. (Eng.), Director of the Institute of Advanced Technologies and Industrial Programming, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: pushkin@mirea.ru. RSCI SPIN-code 9901-4887, https://orcid.org/0000-0003-3684-550X

Об авторах

Сигов Александр Сергеевич, академик РАН, д.ф.-м.н., профессор, президент ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: sigov@mirea.ru. ResearcherID L-4103-2017, Scopus Author ID 35557510600, SPIN-код РИНЦ 2869-5663, https://www.researchgate.net/profile/A_Sigov

Лазаренко Евгений Русланович, заместитель руководителя Федерального агентства по техническому регулированию и метрологии (Росстандарт) (125039, Россия, Москва, Пресненская наб., д. 10, стр. 2). E-mail: elazarenko@rst.gov.ru.

Голованова Наталия Борисовна, д.э.н., профессор, заместитель первого проректора ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: golovanova@mirea.ru. Scopus Author ID 57191447039, SPIN-код РИНЦ 7197-9948, https://orcid.org/0000-0002-9901-8897

Минаева Ольга Александровна, д.т.н., заведующий кафедрой метрологии и стандартизации Института перспективных технологий и индустриального программирования ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: minaeva_o@mirea.ru. Scopus Author ID 6603019847, SPIN-код РИНЦ 5081-1976, https://orcid.org/0000-0001-9465-3210

Аневский Сергей Иосифович, д.т.н., профессор кафедры метрологии и стандартизации Института перспективных технологий и индустриального программирования ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: anevskij@mirea.ru. https://orcid.org/0000-0001-9197-0034

Минаев Роман Владимирович, к.т.н., начальник научно-исследовательского отдела ООО «Электростекло» (119571, Россия, Москва, ул. 26-ти Бакинских Комиссаров, д. 5). E-mail: minaev@eltktrosteklo.ru. Scopus Author ID 22235214600, https://orcid.org/0009-0009-6335-5531

Пушкин Павел Юрьевич, к.т.н., директор Института перспективных технологий и индустриального программирования ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: pushkin@mirea.ru. SPIN-код РИНЦ 9901-4887, https://orcid.org/0000-0003-3684-550X

Translated from Russian into English by Lyudmila O. Bychkova
Edited for English language and spelling by Thomas A. Beavitt

Analytical instrument engineering and technology

Аналитическое приборостроение и технологии

UDC 621.383.2.032.12+681.586.7
https://doi.org/10.32362/2500-316X-2023-11-5-81-93

RESEARCH ARTICLE

DC/DC converter to power spectral lamps

Mikhail Yu. Nikolshin [1], Alexander V. Frunze [2], Vladimir K. Bityukov [1, @]

[1] *MIREA – Russian Technological University, Moscow, 119454 Russia*
[2] *TERMOKONT, Moscow, 119361 Russia*
[@] *Corresponding author, e-mail: bitukov@mirea.ru*

Abstract

Objectives. The paper describes the creation of a DC/DC converter for powering hollow cathode lamps widely used currently as highly stable sources of spectral lines in spectral absorption analyzers and other applications. Typically, mains power supplies are used for such lamps, since installations using hollow cathode lamps are manufactured as stationary. However, there are no principal obstacles to manufacturing portable versions by simply substituting the power supply. However, special attention should in this case be paid to the power supply of the spectral lamp itself, since the amplitude stability of the radiation depends on the smoothness of its supply voltage. Therefore, the development of the pulse DC/DC converter with high efficiency and low rippling is a relevant and expedient problem.

Methods. The set task is solved by methods of mathematical calculations, circuit simulation in *LTSpice XVII* computer-aided design system, and experimental verification.

Results. The structural and principal electrical circuit of a prototype converter is developed on the basis of a topological analysis of pulse DC/DC converters, along with its calculations and simulation, and a printed circuit board. The developed autonomous high-voltage DC/DC converter has a low ripple level (~250 mV) of the output voltage (~491 V) at a load current of ~20 mA to ensure highly stable radiation.

Conclusions. The possibility of obtaining a high voltage when using the topology of a step-up DC/DC converter with a choke is demonstrated. The experimental verification confirmed the correctness of calculations and modeling of a high-voltage DC/DC converter for powering hollow cathode spectral lamps.

Keywords: hollow cathode spectral lamp, gas discharge lamp, pulse converter, DC/DC converter, high-voltage converter, simulation

• **Submitted:** 27.01.2023 • **Revised:** 07.05.2023 • **Accepted:** 13.07.2023

For citation: Nikolshin M.Yu., Frunze A.V., Bityukov V.K. DC/DC converter to power spectral lamps. *Russ. Technol. J.* 2023;11(5):81–93. https://doi.org/10.32362/2500-316X-2023-11-5-81-93

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

DC/DC-преобразователь для питания спектральных ламп

М.Ю. Никольшин [1],
А.В. Фрунзе [2],
В.К. Битюков [1, @]

[1] *МИРЭА – Российский технологический университет, Москва, 119454 Россия*
[2] *АНО НТП «Термоконт», Москва, 119361 Россия*
[@] *Автор для переписки, e-mail: bitukov@mirea.ru*

Резюме

Цели. Цель работы – создание DC/DC-преобразователя для питания ламп с полым катодом, которые в настоящее время широко используются в качестве высокостабильных источников спектральных линий в установках спектрального абсорбционного анализа и в иных случаях. Зачастую для таких ламп используются сетевые источники питания, т.к. установки с использованием ламп с полым катодом выполняются в стационарном исполнении. Однако препятствий принципиального характера для выполнения такого рода установок в переносном варианте нет. Для этого, в первую очередь, следует отказаться от привязки питания к сети переменного тока. Особое внимание при этом должно быть обращено на питание самой спектральной лампы, т.к. от пульсаций ее питающего напряжения зависит амплитудная стабильность излучения. Следовательно, разработка импульсного DC/DC-преобразователя с высоким КПД и малыми пульсациями является актуальной и целесообразной проблемой.

Методы. Поставленная задача решена методами математических расчетов, схемотехнического моделирования в системе автоматизированного проектирования *LTSpice XVII* и экспериментальной проверки.

Результаты. Выполнен анализ топологий импульсных DC/DC-преобразователей, разработаны структурная и принципиальная электрические схемы преобразователя, проведены их расчеты и моделирование, разработана печатная плата. Для обеспечения высокостабильного излучения разработан и создан автономный высоковольтный DC/DC-преобразователь, имеющий малый уровень пульсаций (~250 мВ) выходного напряжения (~491 В) при токе нагрузке ~20 мА.

Выводы. Показана принципиальная возможность получения высокого напряжения при использовании топологии повышающего DC/DC-преобразователя с дросселем. Экспериментальная проверка подтвердила корректность расчетов и моделирования высоковольтного DC/DC-преобразователя для питания спектральных ламп с полым катодом.

Ключевые слова: спектральная лампа с полым катодом, газоразрядная лампа, импульсный преобразователь, DC/DC-преобразователь, высоковольтный преобразователь, моделирование

• **Поступила:** 27.01.2023 • **Доработана:** 07.05.2023 • **Принята к опубликованию:** 13.07.2023

Для цитирования: Никольшин М.Ю., Фрунзе А.В., Битюков В.К. DC/DC-преобразователь для питания спектральных ламп. *Russ. Technol. J.* 2023;11(5):81–93. https://doi.org/10.32362/2500-316X-2023-11-5-81-93

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

Hollow cathode lamps are used to obtain highly stable spectra of certain wavelengths [1–5]. When a discharge occurs in a lamp of this type, its cathode and inert filler gas start emitting energy in the ultraviolet and visible spectral ranges. The wavelengths of radiation are determined by the composition of the filler gas and cathode alloy. For example, LF-6M lamp with zinc-coated cathode has the main spectral line of 229.6 nm.

Since a hollow cathode lamp is a gas discharge lamp according to its principle of operation, it follows that the essential nature of its discharge occurrence does not differ from that of other gas discharge lamps, such as indicator lamps with neon filling and gas-filled relays. However, indicators can tolerant a significant level of pulsations in their power supply, which is unacceptable for spectral lamps, since the pulsations of current flowing through the lamp cause changes in the amplitude of spectral components. Considering that spectral lamps are used in instrumentation, the intensity of the formed spectral components must be as stable as possible. Thus, the stability of the radiation intensity depends on both the properties of the lamp itself and the parameters of the supply voltage. However, a detailed consideration of the properties of the lamp affecting the stability of the radiation intensity, such as emission, purity of glass, durability, etc., are beyond the scope of the present paper.

The stability of power supply parameters can be achieved by using a negative feedback system [6]. For devices with spectral lamps, this may be achieved by introducing a negative feedback loop either for the lamp current consumption or for the level of the emitted luminous flux. Here, the reduction of the lamp emission ripple is largely related to the reduction of ripples in its supply voltage.

As a rule, devices using spectral discharge lamps are mains powered, which determines the stationary nature of their use. Such a power supply is traditionally carried out using a continuous regulation scheme according to which devices permit a relatively stable voltage having small ripples at the output. Moreover, issues of their efficiency and mass-dimensional parameters are not relevant. One of the implementations of such power supply for spectral lamps wherein the mains voltage multiplication principle is used for obtaining high voltage is proposed in [7].

However, the portable form-factor of using devices with a spectral lamp is also relevant and fundamentally possible. For such autonomous devices, it is advisable to use a switching power supply due to its smaller mass-size parameters in comparison with linear sources of secondary power supply.

CONVERTER PARAMETERS AND ITS BLOCK DIAGRAM

Specifications for output voltage and load current of the DC/DC converter are determined by the parameters of manufactured Russian hollow cathode lamps shown in Table 1.

Table 1. Parameters of ignition voltage and amperage of Russian hollow cathode lamps

Lamp type	Ignition voltage, V	Lamp amperage, mA
LT-2	450	20
LV-2	300	20
LT-6M	500	12

The following basic requirements to the pulse converter for hollow cathode lamps may be formulated: output voltage ~500 V; load current ≥0.02 A; output voltage ripple at current load 20 mA ≤0.5 V; and input voltage ~15 V.

These parameters may be obtained using the DC/DC converter based on the classical step-up converter scheme (Fig. 1) [8, 9].

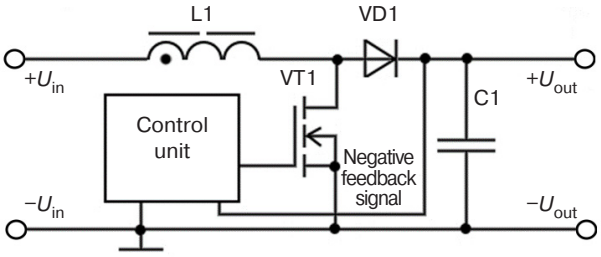


Fig. 1. The power section schematic diagram of the step-up DC/DC converter. Here and in the following figures, the designations of scheme elements correspond to the designations adopted in GOST 2.710-81[1]

The output voltage of the converter is regulated by pulse-width modulation (PWM) or pulse-frequency modulation (PFM). The features of these modulations are considered in [8–12].

The control unit changes the pulse duty factor—or ripple frequency—when the specified output voltage threshold is reached. The main element of the control unit circuit is the one that changes the duty factor upon reaching a certain output voltage threshold. The widespread integrated timer KR1441VI1[2] (TLC55 chip analog[3] being the NE555 CMOS version[4]), which

[1] GOST 2.710-81. *Unified system for design documentation. Alpha-numerical designations in electrical diagrams*. Moscow: Izd. Standartov; 1985 (in Russ.).

[2] Manufactured by Mikron, Zelenograd, Russia.

[3] TLC555 datasheet. https://datasheet.lcsc.com/lcsc/2008261939_HGSEMI-TLC555N_C725329.pdf. Accessed 20.05.2023.

[4] A complementary metal-oxide semiconductor (CMOS) is the integrated circuit technology and circuit design based on *p*- and *n*-channel transistors with insulated gate.

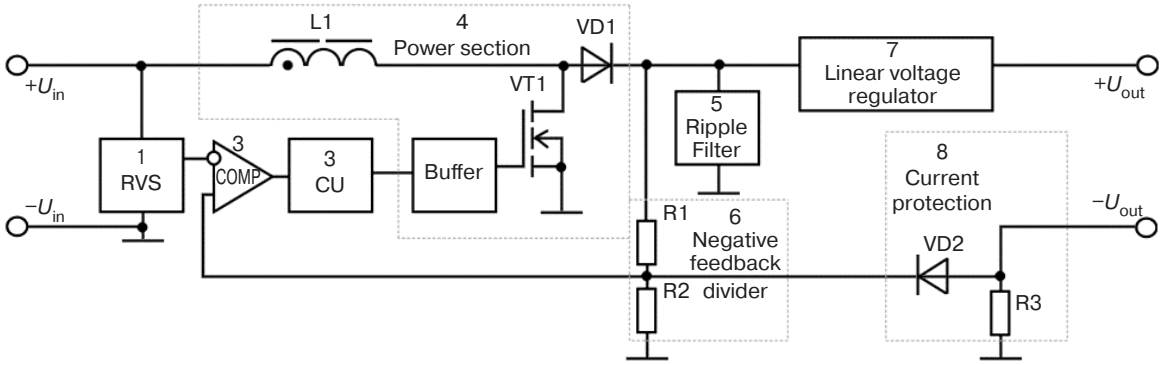


Fig. 2. Structural diagram of the DC/DC converter

can be switched into stable multivibrator mode, i.e., a rectangular pulse generator having a given duty cycle, may be used for this purpose. The adjustment of output voltage in the circuit is provided by changing current between the control pin (pin 5 for 8-pin case) and the circuit common wire. This would change the frequency of the output ripples. The bipolar transistor is used as the regulating element that changes the control pin current.

The circuit clearly recording the fact of reaching the specified voltage is required as the sensor that would detect the voltage of 500 V at the converter output. This is due to the fact that the control pin switch has an active mode area when it is not fully open, as well as a technological opening voltage spread. For example, one bipolar transistor may be fully open at 0.76 V, another at 0.78 V, the third at 0.75 V, etc. The voltage spread would cause the stabilization voltage spread, i.e., the transistor opening should be provided by a third-party circuit having the lowest actuation voltage spread.

A comparator, for example, K554SA301A[5] (LM311 analog[6]), may be used as such sensor to whose inverting input the reference voltage source (RVS) implemented on a parametric voltage regulator is connected, while to the non-inverting input the midpoint of the output voltage divider is connected, designed so that the voltage equal to the RVS output voltage would be at its midpoint at the input voltage of 500 V. When the voltage at the non-inverting input is exceeded, the comparator closes the built-in output transistor to set a high voltage level at the output. This high voltage opens the transistor connected to the control pin of the KR1441VI1 chip thus stopping the generation. When the output voltage drops below 500 V, the comparator changes its state due to the reference voltage level exceeding the divider output voltage, while the transistor closes causing the generation of ripples again.

The structural diagram of the converter is shown in Fig. 2.

The functionality of the DC/DC converter blocks marked by numbers in Fig. 2 is as follows:

1. RVS. This is designed to form stable voltage being the reference voltage for actuating comparator 2 upon reaching the set output voltage.

2. Comparator (COMP). As long as the voltage at the output of the negative feedback divider 6 is less than the voltage generated by RVS, the comparator has low voltage level at its output. The low voltage level at the control unit 3 input is responsible for ripple generation. When the voltage at the negative feedback divider 6 output exceeds the RVS level, the comparator sets high voltage level at the output which prohibits the control unit 3 from generating ripples.

3. Control unit (CU). This generates rectangular pulses that control the power section 4 of the converter by switching MOSFET VT1[7] on or off.

4. Power section. This consists of the choke L1 on which the induction EMF is induced, MOSFET VT1 commutating the right output of the choke L1 either to diode VD1 or to "ground", the rectifier diode VD1, and the buffer stage. The buffer stage is required for matching the low-current output stage of the control unit 3 with the high gate charge current of MOSFET VT1 when it opens and closes in a short time interval.

5. Ripple filter. This is designed for smoothing the output voltage.

6. Negative feedback divider. The divider division ratio determines the output voltage: when the voltage at the output of the negative feedback divider connected to the non-inverting input of comparator 2 becomes higher than the output voltage of RVS 1 (connected to its inverting input), comparator 2 would set high level while control unit 3 would stop generating ripples, thus stopping conversion until the comparator output level becomes low again. This is achieved only if the voltage at the output of divider 6 supported by capacitive filter 5 does not become lower than the voltage of RVS 1 again.

[5] Manufactured by Mikron, Zelenograd, Russia.

[6] LM311 datasheet. https://www.onsemi.com/pdf/datasheet/lm211-d.pdf. Accessed May 20, 2023.

[7] The metal-oxide-semiconductor field-effect transistor (MOSFET) is a field-effect transistor with insulated gate.

The divider output voltage U_{div} is determined by the following formula:

$$U_{div} = U_{in} \frac{R_2}{R_1 + R_2}, \quad (1)$$

where U_{in} is the input voltage of the divider (it is the converter output voltage in this circuit), while R_1 and R_2 are resistances of the negative feedback divider arms.

7. Linear stabilizer. For reducing the output voltage ripple level, a compensating voltage regulator with continuous regulation is used. Its scheme is shown in Fig. 3 (elements R16, VD6–VD10, VT7, and C7). The stabilizer output voltage is equal to the sum of stabilization voltages of stabilitrons VD6–VD10 plus the base-emitter voltage of transistor VT7 (about 0.5–0.7 V). Calculations have shown that, at a pulse converter output voltage of 530 V, a linear stabilizer output voltage of 500 V, and a load current 20 mA, power of 0.6 W would be dissipated on the control transistor.

8. Protection. When the current exceeds the nominal value determined by the resistance of resistor R3, the voltage at the non-inverting input of comparator 2 becomes higher than the voltage level of RVS 1, which would cause the control unit 3 to stop generating ripples. This is essentially a parallel feedback channel. The diode VD2 is designed to prevent the current flow from the negative feedback divider 6 through resistor R3. The resistor R3 resistance is calculated based on the switching voltage of the comparator U_{RVS}, the load current flow at which protection $I_{protect}$ actuates, and the direct voltage drop $U_{VD2\ dir}$ across diode VD2 by the following formula:

$$R_3 = \frac{U_{RVS} + U_{VD2\ dir}}{I_{protect}}. \quad (2)$$

For the protection actuation current of 25 mA, the reference voltage of 3.3 V and the direct voltage drop across the diode of 0.6 V, the resistor R3 resistance is 156 Ω. This resistor is included in series with the load, and it would inevitably drop some of the output voltage. Operating voltage of spectral lamps ranges within 250–300 V, while the recommended current for maintaining the discharge is up to 2/3 of the maximum, i.e., 8–12 mA. It should be noted that the service life of spectral lamps is normalized in milliampere-hours, i.e., reducing the current increases not only the empirical, but also the documented service life of the lamp. The equivalent resistance of the lamp, accordingly, would be in the range from 20.8 to 37.5 kΩ. In the worst case, i.e., when the tube resistance is minimal, the voltage drop across R2 is calculated by formula (1) is 3.7 V, i.e., less than 1% of output voltage, which is totally unnoticeable in practice. If, however, the output voltage should precisely match the data sheet, it can be adjusted by setting the output voltage not at idle but under load, which may be a resistor of equivalent resistance to that of a particular type of lamp.

It is worth detailing the selection of the type and rating of choke L1. The step-up converter has both continuous-current and discontinuous-current operating modes. The discontinuous-current mode means that the current flows through the choke L1 winding to all converter operation phases, which varies but never goes to zero. In discontinuous-current mode, the current through the choke L1 winding has time to drop to zero when the key transistor closes. In terms of reducing the ripple level, discontinuous current mode is preferable. However, the choke calculation according to the method [8] shows that the inductance for discontinuous currents mode for PWM would be in a range from 1.5 to 5.0 mH (depending on the converter frequency). Due to the considerable size, mass, and cost of such a choke, the mode of discontinuous currents with inductance of the choke of the order of 50–150 μHY in the converter. The inevitable increase in ripple was leveled by using the linear voltage regulator.

SCHEMATIC CIRCUIT DIAGRAM

The schematic circuit diagram of the DC/DC converter for powering spectral lamps is shown in Fig. 3.

Elements L1, C1, and C2 form the input filter. Choke L1 smooths the input current, thus reducing the pulse interference level of the converter power supply.

The master oscillator is made on elements DD1, R2, R4, and C3. The ripple frequency before transiting to the stabilization mode is 27.7 kHz, while the duty factor is 0.8.

The power section of the converter is built on elements R3, VT2–VT6, R6–R8, and L2. It is noteworthy the schematic feature of the power section construction that is parallel inclusion of MOSFETs VT5, VT6, and VT7 operating in the cutoff mode. The KP7173A[8] is one of the few Russian MOSFETs suitable[9] for this application in terms of drain-source voltage. Unfortunately, the resistance of its open channel leaves much to be desired and can reach 2 Ω (according to data on its full analogue STP4NK60Z[10]). Obviously, with the drain current of 4 A and the duty factor of 0.5, the mean square value of power dissipation would be 4 W. This does not take into account losses caused by the transistor operating point being in the active mode during opening and closing of the channel. Such dissipated power is acceptable for the transistor but it forces to use coolers for it, thus reducing the converter efficiency significantly. One way to solve this problem is connecting several transistors in parallel.

[8] Reference data on KP7173A. https://integral.by/sites/default/files/pdf/kp7173.pdf. Accessed May 20, 2023.

[9] Manufactured by Integral, Minsk, Belarus.

[10] STP4NK60Z datasheet: https://www.st.com/resource/en/datasheet/stp4nk60z.pdf. Accessed May 20, 2023.

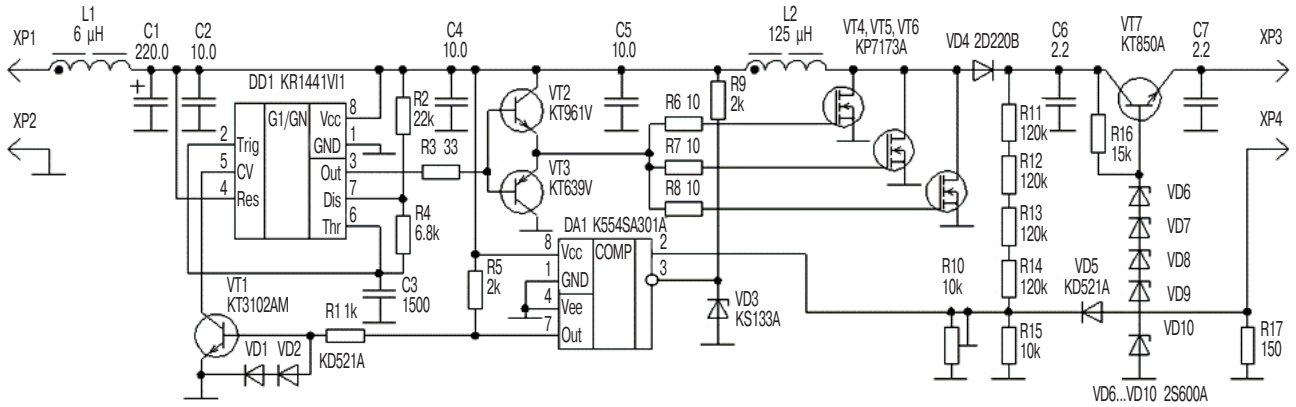


Fig. 3. Schematic circuit diagram of the DC/DC converter

Since, the transistor with the lowest resistance would be more heated because of the greater current flow even in the case of (inevitable) spread of channel resistances, parallel connection of MOSFETs is acceptable. When the temperature of the MOSFET crystal increases, the resistance of its open channel increases, thereby causing a decrease in current.

