DISCRETE AND CONTINUOUS MODELS AND APPLIED COMPUTATIONAL SCIENCE

Volume 34 Number 2 (2026)

Founded in 1993

Founder: PEOPLES' FRIENDSHIP UNIVERSITY OF RUSSIA NAMED AFTER PATRICE LUMUMBA

DOI: 10.22363/2658-4670-2026-34-2

Edition registered by the Federal Service for Supervision of Communications, Information Technology and Mass Media
Registration Certificate: ПИ № ФС 77-76317, 19.07.2019

ISSN 2658-7149 (Online); 2658-4670 (Print)
4 issues per year.
Language: English.

Publisher

Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University).

Indexed by

- Scopus (https://www.scopus.com),
- Ulrich's Periodicals Directory (http://www.ulrichsweb.com),
- Directory of Open Access Journals (DOAJ) (https://doaj.org),
- Russian Index of Science Citation (https://elibrary.ru),
- RCNI White list.

Aim and Scope

Discrete and Continuous Models and Applied Computational Science arose in 2019 as a continuation of RUDN Journal of Mathematics, Information Sciences and Physics. RUDN Journal of Mathematics, Information Sciences and Physics arose in 2006 as a merger and continuation of the series "Physics", "Mathematics", "Applied Mathematics and Computer Science", "Applied Mathematics and Computer Mathematics".

Discussed issues affecting modern problems of physics, mathematics, queuing theory, the Teletraffic theory, computer science, software and databases development.

It's an international journal regarding both the editorial board and contributing authors as well as research and topics of publications. Its authors are leading researchers possessing PhD and PhDr degrees, and PhD and MA students from Russia and abroad. Articles are indexed in the Russian and foreign databases. Each paper is reviewed by at least two reviewers, the composition of which includes PhDs, are well known in their circles. Author's part of the magazine includes both young scientists, graduate students and talented students, who publish their works, and famous giants of world science.

The Journal is published in accordance with the policies of COPE (Committee on Publication Ethics).

The editors are open to thematic issue initiatives with guest editors. Further information regarding notes for contributors, subscription, and back volumes is available at http://journals.rudn.ru/miph

E-mail: miphj@rudn.ru.

Editorial board

Computer Design: *Anna V. Korolkova*, *Dmitry S. Kulyabov*
English Text Editors: *Nikolay E. Nikolaev*, *Ivan S. Zaryadov*, *Konstantin P. Lovetskiy*
Address of editorial board: 3 Ordzhonikidze St, 115419, Moscow, Russia, +7 (495) 955-07-16, e-mail: publishing@rudn.ru
Editorial office: +7 (495) 952-02-50, e-mail: miphj@rudn.ru, site: http://journals.rudn.ru/miph

---

Approved for printing: 15.06.2026. Published: 25.06.2026.
Paper size 70×100/16. Offset paper. Offset printing. Typeface “Adobe Source.”
Conventional printed sheets: 10.32. Printing run 500 copies. Open price. The order: 542.
PEOPLES' FRIENDSHIP UNIVERSITY OF RUSSIA NAMED AFTER PATRICE LUMUMBA
6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation
Printed at RUDN Publishing House:
3 Ordzhonikidze St, Moscow, 115419, Russian Federation,
+7 (495) 955-08-61; e-mail: publishing@rudn.ru

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2026, 34 (2)
http://journals.rudn.ru/miph

Contents

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2026, 34 (2) 155–159
http://journals.rudn.ru/miph

Editorial

DOI: 10.22363/2658-4670-2026-34-2-155-159

EDN: JEDOZP

ISO rules for typesetting mathematics

Dmitry S. Kulyabov[1,2], Anna V. Korolkova[1], Leonid A. Sevastianov[1], Yuri P. Rybakov[1]

[1] *RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation*

[2] *Joint Institute for Nuclear Research, 6 Joliot-Curie St, Dubna, 141980, Russian Federation*

Abstract. Using ISO standards for mathematical notation is important for improving the quality of scientific paper presentation. This article describes the key requirements of ISO standards for mathematical notation. It also provides recommendations for the notation of differential operators.

Key words and phrases: typesetting mathematics standards, ISO, LaTeX

For citation: D. S. Kulyabov, A. V. Korolkova, L. A. Sevastianov, Y. P. Rybakov "ISO rules for typesetting mathematics," *Discrete and Continuous Models and Applied Computational Science*, vol. 34, no. 2, pp. 155–159, 2026. DOI: 10.22363/2658-4670-2026-34-2-155-159. EDN: JEDOZP.

1. ISO standards

When typesetting mathematics, standards must be followed.

ISO standards represent strict rules for the typesetting of mathematical data in the fields of physics and applied sciences. ISO standards are not free, but they are the reference standard for all typography guidelines.

Standards:

- ISO 31-11:1992 (https://en.wikipedia.org/wiki/ISO_31-11). Superseded by ISO 80000-1:2009, ISO 80000-2:2019.
- ISO 80000-1:2009/Amd 1:2021 — Quantities and units — Part 1: General (https://www.iso.org/standard/76921.html, https://en.wikipedia.org/wiki/ISO/IEC_80000) [1]. General principles: number notation, decimal sign, digit grouping, rules for units of measurement.
- ISO 80000-2:2019 — Quantities and units — Part 2: Mathematics (https://www.iso.org/standard/64973.html) [2]. The basic standard for mathematical typesetting rules: italic/roman font, notations for functions, derivatives, imaginary units, etc.

2. Brief description of the standard

The essence of the standard is that the mathematical role of a symbol should be immediately clear from its design.

2.1. Variables

The word *variable* is used to denote a mathematical entity representing variable data.

Variables are denoted by a single letter. Abbreviations are not permitted. [1]. Some abbreviations have become widely popular and used, but they are prohibited by ISO rules.

Variables are displayed in italics. If capital Greek letters denote a variable, they are also italicized (remember that in standard LaTeX, they are upright by default).

2.2. Quantities

The word *quantity* is used to denote any physical entity that can be measured according to metrological practice.

All quantity symbols should be italicized; oblique fonts are permitted, but italicized serif fonts should be preferred unless ISO regulations specify a sans-serif font:

- mathematical constants: numbers whose value does not change should be upright (e.g., e, i, π);
- differential symbol d should be upright to avoid confusion with the physical quantity d;
- number e should be upright to avoid confusion with the elementary electric charge e;
- The imaginary unit j in electrical engineering (i in other sciences) should be written in roman font to avoid confusion with the electric current density j or electric current i;
- The transcendental number $\pi = 3.14159$ should be distinguished from the angle π, etc.

2.3. Other designations

All characters that do not represent quantities should be set in a roman font, preferably a serif font, except where ISO rules require a sans-serif font:

- the rule includes numbers and their digits, and characters representing constant numerical values;
- mappings are not quantities or variables: for example, in the expression V_i, the subscript i is variable because it represents the ith element in a sequence such as V_0, V_1, V_2,...;
- if the subscript is an addition, such as "i" meaning "input", it is written in roman: V_i.

Matrices are typed in bold italic:

- single-column matrices are vectors and should be treated like any other matrix;
- multi-row and multi-column matrices are typed in uppercase;
- lowercase letters are reserved for vectors;
- geometric vectors, which are typed using uppercase or lowercase italic letters with an arrowhead, are not considered by ISO rules.

Tensors should be typed in an italic, bold, sans-serif font.

Labels for geometric objects such as points, segments, and angles (not their dimensions) should be in sans-serif fonts. The same rule applies to labels used in sketches and drawings representing mechanisms, electrical circuits, and the like, when the label refers to the object rather than its dimension.

Notes for general (non-standard) functions are always italicized (e.g., x, $f(x)$). Standard functions (e.g., sin, cos, ln) are typed in roman font.

[1] In programming, identifiers are called variables in the sense that they can represent variable data, but computer programs are not mathematics.

3. Software used

3.1. Packages for XeLaTeX and LuaLaTeX

For XeLaTeX and LuaLaTeX, you can use the `unicode-math` package (https://ctan.org/pkg/unicode-math). This package provides full control over fonts in formulas for XeLaTeX and LuaLaTeX. It allows you to choose a font where all symbols are natively ISO-compliant (for example, constants are direct). When using `unicode-math`, many older math packages (`amssymb`, `mathrsfs`, etc.) become unnecessary, since their symbols are now directly accessible.

3.2. Packages for pdfLaTeX

Since it is not possible to implement all the features using a standard font in this case, it is necessary to use a set of packages that provide similar features.

- `PM-ISOmath`[2] (https://www.ctan.org/pkg/pm-isomath). Instead of loading new mathematical alphabets, the package defines a number of user-defined macros and workarounds for properly formatting characters in pdfLaTeX without exhausting the limited allocations for mathematical groups. Focuses on the core user-defined macros needed for standard ISO-compliant characters.
- `isomath` (https://www.ctan.org/pkg/isomath). Loads and defines special new mathematical alphabets and font slots for bold italic and bold italic sans-serif. May result in the error *too many mathematical alphabets*. Configures full mathematical alphabets for Latin and Greek characters. Provides explicit semantic markup commands for vectors, matrices, and tensors.
- `mismath`. The package implements a set of macros for ISO-compliant typesetting.

3.3. Additional packages

- `siunitx`. For the set of units of measurement in accordance with ISO requirements [3].
- `derivative`. A package for setting derivatives and differentials. Since these operators are widely used, we will examine this package in more detail.

4. Derivative package

The `derivative` package is a modern solution for typesetting derivatives, differentials, and integrals in LaTeX. It is designed as a direct, safe, and extremely convenient replacement for the **`\dv`** and **`\pdv`** commands from the legacy `physics` package.

CTAN: https://www.ctan.org/pkg/derivative.

Repository: https://github.com/sjelatex/derivative

The package is written in LaTeX3. It works in any engine (pdfLaTeX, XeLaTeX, LuaLaTeX) and does not require encoding or font changes. The package uses semantic markup. The author specifies what needs to be obtained (ordinary derivative, quotient, differential), and the package takes care of proper spacing, argument order, and correct display.

4.1. Basic commands and basic syntax

The core of the package consists of four families of commands similar to those used in the `physics` package, but with more logical and predictable behavior.

[2] The Poor Man ISO math bundle.

Ordinary derivatives are entered using **\odv** (ordinary derivative).

Ordinary derivatives

```
\(\odv{f}{x}\)\\
\(\odv[n]{f}{x}\)\\
\(\odv[delims-eval=.|]{f}{x}_{x=0}\)\\
\(\odv*{f}{x}\)
```

$$\frac{\mathrm{d}f}{\mathrm{d}x}$$
$$\frac{\mathrm{d}^n f}{\mathrm{d}x^n}$$
$$\frac{\mathrm{d}f}{\mathrm{d}x}_{x=0}$$
$$\frac{\mathrm{d}}{\mathrm{d}x}f$$

Partial derivatives are defined by the operator **\pdv** (partial derivative).

Partial derivatives

```
\(\pdv{f}{x}\)\\
\(\pdv[2]{f}{x}\)\\
\(\pdv{f}{x}{y}\)\\
\(\pdv[order={2,1}]{f}{x,y}\)\\
\(\pdv[order={n}]{f}{x}\)\\
\(\pdv[order={2,3}]{f}{x,y}\)\\
\(\pdv{f}{x,y,z,k,l}\)
```

$$\frac{\partial f}{\partial x}$$
$$\frac{\partial^2 f}{\partial x^2}$$
$$\frac{\partial f}{\partial x}y$$
$$\frac{\partial^3 f}{\partial x^2\,\partial y}$$
$$\frac{\partial^n f}{\partial x^n}$$
$$\frac{\partial^5 f}{\partial x^2\,\partial y^3}$$
$$\frac{\partial^5 f}{\partial x\,\partial y\,\partial z\,\partial k\,\partial l}$$

If the order of the derivative is not explicitly specified, the package automatically calculates the sum of the orders for mixed derivatives, assuming the order of each variable is equal to 1.

Differentials are specified by the operators **\odif** (ordinary differential) and **\pdif** (partial differential).

Differentials

```
\(\odif{x}\)\\
\(\pdif{x}\)\\
\(\odif{f(x)}\)\\
\(\odif[2]{x}\)\\
\(\int f(x) \odif{x}\)
```

$$\mathrm{d}x$$
$$\partial_x$$
$$\mathrm{d}f(x)$$
$$\mathrm{d}^2 x$$
$$\int f(x)\,\mathrm{d}x$$

Author Contributions: The contributions of the authors are equal. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analysed during this study. Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Declaration on Generative AI: The authors have not employed any Generative AI tools.

References

[1] *Quantities and units – Part 1: General*, ISO 80000-1:2022, Geneva, Switzerland: ISO, Nov. 2, 2023.
[2] *Quantities and units – Part 2: Mathematics*, ISO 80000-2:2019, Geneva, Switzerland: ISO, Sep. 15, 2019.
[3] D. S. Kulyabov, A. V. Korolkova, L. A. Sevastianov, and Y. P. Rybakov, "Physical dimensional quantities typesetting," *Discrete and Continuous Models and Applied Computational Science*, vol. 34, no. 1, pp. 5–11, 2026. DOI: 10.22363/2658-4670-2026-34-1-5-11

Information about the authors

Kulyabov, Dmitry S.—Professor, Doctor of Sciences in Physics and Mathematics, Professor of Department of Probability Theory and Cyber Security of RUDN University; Senior Researcher of Laboratory of Information Technologies, Joint Institute for Nuclear Research (e-mail: kulyabov-ds@rudn.ru, ORCID: 0000-0002-0877-7063, ResearcherID: I-3183-2013, Scopus Author ID: 35194130800)

Korolkova, Anna V.—Docent, Candidate of Sciences in Physics and Mathematics, Associate Professor of Department of Probability Theory and Cyber Security of RUDN University (e-mail: korolkova-av@rudn.ru, ORCID: 0000-0001-7141-7610, ResearcherID: I-3191-2013, Scopus Author ID: 36968057600)

Sevastianov, Leonid A.—Professor, Doctor of Sciences in Physics and Mathematics, Professor of Department of Computational Mathematics and Artificial Intelligence of RUDN University (e-mail: sevastianov-la@rudn.ru, ORCID: 0000-0002-1856-4643, ResearcherID: B-8497-2016, Scopus Author ID: 8783969400)

Rybakov, Yuri P.—Professor, Doctor of Sciences in Physics and Mathematics, Professor of the Institute of Physical Research and Technologies of RUDN University (e-mail: rybakov-yup@rudn.ru, ORCID: 0000-0002-7744-9725, ResearcherID: S-4813-2018, Scopus Author ID: 16454766600)

RUSSIAN METADATA

DOI: 10.22363/2658-4670-2026-34-2-155-159 EDN: JEDOZP

Правила ISO для набора математики

Д. С. Кулябов[1,2], А. В. Королькова[1], Л. А. Севастьянов[1], Ю. П. Рыбаков[1]

[1] *Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация*

[2] *Объединённый институт ядерных исследований, ул. Жолио-Кюри, д. 6, Дубна, 141980, Российская Федерация*

Аннотация. Использование стандартов ISO для набора математики важно для повышения качества представления научных работ. В статье описываются основные требования стандартов ISO для набора математики. Кроме того, даются рекомендации по набору дифференциальных операторов.

Ключевые слова: стандарты набора математики, ISO, LaTeX

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2026, 34 (2) 160–174
http://journals.rudn.ru/miph

Research article

UDC 519.7
PACS 07.05.Tp
DOI: 10.22363/2658-4670-2026-34-2-160-174 EDN: JHMSMG

Modeling authorship attribution for short texts under training corpus degradation

Viktor M. Khvostenko[1], Andrey P. Kovalenko[2], Sergey Yu. Melnikov[1,3], Roman V. Meshcheryakov[1,4]

[1] *HSE University, 11 Pokrovsky Bulvar, Moscow, 109028, Russian Federation*

[2] *Academy of Cryptography of the Russian Federation, post office box 100, Moscow, 119331, Russian Federation*

[3] *RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation*

[4] *Institute of Control Science of RAS, ICS RAS, 65 building 2 Profsoyuznaya St, Moscow, 117342, Russian Federation*

(received: January 12, 2026; revised: February 10, 2026; accepted: February 20, 2026)

Abstract. Attributing authorship of short texts is a complex task, and its accuracy heavily depends on the quality and quantity of training data. In many practical scenarios, however, this data is subject to degradation. This study examines how two specific types of degradation – textual corruption and reduction in corpus size – affect the performance of six standard classifiers for authorship attribution of English-language tweets. We model textual corruption as random and independent character-level distortions applied to the texts in the training corpora. The test set (the texts whose authorship is to be attributed) remains unchanged. The experimental setup involves 50 authors, with 200 test texts. The volume of training texts per author varies from 800 to 50 (across 16 incremental steps), and the distortion level ranges from 0 to 20% (6 gradations). For each combination of corpus size and distortion level, we train and test six classifiers: Support Vector Machines (SVM), Logistic Regression, Naive Bayes, Random Forest, Decision Tree, and k-Nearest Neighbors (kNN). Authorial style is represented using Bag of Words (BOW), TF-IDF, word token N-grams, and their combinations. The results show that for most classifiers, degradation in the training sets leads to comparable reductions in attribution accuracy. The highest overall performance was achieved by the Support Vector Machine with TF-IDF features. It proved to be the most robust method both when the number of training texts was reduced and under high distortion levels. In the absence of distortions, SVM with TF-IDF achieved an accuracy of 0.92; at distortion levels of 5 and 20%, its accuracy reached 0.89 and 0.75, respectively. Logistic Regression with Bag of Words ranked second, delivering accuracies of 0.9, 0.87, and 0.74 under the same conditions. The kNN method demonstrated the lowest accuracy among the classifiers examined.

Key words and phrases: Authorship Attribution, Short Texts, Training Corpus Degradation, Textual Corruption, Classifier Performance

For citation: V. M. Khvostenko, A. P. Kovalenko, S. Y. Melnikov, R. V. Meshcheryakov "Modeling authorship attribution for short texts under training corpus degradation," *Discrete and Continuous Models and Applied Computational Science*, vol. 34, no. 2, pp. 160–174, 2026. DOI: 10.22363/2658-4670-2026-34-2-160-174. EDN: JHMSMG.

1. Introduction

Research into automatic authorship attribution dates back to the mid-20th century. Over this period, a wide range of methods has been developed – from statistical author profiles to modern large language models. Today, the task remains highly relevant due to the proliferation of digital communication channels (messengers, email, forums, blogs), which often lack robust user verification and are susceptible to identity forgery.

Contemporary studies in authorship attribution increasingly focus on more challenging variants of the classical problem. These include the analysis of short texts [1], attribution under conditions of textual distortion [2], and scenarios with a large number of candidate authors [3]. Each of these complications has been shown to significantly reduce attribution accuracy.

However, the combined effect of two simultaneous challenges—reduced training corpus size and textual distortions within those corpora—remains underexplored. This paper addresses this gap by investigating authorship attribution for short texts under such conditions of training data degradation.

We focus on random character-level distortions (specifically, symbol substitutions) and consider six standard classifiers: SVM, Logistic Regression, Naive Bayes, Random Forest, Decision Tree, and kNN. As stylistic features, we employ token N-grams ($N = 1, 2$) and TF-IDF values computed on various combinations of N-grams—features commonly used in authorship attribution studies.

The aim of this work is to analyze how the accuracy of these classifiers degrades when the training data is subject to random distortions and/or when the volume of the training corpus is reduced.

1.1. Negative effects arising from text distortions

Text distortions can arise from various sources, including typos, optical character recognition (OCR) errors [4], speech recognition errors [5], machine translation inaccuracies [6], and imperfections from decrypting certain ciphers [7]. Such distortions can significantly impair both the human readability of a text and the performance of automated processing systems.

A review by [8] examines how traditional text analysis methods—such as information retrieval, topic identification, summarization, and named entity extraction—are affected by distorted input. More recently, the impact of distortions on large language model (LLM) performance has been explored in [9].

The specific challenge of summarizing documents containing OCR errors was addressed in [10]. The study demonstrated that while the "Cut-and-Paste Text Summarization" method [11] performs well on clean text, its effectiveness drops sharply when applied to distorted text. The quality of the generated summary degrades rapidly with even low levels of noise, with tokenization errors and incorrect sentence boundary detection causing the most significant harm.

Errors introduced at the character level can also propagate through processing pipelines. In [4], the authors trace how OCR errors (character replacements, insertions, and deletions) lead to incorrect tokenization. This, in turn, causes faulty sentence segmentation and ultimately results in errors during part-of-speech tagging. Thus, in a pipeline-based approach, even minor initial distortions can amplify at each subsequent stage.

Interestingly, the impact of distortions varies by task. While [12] reports that OCR errors and synthetic character-level distortions have only a minor effect on document clustering accuracy, they significantly reduce the performance of topic modeling with Latent Dirichlet Allocation.

1.2. On the attribution of authorship of distorted texts

A number of studies have investigated how distortions affect authorship attribution under various conditions. In [13], the problem is considered in a scenario where the training and test texts differ in genre and subject matter—a difference that negatively impacts identification accuracy. The study demonstrates that stylistic features based on character N-grams (particularly for $N = 3$) are more robust to such cross-genre and cross-topic variations than word-based features.

The robustness of character N-grams to textual noise itself was explored in [14] using the "Burrows' delta" method [15]. The authors employed frequency vectors of the most frequent words (MFW) and character N-grams as stylistic features, introducing random character-level distortions at noise levels ranging from 1 to 100% (in 1% increments). Experiments on relatively long historical texts in several European languages revealed a trade-off: higher attribution accuracy at low noise levels comes at the cost of reduced noise immunity, and vice versa. Specifically, increasing the length of MFW vectors improves accuracy when noise is minimal, while shorter vectors perform better as distortion levels rise. The study also found that vectors of the most frequent character 3-grams serve as reliable stylistic markers even under heavy distortion—in some corpora, authorial signal remained traceable up to a 60% noise level.

A different approach to distortion—optical character recognition (OCR) of noisy images—is examined in a series of studies [16–18]. The authors constructed a corpus of 25 English texts (approximately 850 words each) written by five authors. Printed pages were saved as high-resolution JPG images, then artificially degraded using MATLAB's "Salt and Pepper noise" filter [19] at levels from 1 to 7% (in 1% increments). The noisy images were subsequently processed by an OCR system, and authorship attribution experiments were conducted on the resulting texts. Four classifiers were tested, using words and character N-grams ($N = 1$–4) as stylistic features. Unsurprisingly, attribution accuracy decreased as noise increased, with classifiers based on 4-gram characters achieving the highest overall performance.

Finally, a real-world application of authorship attribution under distortion is presented in [2], which analyzes letters by the Brothers Grimm—Jacob (1785–1863) and Wilhelm (1786–1859). The corpus of 85 letters was digitized using the Transkribus system [20] for printed and handwritten text recognition. The proportion of words correctly recognized (matching expert manual transcription) was 88% for printed texts and 81% for handwritten ones. The authors argue that this level of accuracy is sufficient for practical authorship attribution, at least in this binary classification task. Using a support vector machine with a linear kernel (implemented in Scikit-learn [21]), they achieved better-than-chance attribution accuracy (above 50%) on texts where at least 20% of characters were undistorted.

2. Description of experiments

2.1. Corpus Description

The experiments were conducted on the English-language Twitter corpus introduced in [22]. This dataset has been widely used for benchmarking authorship attribution models (see, e.g., [23, 24]). It comprises tweets collected in 2009–2010 from 7,000 globally popular accounts; each account is considered a distinct author. From this pool of 7,000 authors, 50 were randomly selected for the present study, and 1,000 messages were retained for each.

Representative examples from the corpus are shown as follows. The first three texts are from the same author, texts 4 and 5 from another, texts 6 and 7 from a third, and texts 8, 9, and 10 each have different authors.

1. Now playing: Modern Talking - Angie's Heart. Tune in: http://stream.laut.fm/eurodance.m3u.
2. Now playing: A-ha - Take on me. Tune in: http://stream.laut.fm/eurodance.m3u.
3. Now playing: Mozzart - Malice And Vice. Tune in: http://stream.laut.fm/eurodance.m3u.
4. Build your Sales Marketing for Free. Utilize this permission emailing tool to successfully communicate with clients. http://ow.ly/pROT.
5. Build your corporate brand or a new brand through a Free marketing and sending permission based email solution. http://ow.ly/zZopC.
6. New blog post: Stop Smoking by Hypnosis at Self Hypnosis Stop Smoking http://bit.ly/2bRcbm.
7. New blog post: Stop Smoking Habitrol Gum | Little Green Papoose http://bit.ly/8a1JmE.
8. Obama says to reduce role of U.S. nuclear weapons (Reuters) http://bit.ly/9XIGY8.
9. The time in Hollywood, CA, is 6:05:03 am.
10. @julio_zin *smirks and smiles slyly* are you now?

2.2. Data Preparation and Experimental Design

The corpus was split into training and test sets using an 80/20 ratio. For each of the 50 authors, 800 messages were allocated as potential training data and 200 messages as test data.

All texts underwent preliminary preprocessing using the NLTK and Scikit-learn libraries, which included removal of hyperlinks and hashtags, word tokenization, stopword removal, and lemmatization.

The experimental workflow comprised eight stages, illustrated in Figure 1. After preprocessing, features were extracted from the original (clean) test texts. These features served as the reference for evaluation.