The calculation of the required number of key transistors is dictated by the current flowing through them as well as their heating.

The maximum transition temperature of the KP7173A transistor is +150°C; the "crystal-to-package" thermal resistance is 1.78 K/W. The "crystal-to-medium" thermal resistance for the KT-28-2 (TO-220) package wherein the KP7173A transistor crystal is mounted, according to various data, ranges from 50°C/W to 70°C/W. When using the converter under UHL 4.2 conditions[11] (maximum air temperature is +40°C), taking a margin of +25°C for heating inside the item, and selecting the maximum transition temperature of the transistor of +100°C, the permissible overheating of the package is +35°C. In the case of the "crystal-to-medium" thermal resistance of 70°C/W, a maximum power dissipation of 0.5 W is obtained. Hence, three transistors (power dissipation on each is 0.45 W, crystal temperature is 96.5°C, and package temperature is 54.2°C) would be sufficient in terms of thermal conditions without using coolers[12].

Transistors VT2 (KT639V[13]) and VT3 (KT961V[14]) are driver transistors for field-effect transistors VT4, VT5, and VT6, since the MOSFET gate is a capacitance which has to be rapidly charged and discharged for opening and closing the transistor. The total charge Q of the gate [13] for the STP4NK60Z transistor is 26 nC, as stated in its technical documentation. Knowing the total charge and the necessary time T to communicate this charge, the current intensity I is determined by the following formula:

$$I = \frac{Q}{T}. \quad (3)$$

For the charge of 78 nC (three transistors connected in parallel) and the charge time of 100 ns, the charge current is 780 mA[15]. Since the output of the KR1441VI1 chip is not designed for such current, a buffer stage is required. At the same time, the current consumed by the buffer stage at the input would be h_{21E} times less than the output current. According to [14], the minimum current transfer coefficient h_{21E} for the KT961V and KT639V transistors is 100, which means that the maximum output chip current would be reduced to 7.8 mA. Resistor R5 serves to limit current while resistors R9–R11, in addition to current limiting, also suppress resonant oscillations in the "capacitance-'gate-source'-drain inductance" circuit [15]. Choke D13–14 with windings connected in series (125 μHY, 4 A, 0.05 Ω, and 2 MHz) is used as storage choke L2. The output voltage rectification is carried out by the 2D220B-type diode VD4[16].

The output voltage stabilization circuit is made on the comparator DA1, transistor VT1 and its harness, parametric voltage regulator (RVS) on stabilitron VD3 (3.3 V), and divider R11–R15. The divider R11–R15 upper arm consists of four resistors connected in series for increasing the maximum allowable applied voltage. Resistor R10 serves as a trimmer, rotating its slider to set the output voltage to 530 V. In case of its interruption the output voltage

[11] GOST 15050-69. *Machines, instruments and other industrial products. Modifications for different climatic regions. Categories, operating, storage and transportation conditions as to environment climatic aspects influence*. Moscow: Standardinform; 2010 (in Russ.).

[12] At the thermal resistance of 50°C/W, the crystal temperature is 87.5°C while the body temperature is 49.2°C.

[13] Manufactured by GRUPPA KREMNI EL, Bryansk, Russia.

[14] Manufactured by Integral, Minsk, Belarus.

[15] The limiting case is 26 nC of gate charge; the typical value is 18.8 nC.

[16] Manufactured by the Scientific and Production Enterprise TOMILINO ELECTRONIC PLANT, Tomilino, Russia.

Table 2. Power dissipation of parallel connected transistors

Number of transistors, items	Equivalent resistance, Ω	Total power dissipation at 4 A and duty factor 0.5, W	Power dissipation per transistor in case of equal channel resistances, W
3	0.667	1.35	0.45
4	0.5	1.00	0.25
5	0.4	0.80	0.20

Table 3. Components used and their analogs

Type by scheme	Model	Analog (full or partial)	Comments
KR1441VI1	TLC555	Full	–
K554SA301A	LM311	Full	–
KP7173A	SPP20N60C3	Partial	The 1.4-ohm resistor is added to the drain for approaching $R_{DS\ on}$ of KP7173A, where $R_{DS\ on}$ is the drain-source resistance in the open state of the transistor
2D220B	DI_US1J[1]	Partial	Matches in capacitance, reverse voltage U_{rev}, and direct current I_{dir}, but has a shorter reverse recovery time (75 ns vs. 600 ns)
2S600A	PH_BZX585-B75[2], PH_BZX585-B51[20]	Full	The model uses an assembly of 75 V and 51 V stabilitrons for total stabilization voltage of 601 V
KS133A[3]	BZX84C3V3[19]	Full	–
KT961V	BD139[4]	Full	–
KT639V	BD140[22]	Partial	Matches in the collector-emitter voltage U_{CE}, collector current I_C, and cutoff frequency F_{CF}
KT850A	BD139	Partial	Matches in U_{CE}, I_C, and F_{CF}
KT3102A[5]	2N3904p[6]	Partial	Has lowest h_{21E} but it does not matter in this particular case

[1] Manufactured by Vishay, USA.
[2] Manufactured by Nexperia, Netherlands.
[3] Manufactured by Novosibirsk Semiconductor Device Plant, Novosibirsk, Russia.
[4] Manufactured by ST Microelectronics, Switzerland.
[5] Manufactured by GRUPPA KREMNI EL, Bryansk, Russia.
[6] Manufactured by Diotec Semiconductor, Germany.

would be set to 160 V. When the converter is switched on, the output voltage starts increasing; therefore, the divider output voltage starts increasing as well. Once it equals or exceeds the reference voltage of 3.3 V, the comparator would set a high level at the output by opening transistor VT1 and closing pin 5 of chip DD1 to the common wire to prevent ripple generation. The output voltage would be maintained by the capacitive filter. As soon as the voltage at the divider output drops below 3.3 V, the comparator would set low level, closing transistor VT1, and thus enabling ripple generation again.

The output voltage capacitive filter is made on capacitor C6. The parallel connection of ten 0.22 μF × 630 V ceramic capacitors of size 2220 with total capacitance of 2.2 μF is used.

The linear stabilizer is made on transistor VT7 KT850A[17], stabilitrons VD6-VD10 2S600A[18], and the current set resistor R16. The series connection of five diodes is used for obtaining stabilization voltage of 500 V, since the Russian industry does not produce 500 V stabilitrons.

The current protection is made on elements R18 and VD7. When the current through the load (hence, through resistor R17) exceeds 25 mA, the voltage drop across resistor R17 minus the voltage drop across diode VD5 would be 3.3 V. This will cause the comparator DD2 to set the output to the high level, thus stopping ripple generation. Diode VD5 is designed to prevent current flowing from the negative feedback divider to resistor R17.

CONVERTER SIMULATION

The simulation has been performed within the *LTSpice XVII* software [16]. The active component models given in Table 3 have been used.

The model schematic is shown in Fig. 4.

[17] Manufactured by GRUPPA KREMNI EL, Bryansk, Russia.
[18] Manufactured by Novosibirsk Semiconductor Device Plant, Novosibirsk, Russia.

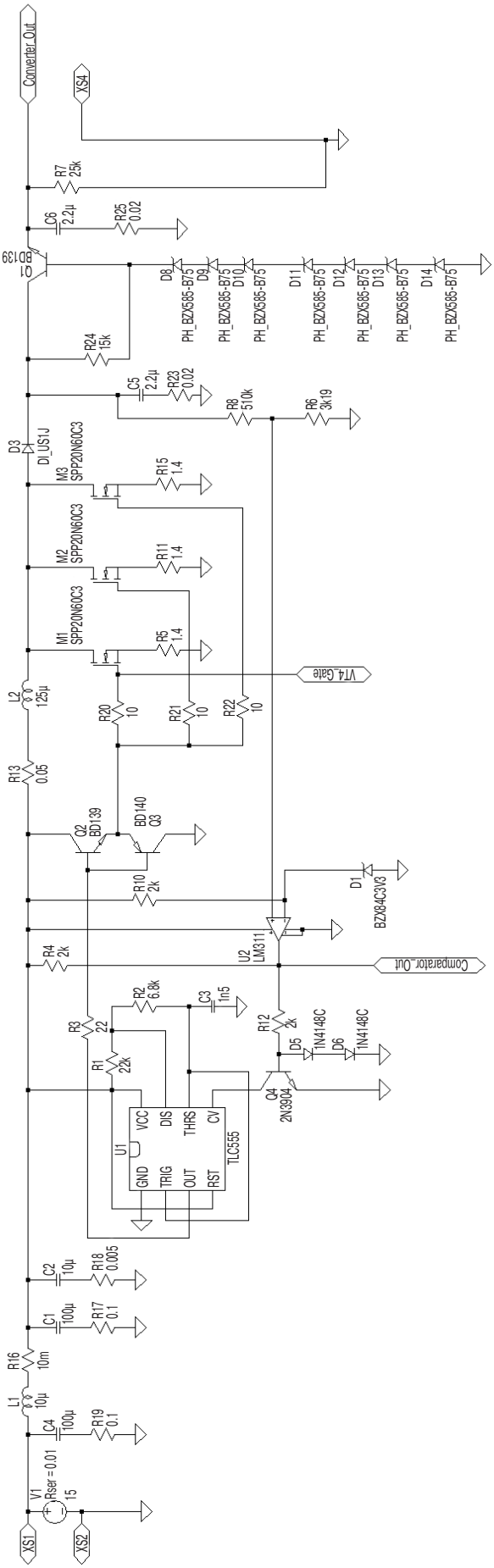


Fig. 4. Schematic diagram of the converter in *LTSpice*

Simulation has been performed using the .tran 0 500m 1u uic directive: transient analysis; start time of the data acquisition is 0 s; end time of the data acquisition is 0.5 s; simulation step is 1 μs; and transients in reactivities are included.

Main characteristic diagrams are shown in Fig. 5.

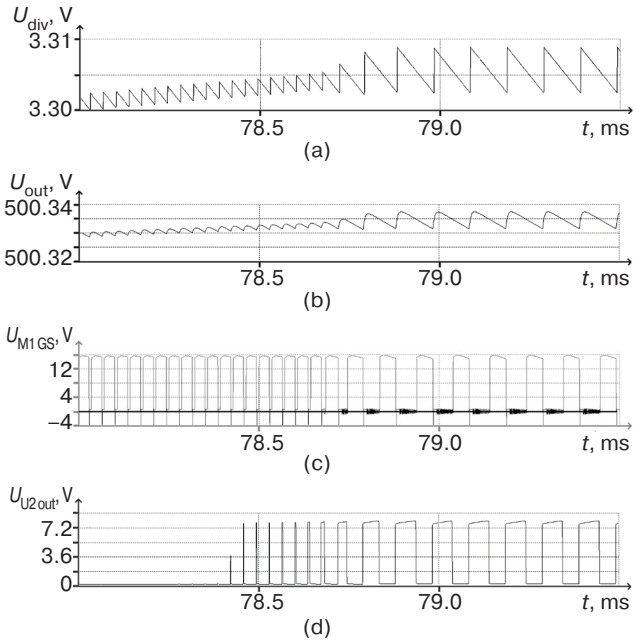


Fig. 5. Characteristic diagrams of the converter operation: (a) feedback divider output voltage U_{div}, (b) converter output voltage U_{out}, (c) gate voltage of field-effect transistor M1 $U_{M1\,GS}$, (d) comparator output voltage U2 $U_{U2\,out}$

The simulative converter reaches operating mode about 79 ms after start. At this time, the voltage at the feedback divider output U_{div} reaches the level of the comparator reference voltage. The comparator output $U_{U2\,out}$ is set to the "high state" level. The high signal level at the comparator output, which is formed when the output voltage reaches 500 V, causes generation to stop at the TLC555 chip output.

Fig. 6. Assembled converter board

ASSEMBLY AND EXPERIMENTAL VERIFICATION OF THE CONVERTER

The converter is assembled on a double-sided printed circuit board (PCB) made of foil fiberglass. The pattern of the printed tracks is made completely on the bottom side of the board. The upper side of the board is used for the common wire; thus, its resistance and inductance are minimized, which is important in pulsed high-frequency circuits [17]. The mounting is mixed. The PCB image is shown in Fig. 6.

The test bench of the converter consists of a Element PSN-305D laboratory power supply unit[19] providing power for the converter; a Fluke 17B+ handheld multimeter[20] in the voltmeter mode for measuring the output voltage; and a Uni-T UT802 bench multimeter[21] in the milliammeter mode for measuring the output current. A Rigol MSO 4024 digital oscilloscope[22] is used for taking voltage waveforms. The test bench images are shown in Figs. 7 and 8. The oscillogram of pulsations at the converter output under load is shown in Fig. 9.

[19] Manufactured by Element, China.
[20] Manufactured by Fluke, China.
[21] Manufactured by Uni-Trend Group, China.
[22] Manufactured by Rigol Technologies, China.

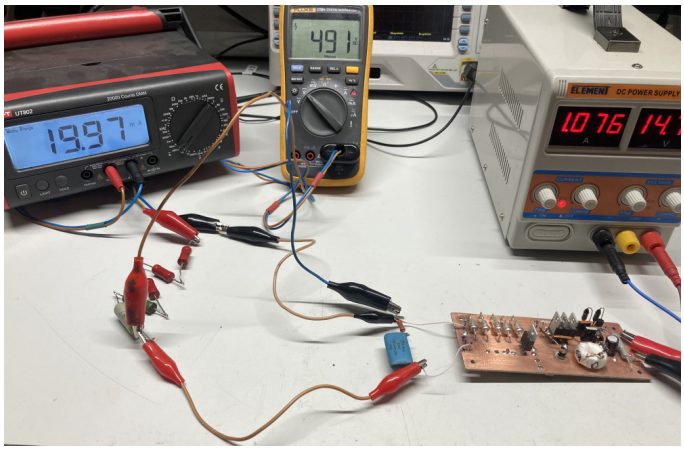


Fig. 7. Checking the load capacity of the converter

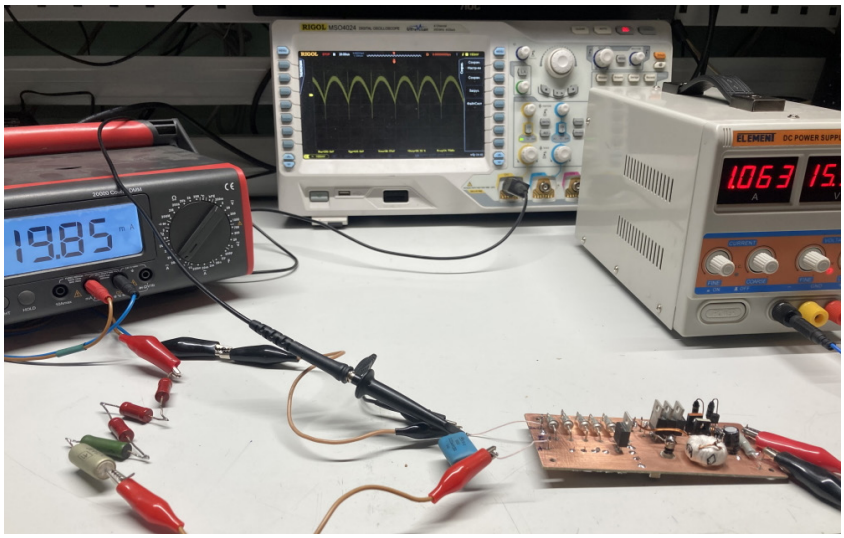


Fig. 8. Checking the ripple level at the converter output under load

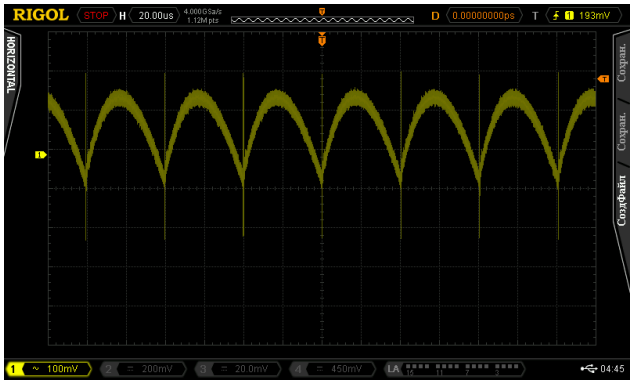


Fig. 9. Oscillogram of the output voltage ripple under load

The load capacity of the converter is confirmed by the output voltage of 491 V at a current of 20 mA. The compound resistor with total resistance of 25 kΩ is used as the load.

For checking the ripple level, the oscilloscope is connected to the converter output through the differential RC-chain with ratings of 0.022 μF and 100 kΩ. The cutoff frequency of the high-pass filter formed by such circuit is 72 Hz at the 3 dB level.

The peak ripple level is 250 mV (0.05% over 500 V), which confirms the above conclusions about the low ripple level of the converter under consideration.

CONCLUSIONS

In the paper, the results of the development, modeling, circuit assembly, and experimental study of the specialized pulse DC/DC converter for powering spectral tubes are presented. Development of the circuit diagram is carried out with the expectation for a minimum number of tuning operations during assembly (as shown, this operation may be neglected), as well as the absence of a necessity to use a heat sink, and minimized use of imported element base. Simulation of the DC/DC converter is done partly on approximated analogs, whereby additional elements are introduced into the circuit for approximating model parameters to actually used components. The PCB for the converter is traced with allowance for minimizing the inductance and resistance of the high-frequency conductors and the common wire. The converter performance (including with an equivalent load for achieving the required output current) and the compliance of measurements with simulation results confirm the correctness of calculations and simulation. The converter provides 491 V at the current of about 20 mA and peak ripple of 250 mV (0.05%).

Authors' contribution. All authors equally contributed to the research work.

REFERENCES

1. Moskalev B.I. *Razryad s polym katodom* (*Discharge with Hollow Cathode*). Moscow: Energiya; 1969. 184 p. (in Russ.).
2. Maksimov D.E., Rudnevskii N.K. *Spektral'nyi analiz s primeneniem razryada v polom katode* (*Spectral Analysis Using Hollow Cathode Discharge*). Gor'kii: GGU; 1979. 119 p. (in Russ.).
3. Maksimov D.E., Rudnevskii N.K., Rudnevskii A.N., Shabanova T.M. *Spektral'nyi analiz s primeneniem razryada v polom katode* (*Spectral Analysis Using Hollow Cathode Discharge*). Gor'kii: GGU; 1983. 71 p. (in Russ.).
4. Garmash A.V. *Vvedenie v spektroskopicheskie metody analiza. Opticheskie metody analiza* (*Introduction to Spectroscopic Methods of Analysis. Optical Methods of Analysis*). Moscow: RAN VKhK; 1995. 38 p. (in Russ.).
5. Bityukov B.K., Gorbunov R.A., Simachkov D.S., Frunze A.V. A Highly Stable Source of Spectral Lines. *Opt. Spectrosc*. 2019;126(4):450–453. https://doi.org/10.1134/S0030400X19040040
[Original Russian Text: Bityukov B.K., Gorbunov R.A., Simachkov D.S., Frunze A.V. A Highly Stable Source of Spectral Lines. *Optika i spektroskopiya*. 2019;126(4):528–532 (in Russ.). http://doi.org/10.21883/OS.2019.04.47525.50-18]
6. Kaganov V.I., Bityukov V.K. *Osnovy radioelektroniki i svyazi* (*Fundamentals of Radioelectronics and Communications*). Moscow: Goryachaya liniya – Telekom; 2012. 542 p. (in Russ.).
7. Skvortsov E.A., Simachkov D.S. High-stability radiation device based on hollow cathode lamps. In: *Actual Problems and Perspectives for the Development of Radio Engineering and Infocommunication Systems* (*Radioinfocom 2021*): *Proceedings of the 5th International Scientific and Practical Conference*. Moscow: RTU MIREA; 2021. P. 405–408 (in Russ.). Available from URL: https://www.elibrary.ru/aelqox
8. Bityukov V.K., Simachkov D.S., Babenko V.P. *Istochniki vtorichnogo elektropitaniya* (*Secondary Power Sources*). Moscow: Infra-Inzheneriya; 2020. 376 p. (in Russ.).
9. Babenko V.P., Bityukov V.K., Simachkov D.S. DC/DC Buck-Boost Converter with Single Inductance. *Russ. Microelectron.* 2021;50(6):471–480. https://doi.org/10.1134/S1063739721060044
10. Sergeev B.S. *Skhemotekhnika funktsional'nykh uzlov istochnikov vtorichnogo elektropitaniya* (*Schematics of Functional Units of Secondary Power Sources. Reference-book*). Moscow: Radio i svyaz'; 1992. 224 p. (in Russ.).
11. Meleshin V.I. *Tranzistornaya preobrazovatel'naya tekhnika* (*Transistor Converter Technology*). Moscow: Tekhnosfera; 2005. 632 p. (in Russ.).
12. Naivel't G.S., Mazel' K.B., Khusainov Ch.I., et al. *Istochniki elektropitaniya radioelektronnoi apparatury* (*Power Supply Sources REA: reference-book*). Naivelt G.S. (Ed.). Moscow: Radio i svyaz'; 1985. 576 p. (in Russ.).

СПИСОК ЛИТЕРАТУРЫ

1. Москалев Б.И. *Разряд с полым катодом*. М.: Энергия; 1969. 184 с.
2. Максимов Д.Е., Рудневский Н.К. *Спектральный анализ с применением разряда в полом катоде.* Горький: ГГУ; 1979. 119 с.
3. Максимов Д.Е., Рудневский Н.К., Рудневский А.Н., Шабанова Т.М. *Спектральный анализ с применением разряда в полом катоде*. Горький: ГГУ; 1983. 71 с.
4. Гармаш А.В. *Введение в спектроскопические методы анализа. Оптические методы анализа*. М.: РАН ВХК; 1995. 38 с.
5. Битюков В.К., Горбунов Р.А., Симачков Д.С., Фрунзе А.В. Высокостабильный источник спектральных линий. *Оптика и спектроскопия.* 2019;126(4):528–532. http://doi.org/10.21883/OS.2019.04.47525.50-18
6. Каганов В.И., Битюков В.К. *Основы радиоэлектроники и связи.* М.: Горячая линия-Телеком; 2012. 542 с.
7. Скворцов Е.А., Симачков Д.С. Устройство по созданию высокостабильного излучения на основе ламп с полым катодом. В сб.: *Актуальные проблемы и перспективы развития радиотехнических и инфокоммуникационных систем «Радиоинфоком – 2021»: Сборник научных статей V международной научно-практической конференции*. М.: РТУ МИРЭА; 2021. С. 405–408. URL: https://www.elibrary.ru/aelqox
8. Битюков В.К., Симачков Д.С., Бабенко В.П. *Источники вторичного электропитания*. М.: Инфра-Инженерия; 2020. 376 с.
9. Бабенко В.П., Битюков В.К., Симачков Д.С. Понижающе-повышающий DC/DC преобразователь с единственной индуктивностью. *Микроэлектроника.* 2022;51(1):60–70. https://doi.org/10.31857/S0544126921060041
10. Сергеев Б.С. *Схемотехника функциональных узлов источников вторичного электропитания: Справочник*. М.: Радио и связь; 1992. 224 с.
11. Мелешин В.И. *Транзисторная преобразовательная техника*. М.: Техносфера; 2005. 632 с.
12. Найвельт Г.С., Мазель К.Б., Хусаинов Ч.И. и др. *Источники электропитания радиоэлектронной аппаратуры: Справочник*; под ред. Г.С. Найвельта. М.: Радио и связь; 1985. 576 с.
13. Бабенко В.П., Битюков В.К. Измерение заряда затвора для ключей на мощных MOSFET транзисторах. В сб.: *Современные проблемы профессионального образования: опыт и пути решения. Материалы Второй всероссийской научно-практической конференции с международным участием.* Иркутск: Иркутский государственный университет путей сообщения; 2017. С. 37–41. URL: https://www.elibrary.ru/zebetl
14. Зайцев А.А., Миркин А.И., Мокряков В.В. *Полупроводниковые приборы. Транзисторы средней и большой мощности*: *Справочник;* под ред. А.В. Голомедова. М.: КУбК-а; 1995. 640 с.
15. Новиков П. Затворный резистор. Часть 1. *Силовая электроника*. 2018;6(75):4–8.
16. Володин В.Я. *LTspice: компьютерное моделирование электронных схем.* СПб.: БХВ-Петербург; 2010. 400 с.