For the training data, two types of degradation were introduced. First, synthetic random distortions of the "symbol substitution" type were applied using the method described in [25]. For a given distortion probability of $p \in (0, 1)$, the text was scanned character by character. Each character was independently subjected to replacement with probability p; if replaced, it was substituted by a randomly selected Latin character. The distortion parameter p was chosen such that 0, 1, 3, 5, 10, and 20% of the text characters were distorted. Thus, for each original training text, six versions with different distortion levels were generated. The second type of degradation involved varying the number of training texts per author from 50 to 800 in increments of 50, yielding 16 distinct training set sizes.

These two factors were combined factorially, producing a set of training set variants. For each such variant, stylistic features were extracted. The following feature types were used:

- Bag of Words (BOW)—token counts;
- N-grams—word tokens and bigrams;
- TF-IDF—calculated on five different N-gram combinations: tokens; tokens + bigrams; tokens + bigrams + trigrams; bigrams + trigrams; and tokens + bigrams + trigrams + 4-grams.

Six classifiers from the Scikit-learn library were employed (six classifier variants): Support Vector Machines (SVM), Logistic Regression, Naive Bayes, Random Forest, Decision Tree, and k-Nearest Neighbors (kNN). Default parameters were used for all classifiers, except kNN where $k = 1$ was set.

For each combination of distortion level, corpus size, feature type, and classifier, the model was trained on the degraded training data and subsequently tested on the clean test set. Accuracy was used as the primary evaluation metric, defined as the ratio of correctly attributed authors to the total number of predictions.

Testing the authorship attribution models was carried out in eight stages, as shown in Figure 1.

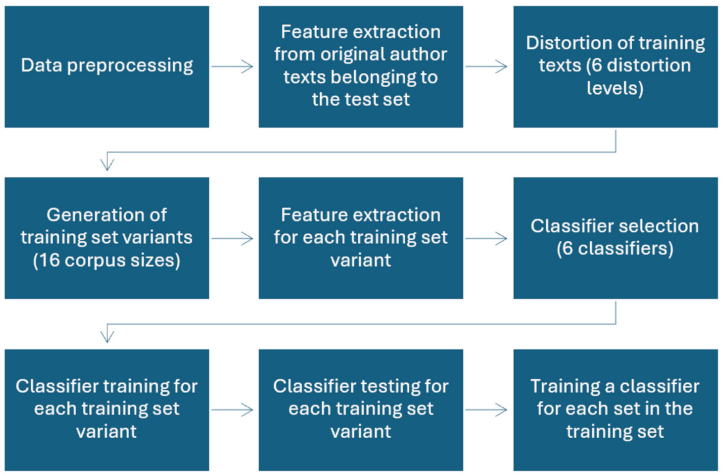

Figure 1. Stages of authorship attribution model testing

3. Results

The experimental results are presented in Figure 2 to 13. The figures show attribution accuracy for six different feature sets, color-coded as follows:

- Blue: Bag of Words (BOW);
- Orange: TF-IDF on tokens;
- Green: *N*-grams (tokens and token bigrams);
- Red: TF-IDF on tokens, bigrams, and trigrams;
- Purple: TF-IDF on bigrams and trigrams;
- Brown: TF-IDF on tokens, bigrams, trigrams, and 4-grams.

For each classifier, two figures are provided: one showing performance on clean training data as a function of training set size, and another showing performance under increasing distortion levels (1, 3, 5, 10, and 20%).

All classifiers were implemented using the Scikit-learn library with default parameters, unless otherwise specified. The kNN classifier required additional parameter tuning, described in Section 3.6.

3.1. Support vector machine (linear kernel)

The linear SVM classifier was implemented using the LinearSVC package from Scikit-learn with default parameters.

3.2. Logistic regression

The logistic regression classifier was implemented using the LogisticRegression package from Scikit-learn with default parameters.

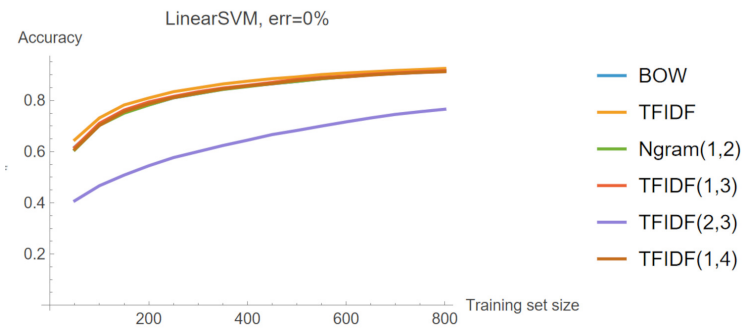

Figure 2. Accuracy of the linear SVM classifier as a function of training set size (clean training data)

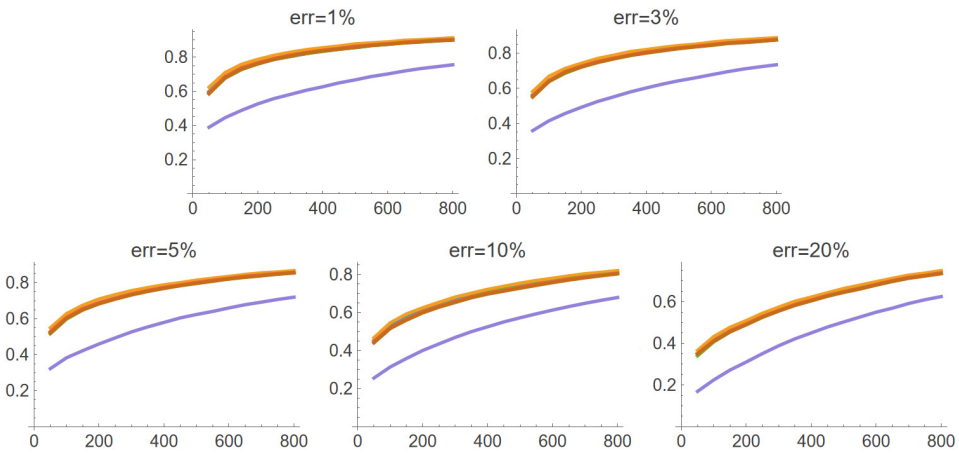

Figure 3. Accuracy of the linear SVM classifier as a function of training set size under different distortion levels (1, 3, 5, 10, and 20%)

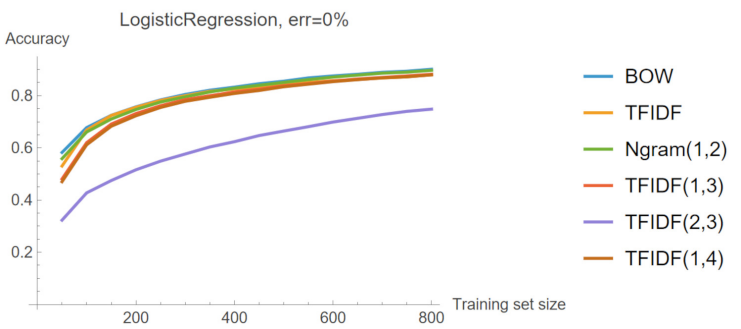

Figure 4. Accuracy of the logistic regression classifier as a function of training set size (clean training data)

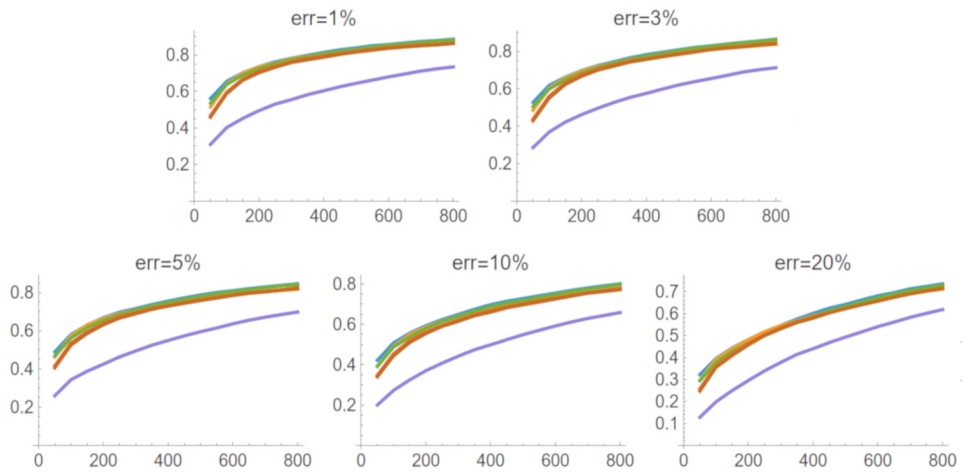

Figure 5. Accuracy of the logistic regression classifier as a function of training set size under different distortion levels (1, 3, 5, 10, and 20%)

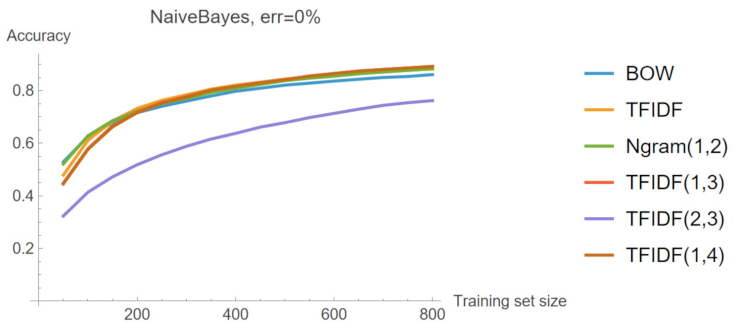

Figure 6. Accuracy of the Naive Bayes classifier as a function of training set size (clean training data)

3.3. Naive bayes

The Naive Bayes classifier was implemented using the MultinomialNB package from Scikit-learn with default parameters.

3.4. Random forest

The Random Forest classifier was implemented using the RandomForestClassifier package from Scikit-learn with default parameters.

3.5. Decision tree

The Decision Tree classifier was implemented using the DecisionTreeClassifier package from Scikit-learn with default parameters.

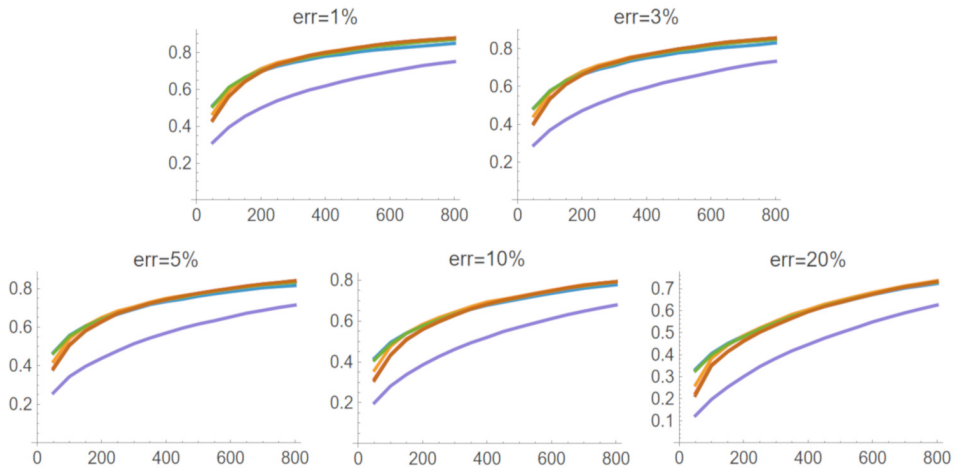

Figure 7. Accuracy of the Naive Bayes classifier as a function of training set size under different distortion levels (1, 3, 5, 10, and 20%)

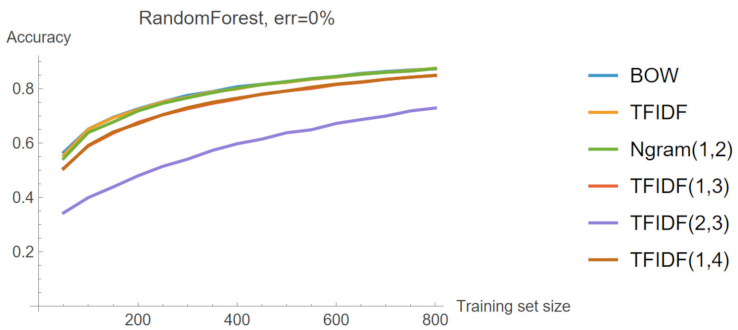

Figure 8. Accuracy of the Random Forest classifier as a function of training set size (clean training data)

3.6. k-Nearest Neighbors (kNN)

The kNN classifier was implemented using the KNeighborsClassifier package from Scikit-learn with the Euclidean distance metric. Since the performance of kNN depends critically on the choice of k, preliminary experiments were conducted to select an appropriate value.

Table 2 presents the accuracy of the kNN classifier for different values of k (from 1 to 10) across all six distortion levels. These results were obtained using TF-IDF features on tokens, which proved to be the best-performing feature set for the maximum training corpus size (see Table 2 below). The highest accuracy for each distortion level is highlighted in bold.

Based on these results, $k = 1$ was selected for all subsequent experiments. Figure 12 and 13 show the accuracy of the kNN classifier (with $k = 1$) as a function of training set size for clean and distorted training data, respectively.

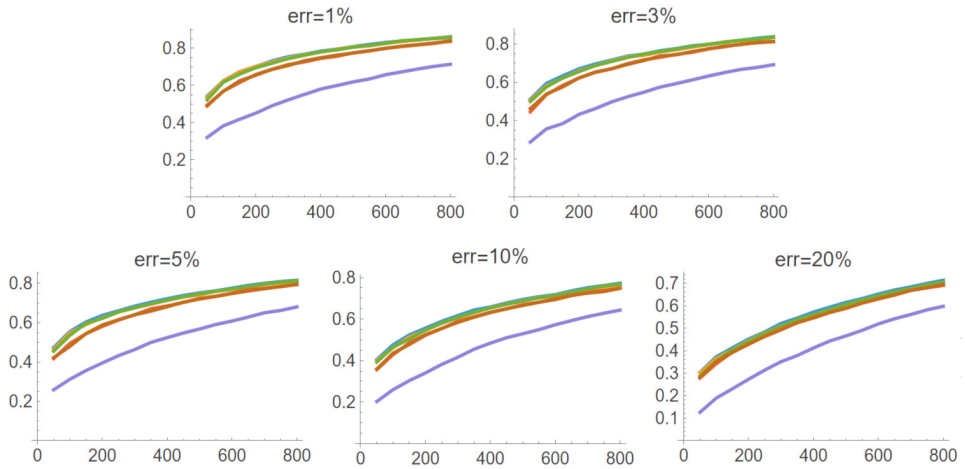

Figure 9. Accuracy of the Random Forest classifier as a function of training set size under different distortion levels (1, 3, 5, 10, and 20%)

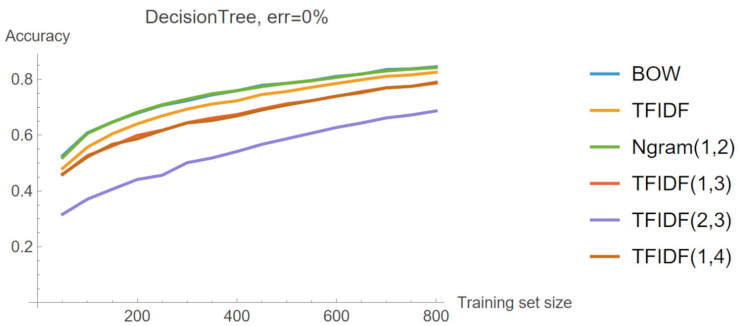

Figure 10. Accuracy of the Decision Tree classifier as a function of training set size (clean training data)

Table 1

Accuracy of the kNN classifier for the maximum training corpus size (800 texts per author) with $1 \leqslant k \leqslant 10$

k	1	2	3	4	5	6	7	8	9	10
Err=0%	**0.83**	0.76	0.74	0.73	0.72	0.71	0.71	0.71	0.72	0.72
Err=1%	**0.82**	0.74	0.72	0.70	0.70	0.69	0.69	0.69	0.69	0.69
Err=3%	**0.79**	0.71	0.68	0.66	0.65	0.64	0.64	0.64	0.64	0.64
Err=5%	**0.77**	0.68	0.65	0.63	0.61	0.6	0.6	0.6	0.6	0.6
Err=10%	**0.72**	0.62	0.59	0.56	0.54	0.52	0.52	0.52	0.51	0.51
Err=20%	**0.66**	0.54	0.49	0.45	0.42	0.41	0.4	0.39	0.39	0.39

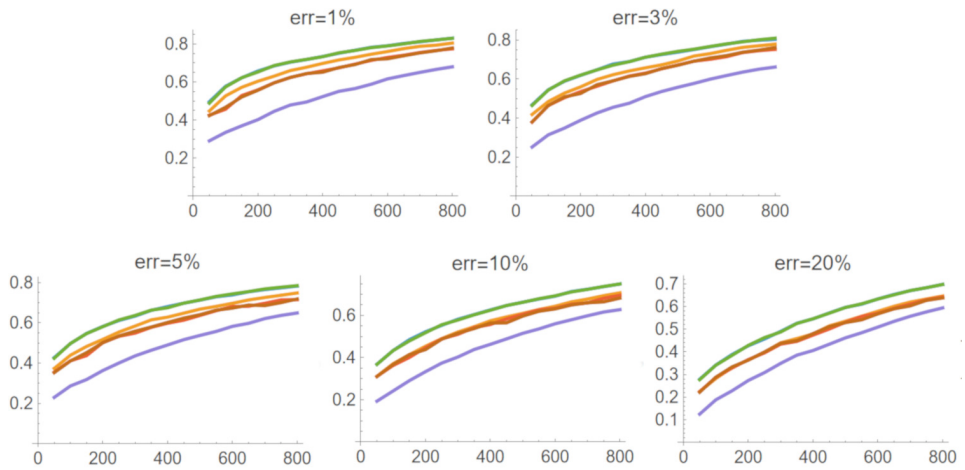

Figure 11. Accuracy of the Decision Tree classifier as a function of training set size under different distortion levels (1, 3, 5, 10, and 20%)

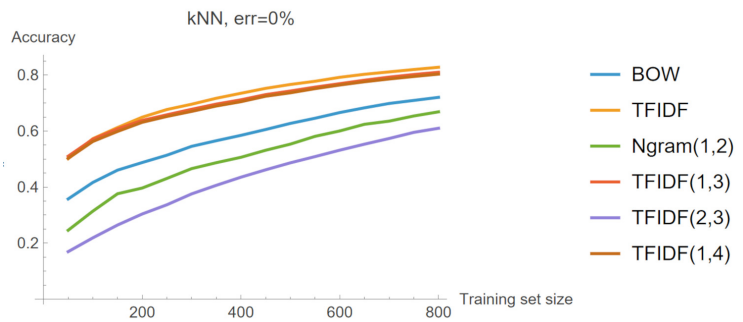

Figure 12. Accuracy of the kNN classifier ($k = 1$) as a function of training set size (clean training data)

4. Analysis and discussion of results

Table 2 summarizes the best attribution accuracy achieved by each classifier for three distortion levels (0%, 5%, and 20%) and five training set sizes (800, 600, 400, 200, and 50 texts per author). The color coding indicates which stylistic feature set yielded the best performance for each classifier under each condition. The feature sets are: Bag of Words (BOW), TF-IDF on tokens, N-grams (1,2), TF-IDF on tokens+bigrams+trigrams (denoted TF-IDF(1,3)), and TF-IDF on tokens+bigrams+trigrams+4-grams (denoted TF-IDF(1,4)).

For the baseline scenario (clean training data with full corpus size), our results closely match those reported in [24], confirming the validity of our experimental setup.

A notable pattern emerges from Figure 2–13: the TF-IDF(2,3) feature set—comprising only bigrams and trigrams of tokens—consistently underperforms compared to all other feature sets. This can be attributed to two factors. First, all other feature sets include unigrams (individual tokens), which appear to be highly informative for authorship attribution. Second, token N-grams are particularly

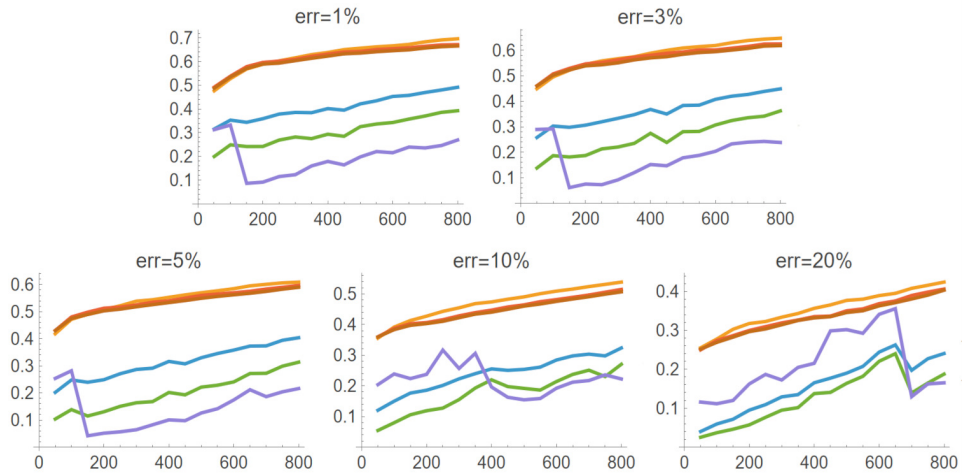

Figure 13. Accuracy of the kNN classifier ($k = 1$) as a function of training set size under different distortion levels (1, 3, 5, 10, and 20%)

vulnerable to character-level distortions, as random symbol substitutions can easily disrupt word boundaries and token sequences.

Support Vector Machine (SVM) demonstrated the highest overall accuracy across nearly all parameter combinations tested. This finding aligns with the well-established suitability of SVM for high-dimensional, sparse data typical of text classification tasks.

Logistic Regression consistently ranked second, delivering comparable performance to SVM in many scenarios, particularly when using Bag of Words features (as noted in Section 3.2).

Naive Bayes, Random Forest, and Decision Tree showed intermediate performance, with accuracy declining predictably as distortion levels increased and training set sizes decreased.

k-Nearest Neighbors (kNN) proved to be the least effective classifier in our experiments, exhibiting both low accuracy and high variability. This poor performance is likely due to the well-known sensitivity of distance-based methods to the high dimensionality of feature spaces—a phenomenon often referred to as the "curse of dimensionality."

An interesting observation is that $k = 1$ consistently yielded the best results for the kNN classifier (see Table 2). While this might appear counterintuitive—since smaller k values typically increase sensitivity to noise—the structure of the Twitter corpus provides an explanation. As illustrated in Table 1, many authors in the dataset exhibit highly formulaic writing patterns (e.g., automated "Now playing..." posts or templated promotional messages). This creates a situation where each author's texts have many close neighbors within their own collection, making the single nearest neighbor a surprisingly effective choice. This phenomenon bears some resemblance to the "overfitting" effects discussed in [26], though in this case the simplicity of the 1-*NN* rule proves advantageous given the repetitive nature of the data.

4.1. Conclusion

This study investigated authorship attribution for short texts under training data degradation, combining two challenging conditions: random character-level distortions and reduction in corpus size. Using a Twitter corpus of 50 authors with 1,000 texts each, training data were varied from 800 to

Table 2

Single division error with leading monomial weight of 95%

	800	600	400	200	50
			Err=0%		
SVM	0.92	0.91	0.88	0.83	0.65
LR	0.9	0.88	0.85	0.78	0.58
NB	0.89	0.87	0.83	0.76	0.53
RF	0.88	0.86	0.82	0.75	0.57
DT	0.84	0.82	0.78	0.71	0.53
kNN	0.72	0.7	0.67	0.63	0.51
			Err=5%		
SVM	0.89	0.87	0.83	0.77	0.58
LR	0.87	0.84	0.8	0.72	0.53
NB	0.86	0.84	0.79	0.71	0.49
RF	0.84	0.81	0.76	0.7	0.51
DT	0.81	0.78	0.73	0.65	0.47
kNN	0.65	0.63	0.6	0.56	0.46
			Err=20%		
SVM	0.75	0.71	0.64	0.54	0.37
LR	0.74	0.7	0.62	0.52	0.32
NB	0.74	0.7	0.63	0.52	0.33
RF	0.71	0.67	0.59	0.48	0.3
DT	0.7	0.65	0.57	0.46	0.28
kNN	0.42	0.39	0.37	0.32	0.25
	BOW	TFIDF	Ngram (1,2)	TFIDF (1,3)	TFIDF (1,4)

50 texts per author across six distortion levels (0–20%), while test texts remained unchanged. Six classifiers—SVM, Logistic Regression, Naive Bayes, Random Forest, Decision Tree, and kNN—were evaluated using BOW and TF-IDF features on various N-gram combinations. With the exception of kNN, all classifiers exhibited comparable accuracy degradation patterns, and those performing well on complete data maintained their advantage under degraded conditions. SVM with TF-IDF features proved most robust across all training set sizes and distortion levels, achieving accuracies of 0.92, 0.89, and 0.75 at 0, 5, and 20% distortion respectively, followed closely by Logistic Regression with BOW (0.90, 0.87, and 0.74). Decision Tree and Random Forest showed greater vulnerability to data reduction and distortion, while kNN yielded the poorest and most unstable results, likely due to the high dimensionality of the feature space. Feature sets including unigrams consistently outperformed those based solely on higher-order N-grams, underscoring the importance of token-level information

for attribution of short, noisy texts. These findings quantify the robustness of common attribution methods and identify SVM with TF-IDF as the most reliable choice when training data quality cannot be guaranteed.