13. Babenko V.P., Bityukov V.K. Gate charge measurement for high power MOSFET switches. In: *Modern Problems of Vocational Education: Experience and Solutions. Materials of the Second All-Russian scientific-practical conference with international participation*. 2017. Irkutsk: Irkutsk State Transport University; 2017. P. 37–41 (in Russ.).
14. Zaitsev A.A., Mirkin A.I., Mokryakov V.V. *Poluprovodnikovye pribory. Tranzistory srednei i bol'shoi moshchnosti* (*Semiconductor Devices. Medium and High Power Transistors: Handbook*). Golomedov A.V. (Ed.). Moscow: KUbK-a; 1995. 640 p. (in Russ.).
15. Novikov P. Gate resistor. Part 1. *Silovaya elektronika* = *Power Electronics*. 2018;6(75):4–8 (in Russ.).
16. Volodin V.Ya. *LTspice: komp'yuternoe modelirovanie elektronnykh skhem* (*LTspice: Computer Simulation of Electronic Circuits*). St. Petersburg: BKhV-Peterburg; 2010. 400 p. (in Russ.).
17. Barnes J.R. *Electronic System Design: Interference and Noise Control Techniques*. Transl. from Engl. Moscow: Mir; 1990. 238 p. (in Russ.).
[Barnes J.R. *Electronic System Design: Interference and Noise Control Techniques*. New Jersey: Prentice Hall; 1987. 244 p.]
17. Барнс Дж. *Электронное конструирование: Методы борьбы с помехами*: пер. с англ. М.: Мир; 1990. 238 с.

About the authors

Mikhail Yu. Nikolshin, Postgraduate Student, Department of Radio Wave Processes and Technology, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: florenica@yandex.ru. RSCI SPIN-code 6339-2251, https://orcid.org/0009-0004-5353-7781

Alexander V. Frunze, Dr. Sci. (Eng.), Director of TERMOKONT Autonomous Non-Profit Organization for Research, Development and Implementation of Scientific and Technical Devices (42, Ozernaya ul. Moscow, 119361 Russia). E-mail: alex.fru@mail.ru. Scopus Author ID 6507641791, RSCI SPIN-code 7558-6141, https://orcid.org/0009-0006-8602-3650

Vladimir K. Bityukov, Dr. Sci. (Eng.), Professor, Department of Radio Wave Processes and Technology, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: bitukov@mirea.ru. Scopus Author ID 6603797260, ResearcherID Y-8325-2018, RSCI SPIN-code 3834-5360, https://orcid.org/0000-0001-6448-8509

Об авторах

Никольшин Михаил Юрьевич, аспирант, кафедра радиоволновых процессов и технологий Института радиоэлектроники и информатики ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: florenica@yandex.ru. SPIN-код РИНЦ 6339-2251, https://orcid.org/0009-0004-5353-7781

Фрунзе Александр Вилленович, д.т.н., директор, Автономная некоммерческая организация по исследованию, разработке и внедрению научно-технических приборов «Термоконт» (119361, Россия, Москва, ул. Озерная, д. 42). E-mail: alex.fru@mail.ru. Scopus Author ID 6507641791, SPIN-код РИНЦ 7558-6141, https://orcid.org/0009-0006-8602-3650

Битюков Владимир Ксенофонтович, д.т.н., профессор, профессор кафедры радиоволновых процессов и технологий Института радиоэлектроники и информатики ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: bitukov@mirea.ru. Scopus Author ID 6603797260, ResearcherID Y-8325-2018, SPIN-код РИНЦ 3834-5360, https://orcid.org/0000-0001-6448-8509

Translated from Russian into English by Kirill V. Nazarov
Edited for English language and spelling by Thomas A. Beavitt

Mathematical modeling

Математическое моделирование

UDC 681.5.015
https://doi.org/10.32362/2500-316X-2023-11-5-94-105

RESEARCH ARTICLE

On adaptive identification of systems having multiple nonlinearities

Nikolay N. Karabutov @

MIREA – Russian Technological University, Moscow, 119454 Russia
@ *Corresponding author, e-mail: karabutov@mirea.ru*

Abstract

Objectives. The solution to the relevant problem of identifying systems with multiple nonlinearities depends on such factors as feedback, ways of connecting nonlinear links, and signal properties. The specifics of nonlinear systems affect control systems design methods. As a rule, the basis for the development of a mathematical model involves the linearization of a system. Under conditions of uncertainty, the identification problem becomes even more relevant. Therefore, the present work sets out to develop an approach to the identification of nonlinear dynamical systems under conditions of uncertainty. In order to obtain a solution to the problem, an adaptive identification method is developed by decomposing the system into subsystems.

Methods. Methods applied include the adaptive identification method, implicit identified representation, S-synchronization of a nonlinear system, and the Lyapunov vector function method.

Results. A generalization of the excitation constancy condition based on fulfilling the S-synchronizability for a nonlinear system is proposed along with a method for decomposing the system in the output space. Adaptive algorithms are obtained on the basis of the second Lyapunov method. The boundedness of the adaptive system trajectories in parametric and coordinate spaces is demonstrated. Approaches for self-oscillation generation and nonlinear correction of a nonlinear system are considered along with obtained exponential stability conditions for the adaptive system

Conclusions. Simulation results confirm the possibility of applying the proposed approach to solving the problems of adaptive identification while taking the estimation of the structural identifiability (S-synchronization) of the system nonlinear part into account. The influence of the structure and relations of the system on the quality of the obtained parametric estimates is investigated. The proposed methods can be used in developing identification and control systems for complex dynamic systems.

Keywords: adaptive identification, identifiability, stability, excitation constancy, Lyapunov vector function, self-oscillation

• **Submitted:** 06.02.2023 • **Revised:** 03.05.2023 • **Accepted:** 21.07.2023

For citation: Karabutov N.N. On adaptive identification of systems having multiple nonlinearities. *Russ. Technol. J.* 2023;11(5):94–105. https://doi.org/10.32362/2500-316X-2023-11-5-94-105

Financial disclosure: The author has no a financial or property interest in any material or method mentioned.

The author declares no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Об адаптивной идентификации систем с несколькими нелинейностями

Н.Н. Карабутов [@]

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[@] *Автор для переписки, e-mail: karabutov@mirea.ru*

Резюме

Цели. Задача идентификации систем с несколькими нелинейностями является актуальной. Решение этой задачи зависит от наличия обратных связей, способов соединения нелинейных звеньев, свойств сигналов. Специфика нелинейных систем накладывает отпечаток на методы синтеза систем управления. В условиях полной априорной определенности обычно применяют линеаризацию систем. Если существует априорная неопределенность, то задача синтеза системы идентификации обеспечения усложняется. Целью настоящей работы является разработка подхода к идентификации нелинейных динамических систем с несколькими нелинейностями. Для решения проблемы применяется подход, основанный на декомпозиции системы на ряд подсистем и разработке метода адаптивной идентификации, использующего только доступную информацию о системе и измерениях. Необходимо оценить частотные свойства сигналов, которые должны гарантировать оценку параметров системы и обеспечивать структурную идентифицируемость нелинейностей в системе; оценить работоспособность синтезированной адаптивной системы.

Методы. Применяются метод адаптивной идентификации системы, неявное идентификационное представление, S-синхронизация нелинейной системы, метод векторных функций Ляпунова.

Результаты. Введено условие постоянства возбуждения переменных состояния с учетом S-синхронизируемости нелинейной части системы. Дано обобщение условия постоянства возбуждения. Предложен способ декомпозиции системы в выходном пространстве. Получены адаптивные алгоритмы на основе второго метода Ляпунова. Доказана ограниченность траекторий адаптивной системы в параметрическом и координатном пространствах на основе векторных функций Ляпунова. Получены условия, гарантирующие экспоненциальную устойчивость траекторий системы. Рассмотрены системы генерации автоколебаний и нелинейной коррекции нелинейной системы.

Выводы. Результаты моделирования подтвердили возможность применения предлагаемого подхода для решения задач адаптивной идентификации с учетом оценки структурной идентифицируемости (S-синхронизируемости) нелинейной части системы. Исследовано влияние структуры и связей системы на качество получаемых параметрических оценок. Предлагаемые методы могут использоваться при разработке систем идентификации и управления сложными динамическими системами.

Ключевые слова: адаптивная идентификация, идентифицируемость, устойчивость, постоянство возбуждения, векторная функция Ляпунова, автоколебания

• **Поступила:** 06.02.2023 • **Доработана:** 03.05.2023 • **Принята к опубликованию:** 21.07.2023

Для цитирования: Карабутов Н.Н. Об адаптивной идентификации систем с несколькими нелинейностями. *Russ. Technol. J.* 2023;11(5):94–105. https://doi.org/10.32362/2500-316X-2023-11-5-94-105

Прозрачность финансовой деятельности: Автор не имеет финансовой заинтересованности в представленных материалах или методах.

Автор заявляет об отсутствии конфликта интересов.

INTRODUCTION

A number of studies [1–11] are devoted to the identification of systems having several nonlinearities (SSN). In [1], the case when a system comprises several nonlinearities in series is considered and an identification method proposed. The detection of nonlinearity is performed using sinusoidal tests. A similar approach based on the functional description method is applied in [2] for parametric identification of a system having two nonlinear elements in different locations. In [3], an approach to estimating the parameters of the transfer function of the second-order nonlinear system containing two nonlinearities is proposed. In this case, the harmonic linearization of nonlinearities is performed beforehand. In [4], in noting difficulties of SSN identification, a proposed approach to the parameter estimation is based on function approximation. Various methods based on the nonlinearity approximation are considered in [5–7]. In [7, 8], approaches to identifying discrete systems with feedback are studied. Identifiability conditions are obtained by applying the least squares method [8].

In [9], the identification of the system having nonlinear mechanical oscillations is considered. The proposed model, which has a "gray box" appearance, is based on the application of nonlinear basis functions using a limited number of measured output variables. Other approaches to identification are considered in [10–14]. These are based on considering physical laws when selecting the model structure [11], applying frequency methods for the feedback system [12], along with learning theory [13].

The review [15] is devoted to the analysis of methods for identifying nonlinear processes in structural dynamics. Here, modifications of the frequency approach are mainly applied. Disadvantages of approaches based on linearization, harmonic balance, and the restoring force surface method are noted. In [16], regression analysis and Hilbert transform are applied. In [17], chaos theory methods are used for identifying bifurcation processes.

Various approaches based on estimating parameters of the transfer function of a feedback system are presented in [18–21]. The identification of the feedback system is often reduced to identifying an open system. In [22, 23], difficulties in identifying the feedback system are noted.

Thus, it follows from the presented review that frequency methods are generally used for identifying a nonlinear system. In some cases, approaches to estimating the nonlinearity structure are proposed. Typically, different procedures of nonlinearity linearization from the given class are applied. However, the identification of systems having multiple nonlinearities has received less attention: this is possibly due to the inherent complexity of such systems. In these cases, different approaches and identification methods based on the localization of nonlinearities are used.

The problem of SSN identification is complex and insufficiently studied. It requires solving a number of problems whose consideration is given below.

PROBLEM STATEMENT

We shall consider system $\mathbf{S_F}$

$$\dot{\mathbf{X}}(t) = \mathbf{A}\mathbf{X}(t) + \mathbf{D}\mathbf{F}_1(\mathbf{X},t) + \mathbf{B}\mathbf{U}(t), \tag{1}$$

$$\mathbf{L}\mathbf{Y}(t) = \mathbf{C}\mathbf{X}(t) + \mathbf{F}_2(\mathbf{X},t), \tag{2}$$

where $\mathbf{X} \in \mathbb{R}^m$ is the state vector; $\mathbf{A} \in \mathbb{R}^{m\times m}$ is the state matrix; $D \in \mathbb{R}^{m\times q}$; $F_1(X,t): \mathbb{R}^m \to \mathbb{R}^q$ is the nonlinear vector function; $U \in \mathbb{R}^k$ is the input (control) vector; $B \in \mathbb{R}^{m\times k}$; $Y \in \mathbb{R}^n$ is the output vector; $C \in \mathbb{R}^{n\times m}$, $F_2(X,t): \mathbb{R}^m \to \mathbb{R}^n$ is the disturbance (measurement errors) vector; $\mathcal{L}$ is the operator defining the way of forming vector $\mathbf{Y}$; t is time. In some cases, $\mathcal{L}$ may be the differential operator indicating dynamic properties of the measurement system.

We shall consider the following data set:

$$\mathbb{I}_o = \left\{Y(t), U(t), t \in \left[t_0, t_N\right]\right\}, t_N < \infty. \tag{3}$$

Assumption 1. Elements $\varphi_{1,i}(x_j) \in \mathbf{F}_1$, $\varphi_{2,i}(x_j) \in \mathbf{F}_2$ (non-linear functions belonging to $\mathbf{F}$) are smooth single-valued functions.

In some cases, condition $\varphi_i^1(x_j) = \varphi_i^1\left(\varphi_k^1(x_j)\right)$, $i \neq k$ may be satisfied. For estimating the parameters of matrices $\mathbf{A}$, $\mathbf{D}$, $\mathbf{B}$, and $\mathbf{C}$, the following model is applied:

$$\begin{cases} \dot{\hat{\mathbf{X}}}(t) = \hat{\mathbf{A}}(t)\hat{\mathbf{X}}(t) + \hat{\mathbf{D}}(t)\mathbf{F}_1(\mathbf{X},t) + \hat{\mathbf{B}}(t)\mathbf{U}(t), \\ \mathcal{L}\hat{\mathbf{Y}}(t) = \mathbf{C}\mathbf{X}(t) + \mathbf{F}_2(\mathbf{X},t), \end{cases} \tag{4}$$

where $\hat{\mathbf{A}}(t)$, $\hat{\mathbf{D}}(t)$, $\hat{\mathbf{B}}(t)$ are matrices with adjustable parameters.

Problem: for system (1) satisfying assumption 1, construct model (4) on the basis of the analysis of $\mathbb{I}_o$ and find such regularities of adjusting parameters of matrices $\hat{\mathbf{A}}(t)$, $\hat{\mathbf{D}}(t)$, and $\hat{\mathbf{B}}(t)$, that

$$\lim_{t\to\infty} \left\| \hat{\mathbf{Y}}(t) - \mathbf{Y}(t) \right\| \leq \delta_y,$$

where $\|\cdot\|$ is the Euclidean norm, while $\delta_y \geq 0$ specifies the model (4) accuracy.

EXCITATION CONSTANCY CONDITION

The excitation constancy (EC) condition plays an important role in parametric estimation problems. If the system is nonlinear, then the fulfillment of this condition may not be sufficient. As shown in [24, 25], the system should have the property of S-synchronizability in order to take into account the nonlinear properties of the system.

Let there exist: (1) bounded vector $P \in \mathbb{R}^m$ and its corresponding set of frequencies $\Omega_P(\omega)$; (2) a set of allowable input frequencies $\Omega_S(\omega)$ providing S-synchronizability of the system. Then the EC condition for matrix $\mathbf{B_P}(t) = \mathbf{P}(t)\mathbf{P}^{\mathrm{T}}(t)$ has the following form:

$$\mathcal{PE}^{\mathrm{S}}_{\alpha,\bar{\alpha}}: \left(\alpha \mathbf{I}_l \leq \mathbf{B}_P(t) \leq \bar{\alpha}\mathbf{I}_l\right) \& \left(\Omega_{\mathbf{P}}(\omega) \subseteq \Omega_{\mathrm{S}}(\omega)\right) \quad (5)$$

for $\alpha > 0$ and $\forall t \geq t_0$ on some interval $T > 0$, where $\bar{\alpha} > 0$ is some number, while $I_l \in \mathbb{R}^{m\times m}$ is the identity matrix. Usually, $\mathbf{P}(t)$ is the vector of measurements and state variables.

THE STRUCTURAL-PARAMETRIC APPROACH TO IDENTIFICATION

The following sets out the procedure for identifying system $\mathbf{S_F}$ based on the structural-parametric approach (SPA) [26]. Depending on the available a priori information, several stages implementing SPA can be applied. The system (1), (2) has a complex form; moreover, the synthesis of adaptive algorithms requires a priori information about its structure. The composition of the subsystems included in system $\mathbf{S_F}$ is assumed to be known. Hence, based on the dimensionality of the system output vector, matrix $\mathbf{A}$ can be divided into n blocks (subsystems $\mathbf{S}_{\mathbf{F},\,n} \subseteq \mathbf{S_F}$, $\{\mathbf{s}_{\mathbf{F},\,i}\} \in \mathbf{S}_{\mathbf{F},\,n}$, $j < n$). Analyzing the subsystems (blocks), we shall distinguish those containing nonlinearities, i.e., $\mathbf{S}_{\mathbf{F},\,\mathbf{nonlin}} \subseteq \mathbf{S}_{\mathbf{F},\,n}$. Then we shall apply SPA to each $\mathbf{S}_{\mathbf{F},\,\mathbf{nonlin},\,k} \in \mathbf{S}_{\mathbf{F},\,\mathbf{nonlin}}$ element. If subsystem $\{\mathbf{s}_{\mathbf{F},\,i}\} \in \mathbf{S}_{\mathbf{F},\,n} \setminus \mathbf{S}_{\mathbf{F},\,\mathbf{nonlin}}$ does not contain nonlinearities, then the adaptive identification procedure is applied to it.

Remark 1. The structural-parametric approach is based on the S-synchronizability of the system and the fulfillment of condition (5).

Under uncertainty, SPA can be divided into two procedures: (1) structural S_{F} analysis and (2) parametric estimation (adaptive identification). These stages are described in detail in [26].

Remark 2. The structural identifiability (S-synchronizability) of a system is greatly influenced by the means of connecting subsystems and mutual influence of variables. In this case, estimating the identifiability of the system nonlinear elements requires the construction of a diagram of the mutual influence in the system [27]. By analyzing the interrelations, the effect of influencing variables can be excluded to determine the structural identifiability of the system (nonlinearity). Constructing the mutual influence diagram is possible only if condition $\mathcal{PE}^{\mathrm{S}}_{\alpha,\bar{\alpha}}$ is satisfied.

Remark 3. Estimating the S-synchronizability of the system is based on the analysis of a special class of dynamic structures $\mathbf{S}_{ey}$ indicating the nonlinearity structure of the corresponding subsystem $\mathbf{S}_{\mathbf{F},\,\mathbf{nonlin},\,k}$. The method of their construction is described in [24, 25].

Remark 4. The obtained estimates of the nonlinearity structure in (1), (2) comprise the basis for implementing the adaptive parametric identification of system $\mathbf{S_F}$.

Remark 5. If a priori information about the nonlinear properties is known, then the structural analysis stage of system $\mathbf{S_F}$ may be skipped.

ADAPTIVE IDENTIFICATION OF THE $\mathbf{S_F}$ SYSTEM

We shall consider subsystem $\mathbf{s}_{\mathbf{F},\,i} \in \mathbf{S_F}$, $\dim \mathbf{s}_{\mathbf{F},\,i} = n_i$, $\mathcal{L}$ being the linear operator in (2). Let set $\mathbb{I}_{\mathrm{o},i} \subset \mathbb{I}_{\mathrm{o}}$ be known for $\mathbf{s}_{\mathbf{F},\,i}$. Subsystem $\mathbf{s}_{\mathbf{F},\,i}$ is described by the following equations:

$$\begin{cases} \dot{\mathbf{X}}_{\mathbf{s}_{\mathbf{F},i}} = \mathbf{A}_{\mathbf{s}_{\mathbf{F},i}} \mathbf{X}_{\mathbf{s}_{\mathbf{F},i}} + \mathbf{D}_{1,\mathbf{s}_{\mathbf{F},i}} \mathbf{F}_{1,\mathbf{s}_{\mathbf{F},i}}(\mathbf{X}) + \mathbf{B}_{\mathbf{s}_{\mathbf{F},i}} \mathbf{U}_{\mathbf{s}_{\mathbf{F},i}}, \\ y_{\mathbf{s}_{\mathbf{F},i}} = x^1_{\mathbf{s}_{\mathbf{F},i}}. \end{cases} \quad (6)$$

We shall represent (6) in the form of the n_ith order differential equation, as follows:

$$W_{\mathbf{s}_{\mathbf{F},i}}(\upsilon) y_{\mathbf{s}_{\mathbf{F},i}} = \sum_{k=1}^{n_i} \sum_{j=1}^{n_i} \left(d^{k,h}_{1,\mathbf{s}_{\mathbf{F},i}} f^{h,j}_{\mathbf{s}_{\mathbf{F},i}} + b^{k,j}_{\mathbf{s}_{\mathbf{F},i}} u^{j}_{\mathbf{s}_{\mathbf{F},i}} \right), \quad (7)$$

where $\mathbf{X}_{\mathbf{s}_{\mathbf{F},i}} \in \mathbb{R}^{n_i}$ is the state vector of subsystem $\mathbf{s}_{\mathbf{F},\,i}$, $x^1_{\mathbf{s}_{\mathbf{F},i}}$ being the first $\mathbf{X}$ element; $\mathbf{D}_{1,\mathbf{s}_{\mathbf{F},i}}$, $\mathbf{F}_{1,\mathbf{s}_{\mathbf{F},i}}$ and $\mathbf{B}_{\mathbf{s}_{\mathbf{F},i}}$ are matrices of appropriate dimensions; $d^{k,h}_{1,\mathbf{s}_{\mathbf{F},i}} \in \mathbf{D}_{1,\mathbf{s}_{\mathbf{F},i}}$, $f^{j}_{\mathbf{s}_{\mathbf{F},i}} \in \mathbf{F}_{1,\mathbf{s}_{\mathbf{F},i}}$, $b^{k,j}_{\mathbf{s}_{\mathbf{F},i}} \in \mathbf{B}_{\mathbf{s}_{\mathbf{F},i}}$, $u^{j}_{\mathbf{s}_{\mathbf{F},i}} \in \mathbf{U}_{\mathbf{s}_{\mathbf{F},i}}$; $\upsilon = d/dt$, $W_{\mathbf{s}_{\mathbf{F},i}}(\upsilon)$ is the polynomial of n_i degree. Matrix $\mathbf{A}_{\mathbf{s}_{\mathbf{F},i}} \in \mathbb{R}^{n_i \times n_i}$ is the Hurwitz matrix. We shall divide left and right parts of (7) by polynomials of $n_i - 1$ degree

$$H_{\mathbf{s}_{\mathbf{F},i}}(\upsilon) = \prod_{k=1}^{n_i - 1} \left(\upsilon + \mu_k\right), \quad (8)$$

and obtain

$$\dot{y}_{\mathbf{s}_{\mathbf{F},i}} = -\eta_{\mathbf{s}_{\mathbf{F},i}} y_{\mathbf{s}_{\mathbf{F},i}} + \\ + \sum_{k=1}^{n_i-1}\sum_{j=1}^{n_i-1}\left(\tilde{d}^{k,h}_{\mathbf{s}_{\mathbf{F},i}} p_{f^{h,j}_{\mathbf{s}_{\mathbf{F},i}}} + \tilde{b}^{k,j}_{\mathbf{s}_{\mathbf{F},i}} p_{u^{j}_{\mathbf{s}_{\mathbf{F},i}}} \right), \quad (9)$$

where υ, μ_k are positive numbers, $\eta_{\mathbf{s}_{\mathbf{F},i}} > 0$, $\tilde{d}^{k,h}_{\mathbf{s}_{\mathbf{F},i}}$ and $\tilde{b}^{k,j}_{\mathbf{s}_{\mathbf{F},i}}$ depend on parameters of subsystem $\mathbf{s}_{\mathbf{F},i}$ and μ_k, and variables $p_{z^{j}_{\mathbf{s}_{\mathbf{F},i}}}$ ($z = f, u$) satisfy the following equation:

$$\dot{p}_{z^{j}_{\mathbf{s}_{\mathbf{F},i}}} = -\mu_j p_{z^{j}_{\mathbf{s}_{\mathbf{F},i}}} + z^{j}_{\mathbf{s}_{\mathbf{F},i}}. \quad (10)$$

Remark 6. The structure of the right part of (9) is determined by the type of matrix $\mathbf{A}_{\mathbf{s}_{\mathbf{F},i}}$ and polynomial $H_{\mathbf{s}_{\mathbf{F},i}}(\upsilon)$.

The adaptive model for estimating parameters (9), (10) is the following:

$$\dot{\hat{y}}_{\mathbf{s}_{\mathbf{F},i}} = -k_{\mathbf{s}_{\mathbf{F},i}}\left(\hat{y}_{\mathbf{s}_{\mathbf{F},i}} - y_{\mathbf{s}_{\mathbf{F},i}}\right) + \\ + \hat{\kappa}_{\mathbf{s}_{\mathbf{F},i}} y_{\mathbf{s}_{\mathbf{F},i}} + \sum_{k=1}^{n_i-1}\sum_{j=1}^{n_i-1}\left(\hat{\tilde{d}}^{k,h}_{\mathbf{s}_{\mathbf{F},i}} p_{f^{h,j}_{\mathbf{s}_{\mathbf{F},i}}} + \hat{\tilde{b}}^{k,j}_{\mathbf{s}_{\mathbf{F},i}} p_{u^{j}_{\mathbf{s}_{\mathbf{F},i}}} \right), \quad (11)$$

where $k_{\mathbf{s}_{\mathbf{F},i}} > 0$, $\hat{\kappa}_{\mathbf{s}_{\mathbf{F},i}}$, $\hat{\tilde{d}}^{k,h}_{\mathbf{s}_{\mathbf{F},i}}$, and $\hat{\tilde{b}}^{k,j}_{\mathbf{s}_{\mathbf{F},i}}$ are adjustable parameters.

The equation for the identification error is the following:

$$\dot{e}_{\mathbf{s}_{\mathbf{F},i}} = -k_{\mathbf{s}_{\mathbf{F},i}} e_{\mathbf{s}_{\mathbf{F},i}} + \Delta\kappa_{\mathbf{s}_{\mathbf{F},i}} y_{\mathbf{s}_{\mathbf{F},i}} + \\ + \sum_{k=1}^{n_i-1}\sum_{j=1}^{n_i-1}\left(\Delta\tilde{d}^{k,h}_{\mathbf{s}_{\mathbf{F},i}} p_{f^{h,j}_{\mathbf{s}_{\mathbf{F},i}}} + \Delta\tilde{b}^{k,j}_{\mathbf{s}_{\mathbf{F},i}} p_{u^{j}_{\mathbf{s}_{\mathbf{F},i}}} \right), \quad (12)$$

where $\Delta\tilde{d}^{k,h}_{\mathbf{s}_{\mathbf{F},i}} = \hat{\tilde{d}}^{k,h}_{\mathbf{s}_{\mathbf{F},i}} - \tilde{d}^{k,h}_{\mathbf{s}_{\mathbf{F},i}}$, $\Delta\tilde{b}^{k,j}_{\mathbf{s}_{\mathbf{F},i}} = \hat{\tilde{b}}^{k,j}_{\mathbf{s}_{\mathbf{F},i}} - \tilde{b}^{k,j}_{\mathbf{s}_{\mathbf{F},i}}$, $e_{\mathbf{s}_{\mathbf{F},i}} = \hat{y}_{\mathbf{s}_{\mathbf{F},i}} - y_{\mathbf{s}_{\mathbf{F},i}}$, $\Delta\kappa_{\mathbf{s}_{\mathbf{F},i}} = \hat{\kappa}_{\mathbf{s}_{\mathbf{F},i}} - \eta_{\mathbf{s}_{\mathbf{F},i}}$.

We shall consider the Lyapunov function $V_{e,i}\left(e_{\mathbf{s}_{\mathbf{F},i}}\right) = 0.5 e^2_{\mathbf{s}_{\mathbf{F},i}}$. Then the following is obtained for $\dot{V}_{e,i}\left(e_{\mathbf{s}_{\mathbf{F},i}}\right)$:

$$\dot{V}_{e,i}\left(e_{\mathbf{s}_{\mathbf{F},i}}\right) = e_{\mathbf{s}_{\mathbf{F},i}}\left[-k_{\mathbf{s}_{\mathbf{F},i}} e_{\mathbf{s}_{\mathbf{F},i}} + \Delta\kappa_{\mathbf{s}_{\mathbf{F},i}} y_{\mathbf{s}_{\mathbf{F},i}} + \right. \\ \left. + \sum_{k=1}^{n_i-1}\sum_{j=1}^{n_i-1}\left(\Delta\tilde{d}^{k,h}_{\mathbf{s}_{\mathbf{F},i}} p_{f^{h,j}_{\mathbf{s}_{\mathbf{F},i}}} + \Delta\tilde{b}^{k,j}_{\mathbf{s}_{\mathbf{F},i}} p_{u^{j}_{\mathbf{s}_{\mathbf{F},i}}} \right)\right].$$

If variables $p_{f^{h,j}_{\mathbf{s}_{\mathbf{F},i}}}$, $p_{u^{j}_{\mathbf{s}_{\mathbf{F},i}}}$ have property $\mathcal{PE}^{\mathrm{S}}_{\alpha,\bar{\alpha}}$, then the following is obtained from condition $\dot{V}_i\left(e_{\mathbf{s}_{\mathbf{F},i}}\right) < 0$:

$$\Delta\dot{\kappa}_{\mathbf{s}_{\mathbf{F},i}} = -\gamma_{\kappa,\mathbf{s}_{\mathbf{F},i}} e_{\mathbf{s}_{\mathbf{F},i}} y_{\mathbf{s}_{\mathbf{F},i}}, \\ \Delta\dot{\tilde{d}}^{k,h}_{1,\mathbf{s}_{\mathbf{F},i}} = -\gamma_{k,h,\mathbf{s}_{\mathbf{F},i}} e_{\mathbf{s}_{\mathbf{F},i}} p_{f^{h,j}_{\mathbf{s}_{\mathbf{F},i}}}, \quad (13) \\ \Delta\dot{\tilde{b}}^{k,j}_{\mathbf{s}_{\mathbf{F},i}} p_{u^{j}_{\mathbf{s}_{\mathbf{F},i}}} = -\gamma_{k,j,\mathbf{s}_{\mathbf{F},i}} e_{\mathbf{s}_{\mathbf{F},i}} p_{u^{j}_{\mathbf{s}_{\mathbf{F},i}}},$$

where $\gamma_{k,j,\mathbf{s}_{\mathbf{F},i}} > 0$ is the amplification factor.

It is not difficult to obtain algorithms for the parameters of the model (11) from (13).

Thus, the adaptive identification system of the subsystem $\mathbf{S}_{\mathbf{F},i}$ is described by Eqs. (12) and (13). We shall denote it as $\mathbf{AS}_{\mathbf{F},i}$.

We shall consider the Lyapunov function

$$V_{\Delta,i} = 0.5\gamma^{-1}_{\kappa,\mathbf{s}_{\mathbf{F},i}}\left(\Delta\kappa_{\mathbf{s}_{\mathbf{F},i}}\right)^2 + \\ + 0.5\mathrm{Sp}\left(\Delta\mathbf{D}_{1,\mathbf{s}_{\mathbf{F},i}} \boldsymbol{\Gamma}^{-1}_{k,h,\mathbf{s}_{\mathbf{F},i}} \Delta\mathbf{D}^{\mathrm{T}}_{1,\mathbf{s}_{\mathbf{F},i}} \right) + \quad (14) \\ + 0.5\mathrm{Sp}\left(\Delta\mathbf{B}_{\mathbf{s}_{\mathbf{F},i}} \boldsymbol{\Gamma}^{-1}_{k,j,\mathbf{s}_{\mathbf{F},i}} \Delta\mathbf{B}^{\mathrm{T}}_{\mathbf{s}_{\mathbf{F},i}} \right),$$

where $\boldsymbol{\Gamma}_{k,h,\mathbf{s}_{\mathbf{F},i}} = \mathrm{diag}(\gamma_{k,h,\mathbf{s}_{\mathbf{F},i}})$, Sp(·) is the spur of matrix, $\boldsymbol{\Gamma}^{-1}_{k,j,\mathbf{s}_{\mathbf{F},i}} = \mathrm{diag}(\gamma_{k,j,\mathbf{s}_{\mathbf{F},i}})$, $\Delta\mathbf{D}_{1,\mathbf{s}_{\mathbf{F},i}}$ and $\Delta\mathbf{B}_{\mathbf{s}_{\mathbf{F},i}}$ contain elements $\Delta\tilde{d}^{k,h}_{\mathbf{s}_{\mathbf{F},i}}$ and $\Delta\tilde{b}^{k,j}_{\mathbf{s}_{\mathbf{F},i}}$, respectively. Let $\Delta\mathbf{K}_{\mathbf{s}_{\mathbf{F},i}} \triangleq \left[\Delta\mathbf{D}_{1,\mathbf{s}_{\mathbf{F},i}}, \Delta\mathbf{B}_{\mathbf{s}_{\mathbf{F},i}}\right]$ and $V_{\mathbf{s}_{\mathbf{F},i}}(t) = V_e(t) + V_{\Delta,i}(t)$.

Theorem 1. Let: 1) functions $V_{e,i}(t)$, $V_{\Delta,i}(t)$ are positively definite and admit the infinitesimal higher limit at $|e| \to \infty$ and $\left\|\Delta\mathbf{K}_{\mathbf{s}_{\mathbf{F},i}}\right\| \to \infty$, $\left|\Delta\kappa_{\mathbf{s}_{\mathbf{F},i}}\right| \to \infty$; 2) matrix $\mathbf{A}_{\mathbf{s}_{\mathbf{F},i}} \in \mathbb{R}^{n_i \times n_i}$ is the Hurwitz matrix; 3) $p_{f^{h,j}_{\mathbf{s}_{\mathbf{F},i}}}$, $p_{u^{j}_{\mathbf{s}_{\mathbf{F},i}}}$ have property $\mathcal{PE}^{\mathrm{S}}_{\alpha,\bar{\alpha}}$. Then all trajectories of system $\mathbf{AS}_{\mathbf{F},i}$ are bounded, lie in the following domain:

$$G_t = \left\{\left(e_{\mathbf{s}_{\mathbf{F},i}}, \Delta\kappa_{\mathbf{s}_{\mathbf{F},i}}, \Delta\mathbf{K}_{\mathbf{s}_{\mathbf{F},i}}\right) : V_{\mathbf{s}_{\mathbf{F},i}}(t) \le V_{\mathbf{s}_{\mathbf{F},i}}\left(t_0\right)\right\}$$

and the following estimation is valid:

$$2k_{\mathbf{s}_{\mathbf{F},i}} \int_{t_0}^{t} V_{e,i}(\tau)\, d\tau \le V_{\mathbf{s}_{\mathbf{F},i}}(t_0) - V_{\mathbf{s}_{\mathbf{F},i}}(t).$$

Let estimation called A1 be true for $V_{\Delta,i}(t)$:

$$0.5\,\underline{\vartheta}\left\{\mathrm{Sp}\left(\Delta\mathbf{K}_{\mathbf{s}_{\mathbf{F},i}}(t)\Delta\mathbf{K}^{\mathrm{T}}_{\mathbf{s}_{\mathbf{F},i}}(t)\right) + \left(\Delta\kappa_{\mathbf{s}_{\mathbf{F},i}}\right)^2\right\} \le V_{\Delta,i}(t) \le$$

$$\le 0.5\,\overline{\vartheta}\left\{\mathrm{Sp}\left(\Delta\mathbf{K}_{\mathbf{s}_{\mathbf{F},i}}(t)\Delta\mathbf{K}^{\mathrm{T}}_{\mathbf{s}_{\mathbf{F},i}}(t)\right) + \left(\Delta\kappa_{\mathbf{s}_{\mathbf{F},i}}\right)^2\right\},$$

where $\boldsymbol{\Gamma}_{\mathbf{s}_{\mathbf{F},i}} = \boldsymbol{\Gamma}_{k,h,\mathbf{s}_{\mathbf{F},i}} \dot{+} \boldsymbol{\Gamma}_{k,j,\mathbf{s}_{\mathbf{F},i}}$, $\dot{+}$ is the sign of the direct sum of matrices, $\beta_1\left(\boldsymbol{\Gamma}_{\mathbf{s}_{\mathbf{F},i}}\right)$, $\beta_l\left(\boldsymbol{\Gamma}_{\mathbf{s}_{\mathbf{F},i}}\right)$ are the minimal and maximal eigenvalues of matrix $\boldsymbol{\Gamma}_{\mathbf{s}_{\mathbf{F},i}}$, $\underline{\vartheta} = \min\left(\beta_l^{-1}\left(\boldsymbol{\Gamma}_{\mathbf{s}_{\mathbf{F},i}}\right), \gamma^{-1}_{\kappa,\mathbf{s}_{\mathbf{F},i}}\right)$, and $\overline{\vartheta} = \max\left(\beta_1^{-1}\left(\boldsymbol{\Gamma}_{\mathbf{s}_{\mathbf{F},i}}\right), \gamma^{-1}_{\kappa,\mathbf{s}_{\mathbf{F},i}}\right)$.

Theorem 2. Let the following conditions be satisfied: 1) positive definite Lyapunov functions $V_{e,i}(t) = 0.5 e_{\mathbf{s}_{\mathbf{F},i}}(t)$,

$$V_{\Delta,i}(t) = 0.5\gamma^{-1}_{\kappa,\mathbf{s}_{\mathbf{F},i}}\left(\Delta\kappa_{\mathbf{s}_{\mathbf{F},i}}\right)^2 +$$

$$+ 0.5\mathrm{Sp}\left(\Delta\mathbf{D}_{1,\mathbf{s}_{\mathbf{F},i}}\boldsymbol{\Gamma}^{-1}_{k,h,\mathbf{s}_{\mathbf{F},i}}\Delta\mathbf{D}^{\mathrm{T}}_{1,\mathbf{s}_{\mathbf{F},i}}\right) +$$

$$+ 0.5\mathrm{Sp}\left(\Delta\mathbf{B}_{\mathbf{s}_{\mathbf{F},i}}\boldsymbol{\Gamma}^{-1}_{k,j,\mathbf{s}_{\mathbf{F},i}}\Delta\mathbf{B}^{\mathrm{T}}_{\mathbf{s}_{\mathbf{F},i}}\right)$$

admit the infinitesimal higher limit at $\left|e_{\mathbf{s}_{\mathbf{F},i}}(t)\right| \to 0$, $\left|\Delta\kappa_{\mathbf{s}_{\mathbf{F},i}}\right| \to 0$, $\left\|\Delta\mathbf{K}_{\mathbf{s}_{\mathbf{F},i}}\right\| \to 0$; 2) matrix $\mathbf{W}_{\mathbf{s}_{\mathbf{F},i}}(t) = \mathbf{P}^{\mathrm{T}}_{f^{h,j}_{\mathbf{s}_{\mathbf{F},i}},u^{j}_{\mathbf{s}_{\mathbf{F},i}}}(t)\mathbf{P}_{f^{h,j}_{\mathbf{s}_{\mathbf{F},i}},u^{j}_{\mathbf{s}_{\mathbf{F},i}}}(t)$ and $y_{\mathbf{s}_{\mathbf{F},i}}$ are piecewise continuous bounded and $\mathbf{W}_{\mathbf{s}_{\mathbf{F},i}}(t) \in \mathcal{PE}^{\mathrm{S}}_{\alpha,\overline{\alpha}}$, $y_{\mathbf{s}_{\mathbf{F},i}}(t) \in \mathcal{PE}^{\mathrm{S}}_{\alpha_{y_i},\overline{\alpha}_{y_i}}$; 3) the following equality is valid:

$$e_{\mathbf{s}_{\mathbf{F},i}} \sum_{k=1}^{n_i-1}\sum_{j=1}^{n_i-1}\left(\Delta\tilde{d}^{k,h}_{\mathbf{s}_{\mathbf{F},i}} p_{f^{h,j}_{\mathbf{s}_{\mathbf{F},i}}} + \Delta\tilde{b}^{k,j}_{\mathbf{s}_{\mathbf{F},i}} p_{u^{j}_{\mathbf{s}_{\mathbf{F},i}}}\right) =$$

$$= \pi\left\{\mathrm{Sp}\left(\Delta\mathbf{K}_{\mathbf{s}_{\mathbf{F},i}}(t)P^{\mathrm{T}}_{f^{h,j}_{\mathbf{s}_{\mathbf{F},i}},u^{j}_{\mathbf{s}_{\mathbf{F},i}}} P_{f^{h,j}_{\mathbf{s}_{\mathbf{F},i}},u^{j}_{\mathbf{s}_{\mathbf{F},i}}} \Delta\mathbf{K}^{\mathrm{T}}_{\mathbf{s}_{\mathbf{F},i}}(t)\right) + \left(\Delta\hat{e}_{\mathbf{s}_{\mathbf{F},i}} y_{\mathbf{s}_{\mathbf{F},i}}\right)^2 + e^2_{\mathbf{s}_{\mathbf{F},i}}\right\}$$

in domain $O_{\mathrm{v}}(O)$, where $\pi > 0$, $O = \{0,\ 0^{n_i\times n_i}\} \subset \subset \mathbb{R}\times\mathbb{R}^{n_i\times n_i}\times\mathbb{J}_{0,\infty}$, $0^{n_i\times n_i} \in \mathfrak{R}^{n_i\times n_i}$ is a zero matrix, O_{v} is some neighborhood of point O, $t \in [0,\infty] = \mathbb{J}_{0,\infty}$ is time interval; 4) estimation A1 is valid for function $V_{\Delta,i}(t)$; 5) the following system of inequalities is fulfilled:

$$\begin{bmatrix}\dot{V}_{e,i}\\ \dot{V}_{\Delta,i}\end{bmatrix} \le \underbrace{\begin{bmatrix}-k_{\mathbf{s}_{\mathbf{F},i}} & \dfrac{\overline{\alpha}\,\overline{\vartheta}}{2k_{\mathbf{s}_{\mathbf{F},i}}}\\ \dfrac{8}{3}\pi & -\dfrac{3}{8}\pi\alpha_i\end{bmatrix}}_{\mathbf{A_V}}\begin{bmatrix}V_{e,i}\\ V_{\Delta,i}\end{bmatrix}$$

for $\dot{V}_{e,i}, \dot{V}_{\Delta,i}$; 6) the upper solution for $\mathbf{V}_{e_i,\Delta_i}(t) = \left[V_{e,i}(t)\, V_{\Delta,i}(t)\right]^{\mathrm{T}}$ satisfies equation $\dot{\mathbf{S}} = \mathbf{A_V}\mathbf{S}$, if the following inequality is valid:

$$V_{\rho}(t) \le s_{\rho}(t)\ \forall(t \ge t_0)\ \&\ (V_{\rho}(t_0) \le s_{\rho}(t_0)),$$

$\rho = e, i; \Delta, i$ for elements $V_{e,i}(t)$, $V_{\Delta,i}(t)$. Then adaptive system $\mathbf{AS}_{\mathbf{F},i}$ is exponentially stable with the following estimate:

$$\mathbf{V}_{e_i,\Delta_i}(t) \le e^{\mathbf{A_V}(t-t_0)}\mathbf{S}(t_0),$$

if

$$k_{\mathbf{s}_{\mathbf{F},i}} > 0, k_{\mathbf{s}_{\mathbf{F},i}} \ge \frac{4}{3}\sqrt{\frac{2\overline{\alpha}\,\overline{\vartheta}}{\alpha_i}}. \qquad (15)$$

The proof of theorems is based on the approach described in [26].

Theorem 2 shows: if the information matrix

$$\mathbf{W}_{\mathbf{s}_{\mathbf{F},i}}(t) = \mathbf{P}^{\mathrm{T}}_{f^{h,j}_{\mathbf{s}_{\mathbf{F},i}},u^{j}_{\mathbf{s}_{\mathbf{F},i}}}(t)\mathbf{P}_{f^{h,j}_{\mathbf{s}_{\mathbf{F},i}},u^{j}_{\mathbf{s}_{\mathbf{F},i}}}(t)$$

is continuously excitable, then the adaptive system $\mathbf{AS}_{\mathbf{F},i}$ allows obtaining true estimates of the system $\mathbf{S}_{\mathbf{F},i}$ parameters. In this case, the system parameters should satisfy conditions (15).

EXAMPLES

We shall consider the system with nonlinear correction of the nonlinear system. It contains the amplifier with electric motor and the relay control described by function $f_1(u)$. The nonlinear feedback (fb in an index notation) on speed with parabolic characteristic $f_2(x_2)$ is used as a correction device, as follows:

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = x_3, \\ \dot{x}_3 = -a_1 x_2 - a_2 x_3 + b f_1(u), \\ y = x_1, \end{cases} \quad (16)$$

$$f_1(u) = \begin{cases} c, \text{ if } u \ge d, \\ 0, \text{ if } -d \le u \le d, \\ -c, \text{ if } u < -d, \end{cases}$$

where $u = g - x_1 - f_2(x_2)$ is control; $f_2(x_2) = k_{\text{fb}} x_2^2 \operatorname{sign}(x_2)$; g is the system input; $c > 0$, $d > 0$; $k_{\text{fb}} > 0$; x_i is a state variable; and a_i, b are system parameters.

Equation (13) and polynomials in (7), (8) have the following form:

$$\left(\upsilon + \eta - \frac{v_1}{\upsilon + \mu_1} - \frac{v_2}{\upsilon + \mu_2}\right) y = \left(\frac{b_1}{\upsilon + \mu_1} + \frac{b_2}{\upsilon + \mu_2}\right) f_1(u). \quad (17)$$

Equation (17) is obtained as follows. First, system (16) is written in the space ($y(t)$, $f_1(u)$). So, Eq. (7) is obtained. Since the system has the third order, to find the equation for $\dot{y}(t)$, according to the results of the ADAPTIVE IDENTIFICATION OF THE $\mathbf{S_F}$ SYSTEM section, both parts of the obtained Eq. (7) are divided by polynomial $H(\upsilon) = (\upsilon + \mu_1)(\upsilon + \mu_2)$. Then the left and right parts of (7) are decomposed into prime fractions, and Eq. (17) is derived.