Author Contributions: Conceptualization, Melnikov S.Yu.; methodology, Melnikov S.Yu.; software, Khvostenko V.M.; validation, Kovalenko A.P.; formal analysis, Meshcheryakov R.V.; research, Khvostenko V.M.; resources, Kovalenko A.P.; writing—initial draft preparation, Khvostenko V.M.; writing—review and editing, Melnikov S.Yu.; visualization, Khvostenko V.M.; supervision, Meshcheryakov R.V.; project administration, Melnikov S.Yu.; funding acquisition, Meshcheryakov R.V. All authors have read and agreed with the published version of the manuscript.

Funding: The study was supported by the Russian Science Foundation grant No. 24-11-00340 (https://rscf.ru/project/24-11-00340/).

Data Availability Statement: Data sharing is not applicable.

Conflicts of Interest: The author declares no conflict of interest.

Declaration on Generative AI: The author has not employed any generative AI tools.

References

[1] M. Eder, "Does size matter? Authorship attribution, small samples, big problem," *Digital Scholarship in the Humanities*, vol. 30, pp. 167–182, 2015. DOI: 10.1093/llc/fqt065

[2] G. Franzini, M. Kestemont, G. Rotari, M. Jander, J. Ochab, J. Byszuk, and M. Eder, "Attributing authorship in the noisy digitized correspondence of Jacob and Wilhelm Grimm," *Frontiers in Digital Humanities*, vol. 5, p. 4, 2018. DOI: 10.3389/fdigh.2018.00004

[3] K. Luyckx and W. Daelemans, "Authorship attribution and verification with many authors and limited data," in *Proceedings of the 22nd International Conference on Computational Linguistics (COLING 2008)*, 2008, pp. 513–520. DOI: 10.3115/1599081.1599147

[4] D. Lopresti, "Optical character recognition errors and their effects on natural language processing," in *Proceedings of the 3rd Workshop on Analytics for Noisy Unstructured Text Data*, 2009, pp. 115–122. DOI: 10.1145/1568293.1568312

[5] M. Kubis, P. Szymański, M. Ziółko, and K. Wróbel, "Open challenge for correcting errors of speech recognition systems," in *Language and Technology Conference*, 2019, pp. 322–337. DOI: 10.1007/978-3-030-33241-5_27

[6] A. A. Goncharov, N. V. Buntman, and V. A. Nuriyev, "Errors in machine translation: Classification problems," *Sistemy i Sredstva Informatiki*, vol. 29, no. 3, pp. 92–103, 2019, (in Russian). DOI: 10.14357/08696527190308

[7] E. Dawson and L. Nielsen, "Automated cryptanalysis of XOR plaintext strings," *Cryptologia*, vol. 20, no. 2, pp. 165–181, 1996. DOI: 10.1080/0161-119691884871

[8] L. V. Subramaniam, S. Roy, T. A. Faruquie, and S. Negi, "A survey of types of text noise and techniques to handle noisy text," in *Proceedings of the 3rd Workshop on Analytics for Noisy Unstructured Text Data*, 2009, pp. 115–122. DOI: 10.1145/1568293.1568312

[9] G. Sperduti and A. Moreo, *Misspellings in natural language processing: A survey*, 2025. arXiv: 2501.16836.

[10] H. Jing, D. Lopresti, and C. Shih, "Summarizing noisy documents," in *Proceedings of the Symposium on Document Image Understanding Technology*, 2003, pp. 111–119.

[11] N. Moratanch and S. Chitrakala, "A survey on extractive text summarization," in *International Conference on Computer, Communication and Signal Processing (ICCCSP)*, 2017, pp. 1–6. DOI: 10.1109/ICCCSP.2017.7944061

[12] D. D. Walker, W. B. Lund, and E. K. Ringger, "Evaluating models of latent document semantics in the presence of OCR errors," in *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, 2010, pp. 1–10.

[13] E. Stamatatos, "On the robustness of authorship attribution based on character n-gram features," *Journal of Law and Policy*, vol. 21, p. 7, 2013.

[14] M. Eder, "Mind your corpus: Systematic errors in authorship attribution," *Literary and Linguistic Computing*, vol. 28, no. 4, pp. 603–614, 2013. DOI: 10.1093/llc/fqt039

[15] N. Mamaev, X. Piotrowska, M. Marusenko, and A. Ronzhin, "Burrows's delta for authorship attribution of Russian literary texts," in *CEUR Workshop Proceedings*, vol. 2233, 2017.

[16] H. Sayoud et al., "Automatic authorship attribution of noisy documents," in *Proceedings of the FLAIRS Conference*, 2017, pp. 202–205.

[17] H. Benzerroug and S. Khennouf, "Author identification of corrupted OCR-based texts," *HDSKD Journal*, vol. 3, pp. 91–99, 2017.

[18] Z. Hamadache and H. Sayoud, "Authorship attribution of noisy text data with a comparative study of clustering methods," *International Journal of Knowledge and Systems Science*, vol. 9, no. 2, pp. 45–69, 2018. DOI: 10.4018/IJKSS.2018040103

[19] N. Bindal, P. Singh, V. Singh, and D. Gupta, "A systematic review of state-of-the-art noise removal techniques in digital images," *Multimedia Tools and Applications*, vol. 81, pp. 31 529–31 552, 2022. DOI: 10.1007/s11042-022-12980-7

[20] S. Colutto, P. Kahle, S. Guenter, and G. Muehlberger, "Transkribus: A platform for automated text recognition and searching of historical documents," in *15th International Conference on eScience (eScience)*, 2019, pp. 463–466. DOI: 10.1109/eScience.2019.00073

[21] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[22] R. Schwartz, O. Tsur, A. Rappoport, and M. Koppel, "Authorship attribution of micro-messages," in *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, 2013, pp. 1880–1891.

[23] C. Suman, S. Saha, and P. Bhattacharyya, "Authorship attribution of microtext using capsule networks," *IEEE Transactions on Computational Social Systems*, vol. 9, no. 4, pp. 1038–1047, 2021. DOI: 10.1109/TCSS.2021.3073690

[24] T. Alsanoosy, B. Shalbi, and A. Noor, "Authorship attribution for English short texts," *Engineering Technology and Applied Science Research*, vol. 14, no. 5, pp. 16 419–16 426, 2024. DOI: 10.48084/etasr.7794

[25] V. M. Khvostenko et al., "Two-parameter model of synthetic distortions in the problem of assessing the readability of distorted texts," *The European Physical Journal Special Topics*, vol. 234, pp. 3865–3870, 2025. DOI: 10.1140/epjs/s11734-025-01567-8

[26] K. P. Murphy, *Probabilistic Machine Learning: An Introduction*. MIT Press, 2022.

Information about the authors

Khvostenko, Viktor M.—Researcher of Cybersecurity CPS Department of HSE University (e-mail: vkhvostenko@hse.ru, ORCID: 0009-0000-1985-1401)

Kovalenko, Andrey P.—Professor, Doctor of Technical Sciences, Full member of Academy of Cryptography of the Russian (e-mail: kovalenko.ap@cryptoacademy.gov.ru, ORCID: 0009-0007-8777-8622)

Melnikov, Sergey Yu.—Doctor of Physical and Mathematical Sciences, Professor of Department of Probability Theory and Cyber Security of RUDN University; Chief Researcher of Cybersecurity CPS Department of HSE University (e-mail: melnikov-syu@rudn.ru, ORCID: 0000-0002-9023-9896)

Meshcheryakov, Roman V.—Professor, Doctor of Technical Sciences, Chief Researcher of Cybersecurity CPS Department of HSE University; Head of the Laboratory, Institute of Control Science of RAS (e-mail: mrv@ieee.org, ORCID: 0000-0002-1129-8434)

RUSSIAN METADATA

УДК 519.7
PACS 07.05.Tp
DOI: 10.22363/2658-4670-2026-34-2-160-174 EDN: JHMSMG

Моделирование атрибуции авторства коротких текстов в условиях деградации обучающего корпуса

В. М. Хвостенко[1], А. П. Коваленко[2], С. Ю. Мельников[1,3], Р. В. Мещеряков[1,4]

[1] *Национальный исследовательский университет «Высшая школа экономики», Покровский бульвар, д. 11, Москва, 109028, Российская Федерация*

[2] *Академия криптографии Российской Федерации, а/я 100, Москва, 119331, Российская Федерация*

[3] *Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация*

[4] *Институт проблем управления им. В. А. Трапезникова Российской академии наук, ул. Профсоюзная, д. 65, стр. 2, Москва, 117342, Российская Федерация*

Аннотация. Точность атрибуции авторства текстов критически зависит от качества и объема обучающих данных. Изучается, как два типа деградации обучающих авторских коллекций, искажение текста и сокращение размера корпуса, влияют на точность шести стандартных классификаторов при атрибуции авторства англоязычных твитов. Рассматриваются случайные независимые искажения типа «замена символов», применяемые к текстам в обучающих корпусах. Тестовая выборка остается неизменной. Эксперименты проводились на текстах 50 авторов (по 200 тестовых текстов). Количество обучающих текстов одного автора варьировалось от 800 до 50 (16 промежуточных шагов), а уровень искажения – от 0 до 20% (6 градаций). Для каждой комбинации размера корпуса и уровня искажения обучались и тестировались шесть классификаторов: метод опорных векторов (SVM), логистическую регрессию, наивный байесовский классификатор, случайный лес, дерево решений и метод k-ближайших соседей (kNN). В качестве признаков авторского стиля использовались мешок слов (BoW) и TF-IDF для токенов и N-грамм. Установлено, что для большинства классификаторов деградация обучающих данных приводит к сопоставимому снижению точности атрибуции. Наилучшие результаты показал метод опорных векторов с TF-IDF. Он оказался наиболее устойчивым как при сокращении количества обучающих текстов, так и при высоких уровнях искажений. В отсутствие искажений метод SVM с TF-IDF достиг точности 0.92; при уровнях искажений 5 и 20% его точность составила 0.89 и 0.75 соответственно. Логистическая регрессия с с мешком слов заняла второе место, показав точность 0.9, 0.87 и 0.74 в тех же условиях. Метод kNN продемонстрировал наименьшую точность среди рассмотренных классификаторов.

Ключевые слова: атрибуция авторства, короткие тексты, ухудшение качества обучающего корпуса, искажение текста, производительность классификатора

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2026, 34 (2) 175–186
http://journals.rudn.ru/miph

Research article

UDC 519.872
PACS 02.50.Ga
DOI: 10.22363/2658-4670-2026-34-2-175-186 EDN: JEHUQU

A queueing-inventory model for energy-constrained UAV base stations in non-terrestrial networks

Alexander E. Askerov[1], Sergey D. Bogolyubov[1], Kseniia A. Leonteva[1], Anastasia S. Vlaskina[1], Irina A. Kochetkova[1,2]

[1] *RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation*

[2] *Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, 44-2 Vavilova St, Moscow 119333, Russian Federation*

(received: March 25, 2026; revised: April 25, 2026; accepted: April 30, 2026)

Abstract. Unmanned aerial vehicle base stations are envisioned as key enablers of sixth-generation non-terrestrial networks, offering on-demand aerial coverage for remote areas, mass events, and emergency scenarios; however, their operational continuity is fundamentally limited by finite onboard battery capacity, a constraint that existing analytical models address inadequately by either ignoring energy dynamics or treating battery charge as a static parameter. The purpose of this study is to develop an analytical framework that jointly captures stochastic user service and discrete battery dynamics under a controlled recharging discipline. We propose a queueing-inventory model in which the unmanned aerial vehicle base station is represented as a loss system with a discrete energy resource and a threshold-based recharging policy; users arrive according to a Poisson process, service times are exponentially distributed, each served connection consumes one energy unit, and recharging is triggered when the battery level reaches a critical threshold. The system is formulated as a two-dimensional continuous-time Markov chain, and explicit expressions are obtained for the mean number of active users, the mean battery reserve, and the user blocking probability. Numerical experiments across three operational scenarios — remote area monitoring, mass event coverage, and emergency response — reveal that the threshold parameter governs a fundamental trade-off between user throughput and recharging overhead, providing a tractable tool for selecting unmanned aerial vehicle specifications at the design stage of sixth-generation communication systems.

Key words and phrases: queueing-inventory system, loss queueing system, unmanned aerial vehicle base station, non-terrestrial network, threshold-based recharging, discrete energy resource, continuous-time Markov chain, blocking probability, sixth-generation networks

For citation: A. E. Askerov, S. D. Bogolyubov, K. A. Leonteva, A. S. Vlaskina, I. A. Kochetkova "A queueing-inventory model for energy-constrained UAV base stations in non-terrestrial networks," *Discrete and Continuous Models and Applied Computational Science*, vol. 34, no. 2, pp. 175–186, 2026. DOI: 10.22363/2658-4670-2026-34-2-175-186. EDN: JEHUQU.

1. Introduction

Non-terrestrial networks are recognized as a critical component of next-generation communication infrastructure, enabling the integration of space-borne and airborne platforms into terrestrial networks to provide ubiquitous connectivity [1]. Within this framework, unmanned aerial vehicles (UAVs) have emerged as key elements due to their mobility, deployment flexibility, low-altitude operation providing line-of-sight connectivity, and relatively low operational costs [2]. However, line-of-sight availability is not guaranteed in dense urban environments, where ground and rooftop-mounted obstacles introduce non-trivial blockage effects that must be accounted for in deployment planning [3]. The transition to sixth-generation (6G) networks necessitates a fundamental change of perspective on non-terrestrial architectures, positioning UAV-based platforms as indispensable components of the future communication landscape [4]. Indeed, UAVs are now recognized alongside terahertz communication and intelligent reconfigurable surfaces as key enabling technologies for 6G wireless networks [5]. The integration of UAVs with 6G networks unlocks new possibilities for applications ranging from real-time environmental monitoring and cargo delivery to emergency communications and secure relay services [6].

Despite these promising prospects, UAV adoption faces a fundamental challenge: limited energy autonomy. The operational continuity of UAVs is strictly constrained by the finite capacity of their onboard batteries. Flight times of modern UAVs typically range from 10 to 60 minutes due to size, weight, and power limitations [7]. Simply increasing battery capacity creates a paradoxical effect — larger batteries increase weight, requiring more energy for lift and potentially degrading aerodynamic performance [8]. When serving as mobile radio access nodes, additional payload equipment further increases energy consumption, reducing operational time. In remote deployment scenarios lacking ground infrastructure, UAVs must travel significant distances between service areas and recharging stations before battery depletion [9].

These constraints necessitate analytical frameworks that capture the stochastic interaction between user traffic, energy consumption, and recharging discipline. Existing queueing models for UAV networks typically abstract away energy constraints or treat battery charge as a static parameter, failing to capture the dynamic interplay between stochastic service and battery replenishment. Conversely, energy consumption models for UAVs lack the stochastic service perspective necessary for quality-of-service analysis in 6G scenarios. While substantial research has addressed each stream independently, as reviewed in Section 2, their synthesis remains underexplored.

To address this gap, we adopt the formalism of queueing-inventory systems, which provides a unified mathematical framework for jointly modeling service processes and resource replenishment [10]. The UAV battery charge is treated as a discrete resource consumed by stochastic user connections and replenished exclusively through controlled charging from a ground station. The main contributions of our study are as follows:

- We propose a queueing-inventory model for a UAV base station that jointly captures stochastic user arrivals and discrete battery dynamics under a threshold-based recharging policy.
- We formulate the system as a two-dimensional continuous-time Markov chain and derive explicit expressions for three key performance indicators: the mean number of active users, the mean battery reserve, and the user blocking probability.
- We conduct a numerical study across three representative operational scenarios — remote area monitoring, mass event coverage, and emergency response — and quantify the fundamental trade-off between user throughput and recharging overhead governed by the critical threshold parameter.

The remainder of this paper is organized as follows. Section 2 provides a concise review of related work on UAV energy modeling and queueing theory applications. Section 3 describes the system model, including the network scenario, user arrival process, and energy management policy. Section 4 presents the analytical model, where we define the state space, transition rates, and global balance equations, and derive key performance measures. Section 5 contains the numerical results and discussion for three representative operational scenarios. The final section concludes the paper and provides an outlook for further investigations.

2. Related work

This section reviews three streams of related work directly relevant to the proposed model: UAV energy modeling, queueing-theoretic analysis of UAV networks, and queueing-inventory frameworks for resource-constrained systems.

Energy constraints remain a primary operational limitation for UAV systems. Beyond the fundamental size, weight, and power trade-offs [7], propulsion efficiency varies significantly with flight parameters, and payload configuration substantially impacts overall consumption patterns [8]. Empirical models based on battery discharge experiments have established quantitative relationships between flight characteristics and energy consumption rates [11], while generalized propulsion models capture the nonlinear effects of velocity and acceleration on energy usage [8]. Energy-aware path planning has been shown to extend operational time by 15–30% compared to baseline approaches, though gains depend heavily on environmental conditions [9]. Radio-frequency energy harvesting techniques have also been explored as a means of sustaining UAV operations [12], yet such solutions remain limited by low power density and environmental variability. Trajectory optimization combining metaheuristic and reinforcement learning methods has demonstrated adaptive energy management under dynamic conditions [13]. More recently, privacy-preserving federated learning frameworks have been applied to UAV path optimization, enabling autonomous data-driven route planning without centralized data sharing [14]. Diffusion approximation has been applied to model stochastic energy changes in harvesting batteries [15], offering a complementary perspective to the discrete-state approach adopted in the present work.

Queueing-theoretic approaches have been employed to analyze UAV network performance under stochastic traffic conditions. The interaction between signal interference and queue dynamics in unlicensed spectrum has been studied, yielding models for distributed transmission control that account for both channel conditions and buffer states [16, 17]. Digital twin architectures integrated with queueing analysis have shown potential for real-time performance optimization [18]. For multi-UAV systems, Markov models capturing controlled degradation strategies in swarm delivery scenarios have been developed [19], virtual coordinate-based intra-swarm routing protocols have been proposed for GPS-denied environments [20], and deep reinforcement learning has been applied to dynamic resource allocation under quality-of-service constraints [21]. Cognitive communication frameworks with optimized downlink power control have also been investigated [22]. Joint resource allocation and path planning in mobile edge computing scenarios have been examined to minimize latency and energy expenditure simultaneously [23]. While these works address various aspects of UAV network performance, they do not incorporate discrete energy inventory dynamics into the queueing model.

The queueing-inventory framework simultaneously captures service processes and resource dynamics, making it a natural fit for systems in which service capacity depends on a replenishable resource. A comprehensive survey of inventory systems with positive service time established the theoretical foundations for this class of models [10]. The applicability of the framework to energy-constrained systems has been demonstrated through analysis of queues with controlled

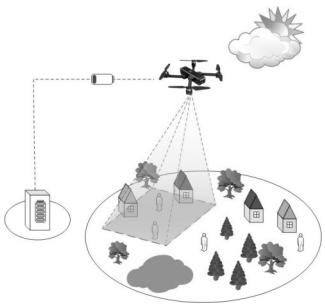

Figure 1. Schematic representation of the UAV base station system

replenishment under general input processes [24]. Queueing-inventory systems under catastrophic inventory depletion and various replenishment policies have been investigated [25], illustrating the versatility of inventory-based models for capturing complex resource dynamics.

Despite this body of work, no existing model jointly addresses stochastic user traffic and discrete battery dynamics under a threshold-driven recharging policy in a UAV base station context. This gap motivates the queueing-inventory formulation proposed in the present paper.

3. System model

The system under consideration is a UAV-mounted base station deployed in environments with limited or absent ground infrastructure, such as remote areas, mass public events, or emergency response zones. The UAV provides wireless connectivity to users in its coverage area while operating autonomously on battery power. The system comprises three main components: a rechargeable battery that supplies energy for user service, a transceiver unit that handles incoming user connections, and a controller that governs system operation by monitoring both the battery level and the number of active connections. The transceiver unit can handle up to N user connections simultaneously, reflecting the radio resource limitations of the UAV. The battery has a finite capacity divided into K identical energy slots, where each slot represents the energy required to serve one user connection to completion. The critical threshold s ($0 \leqslant s < K$) defines the minimum battery level at which the UAV must cease serving users and initiate recharging. A schematic representation of the system architecture is shown in Figure 1.

Users arrive at the coverage area and request connection to the base station according to a Poisson process with intensity λ. If the number of active connections is less than N and the battery level is above the threshold s, the arriving user is accepted and service begins immediately; otherwise the user is blocked and lost. The system thus operates as a loss system without waiting. Each completed connection consumes exactly one energy slot, causing the battery level to decrease by one unit upon service completion. Service times are independent and exponentially distributed with rate μ.

The controller continuously monitors the battery level. When the charge drops to the critical threshold s, the UAV ceases to accept new users, suspends all ongoing connections, and proceeds to the recharging station, which is assumed to be always available. The recharging process restores the battery from level s back to the maximum capacity K, with recharging time exponentially distributed with rate ν. The UAV remains completely unavailable for user service during the entire recharging cycle and resumes normal operation only after the battery reaches level K. The notation used throughout this paper is summarized in Table 1.

Table 1

Main notation

Parameter	Description
N	Maximum number of simultaneous user connections
K	Maximum battery capacity (energy slots)
s	Critical battery threshold ($0 \leqslant s < K$)
λ	User arrival rate (Poisson process intensity)
μ	Per-connection service rate (exponential service time)
ν	Battery replenishment rate (exponential recharging time)
n	Current number of active connections
k	Current battery level (energy slots)

4. Queueing-inventory model

The state of the system at time t is described by a two-dimensional stochastic process $\{\tilde{n}(t), \tilde{k}(t)\}$, where $\tilde{n}(t)$ is the number of active user connections ($0 \leqslant \tilde{n}(t) \leqslant N$) and $\tilde{k}(t)$ is the current battery level ($s \leqslant \tilde{k}(t) \leqslant K$). The system operates in two modes: the serving mode, in which the UAV accepts user connections with battery level above the threshold s, and the recharging mode, in which the UAV is at the charging station with battery level s and no connections are served. The state space is accordingly defined as:

$$\mathcal{X} = \{(n, s) : 0 \leqslant n \leqslant N-1\} \cup \{(n, k) : 0 \leqslant n \leqslant N, s < k \leqslant K\}.$$

States with $k < s$ are unreachable, since the battery never drops below the recharging threshold. The state (N, s) is also unreachable: the battery reaches level s only upon a service completion, which simultaneously frees one connection slot, so the system transitions to $(N-1, s)$ rather than (N, s). Consequently, all recharging states satisfy $0 \leqslant n \leqslant N-1$. A schematic representation of the model is shown in Figure 2.

The system dynamics are governed by three types of transitions, as illustrated in the transition intensity diagram (Figure 3):

- User arrival: from state (n, k) to $(n+1, k)$ at rate λ, when $n < N$ and $k > s$.
- Service completion: from state (n, k) to $(n-1, k-1)$ at rate $n\mu$, when $n > 0$ and $k > s$. If the resulting battery level equals s, the system enters recharging state $(n-1, s)$.
- Recharging completion: from state (n, s) to (n, K) at rate ν, for $0 \leqslant n \leqslant N-1$.

The nonzero elements of the transition intensity matrix Q are:

$$\begin{aligned}
&Q[(n, s),\ (n, K)] = \nu, \quad 0 \leqslant n \leqslant N-1, \\
&Q[(n, k),\ (n+1, k)] = \lambda, \quad 0 \leqslant n \leqslant N-1, s+1 \leqslant k \leqslant K, \\
&Q[(n, k),\ (n-1, k-1)] = n\mu, \quad 1 \leqslant n \leqslant N, s+1 \leqslant k \leqslant K.
\end{aligned}$$

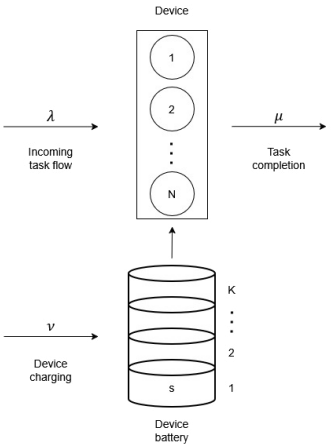

Figure 2. Queueing-inventory model

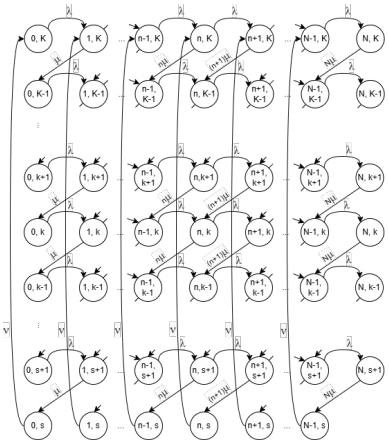

Figure 3. Transition intensity diagram

Let $\pi(n,k) = \lim_{t\to\infty} \mathrm{P}\{\tilde{n}(t) = n, \tilde{k}(t) = k\}$ denote the stationary probability of state (n,k). The global balance equations are:

$$\begin{gathered}
\nu\pi(0,s) = \mu\pi(1,s+1),\\
\nu\pi(n,s) = (n+1)\mu\pi(n+1,s+1), \quad n = 1,\dots,N-1,\\
\lambda\pi(0,k) = \mu\pi(1,k+1), \quad k = s+1,\dots,K-1,\\
(\lambda+n\mu)\pi(n,k) = \lambda\pi(n-1,k) + (n+1)\mu\pi(n+1,k+1), \quad n = 1,\dots,N-1, k = s+1,\dots,K-1,\\
N\mu\pi(N,k) = \lambda\pi(N-1,k), \quad k = s+1,\dots,K,\\
\lambda\pi(0,K) = \nu\pi(0,s),\\
(\lambda+n\mu)\pi(n,K) = \nu\pi(n,s) + \lambda\pi(n-1,K), \quad n = 1,\dots,N-1,
\end{gathered}$$

supplemented by the normalization condition $\sum_{(n,k)\in\mathcal{X}} \pi(n,k) = 1$. The stationary distribution is obtained numerically by solving this linear system.