We assume $\upsilon = d/dt$. The analogue of Eqs. (8) and (9) for system (16) is obtained from (17), as follows:

$$\dot{y} = -\eta y + v_1 p_{y,\mu_1} + v_2 p_{y,\mu_2} + b_1 p_{f_1,\mu_1} + b_2 p_{f_1,\mu_2}, \quad (18)$$

$$\begin{aligned} \dot{p}_{y,\mu_1} &= -\mu_1 p_{y,\mu_1} + y, \\ \dot{p}_{y,\mu_2} &= -\mu_2 p_{y,\mu_2} + y, \\ \dot{p}_{f_1,\mu_1} &= -\mu_1 p_{f_1,\mu_1} + f_1, \\ \dot{p}_{f_1,\mu_2} &= -\mu_2 p_{f_1,\mu_2} + f_1, \end{aligned} \quad (19)$$

where $b_1 = 1.4$, $b_2 = -0.4$, $\eta = 0.35$, $v_1 = -1$, $v_2 = 1.35$, $\mu_1 = 2.05$, $\mu_2 = 2.25$. The parameters of nonlinearity f_1: $c_1 = 2$, $d = 0.5$. Input $g(t) = \sin(0.2t)$.

The adaptive model and algorithms are the following:

$$\dot{\hat{y}} = -k_y e - \hat{\eta} y + \hat{v}_1 p_{y,\mu_1} + \hat{v}_2 p_{y,\mu_2} + \hat{b}_1 p_{f_1,\mu_1} + \hat{b}_2 p_{f_1,\mu_2}, \quad (20)$$

$$\begin{aligned} &\dot{\hat{\eta}} = -\gamma_\eta e y, \dot{\hat{v}}_1 = -\gamma_{v_1} e p_{y,\mu_1}, \\ &\dot{\hat{v}}_2 = -\gamma_{v_2} e p_{y,\mu_2}, \dot{\hat{b}}_1 = -\gamma_{b_1} p_{f_1,\mu_1}, \\ &\dot{\hat{b}}_2 = -\gamma_{b_2} p_{f_1,\mu_2}, \end{aligned} \quad (21)$$

where $e = \hat{y} - y$.

The equation for error $e(t)$ is the following:

$$\dot{e} = -ke - \Delta\eta y + \Delta v_1 p_{y,\mu_1} + \Delta v_2 p_{y,\mu_2} + \Delta b_1 p_{f_1,\mu_1} + \Delta b_2 p_{f_1,\mu_2}, \quad (22)$$

where $k > 0$, $\Delta\sigma = \hat{\sigma} - \sigma$, $\sigma = \eta, v_1, v_2, b_2, b_2$. Coefficients γ_i in (21) vary in the range (0.002; 0.009).

Let $\Delta\mathbf{K} \triangleq \left[\Delta\eta, \Delta v_1, \Delta v_2, \Delta b_1, \Delta b_2\right]^{\mathrm{T}}$. The adjustment law for $\Delta\mathbf{K}$ follows from (21):

$$\Delta\dot{\mathbf{K}} = -\mathbf{\Gamma}_{\mathrm{K}} e \mathbf{P}_y, \quad (23)$$

where $\mathbf{\Gamma}_{\mathrm{K}} = \operatorname{diag}\left(\gamma_\eta, \gamma_{v_1}, \gamma_{v_2}, \gamma_{b_1}, \gamma_{b_2}\right)$,

$$\mathbf{P}_y = \left[y, p_{y,\mu_1}, p_{y,\mu_2}, p_{f_1,\mu_1}, p_{f_1,\mu_2}\right]^{\mathrm{T}}.$$

The boundedness of the system (22), (23) trajectories follows from Theorem 1. The adaptive system results are shown in Figs. 1–4. The adjustment of the model (20) parameters is shown in Figs. 1 and 2. The change in estimation error $e(t)$ is shown in Fig. 3. This change in the error is related to the change in the system output.

It is noted in [28] that by function $f_2(x_2)$, the system is unidentifiable on the set of measurements. In this case, the indirect information about the dependence of $u_1 = \omega - x_1$ on x_2 may be used. This is true since there is a relationship between u_1 and u.

2. We shall consider the self-oscillation generation system consisting of an object (variables y_1, y_2), nonlinear (variable y_3) and linear (variable y_4) converters, and the linear amplifier-converter with nonlinear actuator (variable y_5). Function $f_i(x)$ ($i = 1, 3$) is the saturation function with a dead zone

$$f_i(x) = \begin{cases} c, & \text{if } x \ge d_{2,i}, \\ 2(x - d_{1,i}), & \text{if } d_{1,i} < x < d_{2,i}, \\ 0, & \text{if } -d_{1,i} \le x \le d_{1,i}, \\ 2(x + d_{1,i}), & \text{if } -d_{1,i} < x, \\ -c, & \text{if } x < -d_{2,i}, \end{cases}$$

where $c > 0$, $d_{1,i} > 0$, and $d_{2,i} > 0$ are some numbers,

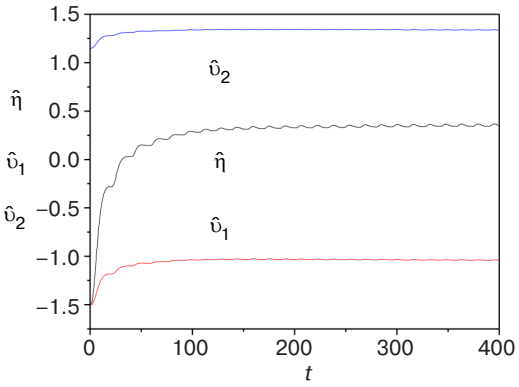


Fig. 1. Adjusting parameters of model (20)

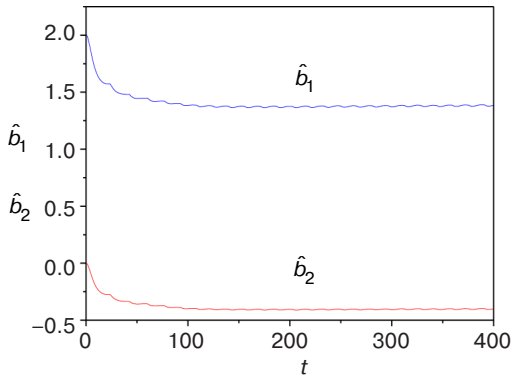


Fig. 2. Adjusting parameters $\hat{b}_i$ of model (20)

$$\dot{\mathbf{Y}} = \mathbf{AY} + \mathbf{DF}(\mathbf{Y})$$

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}, \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & -g & 0 & 0 & k_0 \\ 0 & 0 & -\frac{1}{T_1} & 0 & 0 \\ 0 & \frac{k_2}{T_2} & 0 & -\frac{1}{T_2} & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{T_3} \end{bmatrix}, \quad (24)$$

$$\mathbf{D} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \frac{1}{T_1} & 0 \\ 0 & 0 \\ 0 & -\frac{1}{T_3} \end{bmatrix}, \mathbf{F}(\mathbf{Y}) = \begin{bmatrix} f_1(y_1) \\ f_3(y_3 + y_4) \end{bmatrix},$$

where $T_i > 0$ is a time constant, $g \geq 0$. Variable y_5 is used as the input.

The phase portrait of the object is shown in Fig. 4. It shows that self-oscillations arise in the system. For identifying parameters of system (24), we shall use the ideas of the adaptive observer for the object described above. It is necessary to transform only the first two equations in (24). For this, we represent them

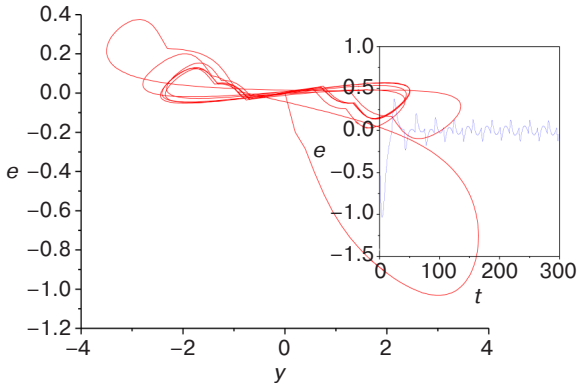


Fig. 3. Change in the estimation error

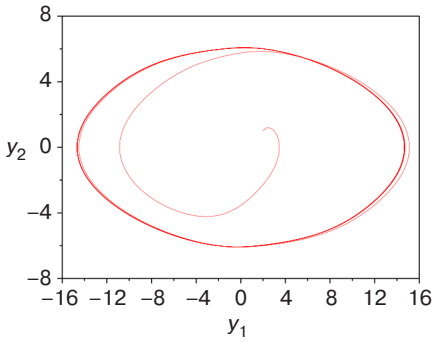


Fig. 4. Phase portrait of object (24)

in the form (38) and divide the resulting equation by υ + μ. Then variables y_1, y_5 are transformed, as follows:

$$\begin{aligned} \dot{p}_{y_1,\mu} &= -\mu p_{y_1,\mu} + y_1, \\ \dot{p}_{y_5,\mu} &= -\mu p_{y_5,\mu} + y_5, \end{aligned} \quad (25)$$

and the following identification representation is obtained:

$$\begin{aligned} \dot{y}_1 &= a_{11}y_1 + a_{12}p_{y_1} + a_{15}p_{y_5}, \\ \dot{y}_3 &= -a_3 y_3 + a_{31}f_1(y_1), \\ \dot{y}_4 &= -a_4 y_4 + a_{42}y_2, \\ \dot{y}_5 &= -a_5 y_5 + a_{53}f_3(y_3 + y_4). \end{aligned} \quad (26)$$

The adaptive system for estimating parameters of system (26) has the following form:

$$\begin{aligned} \dot{\hat{y}}_1 &= -k_1 e_1 + \hat{a}_{11}y_1 + \hat{a}_{12}p_{y_1} + \hat{a}_{15}p_{y_5}, \\ \hat{y}_2 &= +\hat{a}_{11}y_1 + \hat{a}_{12}p_{y_1} + \hat{a}_{15}p_{y_5}, \\ \dot{\hat{y}}_3 &= -k_3 e_3 + \hat{a}_{31}(f_1(y_1) - y_3), \\ \dot{\hat{y}}_4 &= -k_4 e_4 + \hat{a}_{41}y_4 + \hat{a}_{42}\hat{y}_2, \\ \dot{\hat{y}}_5 &= -k_5 e_5 + \hat{a}_{51}(y_5 + f_3(y_3 + y_4)), \end{aligned} \quad (27)$$

where $e_i = \hat{y}_i - y_i$, $i = 1, 3, 4, 5$.

If Lyapunov functions $V_i(e_i) = 0.5e_i^2$ are introduced, then adaptive algorithms for adjusting parameters of system (27) are obtained from condition $\dot{V}_i \le 0$, as follows:

$$\begin{aligned}
&\dot{\hat{a}}_{11} = -\gamma_{11} e_1 y_1, \quad \dot{\hat{a}}_{12} = -\gamma_{12} e_1 p_{y_1}, \\
&\dot{\hat{a}}_{15} = -\gamma_{15} e_1 p_{y_5}, \quad \dot{\hat{a}}_{31} = -\gamma_{31} e_3 \left(f_1(y_1) - y_3\right), \\
&\dot{\hat{a}}_{41} = -\gamma_{41} e_4 y_4, \quad \dot{\hat{a}}_{42} = -\gamma_{42} e_4 \hat{y}_2, \\
&\dot{\hat{a}}_{51} = -\gamma_{51} e_5 \left(y_5 + f_3(y_3 + y_4)\right),
\end{aligned} \tag{28}$$

where $\gamma_{ij} > 0$.

Systems (26) and (27) are simulated with the following parameters: $a_{11} = 0.55$, $a_{12} = -0.6$, $a_{15} = -1.15$, $k_1 = 2$, $\mu = 0.5$, $a_{31} = 2.21$, $k_3 = 0.8$, $k_4 = 1.5$, $a_{41} = 1.15$, $a_{42} = 0.56$, $k_5 = 1.25$, $a_{51} = 1.1$. Parameters of functions f_1, f_3: $d_{11} = 0.5$, $d_{21} = 1.5$, $c = 2$, $d_{13} = 0.25$, $d_{23} = 1.25$. Coefficients γ_{ij} vary in the range (0.001; 0.05).

The adaptive system has the following form:

$$\begin{aligned}
&\dot{e}_1 = -k_1 e_1 + \Delta a_{11} y_1 + \Delta a_{12} p_{y_1} + \Delta a_{15} p_{y_5}, \\
&e_2 = +\Delta a_{11} y_1 + \Delta a_{12} p_{y_1} + \Delta a_{15} p_{y_5}, \\
&\dot{e}_3 = -k_3 e_3 + \Delta a_{31}\left(f_1(y_1) - y_3\right), \\
&\dot{e}_4 = -k_4 e_4 - \Delta a_{41} y_4 + \Delta a_{42} \hat{y}_2, \\
&\dot{e}_5 = -k_5 e_5 - \Delta a_{51}\left(y_5 + f_3(y_3 + y_4)\right),
\end{aligned} \tag{29}$$

$$\begin{aligned}
&\Delta\dot{a}_{11} = -\gamma_{11} e_1 y_1, \quad \Delta\dot{a}_{12} = -\gamma_{12} e_1 p_{y_1}, \\
&\Delta\dot{a}_{15} = -\gamma_{15} e_1 p_{y_5}, \quad \Delta\dot{a}_{31} = -\gamma_{31} e_3 \left(f_1(y_1) - y_3\right), \\
&\Delta\dot{a}_{41} = \gamma_{41} e_4 y_4, \quad \Delta\dot{a}_{42} = -\gamma_{42} e_4 \hat{y}_2, \\
&\Delta\dot{a}_{51} = \gamma_{51} e_5 \left(y_5 + f_3(y_3 + y_4)\right),
\end{aligned} \tag{30}$$

where $\Delta a_{ij}(t) = \hat{a}_{ij}(t) - a_{ij}$.

The results of system (29), (30) are shown in Figs. 5–7. The process of adjusting parameters of the system (27) is shown in Figs. 5, and 6, while the change in the output discrepancies of system (27) is shown in Fig. 7. The results confirm the boundedness of the adaptive system trajectories.

Despite the fulfillment of condition $y_5 \in PE^{S}_{\alpha,\bar{\alpha}}$ for the input of system (24) and the S-synchronizability of the system, it is not possible to ensure the condition of asymptotic stability. This is due to the presence of nonlinearities in the system.

We shall consider the Lyapunov functions

$$V_{e,1-5}(t) = 0.5 \sum_{i=1,\, i\neq 2}^{5} e_i^2(t), \; \mathbf{V}_{e,\Delta} =$$

$$= \left[V_{e,1}, V_{\Delta,1}, V_{e,3}, V_{\Delta,3}, V_{e,4}, V_{\Delta,4}, V_{e,5}, V_{\Delta,5}\right]^{\mathrm{T}},$$

$$V_{\Delta,1-5}(t) = \underbrace{0.5\left(\gamma_{11}^{-1}\Delta a_{11}^2(t) + \gamma_{12}^{-1}\Delta a_{12}^2(t) + \gamma_{15}^{-1}\Delta a_{15}^2(t)\right)}_{V_{\Delta,1}} +$$

$$+ \underbrace{0.5\gamma_{31}^{-1}\Delta a_{31}^2(t)}_{V_{\Delta,3}} + \underbrace{0.5\left(\gamma_{41}^{-1}\Delta a_{41}^2(t) + \gamma_{42}^{-1}\Delta a_{42}^2(t)\right)}_{V_{\Delta,4}} + \tag{31}$$

$$+ \underbrace{0.5\gamma_{51}^{-1}\Delta a_{51}^2(t)}_{V_{\Delta,5}}.$$

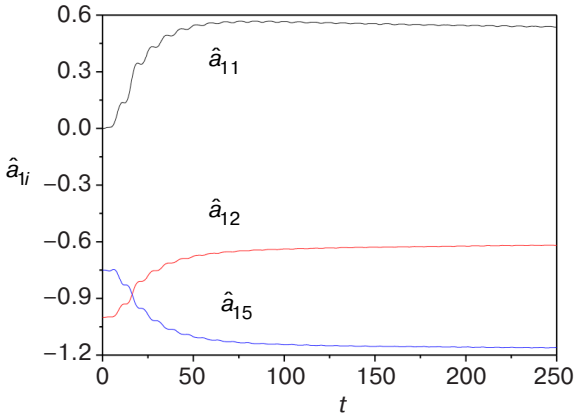


Fig. 5. Adjusting parameters of the models for estimating y_1

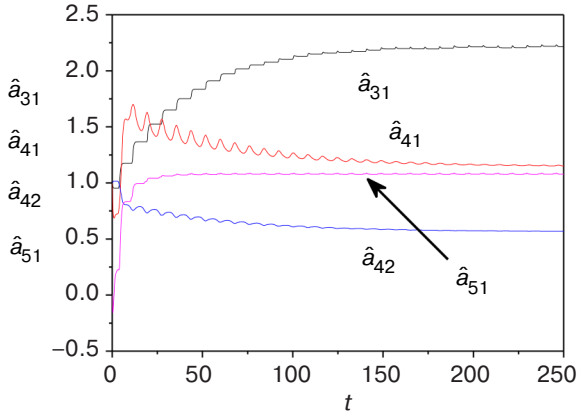


Fig. 6. Adjusting parameters of the models for estimating y_3, y_4, and y_5

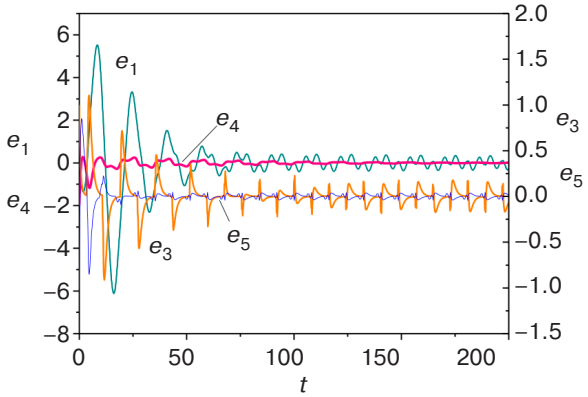


Fig. 7. Change in the discrepancy of outputs of models (27) (green line is e_1, red line is e_4, orange line is e_3, and blue line is e_5)

Theorem 3. Let the following conditions be satisfied:

1) positive definite Lyapunov functions $V_{\Delta,i}(t) = 0.5\gamma_{ij}^{-1}\left(\Delta a_{ij}\right)^2$, $V_{e,i}(t) = 0.5e_i(t)$, $V_{\Delta,1}(t) = 0.5\left(\gamma_{11}^{-1}\Delta a_{11}^2(t) + \gamma_{12}^{-1}\Delta a_{12}^2(t) + \gamma_{15}^{-1}\Delta a_{15}^2(t)\right)$ admit the infinitesimal higher limit at $|e_i(t)| \to 0$, $|\Delta a_{ij}| \to 0$;

2) y_i are piecewise continuous bounded while $\alpha = \max\left(\overline{\alpha}_{y_1}, \overline{\alpha}_{p_{y_1}}, \overline{\alpha}_{p_{y_5}}\right)$, $y_i(t) \in \mathcal{PE}^{\mathrm{S}}_{\alpha_{y_i},\overline{\alpha}_{y_i}}$, $p_{y_1} \in \mathcal{PE}^{\mathrm{S}}_{\alpha_{p_{y_1}},\overline{\alpha}_{p_{y_1}}}$, $p_{y_5} \in \mathcal{PE}^{\mathrm{S}}_{\alpha_{p_{y_5}},\overline{\alpha}_{p_{y_5}}}$, and $i = 1, 3, 4, 5$;

3) for $V_{\Delta,1}(t) = 0.5\Delta\mathbf{G}_1^{\mathrm{T}}(t)\boldsymbol{\Gamma}_1^{-1}\Delta\mathbf{G}_1(t)$, the following inequation is valid:

$$0.5\underline{\vartheta}_1\Delta\mathbf{G}_1^{\mathrm{T}}(t)\Delta\mathbf{G}_1(t) \le V_{\Delta,1}(t) \le 0.5\overline{\vartheta}_1\Delta\mathbf{G}_1^{\mathrm{T}}(t)\Delta\mathbf{G}_1(t),$$

where $\boldsymbol{\Gamma}_1 = \mathrm{diag}\left(\gamma_{11}, \gamma_{11}, \gamma_{15}\right)$, $\underline{\vartheta}_1 = \beta_l^{-1}\left(\boldsymbol{\Gamma}_1\right)$, $\Delta\mathbf{G}_1(t) = \left[\Delta a_{11}(t), \Delta a_{12}(t), \Delta a_{15}(t)\right]^{\mathrm{T}}$, $\overline{\vartheta}_1 = \beta_1^{-1}\left(\boldsymbol{\Gamma}_1\right)$, $\beta_1(\boldsymbol{\Gamma}_1)$, $\beta_l(\boldsymbol{\Gamma}_1)$ are minimal and maximal eigenvalues of matrix $\boldsymbol{\Gamma}_1$;

4) $\alpha_{f_1,y_3} - \underline{\upsilon}_3 \le \left(f_1\left(y_1\right) - y_3\right)^2 \le \overline{\alpha}_{f_1,y_3} + \upsilon_3$, $\underline{\upsilon}_3 \ge 0$, $\upsilon_3 \ge 0$;

5) $\gamma_4 = \max(\gamma_{41}, \gamma_{42})$;

6) equality

$$e_i\Delta a_{ij}\omega_i = \pi_i\left(e_i^2 + \left(\Delta a_{ij}\right)^2\omega_i^2\right)$$

is satisfied in domain $O_{\mathrm{v}}(O)$, where $\pi_i > 0$, $O = \{0^4, 0^n\} \subset \mathbb{R}^4 \times \mathbb{R}^n \times \mathbb{J}_{0,\infty}$, 0^4, 0^n are zero vectors, n is the number of adjustable parameters, O_{v} is some neighborhood of point O, $t \in [0,\infty] = \mathbb{J}_{0,\infty}$;

7) for $V_{e,\Delta}(t)$, matrix system of inequalities $\dot{\mathbf{V}}_{e,\Delta} \le \mathbf{A}_{e,\Delta}\mathbf{V}_{e,\Delta} + \mathbf{B}_{e,\Delta}$ is valid, where $\mathbf{A}_{e,\Delta}$ is the following block diagonal matrix:

$$\mathbf{A}_{e,\Delta} = \mathrm{diag}\left(\mathbf{A}_{e,\Delta,1}, \mathbf{A}_{e,\Delta,3}, \mathbf{A}_{e,\Delta,4}, \mathbf{A}_{e,\Delta,5}\right),$$

$$\mathbf{B}_{e,\Delta} = \left[0 \quad 0 \quad \tilde{\upsilon}_3\vartheta_3 \quad \pi_3\chi \quad 0 \quad 0 \quad 0 \quad 0\right]^{\mathrm{T}},$$

submatrices $\mathbf{A}_{e,\Delta,i}$ have the form similar to $\mathbf{A}_{\mathbf{V}}$ from Theorem 2;

8) the upper solution for $V_{e,\Delta}(t)$ satisfies equation $\dot{\mathbf{S}}_{e,\Delta} = \mathbf{A}_{e,\Delta}\mathbf{S}_{e,\Delta} + \mathbf{B}_{e,\Delta}$, if there exist such functions $s_i(t) \ge 0$, that $V_{e,\Delta,i}(t) \le s_i(t)\ \forall(t \ge t_0)$ & $V_{e,\Delta,i}(t_0) \le s_i(t_0)$, $i = 1, 3, 4, 5$, where $S_{e,\Delta} \in \mathbb{R}^6$, s_i are elements of vector $\mathbf{S}_{e,\Delta}$. Then adaptive system (29), (30) is exponentially dissipative with the following estimation:

$$\mathbf{V}_{e,\Delta}(t) \le e^{\mathbf{A}_{e,\Delta}(t-t_0)}\mathbf{S}(t_0) + \int_{t_0}^{t} e^{\mathbf{A}_{e,\Delta}(t-\tau)}\mathbf{B}_{e,\Delta}d\tau,$$

if $k_1 \ge \frac{2}{3}\sqrt{\frac{\alpha\beta_l\left(\tilde{\mathbf{A}}_1\right)}{\alpha_{P_1}k_1\beta_1\left(\tilde{\mathbf{A}}_1\right)}}$, $k_3 \ge \frac{2}{3}\sqrt{\frac{\overline{\alpha}_{f_1,y_3}}{\left(\alpha_{f_1,y_3} - \underline{\upsilon}_3\right)}}$, $k_4 \ge \frac{2}{3}\sqrt{2\frac{\alpha_{y_1,\hat{y}_2}}{\tilde{\alpha}_{y_1,\hat{y}_2}}}$, $k_5 \ge \frac{2}{3}\sqrt{2\frac{\overline{\alpha}_{y_5} + c^2}{\alpha_{y_5 \cdot f_3}}}$.