Based on the stationary distribution, three key performance metrics are defined. The mean number of active connections is:

$$L = \sum_{n=0}^{N}\sum_{k=s+1}^{K} n\pi(n,k) + \sum_{n=0}^{N-1} n\pi(n,s).$$

The mean battery reserve is:

$$Y_{\text{avg}} = \sum_{n=0}^{N}\sum_{k=s+1}^{K} k\pi(n,k) + \sum_{n=0}^{N-1} s\pi(n,s).$$

The user blocking probability accounts for two sources of blocking — energy depletion and capacity saturation:

$$P_{\text{block}} = \underbrace{\sum_{n=0}^{N-1}\pi(n,s)}_{\text{energy blocking}} + \underbrace{\sum_{k=s+1}^{K}\pi(N,k)}_{\text{capacity blocking}}.$$

These performance metrics are evaluated numerically in Section 5 for varying system parameters N, K, s, λ, μ, and ν.

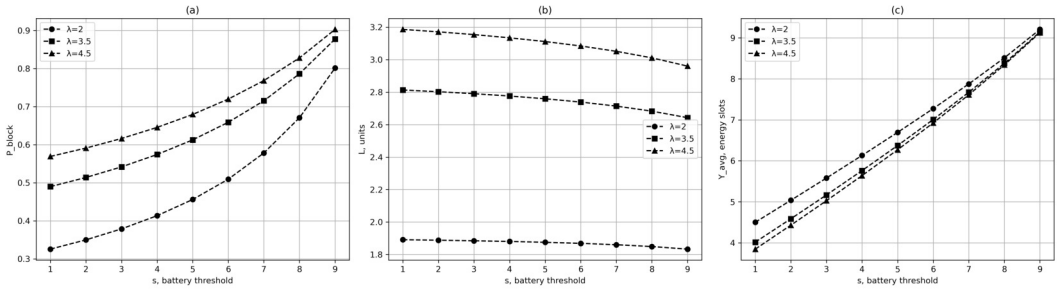

Figure 4. Performance metrics as functions of critical battery threshold s for varying arrival intensities λ: (a) user blocking probability P_{block}, (b) mean number of active connections L, (c) mean battery reserve Y_{avg}

5. Numerical results

Having obtained the stationary distribution $\pi(n,k)$ for all states $(n,k) \in \mathcal{X}$, we evaluate the three performance indicators introduced in Section 4: the mean number of active connections L, the mean battery reserve Y_{avg}, and the user blocking probability P_{block}. The numerical analysis proceeds in two stages. First, we examine the sensitivity of all three metrics to the critical threshold s for varying arrival intensities. Second, we evaluate system behavior across three representative operational scenarios that reflect distinct real-world deployment conditions.

Figure 4 illustrates the dependence of P_{block}, L, and Y_{avg} on the critical threshold s for three values of the arrival intensity ($\lambda = 2,\ 3.5,\ 4.5$). The blocking probability increases monotonically with both s and λ: higher values of s trigger recharging cycles more frequently, reducing the time the station is available for user service, and this effect is amplified under heavier traffic. The mean number of active connections remains relatively stable at low λ but decreases slightly as s increases at higher loads, reflecting the growing fraction of time spent in the recharging mode. The mean battery reserve Y_{avg} grows nearly linearly with s across all arrival intensities, confirming that it is primarily determined by the threshold rather than by traffic load. Collectively, these results demonstrate that s governs a fundamental trade-off: conservative thresholds preserve battery reserves and reduce the risk of energy depletion, while aggressive thresholds sustain higher throughput at the cost of increased blocking.

To evaluate system performance under realistic deployment conditions, we consider three scenarios summarized in Table 2, each characterized by distinct traffic intensity, radio resource capacity, battery capacity, and recharging rate. The ratio K/N reflects the energy-to-capacity balance, while the product $K \times \nu$ characterizes the overall energy throughput of the recharging infrastructure.

In the remote area scenario, the UAV operates under severe resource constraints: a very small number of simultaneous connections ($N = 2$–3), a modest battery ($K = 5$–6), and extremely slow based recharging ($\nu = 0.05$–0.10). Because the system spends most of its time in the recharging mode, the blocking probability is the highest among all three scenarios and remains nearly constant at approximately 0.97 as λ increases, while both the mean number of active connections and the mean battery reserve stay at very low levels. In the mass event scenario, the system has substantially more radio resources ($N = 12$–20), a larger battery ($K = 25$–35), and moderate recharging capability ($\nu = 0.30$–0.60). The blocking probability grows from approximately 0.40 to 0.72 as λ increases, reflecting progressive capacity saturation under heavy traffic, while the mean battery reserve decreases steadily due to continuous high-rate energy consumption. In the rescue and emergency scenario, all system parameters are scaled to their maximum values ($N = 20$–30, $K = 40$–60,

Table 2

System parameters for three operational scenarios

Scenario	λ	N	K	K/N	ν	$K \times \nu$
(A) Remote area	0.5–1	2–3	5–6	2.0–2.5	0.05–0.10	0.3–0.5
(B) Mass events	10–30	12–20	25–35	1.5–2.0	0.30–0.60	10–15
(C) Rescue and emergency	15–25	20–30	40–60	1.3–1.7	0.80–1.20	35–60

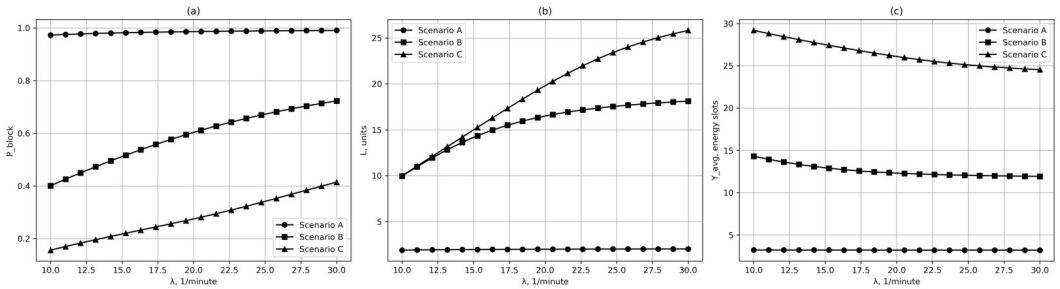

Figure 5. Performance metrics as functions of arrival intensity λ for three operational scenarios: (a) user blocking probability P_{block}, (b) mean number of active connections L, (c) mean battery reserve Y_{avg}

$\nu = 0.80$–1.20) to maximize availability. This generous provisioning yields the lowest blocking probability among the three scenarios, rising from approximately 0.15 to 0.40, and supports the highest mean number of active connections. However, the mean battery reserve exhibits the steepest decline, underscoring the intense energy demand imposed by sustained high-throughput operation.

Figure 5 presents P_{block}, L, and Y_{avg} as functions of λ for each scenario. The results confirm the qualitative distinctions described above and reveal how the interplay between traffic load, battery capacity, and replenishment rate shapes system performance across deployment contexts. The remote area scenario exhibits the highest blocking probability due to its severely limited resources and slow recharging, while the rescue and emergency scenario achieves the lowest blocking probability owing to its maximized capacity and rapid recharging infrastructure.

6. Conclusion

This paper proposed a queueing-inventory model for a UAV base station with a discrete energy resource and a threshold-based recharging policy. The system was formulated as a two-dimensional continuous-time Markov chain, yielding explicit expressions for the mean number of active connections, the mean battery reserve, and the blocking probability. Numerical analysis across three scenarios — remote area monitoring, mass event coverage, and emergency response — confirmed that the threshold parameter s governs a fundamental trade-off between user throughput and recharging overhead, providing a tractable tool for UAV specification selection at the design stage of 6G systems. Future work includes incorporating flight trajectory energy costs, multi-UAV cooperation with handover mechanisms, and heterogeneous traffic models such as batch arrivals or non-exponential service times.

Author Contributions: Conceptualization, supervision, Irina A. Kochetkova; methodology, writing—review and editing, Anastasia S. Vlaskina and Irina A. Kochetkova; software, validation, visualization, writing—original draft preparation, Alexander E. Askerov and Sergey D. Bogolyubov; investigation, Alexander E. Askerov, Sergey D. Bogolyubov, and Kseniia A. Leonteva. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Declaration on Generative AI: During the preparation of this work, the authors used Perplexity AI (Perplexity AI, Inc., 2026) in order to: grammar and spelling check, improve writing style. After using this tool, the authors reviewed and edited the resulting content as needed and take full responsibility for the content of the publication.

References

[1] 3GPP, "Study on New Radio (NR) to support non-terrestrial networks (Release 15)," 3rd Generation Partnership Project (3GPP), Technical Report TR 38.811 V15.4.0, 2020.

[2] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36–42, 2016. DOI: 10.1109/MCOM.2016.7470933

[3] V. Begishev, D. Moltchanov, A. Gaidamaka, and K. Samouylov, "Closed-Form UAV LoS Blockage Probability in Mixed Ground and Rooftop-Mounted Urban mmWave NR Deployments," *Sensors*, vol. 22, no. 3, 2022. DOI: 10.3390/s22030977

[4] M. Giordani and M. Zorzi, "Non-Terrestrial Networks in the 6G Era: Challenges and Opportunities," *IEEE Network*, vol. 35, no. 2, pp. 244–251, 2021. DOI: 10.1109/MNET.011.2000493

[5] W. M. Othman, A. A. Ateya, M. E. Nasr, A. Muthanna, M. ElAffendi, A. Koucheryavy, and A. A. Hamdi, "Key Enabling Technologies for 6G: The Role of UAVs, Terahertz Communication, and Intelligent Reconfigurable Surfaces in Shaping the Future of Wireless Networks," *Journal of Sensor and Actuator Networks*, vol. 14, no. 2, 2025. DOI: 10.3390/jsan14020030

[6] S. Li, B. Duo, X. Yuan, Y.-C. Liang, and M. DI Renzo, "Reconfigurable Intelligent Surface Assisted UAV Communication: Joint Trajectory Design and Passive Beamforming," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 716–720, 2020. DOI: 10.1109/LWC.2020.2966705

[7] L. Gupta, R. Jain, and G. Vaszkun, "Survey of Important Issues in UAV Communication Networks," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 2, pp. 1123–1152, 2016. DOI: 10.1109/COMST.2015.2495297

[8] X. Dai, B. Duo, X. Yuan, and W. Tang, "Energy-Efficient UAV Communications: A Generalized Propulsion Energy Consumption Model," *IEEE Wireless Communications Letters*, vol. 11, no. 10, pp. 2150–2154, 2022. DOI: 10.1109/LWC.2022.3195787

[9] 3GPP, "Enhancement for Unmanned Aerial Vehicles; Stage 1 (Release 17)," 3rd Generation Partnership Project (3GPP), Technical Report TR 22.829 V17.1.0, 2019.

[10] A. Krishnamoorthy, D. Shajin, and V. C. Narayanan, "Inventory with Positive Service Time: A Survey," in *Queueing Theory 2: Advanced Trends*, 2021, pp. 201–237. DOI: 10.1002/9781119755234.ch6

[11] H. V. Abeywickrama, B. A. Jayawickrama, Y. He, and E. Dutkiewicz, "Comprehensive energy consumption model for unmanned aerial vehicles, based on empirical studies of battery performance," *IEEE Access*, vol. 6, pp. 58 383–58 394, 2018. DOI: 10.1109/ACCESS.2018.2875040

[12] G. K. Pandey, D. S. Gurjar, S. Yadav, Y. Jiang, and C. Yuen, "UAV-Assisted Communications With RF Energy Harvesting: A Comprehensive Survey," *IEEE Communications Surveys and Tutorials*, vol. 27, no. 2, pp. 782–838, 2025. DOI: 10.1109/COMST.2024.3425597

[13] R. A. Saeed, E. S. Ali, M. Abdelhaq, R. Alsaqour, F. R. A. Ahmed, and A. M. E. Saad, "Energy Efficient Path Planning Scheme for Unmanned Aerial Vehicle Using Hybrid Generic Algorithm-Based Q-Learning Optimization," *IEEE Access*, vol. 12, pp. 13 400–13 417, 2024. DOI: 10.1109/ACCESS.2023.3344455

[14] W. Min, M. S. A. Muthanna, M. Ibrahim, R. Alkanhel, A. Muthanna, and A. Laouid, "Privacy-preserving federated UAV data collection framework for autonomous path optimization in maritime operations," *Applied Soft Computing*, vol. 173, 2025. DOI: 10.1016/j.asoc.2025.112906

[15] T. Czachórski, E. Gelenbe, and G. S. Kuaban, "Modelling Energy Changes in the Energy Harvesting Battery of an IoT Device," in *Proceedings - IEEE Computer Society's Annual International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunications Systems, MASCOTS*, vol. 2022-October, 2022, pp. 81–88. DOI: 10.1109/MASCOTS56607.2022.00019

[16] M. Ghazikor, K. Roach, K. Cheung, and M. Hashemi, "Exploring the Interplay of Interference and Queues in Unlicensed Spectrum Bands for UAV Networks," in *Conference Record - Asilomar Conference on Signals, Systems and Computers*, 2023, pp. 729–733. DOI: 10.1109/IEEECONF59524.2023.10476884

[17] M. Ghazikor, K. Roach, K. Cheung, and M. Hashemi, "Interference-Aware Queuing Analysis for Distributed Transmission Control in UAV Networks," in *IEEE International Conference on Communications*, 2024, pp. 4524–4529. DOI: 10.1109/ICC51166.2024.10623099

[18] M. A. B. S. Abir and M. Z. Chowdhury, "Digital Twin-based Software-defined UAV Networks Using Queuing Model," in *Proceedings of the 10th International Conference on Signal Processing and Integrated Networks, SPIN 2023*, 2023, pp. 479–483. DOI: 10.1109/SPIN57001.2023.10116319

[19] I. Kliushnikov, H. Fesenko, G. Fedorenro, S. Rudakov, V. Mikhalevskyi, and O. Kompaniiets, "Swarm of Unmanned Aerial Vehicles as a Multi-State Queueing System with Non-Controlled and Controlled Degradation," in *Proceedings of the 2022 IEEE 12th International Conference on Dependable Systems, Services and Technologies, DESSERT 2022*, 2022. DOI: 10.1109/DESSERT58054.2022.10018784

[20] Y. Gaidamaka and K. Samouylov, "UVCS: Unit Virtual Coordinate System for UAV Intra-Swarm Routing in GPS-Denied Environment," *Mathematics*, vol. 11, no. 3, 2023. DOI: 10.3390/math11030694

[21] B. Khamidehi, M. Raeis, and E. S. Sousa, "Dynamic Resource Management for Providing QoS in Drone Delivery Systems," in *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, vol. 2022-October, 2022, pp. 3529–3536. DOI: 10.1109/ITSC55140.2022.9922133

[22] A. R. Rahul, S. R. Sabuj, M. S. Akbar, H.-S. Jo, and M. A. Hossain, "An optimization based approach to enhance the throughput and energy efficiency for cognitive unmanned aerial vehicle networks," *Wireless Networks*, vol. 27, no. 1, pp. 475–493, 2021. DOI: 10.1007/s11276-020-02450-9

[23] M. Li, N. Cheng, J. Gao, Y. Wang, L. Zhao, and X. Shen, "Energy-Efficient UAV-Assisted Mobile Edge Computing: Resource Allocation and Trajectory Optimization," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 3424–3438, 2020. DOI: 10.1109/TVT.2020.2968343

[24] A. Dudin and V. Klimenok, "Analysis of MAP/G/1 queue with inventory as the model of the node of wireless sensor network with energy harvesting," *Annals of Operations Research*, vol. 331, no. 2, pp. 839–866, 2023. DOI: 10.1007/s10479-022-05036-0

[25] S. Ozkar, A. Melikov, and J. Sztrik, "Queueing-Inventory Systems with Catastrophes under Various Replenishment Policies," *Mathematics*, vol. 11, no. 23, 2023. DOI: 10.3390/math11234854

Information about the authors

Askerov, Alexander E.—Bachelor's Student with the Department of Probability Theory and Cyber Security of RUDN University (e-mail: 1132226538@rudn.ru, ORCID: 0009-0001-7688-2527)

Bogolyubov, Sergey D.—PhD Student with the Department of Probability Theory and Cyber Security of RUDN University (e-mail: 1142240051@rudn.ru, ORCID: 0009-0001-9740-2545)

Leonteva, Kseniia A.—Junior Researcher and PhD Student with the Department of Probability Theory and Cyber Security of RUDN University (e-mail: leontyeva-ka@rudn.ru, ORCID: 0009-0000-9253-3426, ResearcherID: LSI-9560-2024, Scopus Author ID: 58849959400)

Vlaskina, Anastasia S.—Candidate of Sciences in Physics and Mathematics, Senior Lecturer with the Department of Probability Theory and Cyber Security of RUDN University (e-mail: vlaskina-as@rudn.ru, ORCID: 0000-0001-6453-814X, ResearcherID: AAC-6696-2020, Scopus Author ID: 57204395118)

Kochetkova, Irina A.—Doctor of Sciences in Physics and Mathematics, Associate Professor with the Department of Probability Theory and Cyber Security of RUDN University; Senior Researcher with the Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences (e-mail: kochetkova-ia@rudn.ru, ORCID: 0000-0002-1594-427X, ResearcherID: E-3806-2014, Scopus Author ID: 35332169400)

RUSSIAN METADATA

УДК 519.872

PACS 02.50.Ga

DOI: 10.22363/2658-4670-2026-34-2-175-186 EDN: JEHUQU

Система массового обслуживания с запасами для беспилотных базовых станций с ограниченным энергоресурсом в неназемных сетях

А. Э. Аскеров[1], С. Д. Боголюбов[1], К. А. Леонтьева[1], А. С. Власкина[1], И. А. Кочеткова[1,2]

[1] *Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация*

[2] *Федеральный исследовательский центр «Информатика и управление» Российской академии наук, ул. Вавилова, д. 44, корп. 2, Москва, 119333, Российская Федерация*

Аннотация. Беспилотные базовые станции рассматриваются как ключевые элементы неназемных сетей шестого поколения, обеспечивающие мобильное покрытие в удалённых районах, при массовых мероприятиях и в чрезвычайных ситуациях; вместе с тем их операционная непрерывность принципиально ограничена конечной емкостью бортового аккумулятора — ограничение, которое существующие аналитические модели либо игнорируют, либо учитывают лишь статически. Целью исследования является разработка аналитической модели, совместно описывающей стохастическое обслуживание пользователей и дискретную динамику заряда батареи в условиях политики подзарядки. Предлагается система массового обслуживания с запасами, в которой беспилотная базовая станция представлена как система с потерями, дискретным энергетическим ресурсом и пороговой политикой подзарядки; пользователи поступают согласно пуассоновскому процессу, времена обслуживания экспоненциальны, каждое соединение потребляет одну единицу заряда, а подзарядка инициируется при достижении критического порога. Система формализована как двумерный марковский случайный процесс; получены выражения

для среднего числа активных пользователей, среднего уровня заряда батареи и вероятности блокировки. Численный анализ трёх сценариев — удаленная территория, массовое мероприятие и аварийное реагирование — показывает, что параметр порога определяет компромисс между пропускной способностью и накладными расходами на подзарядку, предоставляя инструмент для выбора характеристик беспилотных станций.

Ключевые слова: система массового обслуживания с запасами, система с потерями, беспилотная базовая станция, неназемная сеть, пороговая политика подзарядки, дискретный ресурс энергии, марковский случайный процесс, вероятность блокировки, сети шестого поколения

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2026, 34 (2) 187–200
http://journals.rudn.ru/miph

Research article

UDC 004.89, 004.056.57
PACS 07.05.Mh, 07.05.Kf
DOI: 10.22363/2658-4670-2026-34-2-187-200 EDN: JDISSJ

Development of the malicious traffic detection system for mobile applications security

Nikita D. Khalemskii[1], Vitalii A. Beschastnyi[1,2]

[1] *HSE University, 20 Myasnitskaya St, Moscow, 101000, Russian Federation*

[2] *RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation*

(received: March 29, 2026; revised: March 16, 2026; accepted: March 30, 2026)

Abstract. *Background:* The rapid increase in cyberattacks targeting infrastructure, enterprises, and individual users through mobile devices has created an urgent need for effective intrusion detection systems. Traditional signature-based methods are inadequate against zero-day vulnerabilities and advanced persistent threats, prompting the exploration of machine learning approaches for network traffic analysis. *Purpose*: This study aims to develop and evaluate a privacy-preserving, locally operating Android application for real-time malicious traffic detection using machine learning models, addressing the limitations of cloud-dependent security architectures that compromise user privacy and availability. *Methods:* We implemented five machine learning algorithms — XGBoost, LightGBM, Random Forest, Decision Tree, and Logistic Regression — trained on the Network Traffic Android Malware dataset. The models were exported to the ONNX format for local deployment on Android devices. Performance was evaluated using accuracy, precision, recall, AUC score, training time, and model size metrics, with multi-criteria decision analysis (MCDA) applied for comprehensive comparison. *Results:* Gradient boosting algorithms demonstrated superior performance, with LightGBM achieving the highest MCDA score (0.9759), fastest training time (0.32 s), and smallest model size (0.28 MB), while XGBoost attained the highest AUC-score (0.9627). Both models significantly outperformed Random Forest (MCDA: 0.7532), Decision Tree (0.6743), and Logistic Regression (0.1407). *Conclusions:* LightGBM provides an optimal balance between detection accuracy and mobile resource constraints, making it suitable for on-device deployment. The proposed architecture demonstrates that fully local, ML-based traffic analysis is feasible without compromising detection quality, offering a privacy-centric alternative to server-dependent solutions for next-generation mobile intrusion detection systems.

Key words and phrases: signature-based detection, machine learning, anomaly detection, mobile traffic monitoring, real-time data processing, cybersecurity.

For citation: N. D. Khalemskii, V. A. Beschastnyi "Development of the malicious traffic detection system for mobile applications security," *Discrete and Continuous Models and Applied Computational Science*, vol. 34, no. 2, pp. 187–200, 2026. DOI: 10.22363/2658-4670-2026-34-2-187-200. EDN: JDISSJ.

1. Introduction

Today, among consumer devices, smartphones are the most vulnerable to cyberattacks because they are actively and frequently used in almost all aspects of life. The types of malware used by attackers are becoming increasingly complex and unpredictable. Accordingly, a sufficient number of applications have been developed to protect users from these threats. Network traffic analysis is one of the most popular methods for detecting malware in applications. However, the architecture of such software is often based on sending user traffic data to a remote server, where scanning occurs.

Although in such architectures, a decent part of the working logic does not occur on the user's device, thereby not loading the device itself, one of the drawbacks of this architecture is that some user data are processed on a third-party server; therefore, users become highly dependent on the availability and status of the analyzing server. This study proposes the implementation of another application, where traffic analysis occurs completely locally on the user's device. Our approach to local processing aligns with emerging privacy-preserving paradigms, such as federated learning, which has recently been applied to Android malware detection [1] to enable model improvement without data sharing.

Initially, the analysis of network traffic was based on the signature method (comparison with a blacklist of known threats). However, owing to the increase in zero-day vulnerabilities and Advanced Persistent Threat (APT) attacks, this approach is no longer sufficient. Consequently, information security professionals are increasingly turning to machine learning (ML) techniques that enable behavioral analysis to identify anomalies and previously unknown threats. Thus, to develop an effective traffic analyzer for the Android operating system, there has been active research into the exploitation of modern machine learning models on modern datasets with the classification of traffic as malicious and non-malicious. Five ML models were compared: XGBoost [2], LightGBM [3], Random Forest [4], Decision Tree [5] and Logistic Regression [6]. For the purposes of this study on the Android operating system, there were considered Network Traffic Android Malware [7] to be the best option among all datasets. This dataset was selected because it captures the real-world network behavior of Android applications, allowing for the detection of malware that evades traditional static code analysis. The productivity of the models were evaluated via 6 metrics: accuracy, precision, recall, AUC-score, average training time and model size in Open Neural Network Exchange (ONNX) [8] format.

The rest of this document is organized as follows: Section 2 describes the architecture of IDS using traditional and machine learning methods. Then, innovations will be proposed for the implementation of the local operation of such systems on the user's Android smartphone. In Section 3, we observe the ML models and datasets described above more closely and select the ones that will produce the most effective results on the Android operating system (OS). Section 4 describes the preliminary results obtained using the developed solution. Finally, in Section 6 we summarize all the results and draw conclusions.

2. Architecture of ML-based malicious activity detection application

Owing to the rapid digitalization and increasing complexity of cyberattacks, securing network infrastructure has become a top priority for software developers. Accordingly, the search for malicious network activity in programs has become one of the main areas of activity for information security researchers. It requires a deep understanding of traffic filtering mechanisms [9] and the evolution of methods for identifying malicious activity, ranging from classic hard algorithms to adaptive and intelligent systems.