The proof of Theorem 3 is similar to the proof of Theorem 2.

It follows from Theorem 3 that bound properties of system (29), (30) depend on nonlinear properties, feedback, and compliance with the excitation constancy condition. In particular, this applies to block 3.

CONCLUSIONS

In the paper, the approach to adaptive identification of systems with several nonlinearities is proposed. It is based on the transformation of a system in order to exclude unmeasurable state variables. The synthesis of an adaptive identification system is presented. In order to simplify the adaptation process, an approach involving the decomposition of a system into a number of subsystems is proposed. The boundedness of trajectories in the adaptive system is proven. The problem of the S-synchronizability of a system is considered taking into account the modification of the excitation constancy condition of the system information, which is set with allowance for the specifics of structural identifiability of the nonlinear part of the system. The method of Lyapunov vector functions is applied for proving the exponential stability of the identification system.

An approach to the nonlinear correction of a nonlinear system is considered. Adaptive algorithms for estimating system parameters are obtained. The boundedness of system trajectories is shown. Considering a nonlinear self-oscillation generation system with nonlinear feedback, an adaptive system of parametric identification is proposed. The influence of feedback and nonlinearities on the boundedness of trajectories is investigated. The simulation results confirm theoretical conclusions. The proposed methods could be used in developing identification and control systems for complex dynamic systems.

REFERENCES

1. Aykan M., Özgüven H.N. Parametric identification of nonlinearity from incomplete FRF Data using describing function inversion. In: Adam D., Kersch G., Carrella A. (Eds.). *Topics in Nonlinear Dynamics. V. 3. Conference Proceedings of the Society for Experimental Mechanics Series*. New York: Springer; 2012. P. 311–322. https://doi.org/10.1007/978-1-4614-2416-1_25
2. Zhao Z., Li C., Ahlin K., Du. H. Nonlinear system identification with the use of describing functions – a case study. *Vibroengineering PROCEDIA*. 2016;8:33–38. Available from URL: https://www.extrica.com/article/17407/pdf
3. Pavlov Y.N., Nedashkovskii V.M., Tihomirova E.A. Identification of nonlinear dynamic systems possessing some non-linearities. *Nauka i Obrazovanie. MGTU im. N.E. Baumana = Science and Education of the Bauman MSTU*. 2015;07:217–234 (in Russ.).
4. Prawin J., Rao A.R.M., Sethi A. Parameter identification of systems with multiple disproportional local nonlinearities. *Nonlinear Dyn*. 2020;100(1):289–314. https://doi.org/10.1007/s11071-020-05538-1
5. Pai P.F., Palazotto A.N. Detection and identification of nonlinearities by amplitude and frequency modulation analysis. *Mech. Syst. Signal Process*. 2008;22(5):1107–1132. https://doi.org/10.1016/j.ymssp.2007.11.006
6. Safari S., Monsalve J.M.L. Direct optimisation based model selection and parameter estimation using time-domain data for identifying localised nonlinearities. *J. Sound Vib*. 2021;501(2051):16056. https://doi.org/10.1016/j.jsv.2021.116056
7. Noël J.P., Schoukens J. Grey-box State-space Identification of Nonlinear Mechanical Vibrations. *Int. J. Control*. 2018;91(5):1118–1139. https://doi.org/10.1080/00207179.2017.1308557
8. Li C. Closed-loop identification for a class of nonlinearly parameterized discrete-time systems. *Automatica*. 2021;131:109747. https://doi.org/10.1016/j.automatica.2021.109747
9. Chen H., Kurt M., Lee Y., McFarland D., Bergman L., Vakakis A. Experimental system identification of the dynamics of a vibro-impact beam with a view towards structural health monitoring and damage detection. *Mech. Syst. Signal Process*. 2014;46(1):91–113. https://doi.org/10.1016/j.ymssp.2013.12.014
10. Binczak S., Busvelle E., Gauthier J.-P. Jacquir S. Identification of Unknown Functions in Dynamic Systems. Available from URL: https://pageperso.lis-lab.fr/eric.busvelle/papers/bioid.pdf
11. Singh A., Moore K.J. Identification of multiple local nonlinear attachments using a single measurement case. *J. Sound Vib.* 2021;513:116410. https://doi.org/10.1016/j.jsv.2021.116410
12. Van den Hof P.M.J. Closed-loop issues in system identification. In: Sawaragi Y., Sagara S. (Eds.). *System Identification* (*SYSID'97): Proc. 11th IFAC Symp. System Identification*. 1997. *IFAC Proceedings Series*. 1997. V. 3. P. 1547–1560. Available from URL: http://publications.pvandenhof.nl/Paperfiles/Vanden%20Hof_SYSID97_Plenary.pdf
13. Andersson C., Ribeiro A.H., Tiels K., Wahlström N., Schön T.B. Deep Convolutional Networks in System Identification. In: *IEEE 58th Conference on Decision and Control* (CDC). 2019. https://doi.org/10.1109/CDC40024.2019.9030219
14. Gedon D., Wahlström N., Schön T.B., Ljung L. *Deep State Space Models for Nonlinear System Identification*. arXiv:2003.14162v3 [eess.SY]. https://doi.org/10.48550/arXiv.2003.14162
15. Kerschen G., Worden K., Vakakis A., Golinval J.-C. Past, present and future of nonlinear system identification in structural dynamics. *Mech. Syst. Signal Process*. 2006;20(3):505–592. https://doi.org/10.1016/j.ymssp.2005.04.008
16. Holmes P. The dynamics of repeated impacts with a sinusoidally vibrating table. *J. Sound Vib*. 1982;84(2):173–189. https://doi.org/10.1016/S0022-460X(82)80002-3
17. Azeez M.A.F., Vakakis A.F. Numerical and experimental analysis of a continuously overhung rotor undergoing vibro-impacts. *Int. J. Non-Linear Mechanics.* 1999;34(3):415–435. https://doi.org/10.1016/s0020-7462(98)00022-5
18. Van den Hof P.M.J. Closed-loop issues in system identification. *Annu. Rev. Control*. 1998;22:173–186. https://doi.org/10.1016/S1367-5788(98)00016-9
19. Aljamaan I. *Nonlinear Closed-Loop System identification in the Presence of Non-Stationary Noise Source*. Doctoral Thesis. Calgary, Canada: University of Calgary; 2016. Available from URL: http://hdl.handle.net/11023/3070
20. Forssell U., Ljung L. Closed-loop identification revisited. *Automatica.* 1999;35(7):1215–1241. https://doi.org/10.1016/S0005-1098(99)00022-9
21. Mejari M., Piga D., Bemporad A. A bias-correction method for closed-loop identification of Linear Parameter-Varying systems. *Automatica.* 2018;87:128–141. https://doi.org/10.1016/j.automatica.2017.09.014
22. Piga D., Tyth R. A bias-corrected estimator for nonlinear systems with output-error type model structures. *Automatica*. 2014;50(9):2373–2380. https://doi.org/10.1016/j.automatica.2014.07.021
23. Gilson M., Van den Hof P.M.J. On the relation between a bias-eliminated least squares (BELS) and an IV estimator in closed-loop identification. *Automatica*. 2001;37(10):1593–1600. https://doi.org/10.1016/S0005-1098(01)00119-4
24. Karabutov N.N. *Vvedenie v strukturnuyu identifitsiruemost' nelineinykh system* (*Introduction to Structural Identifiability of Nonlinear Systems*). Moscow: URSS/LENAND; 2021. 144 p. (in Russ.). ISBN 978-5-9710-9022-9
25. Karabutov N.N. S-synchronization Structural Identifiability and Identification of Nonlinear Dynamic Systems. *Mehatronika, Avtomatizacija, Upravlenie*. 2020;21(6):323–336 (in Russ.). https://doi.org/10.17587/mau.21.323-336
26. Karabutov N. Structural-parametrical design method of adaptive observers for nonlinear systems. *International Journal of Intelligent Systems and Applications* (*IJISA*). 2018;10(2):1–16. https://doi.org/10.5815/ijisa.2018.02.01
27. Karabutov N. Geometrical frameworks in identification problem. *Intelligent Control and Automation*. 2018;12(2):17–43. https://doi.org/10.4236/ica.2021.122002
28. Karabutov N. Structural Identifiability of Systems with Multiple Nonlinearities. *Contemp. Math*. 2021;2(2):103–172. https://doi.org/10.37256/cm.222021763

About the author

Nikolay N. Karabutov, Dr. Sci. (Eng.), Professor, Department of Problems Control, Institute of Artificial Intelligence, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). Laureate of the Russian Federation State Prize in Science and Technology. E-mail: karabutov@mirea.ru. Scopus Author ID 6603372930, ResearcherID P-5683-2015, SPIN-код РИНЦ 9646-9721, https://orcid.org/0000-0002-3706-7431

Об авторе

Карабутов Николай Николаевич, д.т.н., профессор, кафедра проблем управления Института искусственного интеллекта ФГБОУ ВО МИРЭА – Российский технологический университет (119454, Россия, Москва, пр-т Вернадского, д. 78). Лауреат Государственной премии в области науки и техники. E-mail: karabutov@mirea.ru. Scopus Author ID 6603372930, ResearcherID P-5683-2015, SPIN-код РИНЦ 9646-9721, https://orcid.org/0000-0002-3706-7431

Translated from Russian into English by Kirill V. Nazarov
Edited for English language and spelling by Thomas A. Beavitt

Mathematical modeling

Математическое моделирование

UDC 536.2
https://doi.org/10.32362/2500-316X-2023-11-5-106-117

RESEARCH ARTICLE

New energy effect in non-cylindrical domains with a thermally insulated moving boundary

Eduard M. Kartashov @

MIREA – Russian Technological University, Moscow, 119454 Russia
@ *Corresponding author, e-mail: professor.kartashov@gmail.com*

Abstract

Objectives. To develop mathematical model representations of the energy effect in non-cylindrical domains having a thermally insulated moving boundary; to introduce a new boundary condition for thermal insulation of a moving boundary both for locally equilibrium heat transfer processes in the framework of classical Fourier phenomenology, as well as for more complex locally non-equilibrium processes in the framework of Maxwell–Cattaneo–Lykov–Vernott phenomenology, taking into account the finite rate of heat propagation into analytical thermophysics and applied thermomechanics; to consider an applied problem of analytical thermophysics according to the theory of thermal shock for a domain with a moving thermally insulated boundary free from external and internal influences; to obtain an exact analytical solution of the formulated mathematical models for hyperbolic type equations; to investigate the solutions obtained using a computational experiment at various values of the parameters included in it; to describe the wave nature of the kinetics of the processes under consideration.

Methods. Methods and theorems of operational calculus, Riemann–Mellin contour integrals are used in calculating the originals of complex images with two branch points. A new mathematical apparatus for the equivalence of functional constructions for the originals of the obtained operational solutions, which considers the computational difficulties in finding analytical solutions to boundary value problems for equations of hyperbolic type in the domain with a moving boundary, is developed.

Results. Developed mathematical models of locally nonequilibrium heat transfer and the theory of thermal shock for equations of hyperbolic type in a domain with a moving thermally insulated boundary are presented. It is shown that, despite the absence of external and internal sources of heat, the presence of a thermally insulated moving boundary leads to the appearance of a temperature gradient in the domain and, consequently, to the appearance of a temperature field and corresponding thermoelastic stresses in the domain, which have a wave character. A stochastic analysis of this energy effect forms the basis for a proposed transition of the kinetic energy of a moving thermally insulated boundary into the thermal energy of the domain. The presented model representations of the indicated effect confirmed the stated assumption.

Conclusions. Mathematical models for locally nonequilibrium heat transfer processes and the theory of thermal stresses are developed and investigated on the basis of constitutive relations of the theory of thermal shock for equations of hyperbolic type in a domain with a thermally isolated moving boundary. A numerical experiment is presented to demonstrate the possibility of transiting from one form of analytical solution of a thermophysical problem to another equivalent form of a new type. The described energy effect manifests itself both for parabolic type equations based on the classical Fourier phenomenology, as well as for hyperbolic type equations based on the generalized Maxwell–Cattaneo–Lykov–Vernott phenomenology.

Keywords: moving thermally insulated boundary, temperature field, temperature stresses, equations of hyperbolic type

• **Submitted:** 30.11.2022 • **Revised:** 21.04.2023 • **Accepted:** 21.07.2023

For citation: Kartashov E.M. New energy effect in non-cylindrical domains with a thermally insulated moving boundary. *Russ. Technol. J*. 2023;11(5):106–117. https://doi.org/10.32362/2500-316X-2023-11-5-106-117

Financial disclosure: The author has no a financial or property interest in any material or method mentioned.

The author declares no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Новый энергетический эффект в областях нецилиндрического типа с термоизолированной движущейся границей

Э.М. Карташов [@]

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[@] *Автор для переписки, e-mail: professor.kartashov@gmail.com*

Резюме

Цели. Разработка математически модельных представлений энергетического эффекта в областях нецилиндрического типа с термоизолированной движущейся границей. Введение в аналитическую теплофизику и прикладную термомеханику нового граничного условия теплоизоляции движущейся границы как для локально равновесных процессов теплопереноса в рамках классической феноменологии Фурье, так и для более сложных локально-неравновесных процессов в рамках феноменологии Максвелла – Каттанео – Лыкова – Вернотта, учитывающих конечную скорость распространения теплоты. Рассмотрение прикладной задачи аналитической теплофизики и теории теплового удара для области с движущейся термоизолированной границей, свободной от внешних и внутренних воздействий. Получение точного аналитического решения сформулированных математических моделей для уравнений гиперболического типа. Исследование полученных решений с помощью вычислительного эксперимента при различных значениях, входящих в него параметров. Описание волнового характера кинетики рассматриваемых процессов.

Методы. Использованы методы и теоремы операционного исчисления, контурные интегралы Римана – Меллина при вычислении оригиналов сложных изображений с двумя точками ветвления. С учетом вычислительных трудностей при нахождении аналитических решений краевых задач для уравнений гиперболического типа в области с движущейся границей, развит новый математический аппарат эквивалентности функциональных конструкций для оригиналов полученных операционных решений.

Результаты. Представлено развитие новых математических моделей локально-неравновесного теплопереноса и теории теплового удара для уравнений гиперболического типа в области с движущейся термоизолированной границей. Показано, что, несмотря на отсутствие внешних и внутренних источников теплоты, наличие термоизолированной движущейся границы приводит к появлению в области градиента температуры и, следовательно, к появлению в области температурного поля и соответствующих ему термоупругих напряжений, имеющих волновой характер. Стохастический анализ указанного энергетического эффекта позволил высказать предположение о переходе кинетической энергии движущейся термоизолированной границы в тепловую энергию области. Приведенные модельные представления указанного эффекта подтвердили высказанное предположение.

Выводы. Развиты и исследованы математические модели для локально-неравновесных процессов теплопереноса и теории термических напряжений на основе определяющих соотношений теории теплового удара для уравнений гиперболического типа в области с термоизолированной движущейся границей. Проведен численный эксперимент и показана возможность перехода от одной формы аналитического решения теплофизической задачи к другой эквивалентной форме нового типа. Описанный энергетический эффект проявляется как для уравнений параболического типа на основе классической феноменологии Фурье, так и для уравнений гиперболического типа на основе обобщенной феноменологии Максвелла – Каттанео – Лыкова – Вернотта.

Ключевые слова: движущаяся теплоизолированная граница, температурное поле, температурные напряжения, уравнения гиперболического типа

• **Поступила:** 30.11.2022 • **Доработана:** 21.04.2023 • **Принята к опубликованию:** 21.07.2023

Для цитирования: Карташов Э.М. Новый энергетический эффект в областях нецилиндрического типа с термоизолированной движущейся границей. *Russ. Technol. J.* 2023;11(5):106–117. https://doi.org/10.32362/2500-316X-2023-11-5-106-117

Прозрачность финансовой деятельности: Автор не имеет финансовой заинтересованности в представленных материалах или методах.

Автор заявляет об отсутствии конфликта интересов.

INTRODUCTION

The effect of the concentration gradient emergence in the domain with moving impermeable boundary was encountered by the author for the first time when studying the phenomenon of adsorption reduction of strength and durability of brittle polymers in surface-active media [1]. A review of the literature confirms that, while the above-described phenomenon affects many fields of science and technology, it has yet to be practically described in scientific publications. Considering thermal processes, it will be shown that a temperature gradient exists in a domain having a moving thermally insulated boundary despite the absence of internal or external sources of heat; this is due to the conversion of the kinetic energy of the boundary motion into the thermal energy of the domain. A stochastic analysis of this energy effect for the temperature average based on an analysis of the corresponding dispersion shows the similarity of the dispersion behavior to that arising in the domain of the average value of temperature stresses, creating the risks of cracks and the possible beginning of material destruction [2].

PROBLEM STATEMENT

We shall briefly consider thermophysical problems in domains with moving boundaries (non-cylindrical domains).

A very wide range of issues arise when considering boundary value problems of nonstationary heat conduction in non-cylindrical domains of the type $[0, y(t)]$, $t > 0$ or $[y(t), \infty)$, $t > 0$, where $y(t)$ is continuous function. Similar problems arise in the theoretical study of energy transfer processes related to changes in the aggregate state of matter, as well as in strength theory, dam theory, soil mechanics, oil-reservoir thermic and electrodynamic problems, filtration problems, the theory of zone cleaning of materials, kinetic theory of crystal growth, thermomechanics in the study of thermal shock, etc. [3].

In mathematical terms, boundary transfer problems in the domain with moving boundaries are fundamentally different from classical ones. Due to the dependence of the domain boundary on time, the classical methods of equations of mathematical physics are inapplicable to this type of problems: it is impossible to match the solution of the heat conduction equation with the motion of the domain boundary while remaining within the framework of such methods. This explains why only the simplest cases with a uniformly moving boundary, or partially with a root dependence, have been considered in analytical thermophysics to date.

Let $\overline{\Omega}_t$ be a non-cylindrical domain, whose cross-section by characteristic plane $t = \text{const} \geq t_0 > 0$ is convex domain D_t of change $M(x, y, z)$ with boundary S_t depending on time $t \geq 0$, $\overline{\mathbf{n}}$ is the external normal to S_t, representing a vector continuous at points S_t, so that $\overline{\Omega}_t = \left\{M \in \overline{D}_t = D_t + S_t, t \geq 0\right\}$.

Let $T(M, t)$ be the temperature function satisfying the conditions of the problem, a be thermal conductivity, and f be the source function. Φ_0 is initial temperature, β_1, β_2 are coefficients, while C^0, C^1, C^2 are function classes.

$$\frac{\partial T}{\partial t} = a\Delta T(M,t) + f(M,t),\ M \in D_t,\ t > 0, \quad (1)$$

$$T(M,t)\big|_{t=0} = \Phi_0(M),\ M \in \overline{D}_{t=0}, \quad (2)$$

$$\beta_1 \frac{\partial T(M,t)}{\partial n} + \beta_2 T(M,t) = \varphi(M,t),\ M \in S_t,\ t \geq 0. \quad (3)$$

Here,

$$f(M,t) \in C^0(\overline{\Omega}_t);\ \Phi_0(M) \in C^1(\overline{\Omega}_t);$$
$$\varphi(M,t) \in C^0(S_t \times t \geq 0);\ \beta_1^2 + \beta_2^2 > 0. \quad (4)$$

The desired solution: $T(M,t) \in C^2(\Omega_t) \cap C^0(\overline{\Omega}_t)$, $\mathrm{grad}_M T(M,t) \in C^0(\overline{\Omega}_t)$.

The boundary condition (3) includes the cases of temperature heating, thermal heating, and heating by the medium (or cooling in all three cases). If $\overline{D}$ is a canonical (cylindrical) domain with a fixed boundary S (elastic half-space, infinite plate, cylinder, ball, etc.), the thermal insulation condition of boundary S of domain D is written in the following form:

$$\frac{\partial T(M,t)}{\partial n}\Big|_{M \in S} = 0,\ t > 0 \quad (5)$$

and is a classical boundary condition in analytical thermophysics when formulating appropriate problems for parabolic type equations. However, the presence of a moving boundary fundamentally changes the form of the boundary condition for its thermal insulation; moreover, it is this circumstance that is not generally considered in the literature describing various kinds of applications related to the thermal insulation of the moving boundary.

For deriving the above condition, we shall consider domain $\Omega_t = (0 < z < y(t),\ t > 0)$. Here, $y(t)$ is a continuous-differentiable function; $v(t) = dy(t)/dt$ is the boundary movement rate; $T(z,\ t)$ is temperature field in Ω_t; $F(z,\ t)$ is nonstationary heat source ($F(z,t)/c\rho = f(z,t)$) continuously distributed in Ω_t, where c is heat capacity and ρ is density. We have for Ω_t:

$$\frac{\partial T}{\partial t} = a\frac{\partial^2 T}{\partial z^2} + f(z,t),\ \ (z,t) \in \Omega_t. \quad (6)$$

We shall write the heat balance equation at time $(t + \Delta t)$ considering boundary $z = y(t)$ insulated as follows:

$$-\lambda\frac{\partial T(0,t)}{\partial z}\Delta t + c\rho\Delta t\int_0^{y(t)+\Delta y} f(z,t)dz =$$
$$= c\rho\int_0^{y(t)} \left[T(z,t+\Delta t) - T(z,t)\right]dz + c\rho\int_{y(t)}^{y(t)+\Delta y} T(z,t+\Delta t)dz,$$

where λ is heat conduction.

For the second of the integrals on the right, we apply the average theorem, as follows:

$$-a\frac{\partial T(0,t)}{\partial z}\Delta t + \Delta t\int_0^{y(t)+\Delta y} f(z,t)dz =$$
$$= \int_0^{y(t)} \left[T(z,t+\Delta t) - T(z,t)\right]dz + T(z,t+\Delta t)\Big|_{z=y+\theta\Delta y}\Delta y,$$

where $0 < \theta < 1$. Dividing both parts of the equation by Δt and going to the limit at $\Delta t \to 0$, we obtain the following:

$$-a\frac{\partial T(0,t)}{\partial z} + \int_0^{y(t)} f(z,t)dz = \int_0^{y(t)} \frac{\partial T}{\partial t}dz + v(t)T(z,t)\Big|_{z=y(t)}.$$

We shall substitute under the integral sign the right-hand side of heat conduction Eq. (6) for $\partial T/\partial t$, integrate it, and add similar terms. The final result is the following condition:

$$\frac{\partial T(z,t)}{\partial z}\Big|_{z=y(t)} + \frac{v(t)}{a}T(z,t)\Big|_{z=y(t)} = 0,\ \ t > 0, \quad (7)$$

which represents the thermal insulation condition for the moving boundary. If the boundary movement rate $v(t) = 0$, then we arrive at the condition $(\partial T/\partial n)|_S = 0$ implying thermal insulation of the stationary boundary surface.