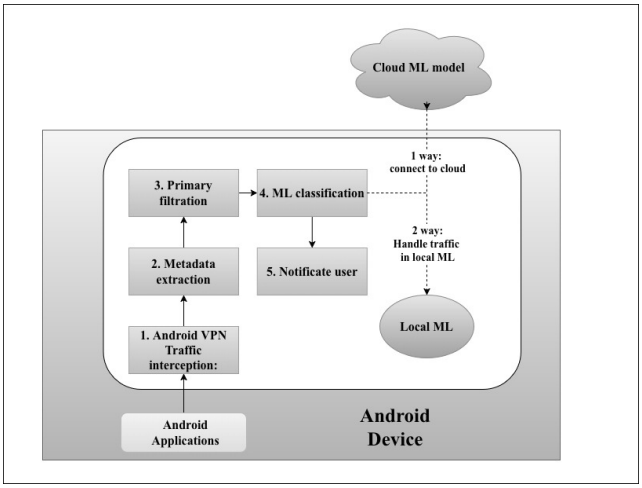

Figure 1. Generic architecture of ML-based traffic analyzer

Traditional approaches to detecting malware in network traffic, as used in intrusion detection systems (IDS), rely on the principle of strict data compliance with predefined rules or patterns. The key method here is signature analysis, which involves searching for data packets in databases of known threat "fingerprints" (hash sums or unique byte sequences). Deep packet analysis (DPI) and static filtering based on ports and protocols were also used. The main advantage of these approaches is their high processing speed and virtually zero false-positive rates for known threats. However, their fundamental flaw lies in their inability to detect attacks, such as zero-day, targeted (APT), and modified versions of malware that do not match existing signatures.

The integration of ML technologies into intrusion detection systems (IDS) has changed the security paradigm and is now considered a key technology for creating the next generation of IDS, which can identify unknown threats by analyzing statistical anomalies [10] and hidden patterns in traffic behavior. This allowed us to move from a pattern-searching approach to a deep analysis of the network traffic anomalies. The integration of supervised and unsupervised learning has enabled systems to classify suspicious activities based on statistical features such as payload entropy, inter-packet arrival times, and message length distribution. Currently, such methods are used in almost all devices and operating systems. However, the architecture differs depending on the application type. For example, the IDS architecture for Android is based on a hybrid model that combines local event monitoring (at the core and framework levels) with cloud analytics to manage complex threat patterns, as shown in Figure 1.

The process of monitoring network traffic on an Android device can be divided into five key stages:

1. Traffic interception: since Android does not give applications direct access to the network interface (promiscuous mode), the main method of interception is to use the VPNService API. The application creates a local VPN tunnel through which all outgoing and incoming traffic on the device passes.
2. Metadata extraction: the system aggregates data into "sessions", and then extracts metadata such as the volume of data transferred, connection duration, number of packets per second, etc.
3. Primary filtration: The mobile device checks domains and IP addresses against local blacklists. If an address is already known to be malicious, access is immediately blocked.

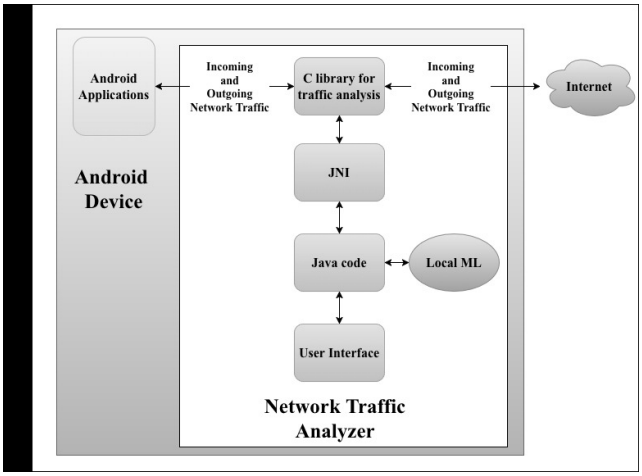

Figure 2. Proposed design of Android application for mobile security

4. ML classification: The extracted metadata is submitted to the machine learning model. Such models can work both locally and on remote servers (in this case, there is already a delayed check to avoid strong delays in the Internet operation on the device).
5. Action: if a threat is detected, the application notifies the user, terminates the connection, or sends a report to the security server.

As previously established, such applications are based on ML models that can operate both locally and on a remote server, analyzing metadata sent to them by the application for processing. In this study, we propose an application design for an Android real-time traffic analyzer. The feasibility of low-latency inference on mobile devices, recently validated by [11] using TSANet, supports our findings that lightweight ML models can operate effectively in real-time scenarios. Our application is based on purely local data processing, without interaction with any third-party server, to preserve user privacy. We based our application on the open-source development PcapDroid [12], which is known for being similar to Wireshark [13] but runs on Android devices (see Figure 2). At the root of this application is a library written in C language, which is responsible for the deep processing of network traffic. The outer shell is implemented in Java, and the interaction between these two components occurs through the Java Native Interface (JNI) [14]. As PcapDroid has its own internal paid functionality for activating the internal firewall, we removed it from our code (as we removed other paid components). To handle network traffic, we imported ML models in the ONNX format, which are run locally on the Android device and interact with the C-based library through JNI.

3. ML-based traffic detection approach

3.1. Machine Learning Models

3.1.1. Extreme Gradient Boosting

Extreme Gradient Boosting (XGBoost) is a scalable, high-performance implementation of a gradient boosting algorithm for decision trees. It is a standard for tabular data and has proven to be a reliable and powerful technology. Although the last release was in 2016, it often outperforms deep neural

networks in terms of the quality and speed of learning. As has been chosen for our research, as it is very effective on binary and multi-class classification tasks. Its advantages in searching for malicious network traffic are as follows:

- high prediction accuracy thanks to built-in L1 and L2 regularization, which prevents overfitting, and optimization based on second derivatives of the loss function;
- processing heterogeneous features: Network flow features can be numerical (duration, number of bytes), categorical (protocol flags), and binary. XGBoost does not require data scaling and works efficiently with mixed data types.
- built-in resilience that allows you to build robust generalization models even on noisy data.

The following disadvantages can be noted:

- model for fine tuning of hyperparameters;
- the complexity of interpreting the model, since the ensemble of trees is difficult to analyze compared to simpler models;

3.1.2. Light Gradient Boosting Machine

The newer Light Gradient Boosting Machine (LightGBM) model was chosen as a modern and highly efficient alternative to XGBoost. Its key advantage lies in the radical optimization of computational efficiency and training speed while maintaining or often surpassing predictive accuracy. This makes it extremely attractive for processing large volumes of network data in real time. Its main advantages are as follows:

- speed and memory efficiency: thanks to Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB) techniques, LightGBM is learning much faster, consuming less memory than XGBoost, especially on large datasets with tens of thousands of features,
- vertical tree growth: Unlike XGBoost and most other algorithms, which grow trees level-wise, LightGBM uses leaf-based growth (leaf-wise). This creates more asymmetric and deeper trees that achieve the same or greater accuracy with significantly fewer nodes;
- support for categorical features: LightGBM can directly process categorical features (e.g., protocol, TCP flags) without the need for preliminary one-hot encoding, which inflates feature space.

The following disadvantages can be noted:

- leaf-wise growth can lead to overfitting on small datasets, requiring careful regularization tuning;
- has a slightly more sophisticated setting of the hyperparameters, as the model is sensitive to them to achieve optimal performance.

3.1.3. Random Forest

In this study, Random Forest (RF), specifically its implementation for classification, RandomForestClassifier, is considered a fundamental and robust baseline ensemble learning algorithm. This method was chosen because it provides high robustness to overfitting, excellent interpretability, and good predictability “out-of-the-box”. Its advantages:

- high resistance to overtraining and noise: effectively averages a large number of “noisy” or overtrained trees, resulting in a stable and reliable model;

– Efficient handling of nonlinear dependencies and interactions: Decision trees are inherently capable of modeling complex nonlinear interactions between features (for example, the relationship between session duration and the number of packets per second), which is typical for various types of network attacks;
– does not require scaling or normalization of features, is robust to multicollinearity, and can work with mixed data types (numeric, categorical after encoding).

The following disadvantages can be noted:

– requires considerable computing resources and time to learn with a large number of trees.

3.1.4. Decision Tree

As part of the analysis of machine learning algorithms for network traffic classification, the Decision Tree (DT) classifier is considered a fundamental and fully interpretable base algorithm. is a non-parametric model that builds classification rules in the form of tree structures. It is one of the oldest ML algorithms. Despite being a classic model, it is rarely used in industrial environments. Although it is historically less effective, we also considered its performance on current datasets because of its interpretability and ability to work with different types of data. Its advantages:

– full interpretability and transparency. The complete classification path of each network event can be traced and understood by humans, which is indispensable for security audits, incident investigations, and regulatory compliance.
– no need to scale features: the algorithm is insensitive to data scale (e.g., the difference between the number of bytes and the session duration);
– detection of nonlinear dependencies and interactions: the tree automatically finds complex combinations of features that are characteristic for different types of attacks.

The following disadvantages can be noted:

– the model is prone to overfitting: without restrictions (pruning or regularization), the tree grows to perfectly fit the training sample, ignoring noise and generalizing poorly to new data;
– small changes in the data (adding/removing one sample) mainly result in drastically different tree structures due to the greedy splitting algorithm;
– emissions and noise can significantly affect splits, especially at the upper levels of the tree, reducing the quality of the model.

3.1.5. Logistic Regression

This is a fundamental algorithm in machine learning. This is a statistical model based on binary classification (which is perfect for our task of binary classification of network traffic). As part of a comparative analysis of machine learning methods for detecting malicious network traffic, logistic regression (LR) occupies a special place as a benchmark linear statistical classifier. Although this algorithm was developed in 1958, it is an “eternal classic” and can show results similar to more modern classification models. Logistic regression models the probability of an object belonging to a positive class (malicious traffic) using a logistic (symmoid) function that converts a linear combination of features into a value in the range $[0, 1]$. Advantages of network securitization:

– high interpretability: A positive weight indicates that increasing the value of a feature increases the likelihood of an attack (all other things being equal), while a negative weight decreases it;
– computational efficiency and scalability: the model is fast and efficient even on very large amounts of data, which makes it suitable for streaming traffic.

The following disadvantages can be noted:

- assumes a linear relationship between features and logits of probabilities, which limits it to nonlinear data;
- does not handle multicollinearity well and may overfit on multidimensional datasets without regularization.

3.2. Training Data

The Network Traffic Android malware dataset contains network features used to detect malicious Android applications. It was created by a Kaggle representative, Christian Urcuqui. The data source for the dataset was the larger DroidCollector [15] dataset, which contains complete network traffic in the pcap format. The Network Traffic Android Malware dataset was selected because it captures real-world behavior that enables the detection of advanced evasion techniques [16]. This is ideal for binary classification tasks. The features in the current dataset were specifically selected from network traffic to effectively detect malicious activity [17]. To train the ML model on the current dataset, it must first be converted by adding additional columns, such as the total number of packets and average packet size in bytes. This is necessary to identify more informative patterns and simplify the model operation.

4. Results

4.1. Metrics of interest

To compare the performance of the ML models, a set of metrics was chosen to evaluate both the quality of the binary classification and their overall dividing capacity. The proportion of correct predictions was estimated using the accuracy metric:

$$\alpha = \frac{TP + TN}{TP + TN + FP + FN},$$

where TP, TN, FP, FN — are the numbers of true positive, true negative, false positive, and false negative classifications, respectively.

For a detailed analysis of the performance of the ML models on the target (positive) class, a family of metrics that are sensitive to imbalance was used. Precision reflects the proportion of correctly detected malicious objects among all objects that the model has identified as malicious:

$$\pi = \frac{TP}{TP + FP}.$$

The recall metric characterizes the model's ability to detect all objects of the target class:

$$\rho = \frac{TP}{TP + FN}.$$

The AUC-Score metric was chosen to evaluate the ranking ability of the models and analyze their quality, regardless of the selected classification threshold. This metric is equal to the probability that a randomly selected positive object (malicious traffic) receives a higher score from the model than a randomly selected negative (harmless traffic) object. This metric was calculated as the area under the ROC curve. This curve is a graph that shows how the trade-off between the True Positive Rate (TPR, same as recall) and False Positive Rate (FPR) changes when the classification threshold is changed, where:

Table 1

Model Performance Comparison

Model	Accuracy	AUC Score	Precision	Recall	Time (s)	Size (MB)
XGBoost	0.8866	0.9627	0.8440	0.8790	0.60	0.59
LightGBM	0.8885	0.9617	0.8406	0.8901	0.32	0.28
RF	0.8706	0.9440	0.8295	0.8519	2.29	2.51
DT	0.8126	0.8883	0.7192	0.8726	0.06	0.03
LR	0.7412	0.8024	0.7382	0.5478	2.17	0.0007

$$TPR = \rho = \frac{TP}{TP + FN},$$

and

$$FPR = \frac{FP}{FP + TN}.$$

To assess the suitability of the model specifically for Android OS, a model weight metric in ONNX format was introduced, which allows model data to be transferred and used within applications running on this operating system. The computational constraints of mobile platforms, as recently reported in [18], impose strict limits on model complexity, which guided our selection of lightweight ONNX-exported models.

This multidimensional approach allows for an objective comparison of models based on different criteria, which in turn allows us to conduct a more detailed and accurate analysis to understand the strengths and weaknesses of each model.

4.2. Models comparison

We will conduct a comprehensive analysis among the best models on this dataset and identify the most suitable ML model for this task. Each has its own advantages and disadvantages.

To perform a mathematically correct comparison, we applied the multi-criteria analysis (MCDA) method [19] using global min-max normalization. We first distributed the weight coefficients for each metric so that their total sum is $\sum_{i=1}^{7} w_i = 1$, as shown in Table 2. We note that the recall metric is assigned a major weight (50%) because the exhaustive identification of malicious traffic (i.e., minimization of FN objects) is the primary task for an IDS, even at the cost of partial blockage of benign traffic.

To ensure the comparability of heterogeneous model characteristics, such as dimensionless accuracy coefficients (such as AUC-score) and physical quantities (weight in MB and time in seconds), a global linear normalization procedure was applied. This approach allows us to eliminate large-scale distortions in the objective function and bring all metrics to a single range $[0, 1]$, where the extreme values correspond to the best and worst results within the entire samples. The use of inverted normalization for resource indicators (weight and time) ensures that all optimization criteria are aligned, which is necessary for the correct calculation of the integral performance indicator using the weighted-sum method.

Table 2

Weight coefficients

Notation	Value	Description
w_1	0.05	AUC-Score
w_2	0.05	Accuracy
w_3	0.25	Precision
w_4	0.5	Recall
w_5	0.1	Model size
w_6	0.05	Training time

Table 3

Model Performance Comparison (Scaled Metrics)

Model name	Accuracy	AUC Score	Precision	Recall	Time (s)	Size (MB)
XGBoost	0.9871	1.0000	1.0000	0.9676	0.7578	0.7652
LightGBM	1.0000	0.9938	0.9728	1.0000	0.8834	0.8887
RF	0.8785	0.8833	0.8838	0.8884	0.0000	0.0000
DT	0.4847	0.5359	0.0000	0.9489	1.0000	0.9883
LR	0.0000	0.0000	0.1522	0.0000	0.0538	1.0000

Accordingly, for quality metrics such as accuracy, AUC-Score, precision, and recall, the norm is applied, where 1.0 is the best result. Therefore, for each model i paired with a dataset j and their corresponding metric $m_{i,j}$, the following formula is applied to calculate the normalized value:

$$n_{ij} = \frac{m_{ij} - min(m_j)}{max(m_j) - min(m_j)},$$

where $max(m_j)$ is the maximum value per metric m among all pairs of the model dataset and $min(m_j)$ is the minimum.

For the computational complexity parameters (model size in ONNX format and training time), an inverse normalization procedure was applied. This approach allows the conversion of minimized target functions into maximized ones, where the value of 1.0 corresponds to the optimal (minimum) feature value in the sample, ensuring that all efficiency vectors are aligned:

$$n_{ij} = \frac{max(m_j) - m_{ij}}{max(m_j) - min(m_j)},$$

where $max(m_j)$ is the maximum value for the metric m among all model-dataset pairs, and $min(m_j)$ is the minimum.

Table 4

Normalized model performance evaluation

Model	XGBoost	LightGBM	Random Forest	Decision Tree	Logistic Regression
MCDA	0.9476	0.9759	0.7532	0.6743	0.1407
F1	0.9852	1.0000	0.8974	0.6768	0.0000
AUC	0.9627	0.9617	0.9440	0.8883	0.8024

Consequently, we obtained the following table with normalized values:

Based on the values defined above, for each pair, we calculated the weighted sum of the normalized values according to the following formula

$$E = \sum_{i=1}^{6} (n_i * w_i),$$

where n_i is the normalized value of metric i, and w_i is its weight, which we defined above. As a result, we obtain the following table:

Based on the obtained data, it can be concluded that the XGBoost model shows higher efficiency than LightGBM. Despite the insignificant numerical gap of 0.1, this model is characterized by the most optimal balance between accuracy and the computational costs.

In conclusion, we can see that the LightGBM and XGBoost gradient boosting algorithms show the best results on the "Network Traffic Android Malware" dataset for the binary classification task, demonstrating their leadership. The results show that the dataset contains tabular data with complex nonlinear and cross-feature interactions, which the gradient boosting algorithms successfully detect and use.

5. Discussion

The primary aim of this study was to develop and evaluate a locally operating machine learning-based malicious traffic detection system for Android devices, addressing the privacy and availability limitations inherent in cloud-dependent security architectures. Our findings demonstrate that lightweight ML models can operate effectively within the stringent computational and storage constraints of mobile platforms while maintaining high detection accuracy.

The implementation of five ML models in the ONNX format, with sizes ranging from 0.0007MB (Logistic Regression) to 2.51MB (Random Forest), confirms that locally deployed models are practical. The MCDA revealed that XGBoost (MCDA score: 0.9476) and LightGBM (0.9759) achieved the best balance between detection performance and resource consumption, validating our hypothesis that gradient boosting algorithms are particularly well-suited for this task.

Our results align with the broader intrusion detection literature, which has consistently demonstrated the superiority of ensemble methods over single classifiers in network traffic analysis. The effectiveness of XGBoost [2] on tabular data with heterogeneous features is directly supported by our findings, as it achieved the highest AUC score (0.9627) among all tested algorithms. Similarly, the authors' claims regarding the computational efficiency of LightGBM [3] are corroborated by our measurements: LightGBM was trained in 0.32 s (versus 0.60 s for XGBoost) while maintaining comparable predictive performance. Our finding that XGBoost achieves superior performance

on network traffic classification is also consistent with recent work [20], which demonstrated that XGBoost outperforms other ensemble algorithms in wireless network traffic evaluation tasks, further confirming its suitability for flow-based analysis.

Notably, our results diverge from those of some earlier studies that reported stronger performance of Random Forest on network traffic datasets. Breiman's [4] original Random Forest implementation, while robust, proved less suitable for real-time mobile deployment in our evaluation owing to its larger memory footprint (2.51 MB) and longer training time (2.29 s), despite achieving acceptable accuracy (0.8706). This discrepancy likely stems from the specific characteristics of the Network Traffic Android Malware dataset and additional constraints imposed by mobile deployment.

The superior performance of gradient boosting algorithms can be explained by several theoretical mechanisms. First, the processing of heterogeneous features — numerical (session duration, byte counts), categorical (protocol flags), and potentially binary — favors tree-based models over linear approaches. XGBoost's built-in L1 and L2 regularization explicitly prevents overfitting to noisy network traffic data, which is particularly valuable given the inherent variability in the real-world mobile network conditions.

The proposed architecture demonstrates that privacy-preserving, fully local traffic analysis is achievable without sacrificing detection quality, which is a critical advancement given the growing regulatory scrutiny of data transfer to third-party servers (e.g., GDPR, CCPA). In addition, the quantitative comparison of the ONNX-exported models provides practical guidance for developers in selecting between accuracy and resource efficiency. The integration with the VPNService-based interception framework of PCAPdroid offers a reusable template for a real-world implementation.

This study has several limitations. First, the evaluation relied exclusively on the Network Traffic Android Malware dataset from Kaggle [7], which, despite its relevance, may not fully represent the diversity of contemporary Android malware families or the behaviors of benign applications. Second, the experiments were conducted under idealized conditions without concurrent background processes that would compete for the CPU and memory on actual user devices. Third, the training time metric, while informative, does not directly translate to inference latency during real-time traffic analysis, which is a critical performance dimension for user experience.

Building on these findings, several research trajectories have emerged as particularly promising:

- Hyperparameter optimization: the current study employed default or minimally tuned hyperparameters. Systematic optimization using Bayesian methods or grid search could potentially improve the performance of XGBoost and LightGBM beyond the reported metrics.
- Cross-dataset validation: the creation of a unified "super" dataset combining features from multiple traffic collections (e.g., CIC-AndMal2017, CICIDS2017) would enable assessment of generalization across different network environments and attack types.
- Adversarial robustness: the vulnerability of gradient boosting models to adversarial examples in the network traffic domain remains unexplored and represents a significant concern for real-world deployment.

In addition, the integration of graph-based architectures, as recently proposed in [21], represents a natural extension of our work that can capture structural relationships in network traffic.

Notably, every claim in this discussion is directly supported by the results presented in Section 4. The conclusion that XGBoost provides the optimal balance (MCDA: 0.9476) follows directly from the normalized metrics in Table 3 and weighted sum calculation in Table 4. The recommendation of gradient boosting over Random Forest is justified by the substantial differences in training time (0.60 vs. 2.29 s) and model size (0.59 vs. 2.51MB) while maintaining superior accuracy (0.8866 vs. 0.8706). No unsupported generalizations were made beyond the experimental evidence.

6. Conclusion

This study demonstrates that a fully local, machine learning-based malicious traffic detection system for Android devices is both feasible and effective, addressing the privacy and availability limitations of cloud-dependent security architecture. Among the five evaluated algorithms — XGBoost, LightGBM, Random Forest, Decision Tree, and Logistic Regression — gradient boosting methods outperformed the others. XGBoost achieved the highest AUC (0.9627), whereas LightGBM offered the fastest training time (0.32s) and smallest model size (0.28MB). Multi-criteria analysis confirmed that LightGBM was the most suitable for real-time mobile deployment (MCDA score 0.9759). Local analysis eliminates cloud dependency, preserves user privacy, and matches or exceeds server-based detection on the Network Traffic Android Malware dataset. Theoretically, ensemble tree-based methods inherently handle heterogeneous and nonlinear traffic features better than linear models. Practically, the ONNX-based architecture provides a ready-to-use blueprint for privacy-centric mobile IDSs. The limitations of this study include the single-dataset evaluation, idealized testing conditions, and lack of direct inference latency and battery impact measurements. Future studies should pursue cross-dataset validation, error analysis of misclassified samples, inference latency benchmarking on physical devices, and adversarial robustness testing. Overall, gradient boosting algorithms offer an optimal compromise between detection accuracy and mobile resource constraints, serving as the foundation for next-generation on-device intrusion detection systems.

Author Contributions: Nikita Khalemskii: methodology, software, formal analysis, investigation, resources, data curation, writing—original draft preparation, and visualization. Vitalii Beschastnyi: conceptualization, validation, writing—review and editing, supervision, project administration, funding acquisition; All authors have read and agreed to the published version of this manuscript.

Funding: The reported study was funded by RSF, projects no. 23-79-10084, https://rscf.ru/en/project/23-79-10084/.

Data Availability Statement: Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Declaration on Generative AI: During the preparation of this work, the authors used Paperpal in order to: grammar and spelling check. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] G. B. H. Vaibhav, M. Bafna, G. Sumathi, and R. Gopi, "Android Malware Detection using Federated Learning," in *Proceedings of the 2025 IEEE International Conference on Artificial Intelligence and Signal Processing (AISP)*, Tirunelveli, India, 2025. DOI: 10.1109/AISP64076.2025.10689155

[2] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, 2016, pp. 785–794. DOI: 10.1145/2939672.2939785

[3] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," in *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, Curran Associates, 2017, pp. 3146–3154.

[4] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. DOI: 10.1023/A:1010933404324

[5] S.-l. developers. "Decision Trees." Scikit-learn documentation, scikit-learn, Accessed: Apr. 24, 2025. [Online]. Available: https://scikit-learn.org/stable/modules/tree.html

[6] S.-l. developers. "Logistic Regression." Scikit-learn documentation, scikit-learn, Accessed: Apr. 24, 2025. [Online]. Available: https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression

[7] C. Urcuqui, *Network Traffic Android Malware*, version 1.0, 2019.

[8] O. Community, *ONNX: Open Neural Network Exchange*, version 1.16, Open standard for machine learning interoperability, 2024.

[9] A. Botvinko and K. Samouylov, "Evaluation of the firewall influence on the session initiation by the SIP multimedia protocol," *Discrete and Continuous Models and Applied Computational Science*, vol. 29, pp. 221–229, Sep. 2021. DOI: 10.22363/2658-4670-2021-29-3-221-229

[10] A. S. Baklashov and D. S. Kulyabov, "Statistical and density-based clustering techniques in the context of anomaly detection in network systems: A comparative analysis," *Discrete and Continuous Models and Applied Computational Science*, vol. 33, no. 1, pp. 27–45, 2025.

[11] S. Zhou, H. Zeng, Y. Lu, Y. Chen, J. Liu, and J. Su, "A Lightweight Embedded Intelligent Threat Detection System Using TSANet on Android Platforms," in *Proceedings of the 2025 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, Raipur, India, 2025. DOI: 10.1109/ANTS63432.2025.10689230

[12] E. Faranda, *PCAPdroid — No-root network monitor*, GitHub, Version 1.7.1, 2024.