Since the late 1960s, systematic publications on hyperbolic transfer models that take into account the finite rate of heat propagation have appeared[1, 2] [4–16]. Nowadays, it is increasingly common practice to distinguish a large class of models based on the following equation:

$$\frac{\partial T(M,t)}{\partial t} = a\Delta T(M,t) - \tau_r\frac{\partial^2 T(M,t)}{\partial t^2},\ (M,t) \in \Omega_t, \quad (8)$$

where τ_r is the thermal flux relaxation time related to the heat propagation rate by relation $v_T = \sqrt{a/\tau_r}$.

The boundary problems for Eq. (8) describe high-intensity heat exchange in pulse and laser devices, laser metal processing, plasma spraying processes, processes occurring in energy channels of nuclear reactors, as well as in a fluidized bed and disperse systems, granular materials, and layered semiconductor structures. The problems also arise in descriptions of electronic heat conduction

[1] Eremin A.V. *Modeling methodology of heat and mass transfer, elastic vibrations and electromagnetic waves with allowance for spatial and temporal nonlocality*. Abstract. Cand. Sci. Thesis (Eng.). Samara; 2021. 30 p. (in Russ.).

[2] Zhukov V.V. *Investigation of internal mechanisms of heat, mass, and momentum transfer with allowance for relaxation phenomena*. Abstract. Cand. Sci. Thesis (Eng.). Samara; 2021. 18 p. (in Russ.).

and high-temperature plasma, in mathematical modeling of thermal decomposition front processes, in catalyst crystals, and the growth of homoepitaxial germanium films during exothermic chemical reactions, etc. In [3], the issues of correct formulation of boundary value problems for Eq. (8) are studied: it is shown that the writing of boundary conditions of the second and third kind significantly differs from (3) for equations of parabolic type. However, the question about the heat isolation of the moving boundary for Eq. (8) remains open. For this, we shall consider the phenomenological Maxwell–Cattaneo–Lykov–Vernott relation [4, 6–7]:

$$\bar{\mathbf{q}}(M,t) = -\lambda\ \mathrm{grad}T(M,t) - \tau_r \frac{\partial \bar{\mathbf{q}}(M,t)}{\partial t},$$

here, $\bar{\mathbf{q}}$ is the vector of heat flux density, which forms the basis for the analytical theory of local nonequilibrium processes of heat transfer in a non-cylindrical domain.

We shall write this equation in the following form:

$$\left(1 + \tau_r \frac{\partial}{\partial t}\right)\bar{\mathbf{q}}(M,t) = -\lambda\ \mathrm{grad}T(M,t),\ M \in D_t,\ t > 0, \quad (9)$$

or

$$\bar{\mathbf{q}}(M,t+\tau_r) = -\lambda\ \mathrm{grad}T(M,t),\ M \in D_t,\ t > 0, \quad (10)$$

using the Maclaurin series Eq. (9) of function $\bar{\mathbf{q}}(M,t+\tau_r)$ in the vicinity of point $\tau_r = 0$.

Equation (10) may be rewritten in the following form:

$$\bar{\mathbf{q}}(M,t) = -\lambda\ \mathrm{grad}T(M,t-\tau_r),\ M \in D_{t-\tau_r},\ t > \tau_r. \quad (11)$$

Using energy equation $c\rho\partial T(M,t)/\partial t = -\mathrm{div}\left[\bar{\mathbf{q}}(M,t)\right]$ and relation (11), Eq. (8) may be written in the following form:

$$\frac{\partial T(M,t)}{\partial t} = a\Delta T(M,t-\tau_r),\ M \in D_{t-\tau_r},\ t > \tau_r. \quad (12)$$

Consider now the domain of interest $z > y(t)$, $t > 0$, wherein (12) is the following:

$$\frac{\partial T(z,t)}{\partial t} = a\frac{\partial^2 T(z,t-\tau_r)}{\partial z^2},\ z > y(t-\tau_r),\ t > \tau_r. \quad (13)$$

Under constant initial conditions, as well as in the absence of internal heat sources and external heating conditions at the thermal insulation of the moving boundary, the following condition is true:

$$c\rho \int_{y(t-\tau_r)}^{\infty} T(z,t)dz = \mathrm{const},\ t > \tau_r. \quad (14)$$

Differentiating both parts of (14) by t and using Eq. (13), the following relation is obtained:

$$\left[\frac{\partial T(z,t-\tau_r)}{\partial z} + \frac{1}{a}\cdot\frac{dy(t-\tau_r)}{\partial t}T(z,t)\right]_{z=y(t-\tau_r)} = 0,\ t > \tau_r, \quad (15)$$

which may be rewritten as:

$$\left[\frac{\partial T(z,t)}{\partial z} + \frac{v(t)}{a}T(z,t+\tau_r)\right]_{z=y(t)} = 0,\ t > 0, \quad (16)$$

where $v(t) = dy/dt$.

Equation (16) is the moving boundary heat isolation condition for locally nonequilibrium heat transfer processes described by hyperbolic type equations. In particular cases (local-equilibrium processes, $\tau_r = 0$) or cylindrical type domain ($v(t) = 0$), the thermal insulation conditions discussed above are obtained.

THE TEMPERATURE GRADIENT EFFECT IN THE DOMAIN WITH A MOVING THERMALLY INSULATED BOUNDARY

In the corresponding model representations of nonstationary heat conduction, boundary conditions (7), (16) create the temperature gradient effect in the domain and consequent appearance of corresponding thermoelastic stresses, which occur despite the absence of external and internal thermal impacts. While formally, there is an idea concerning the impossibility of manifestation of the described effect, analytical solutions of model problems demonstrate the opposite. In [2], it is suggested that the kinetic energy of the moving insulated boundary is converted into the thermal energy of the domain, thus causing heat and thermal effects. In this connection, in formulating the thermal problem for the hyperbolic type equation, we shall consider elastic half-space $z > l + vt$, $t > 0$ with a uniformly moving insulated boundary in the absence of external and internal thermal loads:

$$\frac{\partial T}{\partial t} = a\frac{\partial^2 T}{\partial z^2} - \tau_r\frac{\partial^2 T}{\partial t^2},\ z > l + vt,\ t > 0, \quad (17)$$

$$T(z,t)\big|_{t=0} = T_0,\ \left.\frac{\partial T(z,t)}{\partial t}\right|_{t=0} = 0,\ z \geq l, \quad (18)$$

$$\left[\frac{\partial T(z,t)}{\partial z} + \frac{v}{a}T(z,t+\tau_r)\right]_{z=l+vt} = 0,\ t > 0, \quad (19)$$

$$\left|T(z,t)\right| < \infty,\ z \geq l + vt,\ t \geq 0. \quad (20)$$

The boundary condition (19) may be written in the following form:

$$\left[\frac{\partial T(z,t)}{\partial z}+\frac{v}{a}T(z,t)+\frac{v\tau_r}{a}\cdot\frac{\partial T(z,t)}{\partial t}\right]\Bigg|_{z=l+vt}= \quad (21)$$
$$=0,\ \ t>0.$$

We shall introduce dimensionless variables

$$z'=(z-l)/l;\ \ \tau=at/l^2;\ v_0=vl/a;\ \tau_0=a\tau_r/l^2;$$
$$T^*(z',\tau)=\left[T(z,t)-T_0\right]/T_0$$

and then the moving coordinate system $\xi=z'-v_0\tau,\ \tau>0$, assuming $T^*(z',\tau)=W(\xi,\tau)$.

Relations (17)–(21) have the following form:

$$\frac{\partial W}{\partial\tau}=(1-\tau_0v_0^2)\frac{\partial^2W}{\partial\xi^2}+v_0\frac{\partial W}{\partial\xi}+$$
$$+2v_0\tau_0\frac{\partial^2W}{\partial\xi\partial\tau}-\tau_0\frac{\partial^2W}{\partial\tau^2}=0,\ \xi>0,\ \tau>0, \quad (22)$$

$$W(\xi,\tau)\big|_{\tau=0}=0,\ \ \frac{\partial W(\xi,\tau)}{\partial\tau}\Bigg|_{\tau=0}=$$
$$=v_0\frac{\partial W(\xi,\tau)}{\partial\xi}\Bigg|_{\tau=0}=0,\ \xi\ge0, \quad (23)$$

$$(1-\tau_0v_0^2)\frac{\partial W(\xi,\tau)}{\partial\xi}\Bigg|_{\xi=0}=$$
$$=-v_0\left[W(\xi,\tau)+\tau_0\frac{\partial W(\xi,\tau)}{\partial\tau}+1\right]\Bigg|_{\xi=0},\tau>0, \quad (24)$$

$$\left|W(\xi,\tau)\right|<\infty,\ \ \xi\ge0,\ \ \tau\ge0. \quad (25)$$

In the Laplace image space:

$$\overline{W}(\xi,p)=\int_0^\infty W(\xi,\tau)\exp(-p\tau)d\tau$$

the operational solution of the transformed problem (22)–(25)

$$(1-\tau_0v_0^2)\frac{d^2\overline{W}}{d\xi^2}+v_0(1+2\tau_0p)\frac{d\overline{W}}{d\xi}-$$
$$-\ p(1+\tau_0p)\overline{W}=0,\ \xi>0,$$

$$(1-\tau_0v_0^2)\frac{d\overline{W}}{d\xi}\Bigg|_{\xi=0}=-v_0\left[(1+\tau_0p)\overline{W}+\frac{1}{p}\right]_{\xi=0},$$

$$\left|\overline{W}(\xi,p)\right|<\infty,\ \xi\ge0$$

may be written in the following form:

$$\overline{W}(\xi,p)=\overline{\Psi}_1(\xi,p)\overline{\Psi}_2(\xi,p), \quad (26)$$

where

$$\overline{\Psi}_1(\xi,p)=\frac{v_0}{\left[(-v_0/2)+\sqrt{\tau_0}\sqrt{(p+2\alpha)(p+2\beta)}\right]}\times$$
$$\times\exp\left[-\frac{(v_0/2)\xi}{1-\tau_0v_0^2}\right], \quad (27)$$

$$\overline{\Psi}_2(\xi,p)=\frac{1}{p}\times$$
$$\times\exp\left\{-\left[\frac{\xi\sqrt{\tau_0}}{1-\tau_0v_0^2}\sqrt{(p+2\alpha)(p+2\beta)}+\frac{\tau_0v_0\xi}{1-\tau_0v_0^2}p\right]\right\}, \quad (28)$$

$$\alpha=\frac{1+\sqrt{1-\tau_0v_0^2}}{4\tau_0};\ \ \beta=\frac{1-\sqrt{1-\tau_0v_0^2}}{4\tau_0}.$$

For finding the originals of images (27)–(28), we shall first consider new transformations of operational calculus that are of interest for hyperbolic transfer models. In [3], the original image is

$$\frac{1}{p}\exp\left[-\xi\sqrt{(p+2\alpha)(p+2\beta)}\right]\leftarrow$$
$$\leftarrow\left[\exp(-\rho\xi)+\sigma\xi\int_\xi^t\exp(-\rho\tau)\frac{I_1\left(\sigma\sqrt{\tau^2-\xi^2}\right)}{\sqrt{\tau^2-\xi^2}}d\tau\right]\times \quad (29)$$
$$\times\ \eta(t-\xi)=W_1(\xi,t)\eta(t-\xi).$$

Here, $\sigma=\alpha-\beta$, $\rho=\alpha+\beta$, $I_1(z)$ is the modified Bessel function and $\eta(z)$ is the Heaviside function. On the other hand, calculating the original image on the left in (29) using the Riemann–Mellin contour integral with two branching points by the method developed in [3], we find:

$$\frac{1}{p}\exp\left[-\xi\sqrt{(p+2\alpha)(p+2\beta)}\right]\leftarrow$$
$$\leftarrow\left\{\exp(-2\xi\sqrt{\alpha\beta})-\frac{1}{\pi}\int_0^{2\sigma}\frac{\sin\xi\sqrt{y(2\sigma-y)}}{y+2\beta}\times\right. \quad (30)$$
$$\left.\times\exp\left[-(y+2\beta)t\right]dy\right\}\eta(t-\xi)=W_2(\xi,t)\eta(t-\xi).$$

We shall show that $W_1(\xi,t)=W_2(\xi,t)$.
We have:

$$W_1(\xi,t)=\frac{\partial}{\partial\xi}\left[-\int_{\xi}^{t}\exp(-\rho\tau)I_0(\sigma\sqrt{\tau^2-\xi^2})d\tau\right]= \\ =\frac{\partial}{\partial\xi}\left[-\int_{\xi}^{t}\exp(-\rho\tau)J_0(\sigma\sqrt{\xi^2-\tau^2})d\tau\right]. \quad (31)$$

We shall differentiate both parts of (31) by t:

$$\left[W_1(\xi,t)\right]'_t=\frac{\partial}{\partial\xi}\left[-\exp(-\rho t)J_0(\sigma\sqrt{\xi^2-t^2})\right]= \\ =\frac{\partial}{\partial\xi}\left[-\exp(-2\beta t)\exp(-\sigma t)J_0(\sigma\sqrt{\xi^2-t^2})\right].$$

We shall use further a rather rare integral [3]:

$$\int_0^a\frac{\exp(-px)}{\sqrt{ax-x^2}}\cos c\sqrt{ax-x^2}\,dx= \\ =\pi\exp(-ap/2)J_0\left(\frac{a}{2}\sqrt{c^2-p^2}\right).$$

Hence:

$$\left[W_1(\xi,t)\right]'_t=\frac{1}{\pi}\int_0^{2\sigma}\sin\xi\sqrt{y(2\sigma-y)}\exp\left[-(y+2\sigma)t\right]dy. \quad (32)$$

Integrating both parts of (32) by t and using the finite value theorem $\lim\limits_{t\to\infty}f(t)=\lim\limits_{p\to0}p\overline{f}(p)$ giving $C=\exp(-2\xi\sqrt{\alpha\beta})$ in (30) to find the integration constant, the following is finally obtained:

$$W_1(\xi,t)=\exp(-2\xi\sqrt{\alpha\beta})- \\ -\frac{1}{\pi}\int_0^{2\sigma}\frac{\sin\xi\sqrt{y(2\sigma-y)}}{y+2\beta}\exp\left[-(y+2\beta)t\right]dy=W_2(\xi,t).$$

Thus, the original is found:

$$\frac{1}{p}\exp\left[-\frac{\xi\sqrt{\tau_0}}{1-\tau_0v_0^2}\sqrt{(p+2\alpha)(p+2\beta)}\right]\leftarrow \\ \leftarrow\left\{\exp\left[-\frac{(v_0/2)\xi}{1-\tau_0v_0^2}\right]-\frac{1}{\pi}\int_0^{2\sigma}\frac{1}{y+2\beta}\sin\frac{\xi\sqrt{\tau_0}\sqrt{y(2\sigma-y)}}{1-\tau_0v_0^2}\times\right. \quad (33) \\ \left.\times\exp\left[-(y+2\beta)\tau\right]dy\right\}\times\eta\left(\tau-\frac{\xi\sqrt{\tau_0}}{1-\tau_0v_0^2}\right).$$

Now, using (33), we find the original image $\overline{\Psi}_2(\xi,p)$ (28):

$$\Psi_2(\xi,\tau)= \\ =\left\{\exp\left[-\frac{(v_0/2)\xi}{1-\tau_0v_0^2}\right]-\frac{1}{\pi}\int_0^{2\sigma}\frac{1}{y+2\beta}\sin\frac{\xi\sqrt{\tau_0}\sqrt{y(2\sigma-y)}}{1-\tau_0v_0^2}\times\right. \quad (34) \\ \left.\times\exp\left[-(y+2\beta)\left(\tau-\frac{\tau_0v_0\xi}{1-\tau_0v_0^2}\right)\right]dy\right\}\times\eta\left(\tau-\frac{\xi\sqrt{\tau_0}}{1-v_0\sqrt{\tau_0}}\right).$$

The original image $\overline{\Psi}_1(\xi,p)$ (27) has the following form:

$$\Psi_1(\xi,\tau)=\exp\left(-\frac{(v_0/2)\xi}{1-\tau_0v_0^2}\right)\times \\ \times\left[\frac{v_0\sqrt{\tau_0}}{\sqrt{\pi}}\int_{2\beta}^{2\alpha}\frac{\sqrt{(2\alpha-y)(y-2\beta)}\exp(-y\tau)}{(v_0^2/4)+\tau_0(2\alpha-y)(y-2\beta)}dy\right]. \quad (35)$$

The desired original image $\overline{W}(\xi,p)$ (26) may be written now in the following form:

$$W(\xi,\tau)=\left[\int_{\frac{\xi\sqrt{\tau_0}}{1-v_0\sqrt{\tau_0}}}^{\tau}\Psi_1(\xi,\tau-\tau')\Psi_2(\xi,\tau')d\tau'\right]\times \\ \times\eta\left(\tau-\frac{\xi\sqrt{\tau_0}}{1-v_0\sqrt{\tau_0}}\right). \quad (36)$$

Figure 1 shows curves of the temperature function (36) versus τ in the cross section $\xi=1$ for different v_0 at $\tau=0.25$. The curves in Fig. 1 clearly show the peculiarities of the thermal response of the domain for locally nonequilibrium processes (the analytical solution (36) contains the Heaviside function explaining the delay in the onset of heat propagation in the fixed cross section).

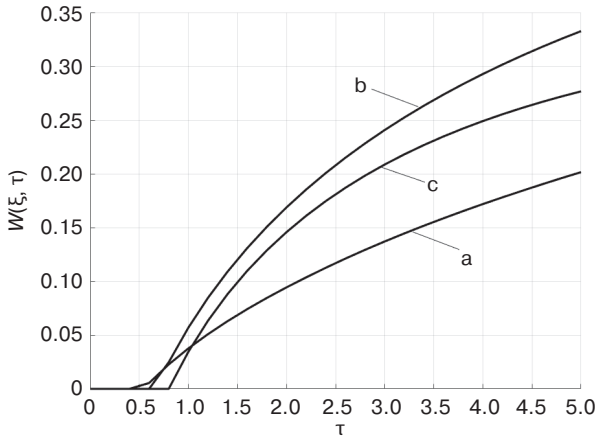


Fig. 1. Dependence of the temperature function $W(\xi,\tau)$ (36) on τ in cross section $\xi=1$ for different v_0: $v_0=0.16$ (a); $v_0=0.5$ (b); $v_0=0.75$ (c) at $\tau_0=0.25$

THERMAL RESPONSE OF THE DOMAIN TO THE TEMPERATURE FIELD $W(\xi, \tau)$

The next step is investigating the thermal response of domain $\overline{\Omega}_t = (z \geq l + vt,\ t \geq 0)$ with a moving thermally insulated boundary in the framework of the model problem (17)–(20). We shall consider, as above, an elastic half-space, which is of practical interest for many areas of science and technology described in [3]. We shall write down the defining relations of dynamic thermoelasticity for domain $\overline{\Omega}_t = \left\{M(x,y,z) \in \overline{D}_t = D_t + S_t,\ t \geq 0\right\}$ with temperature function $T(M, t)$. Let T_0 be the initial temperature at which the domain is in the undeformed and unstressed state; $\sigma_{ij}(M, t)$, $\varepsilon_{ij}(M, t)$, $U_i(M, t)$ $(i = x, y, z)$ are the components of stress tensors, strain, and displacement vectors, respectively, satisfying the basic equations of (unbounded) thermoelasticity (in index notations) [19, 20]:

$$\sigma_{ij,j}(M,t) = \rho \ddot{U}_i(M,t), \tag{37}$$

$$\varepsilon_{ij}(M,t) = (1/2)\left[U_{i,j}(M,t) + U_{j,i}(M,t)\right], \tag{38}$$

$$\begin{gathered}\sigma_{ij}(M,t) = 2\mu\varepsilon_{ij}(M,t) + \\ + \left[\lambda\varepsilon_{ii}(M,t) - (3\lambda + 2\mu)\alpha_T(T(M,t) - T_0)\right]\delta_{ij},\end{gathered} \tag{39}$$

where ρ is density; $\mu = G$, G is shear modulus; $\lambda = 2Gv/(1 - 2v)$ are isothermal Lame coefficients; v is Poisson ratio with $2G(1 + v) = E$, E is Young modulus; α_T is linear thermal expansion coefficient; δ_{ij} is the Kronecker symbol; and $\varepsilon_{ii}(M,t) = \overline{e}(M,t) = U_{i,i}(M,t)$ is volume strain related to the sum of normal stresses $\overline{\sigma}(M,t) = \sigma_{nn}(M,t)$, $n = x, y, z$ by the following relation:

$$\overline{e}(M,t) = \frac{1-2\nu}{E}\overline{\sigma}(M,t) + 3\alpha_T\left[T(M,t) - T_0\right]. \tag{40}$$

For the case of one-dimensional motion $M = M(z, t)$, $z > l + vt$, $t > 0$, we have the following from (37)–(40):

$$\begin{gathered}U_x = U_y = 0,\ U_z = U_z(z,t),\ \varepsilon_{xx} = \varepsilon_{yy} = 0,\\ \varepsilon_{zz}(z,t) = \partial U_z(z,t)/\partial z = [1/(1-v)] \times \\ \times \{[(1-2v)/(2G)]\sigma_{zz}(z,t) + (1+v)\alpha_T T(z,t) - T_0]\},\end{gathered}$$

$$\frac{\partial\sigma_{zz}(z,t)}{\partial z} = \rho\frac{\partial^2 U_z(z,t)}{\partial t^2}.$$

Differentiating this relation by z and substituting the value $\partial U_z(z, t)/\partial z$, the equation of the following form is obtained:

$$\begin{gathered}\frac{\partial^2\sigma_{zz}}{\partial z^2} - \frac{1}{v_p^2}\cdot\frac{\partial^2\sigma_{zz}}{\partial t^2} = \\ = \frac{(1+\nu)}{(1-\nu)}\alpha_T\rho\frac{\partial^2\left[T(z,t) - T_0\right]}{\partial t^2},\ z > l + vt,\ t > 0\end{gathered} \tag{41}$$

with the following boundary conditions:

$$\sigma_{zz}(z,t)\big|_{t=0} = 0,\ \left.\frac{\partial\sigma_{zz}(z,t)}{\partial t}\right|_{t=0} = 0,\ z \geq l, \tag{42}$$

$$\begin{gathered}\sigma_{zz}(z,t)\big|_{z=l+vt} = 0,\ t > 0,\\ \left|\sigma_{zz}(z,t)\right| < \infty,\ z \geq l + vt,\ t \geq 0.\end{gathered} \tag{43}$$

In (41), $v_p = \sqrt{2G(1-\nu)/\left[\rho(1-2\nu)\right]} = \sqrt{(\lambda + 2\mu)/\rho}$ is the expansion wave propagation velocity in the elastic medium close to the speed of sound. The remaining nonzero components of the stress tensor, according to (37)–(39), have the following form:

$$\sigma_{xx}(z,t) = \sigma_{yy}(z,t) = \frac{\nu}{1-\nu}\sigma_{zz}(z,t) - \frac{E\alpha_T\left[T(z,t) - T_0\right]}{1-\nu}.$$

In addition,

$$\varepsilon_{zz}(z,t) = \frac{1-2\nu}{2G(1-\nu)}\sigma_{zz}(z,t) + \frac{(1+\nu)}{(1-\nu)}\alpha_T\left[T(z,t) - T_0\right].$$