[13] T. W. Team, *Wireshark — Network Protocol Analyzer*, version 4.4.2, 2024.

[14] O. Corporation. "Java Native Interface Specification." Java SE 8 Documentation, Oracle, Accessed: Apr. 24, 2025. [Online]. Available: https://docs.oracle.com/javase/8/docs/technotes/guides/jni/

[15] D. Cao, S. Wang, Q. Li, Z. Cheny, Q. Yan, L. Peng, and B. Yang, "DroidCollector: A High Performance Framework for High Quality Android Traffic Collection," in *2016 IEEE Trustcom/BigDataSE/ISPA*, IEEE, 2016, pp. 1753–1758. DOI: 10.1109/TrustCom.2016.0270

[16] J. Tang, S. Zhou, T. Peng, X. Yan, X. Hu, and W. Tian, "DTDroid: Adversarial Packed Android Malware Detection Based on Traffic and Dynamic Behavioral," *IEEE Internet of Things Journal*, vol. 12, no. 3, pp. 2646–2658, 2025. DOI: 10.1109/JIOT.2024.3477442

[17] C. C. U. López, J. S. D. Villarreal, A. F. P. Belalcazar, A. N. Cadavid, and J. G. D. Cely, "Features to Detect Android Malware," in *2018 IEEE Colombian Conference on Communications and Computing (COLCOM)*, IEEE, 2018, pp. 1–6. DOI: 10.1109/ColComCon.2018.8466715

[18] S. Anand, B. Mitra, S. Dey, A. Rao, R. Dhar, and J. Vaidya, "MALITE: Lightweight Malware Detection and Classification for Constrained Devices," *IEEE Transactions on Emerging Topics in Computing*, vol. 13, no. 3, pp. 1099–1112, 2025. DOI: 10.1109/TETC.2025.3566370

[19] E. K. Zavadskas, Z. Turskis, and J. Antuchevičienė, "Multi-Criteria Decision Making (MCDM) Methods and Concepts," *Encyclopedia*, vol. 3, no. 1, p. 6, 2023. DOI: 10.3390/encyclopedia3010006

[20] O. Y. Mohammed and I. A. Saleh, "Prediction Wireless Network Traffic Evaluation Potential Based on Ensemble Algorithms," in *Proceedings of the 2025 IEEE 22nd International Multi-Conference on Systems, Signals & Devices (SSD)*, Monastir, Tunisia, 2025, pp. 914–921. DOI: 10.1109/SSD63744.2025.10689473

[21] Y. K. Sharma, D. S. Tomar, R. K. Pateriya, and S. Solanki, "GNSTAM: Integrating Graph Networks With Spatial and Temporal Signature Analysis for Enhanced Android Malware Detection," *IEEE Access*, vol. 13, pp. 81 326–81 346, 2025. DOI: 10.1109/ACCESS.2025.3582741

Information about the authors

Khalemskii, Nikita D.—Master Student of Telecommunications R&D Institute National Research University Higher School of Economics (HSE University) (e-mail: ndkhalemskiy@edu.hse.ru, ORCID: 0009-0001-2418-3556)

Beschastnyi, Vitalii A.—Candidate of Physical and Mathematical Sciences, assistant professor of Department of Probability Theory and Cybersecurity, RUDN University; Senior Researcher of Telecommunications R&D Institute National Research University Higher School of Economics (HSE University) (e-mail: beschastnyy-va@rudn.ru, ORCID: 0000-0003-1373-4014, ResearcherID: 3293250, Scopus Author ID: 57192573001)

RUSSIAN METADATA

УДК 004.89, 004.056.57
PACS 07.05.Mh, 07.05.Kf
DOI: 10.22363/2658-4670-2026-34-2-187-200 EDN: JDISSJ

Обеспечение безопасности мобильных систем с помощью приложений для обнаружения вредоносного трафика

Н. Д. Халемский[1], В. А. Бесчастный[1, 2]

[1] *Национальный исследовательский университет «Высшая школа экономики», ул. Мясницкая, д. 20, Москва, 101000, Российская Федерация*

[2] *Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация*

Аннотация. *Актуальность:* Стремительный рост кибератак, направленных на объекты инфраструктуры, предприятия и индивидуальных пользователей через мобильные устройства, обусловливает потребность в эффективных системах обнаружения вторжений. Традиционные сигнатурные методы не позволяют противостоять уязвимостям нулевого дня и сложным постоянным угрозам (Advanced Persistent Threats, APT), что стимулирует исследование подходов машинного обучения для анализа сетевого трафика. *Цель:* Исследование направлено на разработку и оценку локально функционирующего Android-приложения, обеспечивающего конфиденциальность данных и выполняющего обнаружение вредоносного трафика в реальном времени на основе моделей машинного обучения. Решение позволяет преодолевать ограничения облачных архитектур безопасности, которые ставят под угрозу приватность пользователей и доступность сервисов. *Методы:* Были проанализированы пять моделей машинного обучения — XGBoost, LightGBM, случайный лес, дерево решений и логистическая регрессия, — обученных на наборе данных Network Traffic Android Malware. Модели были экспортированы в формат ONNX для локального развёртывания на устройствах под управлением Android. Оценка производительности осуществлялась по показателям точности (accuracy), полноты (recall), площади под ROC-кривой (AUC), времени обучения и размеру модели; для комплексного сравнения применялся многокритериальный анализ решений. *Результаты:* Алгоритмы градиентного бустинга продемонстрировали превосходные результаты: LightGBM достиг наивысшей оценки (0,9759), минимального времени обучения (0,32 с) и наименьшего размера модели (0,28 МБ), тогда как XGBoost показал максимальную площадь под ROC-кривой (0,9627). Обе модели значительно превзошли случайный лес (0,7532), дерево решений (0,6743) и логистическую регрессию (0,1407). *Выводы:* LightGBM показала оптимальный баланс между точностью обнаружения и ограниченными ресурсами мобильных устройств, что делает её пригодной для внутриплатформенного развёртывания. Предложенная архитектура демонстрирует, что полностью локальный анализ трафика на основе машинного обучения осуществим без потери качества обнаружения и представляет собой конфиденциально-ориентированную альтернативу серверно-зависимым решениям для систем обнаружения вторжений в мобильных сетях следующего поколения.

Ключевые слова: сигнатурные методы обнаружения, машинное обучение, обнаружение аномалий, мониторинг трафика, обработка даных в реальном времени, кибербезопасность

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2026, 34 (2) 201–213
http://journals.rudn.ru/miph

Research article

UDC 512.624.3, 512.76, 512.542.74, 519.622.2
PACS 02.20.Hj 02.70.Bf
DOI: 10.22363/2658-4670-2026-34-2-201-213 EDN: JDOCAU

On quadratic transformations of the Fano plane

Ilshat T. Dulatov[1], Mikhail D. Malykh[1,2], Anton L. Sevastianov[1], Aleksander V. Zorin[1]

[1] *RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation*

[2] *Joint Institute for Nuclear Research, 6 Joliot-Curie St, Dubna, 141980, Russian Federation*

(received: March 4, 2026; revised: March 25, 2026; accepted: March 30, 2026)

Abstract. The importance of studying quadratic Cremona transformations over algebraically non-closed fields for the theory of Kahan difference schemes for dynamical systems with a quadratic right-hand side is discussed. Cremona transformations of the projective plane over a Galois field of size 2, i.e., the Fano plane, are considered. The notation proposed by J. Rosanes is used to describe quadratic Cremona transformations. The relationship between quadratic transformations and matrix pencils is described. Quadratic transformations without singular points are called regular. The Sage system is used to implement procedures that convert a quadratic transformation into a permutation of 7 points of the Fano plane (an element of the symmetric group S_7) and a permutation of 7 points of the plane into a quadratic transformation. By enumerating all quadratic transformations, it is proved that regular quadratic transformations generate the entire permutation group of the Fano plane. This theorem is analogous to Noether's theorem over an algebraically non-closed field. It is proved that regular quadratic Cremona transformations have even order, and a description of the corresponding permutations of the group S_7 is given. It is shown that a permutation always corresponds to some quadratic Cremona transformation, but this transformation is not always regular. The resulting quadratic transformations can be supplemented by a rule that resolves ambiguities at fundamental points. An example of a quadratic transformation for which such resolution is impossible is given. This leads to a natural classification of quadratic transformations of the Fano plane.

Key words and phrases: projective plane, Galois fields, Cremona transformations, birational map, Kahan's method

For citation: I. T. Dulatov, M. D. Malykh, A. L. Sevastianov, A. V. Zorin "On quadratic transformations of the Fano plane," *Discrete and Continuous Models and Applied Computational Science*, vol. 34, no. 2, pp. 201–213, 2026. DOI: 10.22363/2658-4670-2026-34-2-201-213. EDN: JDOCAU.

1. Introduction

We encountered Cremona transforms while studying difference schemes for dynamical systems [1] obtained using the Kahan method [2], developed in the works of Sanz-Serna [3], Suris [4–6], and McLachlan [7]. In such schemes, each time step is described by a quadratic Cremona transformation. Although the construction itself appears quite classical, the few known results prove of little use. The point is that the coefficients of difference schemes typically belong to algebraically open fields, for example, the field of fractions of the ring $\mathbb{Q}[\Delta t]$, where Δt is the time step (symbolic variable), or even to the field $\mathbb{Q}$. Moreover, an approximate solution found for a fixed rational step is an infinite

sequence of rational numbers. Therefore, for these studies, it would be desirable to have a theory of quadratic Cremona transformations over algebraically open fields. However, research on Cremona transformations, both new and classical, is usually conducted over an algebraically closed field [8, 9].

In the case of algebraically open fields, there is one case that can be studied completely and without much effort. This is the case of finite fields. In this article, we consider quadratic transformations of the projective plane on a Galois field of size 2.

2. Quadratic transformations of the plane

Let us recall the basic definitions from classical algebraic geometry. Two equations

$$f(x, y; \hat{x}, \hat{y}) = 0, \quad g(x, y; \hat{x}, \hat{y}) = 0, \tag{1}$$

connect two pairs of variables, (x, y) and $(\hat{x}, \hat{y})$, and in the non-degenerate case define an algebraic correspondence on the plane P_2: given an initial point (x, y) from this system, one can find several corresponding endpoints $(\hat{x}, \hat{y})$, and vice versa. If $\hat{x}$, $\hat{y}$ can be expressed rationally in terms of x, y, then one speaks of a rational transformation of the plane [10, no. 14], and if, in addition, If x, y can also be expressed rationally in terms of $\hat{x}$, $\hat{y}$, then we speak of a birational transformation (birational map) of the plane or the Cremona transformation [10, no. 16].

In particular, if the equations (1) are linear in both x, y and $\hat{x}$, $\hat{y}$, then the correspondence will be one-to-one, and therefore x, y can be expressed as rational functions of $\hat{x}$, $\hat{y}$ and, conversely, $\hat{x}$, $\hat{y}$ as rational functions of x, y. This gives a birational map, which is traditionally called the quadratic Cremona transformation, since it maps straight lines to lines of the second order [10, no. 17].

Example 1. Bilinear equations

$$\hat{x}x = y, \quad \hat{y}y = 1$$

define a birational transformation of the plane that maps the line

$$a\hat{x} + b\hat{y} + c = 0$$

to a second-order line

$$ay^2 + bx + cxy = 0.$$

We will denote the projective coordinates of a point (x, y) as $(u : v : w)$, assuming that

$$x = \frac{u}{w}, \quad y = \frac{v}{w}.$$

When using matrix notation, we will consider the projective coordinates of a point $(u : v : w)$ as a column $\mathfrak{u}$. With respect to projective coordinates, the equations (1) defining a quadratic transformation can be written as

$$a(\hat{\mathfrak{u}}, \mathfrak{u}) = 0, \quad b(\hat{\mathfrak{u}}, \mathfrak{u}) = 0, \tag{2}$$

where a and b are two bilinear forms. We denote by A and B the matrices of the bilinear forms a and b and write the equations (2) as

$$\hat{\mathfrak{u}}^T A\mathfrak{u} = 0, \quad \hat{\mathfrak{u}}^T B\mathfrak{u} = 0. \tag{3}$$

We can say that every quadratic transformation is defined by a pair of matrices. This notation for quadratic transformations was proposed in the 1870s by J. Rosanes. Unfortunately, this work went unnoticed. It is mentioned in thick reviews spanning half a century, but leaving Rosanes's notation without attention.

Example 2. The transformation

$$\hat{x}x = y, \quad \hat{y}y = 1$$

from Example 1 in homogeneous coordinates is written as

$$\hat{u}u - \hat{w}v = 0, \quad \hat{v}v - \hat{w}w = 0$$

and, therefore, in Rosanes notation is described by a pair of matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Not every pair of matrices defines a birational transformation. For this to be true, it is necessary and sufficient that the matrices A and B be linearly independent.

Generally speaking, a birational transformation, unlike a linear transformation, can have singular points, that is, points $\mathfrak{u}$ that correspond to several points $\hat{\mathfrak{u}}$, and points $\hat{\mathfrak{u}}$ that correspond to several points $\mathfrak{u}$. Nineteenth-century authors called these points fundamental [10, n. 15, 17]. In Rosanes's time, matrix theory was in its infancy, so the connection between fundamental points and the eigenvalue problem was not noted by him. We can formulate it as follows.

Theorem 1. *The quadratic transformation* (3) *is not uniquely invertible only at those points* $\mathfrak{u}$ *that are eigenvectors of the eigenvalue problem.*

$$\mu A\mathfrak{u} = \lambda B\mathfrak{u}, \tag{4}$$

The quadratic transformation (3) *is not uniquely invertible only at those points* $\hat{\mathfrak{u}}$ *that are eigenvectors of the eigenvalue problem.*

$$\mu A^T\mathfrak{u} = \lambda B^T\mathfrak{u}. \tag{5}$$

Here, by an eigenvalue of problem (4) we mean a pair (μ, λ) in which one of the elements is nonzero. Moreover, it is convenient to consider an eigenvalue as a point $(\mu : \lambda)$ on the projective line P_1^1. An eigenvalue $(\mu : \lambda) = (1 : 0)$ exists when the determinant $\det A$ is zero.

The eigenvalues of problems (4) and (5) coincide and are the zeros of the determinant

$$\det(\mu A - \lambda B) \tag{6}$$

Definition 1. *Quadratic transformations generated by matrices such that the determinant* (6) *is not zero at any point* $(\mu : \lambda)$ *of the projective line* P_1^1 *are called regular.*

In the case of algebraically closed fields, this determinant is necessarily zero somewhere, and therefore regular quadratic transformations do not exist. Let us consider such transformations in the simplest case, over a Galois field of size 2, which we will henceforth denote as GF(2).

3. Regular quadratic transformations of the Fano plane

The affine plane over the Galois field GF(2) consists of 4 points, while the projective plane consists of 7 points; it is called the Fano plane [11, n. 1.1, 1.10]. For our calculations, we will use the Sage computer algebra system [12], in which the points of the Fano plane are numbered as shown in Figure 1.

The group of all automorphisms of the Fano plane is the symmetric group S_7 [11, n. 1.12]. A regular quadratic transformation is one-to-one at all points of the Fano plane (Theorem 1), so it permutes the points of the Fano plane and, in this sense, is an element of the group S_7.

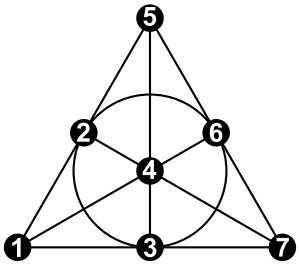

Figure 1. Fano Plane

Example 3. Consider, for example, a transformation that in Rosanes notation is described by a pair of matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

The determinant

$$\det(\mu A - \lambda B) = \mu^3 + \mu^2\lambda + \lambda^3$$

is equal to 1 at all three points of the projective line P_1, which is traversed by the point $(\mu : \lambda)$. Therefore, this pair of matrices defines a regular quadratic transformation of the Fano plane. The first point has coordinates $(0 : 0 : 1)$, so it corresponds to the point $\hat{u}$, whose coordinates are determined from a system of linear equations

$$\begin{pmatrix} \hat{u} & \hat{v} & \hat{w} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0, \quad \begin{pmatrix} \hat{u} & \hat{v} & \hat{w} \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0$$

or

$$\hat{w} = 0, \quad \hat{v} = 0.$$

Thus, the first point maps to point $(1 : 0 : 0)$, the 7th in Figure 1. Searching through all seven points of the Fano plane, we find that the quadratic transformation is a permutation

$$(1, 7, 5)(2, 4).$$

Of course, this search was performed in Sage rather than manually. For convenience, we have placed an archive of Sage calculations in [13]. We call the transformation of a pair of matrices defining a quadratic transformation in Rosanes notation into a permutation of the symmetric group S_7 `rosanes_to_permutation(A,B)`.

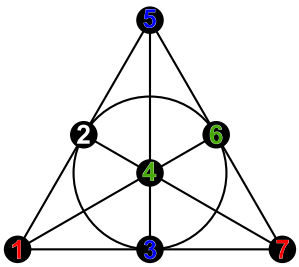

Figure 2. Permutation $(1,7)(3,5)(4,6)$ Fano plane

4. Analog of Noether's theorem

The composition of two quadratic transformations is a birational transformation, but not necessarily quadratic. Therefore, the set of all quadratic transformations is not a group. The classical Noether theorem states that over $\mathbb{C}$, every birational automorphism of a plane can be represented as a product of quadratic transformations. The key point in the proof of Noether's theorem is the identification of fundamental points of the transformation [14, 15], so it seems that it cannot be extended to the case of finite planes.

We compiled a list of all pairs of matrices for which the determinant (6) cannot be factorized over GF(2); it contains 8064 pairs. For each such pair, we computed the corresponding permutation (Example 3) and by simple enumeration established the following theorem.

Theorem 2. *Regular quadratic transformations of the Fano plane generate the entire symmetric group S_7.*

In other words, every automorphism of the Fano plane is a product of permutations that are quadratic transformations. In this sense, the set of all regular permutations is complete; it generates all permutations of points of the Fano plane, which distinguishes it from the set of all possible colliniations of the Fano plane that generate a subgroup of S_7, described in detail in [11, n. 1.12].

5. Classification of regular quadratic transformations as elements of S_7

Since we have enumerated all quadratic transformations in Sage, we can describe the properties of permutations of the symmetric group that are quadratic transformations.

Theorem 3. *Regular quadratic transformations of the Fano plane are permutations of even order.*

To describe regular quadratic transformations as permutations in more detail, recall that the Fano plane has exactly 7 lines, each of which consists of 3 points [11, n. 1.1].

5.1. Second-order permutations

Theorem 4. *A second-order permutation consists of three cycles*

$$(p_1, p_2)(p_3, p_4)(p_5, p_6),$$

where $p_1 \ldots p_6$ are the vertices of a complete quadrilateral. Second-order quadratic transformations generate the entire group S_7.

Example 4. A pair of matrices

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

corresponds to a permutation

$$(1, 7)(3, 5)(4, 6).$$

In Figure 2 three pairs of points are highlighted in color. These points are the vertices of the quadrilateral formed by the four lines 137, 567, 345, and 146.

The square of a second-order permutation is the identity permutation (), which, by Theorem 3, is not a quadratic transformation.

The appearance of the quadrilateral in Theorem 4 is interesting, since this construction is fundamental to the axiomatics of projective geometry [11, n. 1.1].

5.2. Fourth-order permutations

Theorem 1. *Fourth-order transformations are a cycle*

$$(p_1, p_2, p_3, p_4),$$

with three of the four points collinear.

Example 5. The pair of matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

corresponds to the permutation

$$(1, 7, 3, 4),$$

in which the points 1, 7, and 3 lie on the same line. The square of this permutation is

$$(1, 3)(4, 7),$$

it is not among the regular quadratic transformations; its cube is

$$(1, 4, 3, 7),$$

it corresponds to the quadratic transformation

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

5.3. Sixth-order permutations

There are two types of sixth-order permutations. First, there is the cycle

$$(p_1, p_2, p_3, p_4, p_5, p_6),$$

Second,

$$(p_1, p_2, p_3)(p_4, p_5).$$

Not all of them occur; for example, there is no

$$(1, 2, 3, 4, 5, 6).$$

Here, it is significantly more difficult to discern a pattern than in the case of second-order permutations.

6. Constructing a quadratic transformation from a given permutation

Above, in the Example 3, we showed how to determine a permutation from a given quadratic transformation. Let us pose the inverse problem: given a permutation $t \in S_7$, find a quadratic transformation.

So, let a permutation $t \in S_7$ be given. For a pair of matrices A, B defining a quadratic transformation, the following must hold:

$$t(\mathfrak{u})^T A\mathfrak{u} = 0, \quad t(\mathfrak{u})^T B\mathfrak{u} = 0 \quad \forall \mathfrak{u} \in P_2. \tag{7}$$

Let X be a 3×3 matrix, whose elements are symbolic variables. Then

$$t(\mathfrak{u})^T X\mathfrak{u} = 0 \quad \forall \mathfrak{u} \in P_2 \tag{8}$$

represent 7 linear homogeneous equations for finding 9 unknowns $x_{11}, \dots, x_{33}$. The solution to this system is a linear subspace in the space of 3×3 matrices. Since its dimension is greater than or equal to $9 - 7 = 2$, there always exist two linearly independent matrices A and B for which (7) holds. This implies the following theorem.

Theorem 2. *There exists a quadratic transformation that permutes the points of the Fano plane in a given way, but this transformation is not always regular.*

From Theorem 3, not every permutation is a regular quadratic transformation. To prevent the formulation of Theorem 2 from appearing to contradict Theorem 3, its formulation specifically stipulates that the resulting quadratic transformations may not be regular.

Example 6. As an example, consider the permutation

$$(1, 2, 3, 4, 5, 6), \tag{9}$$

which does not correspond to any quadratic transformation (see n. 5.3 above). We implemented the method described above as a function `permutation_to_rosanes(t)` in our Sage program [13]. In this case, it yields the expression X as the matrix

$$\begin{pmatrix} 0 & a_4 & a_4 + a_7 + a_8 \\ a_4 + a_8 & a_4 & a_8 \\ a_7 & a_7 & a_8 \end{pmatrix} \tag{10}$$

depending on 3 parameters, i.e., the solution space of the system (8) has dimension 3. However, no two matrices from this space yield a regular quadratic transformation, since the determinant of the expression (10) is equal to

$$a_8 \cdot (a_4 + a_7) \cdot (a_4 + a_7 + a_8).$$

This nonzero polynomial is zero for all possible values of the parameters a_4, a_7, a_8 in the field GF(2).

According to Theorem 2, every permutation corresponds to some quadratic transformation. Therefore, for some permutations, this transformation inevitably has fundamental points at which the uniqueness of the transformation is violated.

Example 7. Let us return to the permutation (9) from Example 6. Let us take two linearly independent elements in the linear space (10), say,

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \tag{11}$$

and find all possible pairs $(\mathfrak{u}, \hat{\mathfrak{u}})$ satisfying the system (3):

$$\begin{pmatrix} 1 & 1 & 1 & 2 & 3 & 3 & 3 & 4 & 5 & 6 & 7 & 7 & 7 \\ 2 & 4 & 7 & 3 & 2 & 4 & 7 & 5 & 6 & 1 & 2 & 4 & 7 \end{pmatrix}$$

The first point corresponds not to one point, but to three: points No. 2, 4, and 7, which lie on the same line. Similarly, points No. 3 and 7 correspond not to a single point, but to the same line 247. We can say that uniqueness is violated on the special line 137, to each point of which the degenerate transformation assigns all three points of line 247. If we supplement this permutation with a point selection rule:

$$1 \to 2, \quad 3 \to 4, \quad 7 \to 7, \tag{12}$$

then we obtain the original permutation (9).

When solving the (8) system, we always obtain quadratic transformations that can be predefined at the fundamental points to a one-to-one permutation of the Fano plane. However, the ambiguity here is that choosing a different rule yields a different permutation.

Example 8. For example, the quadratic transformation (11) from Example 7 can be supplemented not only with the rule (12), but also, say, with such a rule

$$1 \to 2, \quad 3 \to 7, \quad 7 \to 4.$$

Then we obtain a permutation

$$(1, 2, 3, 7, 4, 5, 6), \tag{13}$$

different from the original permutation (9).

There is no contradiction with Theorem 2 in obtaining multiple permutations. A new permutation obtained with a different choice of the additional rule will correspond to a different matrix space, but the determinant of all these matrices is zero.

Example 9. The permutation (13) obtained in Example (8) instead of the original permutation (9) corresponds to a degenerate quadratic transformation, the pair of matrices of which can be taken from a linear space.

$$\begin{pmatrix} a_4+a_7+a_8 & a_4 & a_4+a_7+a_8 \\ a_4+a_8 & a_4 & a_8 \\ a_7 & a_7 & a_8 \end{pmatrix}.$$

The determinant of this matrix is identically zero.

By choosing a pair of linearly independent matrices in the solution space of the system (8) in various ways, we obtain various quadratic transformations which, after supplementing them with the rules for resolving fundamental points, yield the original permutation.

Example 10. If we choose the basis elements in the three-dimensional linear space (10) differently than in Example 6, we obtain, generally speaking, a different transformation. For example, let us compare with the choice considered in Example 6 a pair of matrices

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

and find possible pairs $(\mathfrak{u}, \hat{\mathfrak{u}})$ satisfying the system (3):

$$\begin{pmatrix} 1 & 2 & 2 & 2 & 3 & 3 & 3 & 4 & 4 & 4 & 5 & 6 & 7 \\ 2 & 1 & 3 & 7 & 1 & 4 & 6 & 5 & 6 & 7 & 6 & 1 & 7 \end{pmatrix}.$$

Now, uniqueness is violated at points 2, 3, and 4, which are not collinear. Each of these points corresponds to points on different lines:

$$2 \to 137, \quad 3 \to 146, \quad 4 \to 567.$$

We can supplement the permutation with a rule for choosing points on these lines, and as a result, we obtain the original permutation (9).

Thus, given a permutation, it is always possible to find a quadratic transformation, but, firstly, this transformation may not be regular, and, secondly, there may be several suitable transformations.

7. Two classes of irregular quadratic transformations

Solutions to the system (8) yield irregular transformations that have a remarkable property by construction: they can be supplemented with a fundamental point resolution rule, after which they become a simple permutation of points on the Fano plane. However, it is easy to find an example in which such a rule cannot be specified.

Example 11. To contrast with the irregular transformation from Example 6, let us take as an example a pair of matrices

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

The determinant

$$\det(\mu A - \lambda B) = -(\mu^2 - \mu\lambda - \lambda^2)(\mu - \lambda)$$

is not equal to zero, for example, when $(\mu : \lambda) = (1 : 0)$. Let us find all possible pairs $(\mathbf{u}, \hat{\mathbf{u}})$ satisfying the system (3):

$$\begin{pmatrix} 1 & 2 & 3 & 3 & 3 & 4 & 5 & 6 & 7 \\ 5 & 5 & 2 & 4 & 7 & 3 & 5 & 6 & 1 \end{pmatrix}.$$

Here line 125 contracts to point No. 5, and point No. 3 is broken into line 247. Since three points of line 125 correspond to a single point, No. 5, there is no way to resolve the singularities with an additional rule.

Thus, irregular quadratic transformations are divided into two classes. The first are those whose existence is guaranteed by Theorem 2. These transformations have singular points, but the ambiguity of the correspondence can be restored by imposing additional rules. The second are quadratic transformations that have an irremovable ambiguity. In other words, for an irregular transformation of class 1, there exists a permutation $t \in S_7$ such that (7) holds, while for transformations of class 2, such a permutation does not exist.

The properties of regular transformations and irregular transformations of class 1 can be deduced from the properties of permutations of the group S_7, which are well studied.

8. Correspondence between the group S_7 and quadratic transformations

There is a certain correspondence between the union of the set of regular transformations and the set of irregular transformations of class 1, on the one hand, and the group S_7, on the other hand.

Given a quadratic transformation, we can find a permutation. For regular transformations, this is done uniquely (Example 3), while for irregular ones it is ambiguous (Example 10).

Given a given permutation t of S_7, a quadratic transformation can be found. If the dimension of the solution space of the system (8) of 7 equations with 9 unknowns is 2, then the matrix pencil $\mu A - \lambda B$ is uniquely determined, but not its basis A, B. However, as is easy to see, any basis of this pencil defines a quadratic transformation equivalent to t. Since we consider two transformations that implement identical permutations of points to be equal, different bases of the same matrix pencil define the same quadratic transformation. Therefore, given a permutation, one can uniquely determine the matrix pencil, and from it, the quadratic transformation. Uniqueness is violated when there are linearly dependent equations among the 7 equations of the system (8) .

9. Results and discussion

The goal of this study was to use a simple example to understand how strongly Cremona's theory of quadratic transformations over algebraically non-closed fields differs from the classical theory developed over $\mathbb{C}$.

Contrary to our expectations, Noether's theorem remained valid on the Fano plane (Theorem 2), but the quadratic transformations generating the entire automorphism group of the Fano plane were regular quadratic transformations of order 2 (Theorem 4). There is no analogue of such transformations over $\mathbb{C}$.

When integrating dynamic systems using the Kahan method, the degrees of the quadratic Cremona transformations are calculated. A numerical solution is an orbit calculated for a single initial point. If it is calculated over a floating-point implementation of the field $\mathbb{R}$, the result is fast, but with poorly controlled roundoff error. Attempting to calculate the n-th power of the Cremona transform over the field $\mathbb{Q}$ or $\mathbb{Q}(\Delta t)$, and especially with initial data specified in symbolic form, leads to exponential growth in the n-degree of expressions [16–18]. This makes direct methods for finding periodic solutions, which in theory should return periodic solutions or guarantee their absence in a finite number of steps [18], practically unusable.

Theorem 2 allows looking at these problems with some optimism: direct substitution of one birational transformation into another in the case of finite fields leads to a very complex expression, which, however, does not at all exclude the possibility of describing this transformation as an element of the group S_7 and quickly calculating its degree. The point is that Cremona transformations admit various representations, some of which are very complicated. Over algebraically non-closed fields, an additional degree of freedom appears in this situation: one can attempt to reduce the calculation of the degree of a Cremona transformation to the degree of a regular transformation.

Quadratic Cremona transformations are purely algebraic objects, so it seems that their invariants must also be algebraic. The fact that invariant curves of a Cremona transformation may not be algebraic was first noted relatively recently [19]. We encountered this problem when applying Kahan's method to the Volterra–Lotka system [3, 20]. In this case, it is clearly seen that the points lie on some curved line, but it can be proven that this line is not algebraic [20, ex. 5][18, n. 3.3]. There are many different ways to extend the concept of an invariant line to finite planes, especially if we allow fixed points and lines of arbitrarily large order to be added to the invariant sets. However, for the case of the Fano plane, we can say what these sets actually look like.

From Theorem 4 it follows that for quadratic transformations of the second order, an invariant line exists and is a quadrilateral, since all 6 of its points can be described as the roots of an equation obtained by multiplying the left-hand sides of the equations of the 4 straight lines that form its sides. It is interesting to note that all six points of the invariant quadrilateral are movable; the transformation leaves none of them fixed. On the other hand, the fourth-order transformation

$$(1, 7, 3, 4),$$

considered in Example 5 preserves the pair of lines 137 and 146, but one of the five points of this pair, namely point 6, remains fixed.

For fourth-order transformations in Theorem 1, the invariant set is a pair of lines: the line on which three of the four points represented by the transformation lie, and any line passing through the remaining point.

Finally, it is impossible not to notice that the main object in the developed theory turned out to be transformations equivalent to permutations of 7 points of the Fano plane. Irregular transformations of the second class (Example 11) represent a more complex object, not single-valued on the Fano plane. How should it be described in terms of combinatorics?

Author Contributions: Conceptualization, M. Malykh; methodology, A. Sevastianov; software, I. Dulatov; validation, M. Malykh and I. Dulatov; writing—original draft preparation, I. Dulatov; writing—review and editing, A. Sevastianov; visualization, I. Dulatov; supervision, M. Malykh; project administration, A. Zorin. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Declaration on Generative AI: The authors have not employed any Generative AI tools.

References

[1] M. Malykh, M. Gambaryan, O. Kroytor, and A. Zorin, "Finite Difference Models of Dynamical Systems with Quadratic Right-Hand Side," *Mathematics*, vol. 12, no. 1, p. 167, 2024. DOI: 10.3390/math12010167

[2] W. Kahan and Ren-Chang Li, "Unconventional Schemes for a Class of Ordinary Differential Equations – With Applications to the Korteweg–de Vries Equation," *Journal of Computational Physics*, vol. 134, pp. 316–331, 1997.

[3] J. M. Sanz-Serna, "An unconventional symplectic integrator of W. Kahan," *Applied Numerical Mathematics*, vol. 16, pp. 245–250, 1994.

[4] M. Petrera and Y. B. Suris, "On the Hamiltonian structure of Hirota-Kimura discretization of the Euler top," *Math. Nachr.*, vol. 283, no. 11, pp. 1654–1663, 2010. DOI: 10.1002/mana.200711162

[5] M. Petrera, A. Pfadler, Y. B. Suris, and Y. N. Fedorov, "On the Construction of Elliptic Solutions of Integrable Birational Maps," *Experimental Mathematics*, vol. 26, no. 3, pp. 324–341, 2017. DOI: 10.1080/10586458.2016.1166354

[6] M. Petrera, J. Smirin, and Y. B. Suris, "Geometry of the Kahan discretizations of planar quadratic Hamiltonian systems," *Proc. R. Soc. A*, vol. 475, p. 20 180 761, 2019. DOI: 10.1098/rspa.2018.0761

[7] E. Celledoni, R. I. McLachlan, B. Owren, and G. R. W. Quispel, "Geometric properties of Kahan's method," *J. Phys. A: Math. Theor.*, vol. 46, p. 025 201, 2013. DOI: 10.1088/1751-8113/46/2/025201

[8] J. Blanc and A. Calabri, "On degenerations of plane Cremona transformations," *Math. Z.*, vol. 282, pp. 223–245, 2016. DOI: 10.1007/s00209-015-1539-z

[9] K. A. Shramov, "Finite groups of automorphisms of algebraic and complex manifolds," Ph.D. dissertation, Steklov Mathematical Institute, 2020.

[10] F. Severi, *Lezioni di geometria algebrica*. Padova: Angelo Graghi, 1908.

[11] F. Kárteszi, *Introduction to Finite Geometries*. Budapest: Academiai Kiado, 1976.

[12] T. S. Developers, W. Stein, D. Joyner, D. Kohel, J. Cremona, and B. Eröcal, *SageMath, version 10.8*, 2025.

[13] M. D. Malykh and I. T. Dulatov, *Fano plane, version 1.0*, 2026.

[14] J. Rosanes, "Ueber diejenigen rationalen Substitutionen, welche eine rationale Umkehrung zulassen," *Journal für die reine und angewandte Mathematik*, vol. 73, pp. 101–114, 1870.

[15] B. C. Mlodziejowski, "Sur le théorème de décomposibilité des transformations de Cremona dans le plan," *Matematicheskii Sbornik*, vol. 29, pp. 269–275, 1915.

[16] J. Alonso, Y. B. Suris, and Kangning Wei, "Dynamical degrees of birational maps from indices of polynomials with respect to blow-ups I. General theory and 2D examples," *arXiv*, vol. 2303.15864, 2023.

[17] J. Alonso, Y. B. Suris, and Kangning Wei, "Dynamical degrees of birational maps from indices of polynomials with respect to blow-ups II. 3D examples," *arXiv*, vol. 2307.09939, 2023.

[18] W. Shiwei, A. V. Zorin, M. A. Konyaeva, M. D. Malykh, and L. A. Sevastianov, "On periodic approximate solutions of ordinary differential equations," *Zapiski nauchnyh seminarov POMI*, vol. 543, pp. 68–81, 2025.

[19] K. V. Rerikh, "Chew-Low equations as Cremona transformations structure of general intgrals," *Joint Institute for Nuclear Research, Dubna. Translated from Teoreticheskayai Matematicbeskaya Fizika*, vol. 0, no. 2, pp. 251–260, 1982.

[20] E. A. Ayrjan, M. M. Gambaryan, M. Malykh, and L. Sevastianov, "On the trajectories of dynamical systems with quadratic right-hand side computed by reversible difference schemes," *Zap. nauch. semin. POMI*, vol. 517, pp. 17–35, 2022.

Information about the authors

Dulatov, Ilshat T.—Phd Student of Department of Computational Mathematics and Artificial Intelligence of RUDN University (e-mail: dulatov-it@rudn.ru, ORCID: 0009-0009-2167-6177)

Malykh, Mikhail D.—DSc., Head of Department of Computational Mathematics and Artificial Intelligence of RUDN University; Senior Researcher of Joint Institute for Nuclear Research (e-mail: malykh-md@rudn.ru, phone: +7(495)9522572, ORCID: 0000-0001-6541-6603, ResearcherID: P-8123-20168, Scopus Author ID: 6602318510)

Sevastianov, Anton L.—Candidate of Physical and Mathematical Sciences, Department of Computational Mathematics and Artificial Intelligence of RUDN University (e-mail: alsevastyanov@gmail.com, ORCID: 0000-0002-0280-485X)

Zorin, Aleksander V.—DSc., Professor of Department of Computational Mathematics and Artificial Intelligence of RUDN University; Senior Researcher of Joint Institute for Nuclear Research (e-mail: zorin-av@rudn.ru, ORCID: 0000-0002-5721-4558, ResearcherID: AAH-4011-2019, Scopus Author ID: 57193219091)

RUSSIAN METADATA

УДК 512.624.3, 512.76, 512.542.74, 519.622.2

PACS 02.20.Hj 02.70.Bf

DOI: 10.22363/2658-4670-2026-34-2-201-213 EDN: JDOCAU

О квадратичных преобразованиях плоскости Фано

И. Т. Дулатов[1], М. Д. Малых[1,2], А. Л. Севастьянов[1], А. В. Зорин[1]

[1] *Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация*

[2] *Объединённый институт ядерных исследований, ул. Жолио-Кюри, д. 6, Дубна, 141980, Российская Федерация*

Аннотация. Указано значение исследования квадратичных преобразований Кремоны над алгебраически незамкнутыми полями для теории разностных схем Кагана для динамических систем с квадратичной правой частью. Рассмотрены преобразования Кремоны проективной плоскости над полем Галуа размера 2, то есть плоскости Фано. Для описания квадратичных преобразований Кремоны использована нотация, предложенная Я. Розанесом. Описана связь между квадратичными преобразованиями и матричными пучками. Квадратичные преобразований без особых точек названы регулярными. В системе Sage реализованы процедуры, которые переводят квадратичное преобразование в перестановку 7 точек плоскости Фано (элемент симметрической группы S_7) и перестановку 7 точек плоскости в квадратичное преобразование. Перебором всех квадратичных преобразований доказано, что регулярные квадратичные преобразования порождают всю группу перестановок плоскости Фано. Эта теорема является аналогом теоремы Нетера над алгебраически незамкнутым полем. Доказано, что регулярные квадратичные преобразования Кремоны имеют чётный порядок, и дано описание соответствующих им перестановок группы S_7. Показано, что перестановке всегда отвечает некоторое квадратичное преобразование Кремоны, но не всегда это преобразование является регулярным. Получающиеся при этом квадратичные преобразования можно дополнить правилом, которое разрешает неоднозначность в фундаментальных точках. Приведён пример квадратичного преобразования, для которого такое разрешение сделать нельзя. Отсюда получена естественная классификация квадратичных преобразований плоскости Фано.

Ключевые слова: проективная плоскость, поля Галуа, преобразования Кремоны, метод Кагана

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2026, 34 (2) 214–225
http://journals.rudn.ru/miph

Research article

UDC 519.872, 519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2026-34-2-214-225 EDN: JFJZIZ

Calculation of modified Hamiltonian in Sage

Mikhail D. Malykh[1,2], Marina A. Konyaeva[1]

[1] *RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation*

[2] *Joint Institute for Nuclear Research, 6 Joliot-Curie St, Dubna, Moscow Region, 141980, Russian Federation*

(received: November 7, 2025; revised: January 15, 2026; accepted: January 20, 2026)

Abstract. The algebraic properties of difference approximations of Hamiltonian systems are investigated. Symplectic schemes exactly preserve linear and quadratic integrals by virtue of Cooper's theorem, but not the total mechanical energy of nonlinear systems. However, it is known that instead of energy, symplectic difference schemes preserve with a given order of approximation a quantity that goes over into the Hamiltonian as the time step tends to zero. The paper presents an algorithm for calculating such a modified Hamiltonian and its implementation in the Sage computer algebra system for a given symplectic difference scheme, the required order of energy conservation, and the Hamiltonian of the original mechanical system. The program successfully reproduces formulas previously derived manually, which confirms its consistency. Numerical experiments show that solutions obtained using symplectic schemes closely coincide with the level lines of the modified Hamiltonian, which emphasizes its role in preserving the qualitative behavior of the system during numerical integration over large time intervals

Key words and phrases: symplectic difference schemes, Hamiltonian systems

For citation: M. D. Malykh, M. A. Konyaeva "Calculation of modified Hamiltonian in Sage," *Discrete and Continuous Models and Applied Computational Science*, vol. 34, no. 2, pp. 214–225, 2026. DOI: 10.22363/2658-4670-2026-34-2-214-225. EDN: JFJZIZ.

1. Introduction

Explicit Runge–Kutta difference schemes do not preserve the most important properties of the dynamical systems they approximate. The most painful case is Hamiltonian dynamical systems, where the application of explicit Runge–Kutta schemes leads to the violation of the conservation law of the total mechanical energy [1]. In the 1990s, the concept of geometric integrators emerged, i.e. difference schemes that inherit the symplectic structure of Hamiltonian systems [1–4]. These difference schemes preserve exactly linear and quadratic integrals of motion [1, Cooper's theorem], but do not preserve more complicated integrals of motion of the system [1, 5].

On the other hand, we can construct difference schemes that precisely preserve motion integrals but do not preserve a symplectic structure. The first such scheme for the many-body problem was proposed by Greenspan [6–9], there are such schemes of arbitrarily large order [10].

For symplectic integrator, the change in the integral of motion on an approximate solution approximating a dynamic system with order r also begins with a term of order r or higher. Using Richardson's method [11–17], it is easy to verify that this term has a non-zero coefficient in the overwhelming majority of nonlinear problems. Therefore, in particular, the change in energy turns out to be quite noticeable in the plots obtained from approximate solutions. We will illustrate this with an example in Section 2.

Nevertheless, the first computer experiments were carried out with Hamiltonian systems having one degree of freedom. In these systems, it was found every time that the points of the approximate solution on the phase plane pq are aligned along some line that is close to the line of constant energy level, but slightly different from it. This allows us to hope that the points of the approximate solution lie on the level lines of the modified Hamiltonian, that is, some function $\tilde{H}(p, q, dt)$, depending on the step dt and transforming into the Hamiltonian $H(p, q)$ of the original system at $dt \to 0$.

In Ref. [18, n. 10.1.2] it was proved that for any symplectic difference scheme approximating a given Hamiltonian system of order r, and any order $m > r$, one can compute a modified Hamiltonian

$$H_m(p, q, dt) = H(p, q) + H^{(r)}(p, q)dt^r + \cdots + H^{(m-1)}dt^{m-1},$$

which is preserved on the approximate solution of order m. The problem of finding H_m is solved symbolically. We formulate it as a computer algebra problem in Section 5 and then present our solution as software for Sage.

Automating the calculation of the modified Hamiltonian allows us to clarify, using specific examples, what exactly we see when we think that the points of the approximate solution line up along some curves. In Section 6 we will see that the points line up along the level lines of the modified Hamiltonian H_{r+1} and it takes considerable effort to discern the change in H_{r+1} on the approximate solution.

2. Conservation of mechanical energy on symplectic schemes

Consider a Hamiltonian system

$$\frac{dp}{dt} = \frac{\partial H}{\partial q}, \quad \frac{dq}{dt} = -\frac{\partial H}{\partial p} \tag{1}$$

on a fixed interval $0 < t < T$. Let the approximate solution

$$(p_0, q_0), (p_1, q_1), \ldots, (p_N, q_N)$$

be found by some difference scheme of order r with step $dt = T/N$, and let $p(t), q(t)$ be the exact solution of this system corresponding to the initial conditions

$$p(0) = p_0, \quad q(0) = q_0.$$

Then

$$p_N - p(T) = \mathcal{O}(dt^r), \quad q_N - q(T) = \mathcal{O}(dt^r).$$

The law of energy conservation is satisfied exactly on the exact solution, therefore the change in energy in the interval $0 < t < T$ is equal to zero:

$$H\big(p(T), q(T)\big) - H(p_0, q_0) = 0.$$

On the approximate solution, the change in energy can be represented as follows:

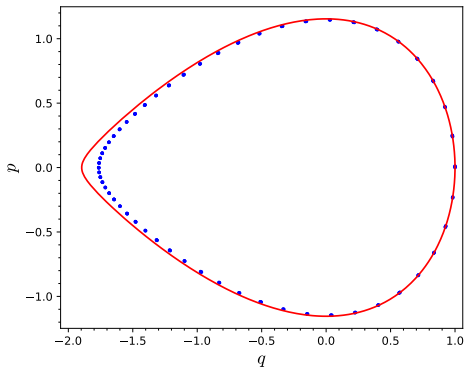

Figure 1. Trajectory in the phase plane for example 1: exact orbit (solid line) and approximate orbit (dots).

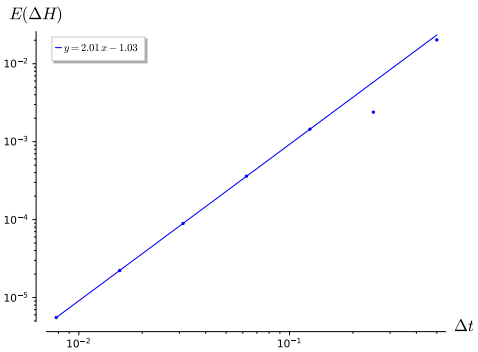

Figure 2. Richardson diagram for energy change in example1

that is, the change in energy begins with a term of order dt^r or higher

$$\Delta H = cdt^r + \dots .$$

We emphasize that here we consider T to be a fixed value, and $dt = T/N$ to be a variable.

The symplecticity of the difference scheme does not contribute to increasing the order of approximation in this equality.

Example 1. For example, let us take the Hamiltonian

$$H = \frac{p^2}{2} + \frac{q^2}{2} + a\,q^3$$

and let us assume for definiteness that $a = 0.166$. With this choice of parameter and initial conditions $p = 0, q = 1$, we obtain a closed trajectory, which is noticeably different from an ellipse. Therefore, we will use these values further. In Figure 1, both the exact orbit , the energy level line, and the approximate orbit found using the midpoint scheme in our fdm for sage [19, 20] system are clearly visible. This scheme is the simplest symplectic scheme of the second order [1]. From the Richardson diagram for the energy change (Figure 2), which is a standard tool in our system for estimating the errors of numerical methods [19], it is clearly seen that

$$\Delta H = cdt^2 + \dots$$

and that $\lg c \simeq 1.03$. Thus, the energy change is a second-order quantity as for any other second-order scheme. However, Figure 1 shows that the points of the approximate solution in the phase plane line up along some line different from the level line for H.

The situation described in Example 1 is typical and this allows us to hope that some expression $H(p, qdt)$, depending on the step dt, is preserved on the approximate solution found using the symplectic scheme:

$$H(p_k, q_k, dt) = H(p_0, q_0, dt).$$

The search for such a modified Hamiltonian must be preceded by a more general study of modified equations.

3. Modified differential equation

Let us consider the system of ordinary differential equations

$$\frac{dx_i}{dt} = f_i(x_1, \dots, x_n), \quad i = 1, \dots, n,$$

which for brevity will be written as

$$\frac{d\mathfrak{x}}{dt} = \mathfrak{f}(\mathfrak{x}). \tag{2}$$

Let dt be a positive constant, which we will interpret as a time step, and let $\hat{\mathfrak{x}}$ be the value of variable $\mathfrak{x}$ at time $t + dt$. The solution of Eq. (1) can be expanded in a Taylor series

$$\begin{aligned} \hat{\mathfrak{x}} &= \mathfrak{x} + \dot{\mathfrak{x}}dt + \frac{1}{2!}\ddot{\mathfrak{x}}dt^2 + \frac{1}{3!}\dddot{\mathfrak{x}}dt^3 + \dots \\ &= \mathfrak{x} + \mathfrak{f}dt + \frac{1}{2!}D(\mathfrak{f})dt^2 + \frac{1}{3!}D^2(\mathfrak{f})dt^3 + \dots, \end{aligned} \tag{3}$$

where the differentiation D is expressed as

$$D = \sum_{i=1}^{n} f_i \frac{\partial}{\partial x_i}.$$

By a difference scheme we mean any system of algebraic equations that relate $\mathfrak{x}$ and $\hat{\mathfrak{x}}$:

$$F_i(x_1, \dots, x_n, \hat{x}_1, \dots, \hat{x}_n, dt), \quad i = 1, \dots, n,$$

which for brevity will be written in the form

$$\mathfrak{F}(\mathfrak{x}, \hat{\mathfrak{x}}, dt) = 0. \tag{4}$$

Definition 1. A difference scheme (4) is said to approximate the differential equation (1) with order of approximation r if substituting in (4) for $\hat{\mathfrak{x}}$ its Taylor series expression (3) yields

$$\mathfrak{F}(\mathfrak{x}, \hat{\mathfrak{x}}, dt) = \mathcal{O}(dt^{r+1}).$$

Recall that what is said in this definition is enough to prove that the error in the determination of $\mathfrak{x}$ in one step is of order dt^{r+1}, and in N steps is of order dt^r [1, 12]. We can save the definition 1 for the case when the right-hand side (1) depends on dt.

Definition 2. A difference scheme (4) is said to approximate the differential equation

$$\frac{d\mathfrak{x}}{dt} = \mathfrak{f}_m(\mathfrak{x}, dt) \tag{5}$$

with order of approximation m if substituting for $\hat{\mathfrak{x}}$ in (4) its Taylor series expression

$$\hat{\mathfrak{x}} = \mathfrak{x} + \mathfrak{f}_m dt + \frac{1}{2!}D_r(\mathfrak{f}_r)dt^2 + \frac{1}{3!}D_r^2(\mathfrak{f}_r)dt^3 + \dots,$$

gives

$$\mathfrak{F}(\mathfrak{x}, \hat{\mathfrak{x}}, dt) = \mathcal{O}(dt^{m+1}).$$

Here and below

$$D = \sum_{i=1}^{n} f_{m,i} \frac{\partial}{\partial x_i}.$$

Statement 1. A difference scheme of order r and a natural number $m > r$ are given. It is required to find a differential equation approximated by this difference scheme with order m.