The function $T(z, t)$ in (41)–(43) satisfies conditions (17)–(20). For solving the problem (41)–(43), we move to the coordinate system (z', τ) by the above relations, assuming that

$$\alpha_0 = v_p l/a,\ S_T = \alpha_T(3\lambda + 2\mu) = \frac{\alpha_T E}{(1-2\nu)},$$

$$\sigma_{z'z'}(z',\tau) = \frac{\sigma_{zz}(z,t)}{S_T T_0}.$$

Omitting intermediate calculations of the transition, we shall further introduce the moving coordinate system: $\xi = z' - v_0\tau$, assuming $\sigma_{\xi\xi}(\xi,\tau) = \sigma_{z'z'}(z',\tau)$, $T^*(z',\tau) = W(\xi,\tau)$. Relations (41)–(43) may be written now in the following form:

$$\begin{gathered}(\alpha_0^2 - v_0^2)\frac{\partial^2\sigma_{\xi\xi}}{\partial\xi^2} + 2v_0\frac{\partial^2\sigma_{\xi\xi}}{\partial\xi\partial\tau} - \frac{\partial^2\sigma_{\xi\xi}}{\partial\tau^2} = \\ = \frac{\partial^2 W}{\partial\tau^2} - 2v_0\frac{\partial^2 W}{\partial\xi\partial\tau} + v_0^2\frac{\partial^2 W}{\partial\xi^2},\ \xi > 0,\ \tau > 0,\end{gathered} \tag{44}$$

$$\left.\begin{aligned}&\sigma_{\xi\xi}(\xi,\tau)\big|_{\tau=0}=\left.\frac{\partial\sigma_{\xi\xi}(\xi,\tau)}{\partial\tau}\right|_{\tau=0}=0,\ \xi\geq 0,\\&\sigma_{\xi\xi}(\xi,\tau)\big|_{\xi=0}=0,\ \tau>0,\ \left|\sigma_{\xi\xi}(\xi,\tau)\right|<\infty,\ \xi\geq 0,\ \tau\geq 0\end{aligned}\right\}. \quad (45)$$

In the Laplace image space $\overline{\sigma}_{\xi\xi}(\xi,p)=\int_0^\infty\sigma_{\xi\xi}(\xi,\tau)\exp(-p\tau)d\tau$, relations (44)–(45) are written in the following form:

$$(\alpha_0^2-v_0^2)\frac{d^2\overline{\sigma}_{\xi\xi}}{d\xi^2}+2v_0p\frac{d\overline{\sigma}_{\xi\xi}}{d\xi}-p^2\overline{\sigma}_{\xi\xi}=$$
$$=\frac{p(p+v_0^2)}{1-\tau_0v_0^2}\overline{W}-\frac{v_0(v_0^2+2p)}{1-\tau_0v_0^2}\frac{d\overline{W}}{d\xi},\ \xi>0, \quad (46)$$

$$\overline{\sigma}_{\xi\xi}(\xi,p)\Big|_{\xi=0}=0,\ \left|\overline{\sigma}_{\xi\xi}(\xi,p)\right|<\infty,\ \xi\geq 0. \quad (47)$$

Here, the following relation is used:

$$\frac{d^2\overline{W}}{d\xi^2}=\frac{p(1+\tau_0p)}{1-\tau_0v_0^2}\overline{W}-\frac{v_0(1+2\tau_0p)}{1-\tau_0v_0^2}\frac{d\overline{W}}{d\xi},$$

derived from the operational form of Eq. (22). For reducing the awkwardness in solving problem (46)–(47), we shall take into account the fact that inertial effects in (41) operate at times of microsecond duration. Then expression $\sqrt{\tau_0p^2+p+v_0^2/4}$ included in general solution (26) may be written in the following form:

$$\sqrt{\tau_0p^2+p+v_0^2/4}\simeq p\sqrt{\tau_0}(1+1/(2\tau_0p))$$

and solution (26) takes the following form:

$$\overline{W}(\xi,p)=\frac{v_0/\sqrt{\tau_0}}{p\left[p+(1-v_0\sqrt{\tau_0})/(2\tau_0)\right]}\times$$
$$\times\exp\left[-\frac{(2\tau_0p+1)\xi}{2\sqrt{\tau_0}(1-v_0\sqrt{\tau_0})}\right].$$

The desired voltage in the image space may be written now as follows:

$$\overline{\sigma}_{\xi\xi}(\xi,p)=$$
$$=\overline{F}(p)\left\{\exp\left[-\overline{\gamma}(p)\xi\right]-\exp\left(-\frac{\xi}{\alpha_0-v_0}p\right)\right\}, \quad (48)$$

$$\overline{F}(p)=\left[\frac{A_{11}}{(p+\gamma_{11})(p+\gamma_{12})}+\frac{A_{12}}{p(p+\gamma_{11})(p+\gamma_{12})}+\right.$$
$$\left.+\frac{A_{13}}{p^2(p+\gamma_{11})(p+\gamma_{12})}\right]-\left[\frac{A_{21}}{(p+\gamma_{11})(p+\gamma_{13})}+\right.$$
$$\left.+\frac{A_{31}}{p(p+\gamma_{11})(p+\gamma_{13})}+\frac{A_{32}}{p^2(p+\gamma_{11})(p+\gamma_{13})}\right];$$

$$A_{11}=\frac{v_0(\alpha_0-v_0)(1+2\tau_0-v_0\sqrt{\tau_0})}{2\alpha_0\sqrt{\tau_0}(\alpha_0\sqrt{\tau_0}-1)(1-\tau_0v_0^2)};$$

$$A_{12}=A_{11}\left[v_0^2+(1+\tau_0v_0^2)/(2\tau_0)\right],$$

$$A_{13}=\frac{v_0^3(\alpha_0-v_0)}{4\alpha_0\sqrt{\tau_0}(1-\tau_0v_0^2)(\alpha_0\sqrt{\tau_0}-1)},$$

$$A_{21}=\frac{v_0(\alpha_0+v_0)(1+2\tau_0-v_0\sqrt{\tau_0})}{2\alpha_0\sqrt{\tau_0}(\alpha_0\sqrt{\tau_0}+1)(1-\tau_0v_0^2)},$$

$$A_{31}=A_{21}\left[v_0^2+(1+\tau_0v_0^2)/(2\tau_0)\right],$$

$$A_{32}=\frac{v_0^3(\alpha_0+v_0)}{4\alpha_0\sqrt{\tau_0}(1-\tau_0v_0^2)(\alpha_0\sqrt{\tau_0}+1)},$$

$$\gamma_{11}=\frac{1-v_0\sqrt{\tau_0}}{2\tau_0},\ \gamma_{12}=\frac{\alpha_0-v_0}{2\sqrt{\tau_0}(\alpha_0\sqrt{\tau_0}-1)},$$

$$\gamma_{13}=\frac{\alpha_0+v_0}{2\sqrt{\tau_0}(\alpha_0\sqrt{\tau_0}+1)},$$

$$\overline{\gamma}(p)=\frac{2\tau_0p+1}{2\sqrt{\tau_0}(1-v_0)\sqrt{\tau_0}}.$$

Now from (48), we find for the desired voltage:

- at $\alpha_0\sqrt{\tau_0}=v_p/v_T>1$

$$\sigma_{\xi\xi}(\xi,\tau)=$$
$$=\begin{cases}0 & \tau<\dfrac{\xi}{\alpha_0-v_0},\\ \sigma_{\xi\xi}^{(2)}(\xi,\tau) & \dfrac{\xi}{\alpha_0-v_0}<\tau<\dfrac{\xi\sqrt{\tau_0}}{1-v_0\sqrt{\tau_0}},\\ \sigma_{\xi\xi}^{(1)}(\xi,\tau)+\sigma_{\xi\xi}^{(2)}(\xi,\tau) & \tau>\dfrac{\xi\sqrt{\tau_0}}{1-v_0\sqrt{\tau_0}};\end{cases} \quad (49)$$

- at $\alpha_0\sqrt{\tau_0} = v_p/v_T < 1$

$$\sigma_{\xi\xi}(\xi,\tau) = \begin{cases} 0, & \tau < \dfrac{\xi\sqrt{\tau_0}}{1-v_0\sqrt{\tau_0}}, \\ \sigma_{\xi\xi}^{(1)}(\xi,\tau), & \dfrac{\xi\sqrt{\tau_0}}{1-v_0\sqrt{\tau_0}} < \tau < \dfrac{\xi}{\alpha_0 - v_0}, \\ \sigma_{\xi\xi}^{(1)}(\xi,\tau) + \sigma_{\xi\xi}^{(2)}(\xi,\tau), & \tau > \dfrac{\xi}{\alpha_0 - v_0}. \end{cases} \quad (50)$$

Here:

$$\sigma_{\xi\xi}^{(1)}(\xi,\tau) = F\left(\tau - \frac{\xi\sqrt{\tau_0}}{1-v_0\sqrt{\tau_0}}\right)\exp\left[-\frac{\xi}{2\sqrt{\tau_0}(1-v_0\sqrt{\tau_0})}\right],$$

$$\sigma_{\xi\xi}^{(2)}(\xi,\tau) = -F\left(\tau - \frac{\xi}{\alpha_0 - v_0}\right),$$

$$F(\tau) = \left[B_{11}\exp(-\gamma_{11}\tau) + B_{12}\exp(-\gamma_{12}\tau) + \frac{A_{13}\tau}{\gamma_{11}\gamma_{12}} + B_{13}\right] - \left[B_{21}\exp(-\gamma_{11}\tau) + B_{31}\exp(-\gamma_{13}\tau) + \frac{A_{32}\tau}{\gamma_{11}\gamma_{13}} + B_{32}\right],$$

$$B_{11} = \frac{\gamma_{11}^2 A_{11} - \gamma_{11}A_{12} + A_{13}}{\gamma_{11}^2(\gamma_{12} - \gamma_{11})},$$

$$B_{12} = \frac{A_{12}\gamma_{12} - A_{11}\gamma_{12}^2 - A_{13}}{\gamma_{12}^2(\gamma_{12} - \gamma_{11})},$$

$$B_{13} = \frac{A_{12}}{\gamma_{11}\gamma_{12}} - \frac{(\gamma_{11} + \gamma_{12})A_{13}}{\gamma_{11}^2\gamma_{12}^2},$$

$$B_{21} = \frac{\gamma_{11}^2 A_{21} - \gamma_{11}A_{13} + A_{32}}{\gamma_{11}^2(\gamma_{13} - \gamma_{11})},$$

$$B_{31} = \frac{\gamma_{13}A_{31} - \gamma_{13}^2 A_{21} - A_{32}}{\gamma_{13}^2(\gamma_{13} - \gamma_{11})},$$

$$B_{32} = \frac{A_{31}}{\gamma_{11}\gamma_{13}} - \frac{(\gamma_{11} + \gamma_{13})A_{32}}{\gamma_{11}^2\gamma_{13}^2}.$$

Figure 2 shows the dependence graph of the dynamic temperature stress (49) on the dimensionless time in cross section $\xi = 1$ at $\tau_0 = 0.25$, $\alpha_0 = 3$ ($\alpha_0\sqrt{\tau_0} = v_p/v_T = 1.5 > 1$; for metals $v_p/v_T > 1$, for polymer glasses $v_p/v_T < 1$) for values $v_0 = 0.16$ and 0.65. The curves show that for locally nonequilibrium processes, accounting for the finite rate of heat propagation results in a significant change in the stress pattern compared to the corresponding curves in the framework of classical Fourier phenomenology [21]. We shall take an arbitrary point (cross section ξ = const). In the beginning, the stresses in it are zero. At the moment of time $\tau = \xi/(\alpha_0 - v_0)$ ($t = (z-l)/v_p$), the longitudinal elastic stress wave which front moves at speed v_p approaches this point. The compression stress changes abruptly and then decreases (increases in absolute value). At the moment of time $\tau = \xi\sqrt{\tau_0}/(1-v_0\sqrt{\tau_0})$, the heat wave which front moves at speed v_T approaches this point (cross section); the stress changes abruptly and then approaches the value close to the quasi-static value.

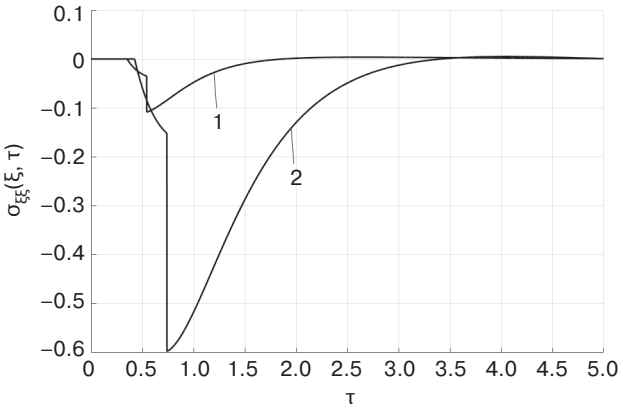


Fig. 2. Dependence of stress $\sigma_{\xi\xi}(\xi, \tau)$ (49) on τ in cross section $\xi = 1$ at $\tau_0 = 0.25$, $v_0 = 0.16$ (1) and $v_0 = 0.65$ (2); $\alpha_0 = 3$

Thus, two waves propagate in a massive solid body (an elastic half-space with a moving thermally insulated boundary), which comprise a thermal wave and an elastic wave; here, the elastic wave front precedes that of the thermal wave. The present author's earlier studies on the effect of heat transfer at the moving boundary of the domain indicate that the dynamic temperature stresses decrease as the heat transfer from the surface of the half-space decreases. If, in the classical case [3], the presence of finite heat transfer from the surface of the half-space boundary results in the disappearance of temperature stress discontinuities, then, in the case of the generalized dynamic thermoelasticity problem [20], the stress character remains the same as at the infinitely large value of the heat transfer coefficient (the first kind boundary condition). It is hoped that this earlier part of the research, being very voluminous in its content, will be published at some point in the future.

CONCLUSIONS

The above model representations provide a basis for the following statement. A new effect of analytical thermophysics and applied thermomechanics is described. In a domain with a moving thermally insulated boundary, the temperature gradient occurs resulting in the appearance of the temperature field and corresponding temperature stresses despite the absence of external and internal heat sources. This is due to the kinetic energy of the moving boundary being converted into the thermal energy of the domain. The above-described effect is manifested both in the classical Fourier phenomenology (parabolic type equations), as well as in the generalized phenomenology for locally nonequilibrium processes (hyperbolic type equations).

REFERENCES

1. Kartashov E.M. Thermal destruction of polymer fibers in the theory of temporary dependence of strength. *Tonk. Khim. Technol. = Fine Chem. Technol*. 2021;16(6):526–540 (in Russ.). https://doi.org/10.32362/2410-6593-2021-16-6-526-540
2. Kartashov E.M., Soloviev I.A. Stochastic interpretation of effect of emergence of the gradient of temperature at the heatisolated moving border. *Izvestiya RAN. Energetika*. 2017;1:119–128 (in Russ.).
3. Kartashov E.M., Kudinov V.A. *Analiticheskie metody teorii teploprovodnosti i ee prilozhenii* (*Analytical Methods of the Theory of Heat Conduction and its Applications*). Moscow: URSS; 2012. 1080 p. (in Russ.). ISBN 978-5-9710-4994-4
4. Vernotte P. Les paradoxes de la theorie continue de l'eguation de la chaleur. *Comptes Rendus. Acad. Sci. Paris*. 1958;246(22):3154–3155.
5. Lykov A.V. *Teoriya teploprovodnosti* (*Theory of Heat Conduction*). Moscow: Vysshaya shkola; 1967. 600 p. (in Russ.).
6. Cattaneo C. Sur une forme de l'eguation de la chaleur eliminant le paradoxe d'une propagation instantance. *Comptes Rendus de l'Académie des Sciences*. 1958;247(4):431–433.
7. Luikov A.V. Application of methods of thermodynamics of irreversible processes to the investigation of heat and mass transfer. *J. Eng. Phys.* 1965;9(3):189–202. https://doi.org/10.1007/BF00828333
   [Original Russian Text: Lykov A.V. Application of methods of thermodynamics of irreversible processes to the investigation of heat and mass transfer. *Inzhenerno-Fizicheskii Zhurnal.* 1965;9(3):287–304 (in Russ.).]
8. Baumeister K., Hamill T. Hyperbolic heat equation. Solving the problem of a semi-infinite body. *Teploperedacha = J. Heat Transfer*. 1969;4:112–119 (in Russ.).
9. Sobolev S.L. Local Non-Equilibrium Transport Models. *Phys. Usp*.1997;40(10):1043. https://doi.org/10.1070/PU1997v040n10ABEH000292
   [Original Russian Text: Sobolev S.L. Local Non-Equilibrium Transport Models. *Uspekhi Fizicheskikh Nauk.* 1997;167(10):1095–1106 (in Russ). https://doi.org/10.3367/UFNr.0167.199710f.1095]
10. Zhou D., Casas-Basquez H., Lebon J. *Rasshirennaya neobratimaya termodinamika* (*Extended Irreversible Thermodynamics*): transl. from Engl. Moscow–Izhevsk: Institut Komp'yuternykh Issledovanii; 2006. 528 p. (in Russ.).

СПИСОК ЛИТЕРАТУРЫ

1. Карташов Э.М. Тепловое разрушение полимерных волокон в теории временной зависимости прочности. *Тонкие химические технологии*. 2021;16(6):526–540. https://doi.org/10.32362/2410-6593-2021-16-6-526-540
2. Карташов Э.М., Соловьев И.А. Стохастический анализ эффекта возникновения градиента температуры при теплоизолированной движущейся границе. *Известия РАН. Энергетика*. 2017;1:119–128.
3. Карташов Э.М., Кудинов В.А. *Аналитические методы теории теплопроводности и ее приложений*. М.: URSS; 2017.1080 с. ISBN 978-5-9710-4994-4
4. Vernotte P. Les paradoxes de la theorie continue de l'eguation de la chaleur. *Comptes Rendus. Acad. Sci. Paris*. 1958;246(22):3154–3155.
5. Лыков А.В. *Теория теплопроводности*. М.: Высшая школа; 1967. 600 с.
6. Cattaneo C. Sur une forme de l'eguation de la chaleur eliminant le paradoxe d'une propagation instantance. *Comptes Rendus de l'Académie des Sciences*. 1958;247(4):431–433.
7. Лыков А.В. Применение методов термодинамики необратимых процессов к исследованию тепло- и массообмена. *Инженерно-физический журнал*. 1965;9(3):287–304.
8. Баумейстер К., Хамилл Т. Гиперболическое уравнение теплопроводности. Решение задачи о полубесконечном теле. *Теплопередача*. 1969;4:112–119.
9. Соболев С.Л. Локально-неравновесные модели процессов переноса. *Успехи физ. наук*.1997;167(10): 1095–1106. https://doi.org/10.3367/UFNr.0167.199710f.1095
10. Жоу Д., Касас-Баскес Х., Лебон Дж. *Расширенная необратимая термодинамика*: пер. с англ. М.- Ижевск: Институт компьютерных исследований; 2006. 528 с.
11. Кудинов В.А., Кудинов И.В. Получение и анализ точного аналитического решения гиперболического уравнения теплопроводности для плоской стенки. *Теплофизика высоких температур*. 2012;50(1): 118–126.
12. Кудинов В.А., Кудинов И.В. Исследование теплопроводности с учетом конечной скорости распространения теплоты. *Теплофизика высоких температур*. 2013;51(2):301–310.
13. Кирсанов Ю.А. *Моделирование теплофизических процессов*. СПб.: Изд-во Политехника; 2022. 230 с.

11. Kudinov V.A., Kudinov I.V. One method of reception of the exact analytical decision of the hyperbolic equation of heat conductivity on the basis of use of orthogonal methods. *High Temp*. 2012;50(1):112–119. https://doi.org/10.1134/S0018151X12010105
[Original Russian Text: Kudinov V.A., Kudinov I.V. One method of reception of the exact analytical decision of the hyperbolic equation of heat conductivity on the basis of use of orthogonal methods. *Teplofizika Vysokikh Temperatur*. 2012;50(1):118–126 (in Russ.).]
12. Kudinov V.A., Kudinov I.V. Studying heat conduction taking into account the finite rate of heat propagation. *High Temp.* 2013;51(2):268–276. https://doi.org/10.1134/S0018151X1204013X
[Original Russian Text: Kudinov V.A., Kudinov I.V. Studying heat conduction taking into account the finite rate of heat propagation. *Teplofizika Vysokikh Temperatur*. 2013;51(2):301–310 (in Russ.).]
13. Kirsanov Yu.A. *Modelirovanie teplofizicheskikh protsessov* (*Modeling of Thermophysical Processes*). St. Petersburg: Politekhnika; 2022. 230 p. (in Russ.).
14. Eremin A.V. Research on fast relaxing temperature excitations caused by ultrashort laser pulses. *Sovremennaya nauka: Aktual'nye problemy teorii i praktiki. Seriya: Estestvennye i tekhnicheskie nauki = Modern Science: Actual Problems of Theory & Practice. Ser.: Natural & Technical Sciences*. 2019;8:47–52 (in Russ.).
15. Eremin A.V. About one method of heat transfer process in solid bodies mathematical modeling. *Perspektivy nauki = Science Prospects*. 2019;7(118):117–119 (in Russ.).
16. Formalev V.F. *Uravneniya matematicheskoi fiziki* (*Equations of Mathematical Physics*). Moscow: URSS; 2021. 648 p. (in Russ.). ISBN 978-5-9710-8380-1

14. Еремин А.В. Исследование быстрорелаксирующих температурных возбуждений, вызываемых сверхкороткими импульсами лазерного излучения. *Современная наука: Актуальные проблемы теории и практики. Серия: Естественные и технические науки*. 2019;8:47–52.
15. Еремин А.В. Об одном методе математического моделирования процесса переноса теплоты в твердых телах. *Перспективы науки*. 2019;7(118):117–119.
16. Формалев В.Ф. *Уравнения математической физики*. М.: URSS; 2021. 648 с. ISBN 978-5-9710-8380-1

About the author

Eduard M. Kartashov, Dr. Sci. (Phys.-Math.), Honored Scientist of the Russian Federation, Honorary Worker of Higher Professional Education of the Russian Federation, Honorary Worker of Science and Technology of the Russian Federation, Honorary Professor of the Lomonosov Moscow State University of Fine Chemical Technology, Laureate of the Golden Medal of the Academy of Sciences of Belarus in Thermophysics, Professor, Department of Higher and Applied Mathematics, M.V. Lomonosov Institute of Fine Chemical Technologies, MIREA – Russian Technological University (86, Vernadskogo pr., Moscow, 119571 Russia). E-mail: professor.kartashov@gmail.com. Scopus Author ID 7004134344, ResearcherID Q-9572-2016, https://orcid.org/0000-0002-7808-4246

Об авторе

Карташов Эдуард Михайлович, д.ф.-м.н., Заслуженный деятель науки РФ, Почетный работник высшего профессионального образования РФ, Почетный работник науки и техники РФ, Почетный профессор МИТХТ им. М.В. Ломоносова, Лауреат Золотой медали АН Беларуси по теплофизике, профессор, кафедра высшей и прикладной математики Института тонких химических технологий им. М.В. Ломоносова ФГБОУ ВО «МИРЭА – Российский технологический университет» (119571, Россия, Москва, пр-т Вернадского, д. 86). E-mail: professor.kartashov@gmail.com. Scopus Author ID 7004134344, ResearcherID Q-9572-2016, https://orcid.org/0000-0002-7808-4246

Translated from Russian into English by Kirill V. Nazarov
Edited for English language and spelling by Thomas A. Beavitt

MIREA – Russian Technological University.
78, Vernadskogo pr., Moscow, 119454 Russian Federation.
Publication date September 30, 2023.
Not for sale.

МИРЭА – Российский технологический университет.
119454, РФ, г. Москва, пр-т Вернадского, д. 78.
Дата опубликования 30.09.2023 г.
Не для продажи.

https://www.rtj-mirea.ru


МИРЭА
Российский
технологический
университет