Analytical methods for studying finite-difference equations are much less developed than methods for studying differential equations. Having learned to solve the problem 1 symbolically, we can reduce the study of the difference scheme to the study of a "modified" differential equation.

Definition 3. Let the difference scheme (4) approximate the differential equation (1) with order r, and the differential equation (5) with order $m > r$. Then we will call the differential equation (5) a modified differential equation of order m, indicating, if necessary, that the original equation is being modified.

4. Algorithm for finding a modified equation

Let the difference scheme approximate the differential equation (1) with order r. Let us find the modified equation

$$\frac{d\mathfrak{x}}{dt} = \mathfrak{f}_{r+1}(\mathfrak{x}, dt) \tag{6}$$

of order $r + 1$. Let us find its right-hand side in the form

$$\mathfrak{f}_{r+1}(\mathfrak{x}, dt) = \mathfrak{f}(\mathfrak{x}) + \mathfrak{g}(\mathfrak{x})dt^r.$$

The solution of the modified equation is given by the Taylor series

$$\hat{\mathfrak{x}} = \mathfrak{x} + (\mathfrak{f} + \mathfrak{g}(\mathfrak{x})dt^r)dt + \frac{1}{2!}D_{r+1}(\mathfrak{f}(\mathfrak{x}) + \mathfrak{g}(\mathfrak{x})dt^r)dt^2 + \dots,$$

where

$$D_{r+1} = \sum_{i=1}^{n}(f_i + g_i dt^r)\frac{\partial}{\partial x_i}.$$

After removing the brackets, we have

$$\begin{aligned}\hat{\mathfrak{x}} = \mathfrak{x} + D(\mathfrak{f})dt + \frac{1}{2!}D(\mathfrak{f})dt^2 + \dots + \frac{1}{r!}D^r(\mathfrak{f})dt^r \\ + \left(\frac{1}{(r+1)!}D^{r+1}(\mathfrak{f}) + \mathfrak{g}(\mathfrak{x})\right)dt^{r+1} + \mathcal{O}(dt^{r+2})\end{aligned}$$

This expression up to terms of the order $r + 1$ coincides with the Taylor series for the solution of the original equation (1), so its substitution into the difference scheme (4) for an arbitrary $\mathfrak{g}$ yields n series in powers of dt that start with dt^{r+1}. The coefficients at dt^{r+1} are linear functions of $\mathfrak{g}$, so the condition

$$\mathfrak{F}(\mathfrak{x}, \hat{\mathfrak{x}}, dt) = \mathcal{O}(dt^{r+2})$$

gives exactly n linear equations for finding $g_1, \dots, g_n$. From this system, $\mathfrak{g}$ is uniquely determined as a set of rational functions $x_1, \dots, x_n$.

Thus, if the difference scheme approximates the original differential equation (1) with order r, then this scheme approximates with order $r + 1$ the equation (6), the right-hand side of which is a polynomial of degree r with respect to dt. To find the modified equation, it is sufficient to compose and solve the appropriate system of linear algebraic equations.

Now it is easy to find the modified equation

$$\frac{d\mathfrak{x}}{dt} = \mathfrak{f}_{r+2}(\mathfrak{x}, dt)$$

of order $r+2$. We will seek for its right-hand side in the form

$$\mathfrak{f}_{r+2} = \mathfrak{f}_{r+1}(\mathfrak{x}, dt) + \mathfrak{g}(\mathfrak{x})dt^{r+1},$$

where $\mathfrak{g}$ is the next correction to the right-hand side of the modified equation, subject to determination.

The solution of the modified equation is given by the Taylor series

$$\hat{\mathfrak{x}} = \mathfrak{x} + D_{r+1}(\mathfrak{f}_{r+1})dt + \cdots + \frac{1}{(r+1)!}D_{r+1}^{r+1}(\mathfrak{f}_{r+1})dt^{r+1} + \left(\frac{1}{(r+2)!}D^{r+2}(\mathfrak{f}_{r+1}) + \mathfrak{g}(\mathfrak{x})\right)dt^{r+2} + \mathcal{O}(dt^{r+3}).$$

This expression up to $(r+2)$-th order terms coincides with the Taylor series for solving the modified equation (6), so substituting it into the difference scheme (4) for an arbitrary $\mathfrak{g}$ yields n series in powers of dt that start with dt^{r+2}. The coefficients of dt^{r+2} are linear functions of $\mathfrak{g}$, so the condition

$$\mathfrak{F}(\mathfrak{x}, \hat{\mathfrak{x}}, dt) = \mathcal{O}(dt^{r+3})$$

gives exactly n linear equations for finding $g_1, \ldots, g_n$. Thus, we find f_{r+2} as a polynomial of degree $r+1$ in dt.

Proceeding in this way, we can reach any given order m.

Theorem 1. *Let the difference scheme approximate the differential equation* (1) *with order r. Then this scheme approximates with order $m > r$ the equation*

$$\frac{d\mathfrak{x}}{dt} = \mathfrak{f}_m(\mathfrak{x}, dt)$$

the right-hand side of which is a polynomial of degree $m-1$ in dt.

5. The problem of finding a modified Hamiltonian

We apply the developed theory to the Hamiltonian system (1). Any difference scheme for this system can be written in the form

$$f(\hat{p}, \hat{q}, p, q, dt) = 0, \quad g(\hat{p}, \hat{q}, p, q, dt) = 0. \tag{7}$$

A difference scheme is called symplectic if there exists a function S such that the equality

$$\hat{p}d\hat{q} = pdq + dS$$

is satisfied on the manifold (7).

Using the method described above, for any difference scheme we can construct a modified system

$$\frac{dp}{dt} = P_m(p, q, dt), \quad \frac{dq}{dt} = Q_m(p, q, dt), \tag{8}$$

approximated by the difference scheme (7) with order $m+1$.

If the difference scheme is symplectic [1], then the modified system is Hamiltonian [18, n. 10.1.2]. We denote its Hamiltonian as H_m. The energy conservation law for the system (8) is satisfied exactly:

$$H_m\big(\tilde{p}(T), \tilde{q}(T)\big) = H(p_0, q_0),$$

where $\tilde{p}(t), \tilde{q}(t)$ is the exact solution of the modified system (8). Therefore, on the approximate solution, the change in energy over the interval $0 < t < T$ is equal to

$$\begin{aligned}\Delta H_m &= H_m(p_N, q_N, dt) - H_m(p_0, q_0, dt)\\ &= H_m(p_N, q_N, dt) - H_m\big(\tilde{p}(T), \tilde{q}(T), dt\big) = \mathcal{O}(dt^m),\end{aligned}$$

i.e., the change in energy starts with a term of order dt^m or higher

$$\Delta H_m = \mathcal{O}(dt^m).$$

Thus, the algorithm for finding a modified equation of a given order (Section 4) allows us to find a function $H_m(p, q, dt)$ that, although not exactly preserved on the approximate solution, is preserved with any given order.

Then the energy conservation law of the modified system must be satisfied on the considered scheme with the order m. Therefore, it must be true

$$\Delta H_m = H_m(p_N, q_N, dt) - H_m(p_0, q_0, dt) = \mathcal{O}(dt^m).$$

Thus, the Hamiltonian of the modified system is the function that is preserved on the difference scheme with the given order m. We will call such a function the modified Hamiltonian of the m-th order.

Statement 2. A symbolic expression H, containing two independent variables p, q, a symplectic difference scheme of order r, and an integer $m > r$ are given. It is required to find a modified Hamiltonian of order m, that is, a symbolic expression of the form

$$H_m(p, q) = H(p, q) + H^{(r)}(p, q)dt^r + \cdots + H^{(m-1)}dt^{m-1},$$

which is preserved with order m.

In Ref. [18, n. 10.1.2], expressions for modified Hamiltonians of order $r+1$ were found for a separate scheme. We implemented an algorithm for solving a more general problem 1 in the Sage computer algebra system [21] and tested its operation on several examples; the program is available in the public repository [22].

Example 2. To verify our program, consider the one-stage split Runge–Kutta scheme

$$\hat{p} - p = -U'(\hat{q})dt, \quad \hat{q} - q = T'(p)dt$$

for a system with Hamiltonian of the form $H = T(p) + U(q)$. This scheme is symplectic and has the first order of approximation [23, p. 3.1]. We define it in our system as follows:

```
var('p,q,dt,pp,qq')
T=function('T')(p)
U=function('U')(q)
H=T+U
scheme=[pp-p + diff(H,q).subs(q=qq)*dt, qq-q-diff(H,p)*dt]
```

Our function `mod_ham` has three arguments: the Hamiltonian, the scheme, and the order. In the case under consideration, the function `mod_ham(H,scheme,2)` returns the modified Hamiltonian

$$H_2 = T(p) + U(q) + \frac{dt}{2}T'(p)U'(q)$$

of the second order, and `mod_ham(H2,scheme,3)` returns the modified Hamiltonian

$$H_3 = T(p) + U(q) + \frac{dt}{2}T'(p)U'(q) + \frac{dt^2}{12}\left(T''(p)U'(q)^2 + T'(p)^2U''(q)\right)$$

of the third order. This coincides with the expressions specified in [18, n. 10.1.1].

Example 3. Midpoint scheme

$$\hat{p} - p = \frac{\partial H}{\partial q}(\overline{p}, \overline{q})dt, \quad \hat{q} - q = -\frac{\partial H}{\partial p}(\overline{p}, \overline{q})dt,$$

where

$$\overline{p} = \frac{\hat{p} + p}{2}, \quad \overline{q} = \frac{\hat{q} + q}{2},$$

is a symplectic scheme and has 2nd order approximation [1, 5]. We define the scheme in Sage as follows:

```
scheme=[pp-p + diff(H,q).subs(q=(q+qq)/2)*dt, \
qq-q-diff(H,p).subs(p=(p+pp)/2)*dt]
```

The modified 3rd order Hamiltonian is

$$H_3 = T(p) + U(q) - \frac{dt^2}{24}\left(T'(p)^2U''(q) + U'(q)^2T''(p)\right) + \mathcal{O}(dt^3)$$

and coincides with the modified 4th order Hamiltonian.

Remark 1. It is worth noting that for a harmonic oscillator

$$H = \frac{p^2}{2m} + \frac{q^2}{2k}$$

it is true

$$T'(p)^2U''(q) + U'(q)^2T''(p) = \frac{p^2}{m^2k} - \frac{q^2}{mk^2} = \frac{2}{mk}H.$$

Thus, the original Hamiltonian is preserved. This is as it should be according to Cooper's theorem [1].

Potentially, our program allows searching for Hamiltonians of any order. However, for the midpoint scheme, it is not possible to calculate a Hamiltonian of the 5th order, since the Sage system cannot calculate the appropriate integrals in the general form.

However, our program supports the specification of the Hamiltonian as any symbolic expression, so when considering particular cases, a specific expression for the energy can be used instead of the general expression. In this case, the integrals can be calculated and we get the opportunity to trace the role of the various terms in (8), not restriciting ourselves to the first one, as is usually done.

6. System with cubic Hamiltonian

Consider the midpoint scheme for the system with cubic Hamiltonian $H = \frac{p^2+q^2}{2} + aq^3$, the consideration of which we began in the Example 1. We define this scheme in the usual way:

```
var('p,q,dt,pp,qq,a')
H=p^2 / 2 + q^2 / 2 + a * q^3
scheme=[pp-p + diff(H,q).subs(q=(q+qq)/2)*dt, \
qq-q-diff(H,p).subs(p=(p+pp)/2)*dt]
```

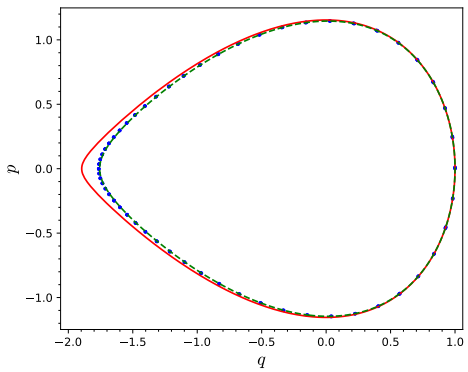

Figure 3. Phase trajectory for $a = 0.166$

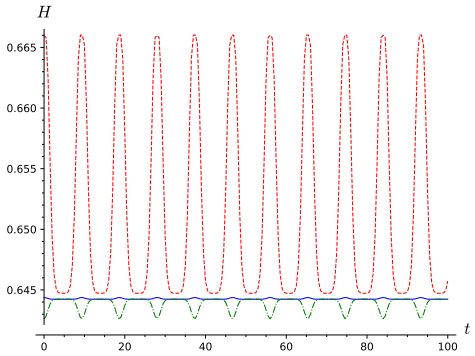

Figure 4. Time dependence of H (dashed), H_3 (dotdash) and H_5 (solid) for $a = 0.166$

First, our function allows us to find

$$H_3 = -\frac{1}{24}\left(9\,a^2q^4 + 6\,ap^2q + 6\,aq^3 + p^2 + q^2\right)dt^2 + \frac{1}{2}\,p^2 + \frac{1}{2}\,q^2 + aq^3,$$

then $H_4 = H_3$, and then

$$H_5 = \frac{1}{160}\left(54\,a^3q^5 + 45\,a^2q^4 + 10\,ap^2q + 12\,aq^3 + \left(30\,a^2p^2 + 1\right)q^2 + p^2\right)dt^4 - \frac{1}{24}\left(9\,a^2q^4 + 6\,ap^2q + 6\,aq^3 + p^2 + q^2\right)dt^2 + \frac{1}{2}\,p^2 + \frac{1}{2}\,q^2 + aq^3.$$

Figure 3 shows the phase trajectory passing through the point $(q, p) = (1, 0)$. The dots represent the approximate solution found using the midpoint scheme. The true trajectory, i.e. the level line of the Hamiltonian H, is shown by the red solid line, and the level line H_3 is shown by the green dotted line. It is clearly seen that with graphical accuracy the points of the approximate solution lie on the curve $H_3(p, q) = \text{const}$, and not on $H(p, q) = \text{const}$. Figure 4 shows the change in the Hamiltonians H, H_3, and H_5 on the approximate solution. It is clearly seen that the amplitude of the change in H is several hundredths, H_3 is several thousandths, and H_5 is not visible in the plot at all.

7. Results

The presented program allows finding a symbolic expression for a modified Hamiltonian of a given order $m > r$ for a given symplectic difference scheme of order r. We tested our program on a split Runge–Kutta scheme, which was previously investigated without involving computer algebra systems.

Then we conducted a series of numerical experiments with a midpoint scheme for a cubic Hamiltonian. In these experiments, it turned out every time that with graphical accuracy the points of the approximate orbit lie on the level lines of the modified Hamiltonian of the 3rd order. It is possible to see that this Hamiltonian is not preserved exactly only on a specially plotted change in energy over time. The change in the modified Hamiltonian of the 5th order over the entire observation period lies within the limits of the calculation error.

8. Discussion

From a practical point of view, this is very successful. The presented algorithm allows an analytical description of the closed line, which inevitably contains all points of the approximate solution, i.e. makes it possible to describe qualitatively the behavior of the approximate solution and its difference from the exact solution. The lines that we see in computer experiments similar to the one described in Example 1 are the level lines of the modified Hamiltonian, the order of which is greater by one than the order of the difference scheme.

However, from the point of view of the theory of geometric integrators, everything looks much more complicated: the proposed algorithm allows us to find Hamiltonians that are preserved with a given order on the approximate solution, while the question of finding the modified Hamiltonian $\tilde{H}(p, q, dt)$, i.e. an expression that is exactly preserved on the approximate solution, remains open.

Author Contributions: Conceptualization, M. Malykh; software, M. Konyaeva; writing, M. Malykh. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Declaration on Generative AI: The authors have not employed any Generative AI tools.

References

[1] E. Hairer, G. Wanner, and C. Lubich, *Geometric Numerical Integration*. Berlin, Heidelberg: Springer, 2006.

[2] Y. B. Suris, "The Canonicity of Mappings Generated by Runge-Kutta Type Methods when Integrating the Systems $\ddot{x} = f(x)$," *Computational Mathematics and Mathematical Physics*, vol. 29, no. 2, pp. 202–211, 1989.

[3] Y. B. Suris, "Hamiltonian Runge–Kutta Type Methods and Their Variational Interpretation," *Mat. Model.*, no. 2, pp. 78–87, 1990.

[4] B. Batgerel, E. G. Nikonov, and I. V. Puzynin, "A Procedure for Constructing Simplectic Numerical Schemes for Solving of Hamiltonian Systems of Equations," *RUDN Journal. Ser. MIF.*, no. 1, pp. 41–58, 2016.

[5] V. P. Gerdt, M. D. Malykh, L. A. Sevastianov, and Y. Ying, "On the properties of numerical solutions of dynamical systems obtained using the midpoint method," *Discrete and Continuous Models and Applied Computational Science*, vol. 27, no. 3, pp. 242–262, 2019. DOI: 10.22363/2658-4670-2019-27-3-242-262

[6] D. Greenspan, "Completely conservative, covariant numerical methodology," *Computers & Mathematics with Applications*, vol. 29, no. 4, pp. 37–43, 1995. DOI: 10.1016/0898-1221(94)00236-E

[7] D. Greenspan, "Completely conservative, covariant numerical solution of systems of ordinary differential equations with applications," *Rendiconti del Seminario Matematico e Fisico di Milano*, vol. 65, pp. 63–87, 1995. DOI: 10.1007/BF02925253

[8] D. Greenspan, *N-Body Problems and Models*. World Scientific, 2004.

[9] E. Graham, G. Jelenić, and M. A. Crisfield, "A note on the equivalence of two recent time-integration schemes for N-body problems," *Communications in Numerical Methods in Engineering*, vol. 18, pp. 615–620, 2002. DOI: 10.1002/cnm.520

[10] Y. Ying, A. Baddour, V. P. Gerdt, M. D. Malykh, and L. A. Sevastianov, "On the Quadratization of the Integrals for the Many-Body Problem," *Mathematics*, vol. 9, no. 24, p. 3208, 2021. DOI: 10.3390/math9243208

[11] L. F. Richardson and J. A. Gaunt, "The deferred approach to the limit," *Phil. Trans. A*, vol. 226, pp. 299–349, 1927.

[12] N. N. Kalitkin, *Numerical methods*. Moscow: Nauka, 1979, In Russian.

[13] N. N. Kalitkin, A. B. Al'shin, E. A. Al'shina, and B. V. Rogov, *Calculations on quasi-uniform grids*. Moscow: Fizmatlit, 2005, In Russian.

[14] A. A. Belov, N. N. Kalitkin, and I. P. Poshivaylo, "Geometrically adaptive grids for stiff Cauchy problems," *Doklady Mathematics*, vol. 93, no. 1, pp. 112–116, 2016. DOI: 10.1134/S1064562416010129

[15] A. A. Belov and N. N. Kalitkin, "Nonlinearity Problem in the Numerical Solution of Superstiff Cauchy Problems," *Mathematical Models and Computer Simulations*, vol. 8, no. 6, pp. 638–650, 2016. DOI: 10.1134/S2070048216060065

[16] A. A. Belov, N. N. Kalitkin, P. E. Bulatov, and E. K. Zholkovskii, "Explicit methods for integrating stiff Cauchy problems," *Doklady Mathematics*, vol. 99, no. 2, pp. 230–234, 2019. DOI: 10.1134/S1064562419020273

[17] A. Baddour and M. D. Malykh, "Richardson-Kalitkin method in abstract description," *Discrete and Continuous Models and Applied Computational Science*, vol. 29, no. 3, pp. 271–284, 2021.

[18] J. M. Sanz-Serna and M. P. Calvo, *Numerical Hamiltonian Problems*. London, Glasgow, New York, Tokyo, Melbourne, Madras: CHAPMAN & HALL, 1994.

[19] A. Baddour, M. M. Gambaryan, L. Gonzalez, and M. D. Malykh, "On Implementation of Numerical Methods for Solving Ordinary Differential Equations in Computer Algebra Systems," *Programming and Computer Software*, vol. 5, pp. 412–422, 2023. DOI: 10.1134/S0361768823020044

[20] M. D. Malykh, W. Shiwei, and Y. Ying, "On a set of tests for numerical methods of integrating differential equations, based on the Calogero system," *Discrete and Continuous Models and Applied Computational Science*, vol. 31, no. 4, pp. 387–398, 2023.

[21] W. A. Stein et al., *Sage Mathematics Software (Version 10.5)*, The Sage Development Team, 2024.

[22] M. D. Malykh, *FDM: a new package for numerical solution of ordinary differential equations in Sage*, RUDN, 2025.

[23] M. N. Gevorkyan, "Specific implementations of symplectic numerical methods," *Bulletin of the Peoples Friendship University of Russia. Series: Mathematics informatics physics.*, no. 1, pp. 77–89, 2013.

Information about the authors

Malykh, Mikhail D.—Doctor of Physical and Mathematical Sciences, head of the department of Mathematical Modeling and Artificial Intelligence, RUDN University, research fellow of MLIT JINR (Dubna) (e-mail: malykh-md@rudn.ru, ORCID: 0000-0001-6541-6603, ResearcherID: P-8123-2016, Scopus Author ID: 6602318510)

Konyaeva, Marina A.—Student of the chair of Mathematical Modeling and Artificial Intelligence, RUDN University (e-mail: 1032217044@pfur.ru)

RUSSIAN METADATA

УДК 519.872, 519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2026-34-2-214-225 EDN: JFJZIZ

Вычисление модифицированного гамильтониана в Sage

М. Д. Малых[1, 2], М. А. Коняева[1]

[1] *Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация*

[2] *Объединённый институт ядерных исследований, ул. Жолио-Кюри, д. 6, Дубна, Московская обл., 141980, Российская Федерация*

Аннотация. Исследуются алгебраические свойства разностных аппроксимаций гамильтоновых системах. Симплектические схемы точно сохраняют линейные и квадратичные интегралы в силу теоремы Купера, но не полную механическую энергию нелинейных систем. Однако известно, что вместо энергии симплектические разностные схемы сохраняют с заданным порядком аппроксимации величину, которая переходит в гамильтониан при стремлении шага по времени к нулю. В работе представлен алгоритм вычисления такого модифицированного гамильтониана и его реализация в системе компьютерной алгебры Sage по заданной симплектической разностной схемы, требуемому порядку сохранения энергии и гамильтониану исходной механической системы. Программа успешно воспроизводит формулы, ранее выведенные вручную, что подтверждает её состоятельность. Численные эксперименты показывают, что решения, полученные с помощью симплектических схем, близко совпадают с линиями уровня модифицированного гамильтониана, что подчеркивает его роль в сохранении качественного поведения системы при численном интегрировании на больших временных интервалах.

Ключевые слова: симплектические разностные схемы, Гамильтоновы системы

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2026, 34 (2) 226–240
http://journals.rudn.ru/miph

Research article

UDC 519.872, 519.217
DOI: 10.22363/2658-4670-2026-34-2-226-240 EDN: JFRCXF

On the behavior of orbits of Vanhaecke system on integral surfaces

Shiwei Wang[1], Mikhail D. Malykh[1,2], Leonid A. Sevastianov[1], Alexander V. Zorin[1]

[1] *RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation*

[2] *Joint Institute for Nuclear Research, 6 Joliot-Curie St, Dubna, 141980, Russian Federation*

(received: November 7, 2025; revised: December 20, 2025; accepted: January 10, 2026)

Abstract. In the 1990s, P. Vanhecke described a Hamiltonian system with two degrees of freedom and a polynomial Hamiltonian integrable in Abelian functions of two variables. This system provides a convenient example of an integrable system (in the sense of Liouville) in which integral curves are wound on a two-dimensional manifold, an algebraic surface in a 4-dimensional phase space. We show that all necessary calculations can be performed in the Sage system. The results of numerical experiments performed in our package FDM for Sage are presented. It is shown that orbits of Vanhecke system can be divided into two classes, periodic and non-periodic. The points of the periodic orbits corresponding to solutions with the same period form an equiperiodic surface. There are infinitely many different algebraic equiperiodic surfaces. The behavior of nonperiodic orbits in numerical experiments is determined not so much by the rationality or irrationality of the ratio of two periods of Abelian functions, but by the possibility of approximating this number with a rational fraction with high accuracy.

Key words and phrases: Abelian functions, dynamical systems, completely integrable Hamiltonian systems

For citation: S. Wang, M. D. Malykh, L. A. Sevastianov, A. V. Zorin "On the behavior of orbits of Vanhaecke system on integral surfaces," *Discrete and Continuous Models and Applied Computational Science*, vol. 34, no. 2, pp. 226–240, 2026. DOI: 10.22363/2658-4670-2026-34-2-226-240. EDN: JFRCXF.

1. Introduction

The idea of the motion nature of conservative dynamical systems was formed in the 1970s [1]. A dynamical system is called transitive if almost any of its orbits fills the entire phase space densely. Transitivity of the system is prevented by conservation laws that hold the orbit on some integral manifold. Therefore, the study of a conservative dynamical system splits into two parts: 1) finding all conservation laws by analytical methods; 2) studying a dynamical system whose restriction to integral manifolds is transitive by statistical methods.

Let us assume that the dynamic system is described by a system of differential equations

$$\frac{d\mathfrak{x}}{dt} = \mathfrak{f}(\mathfrak{x}), \tag{1}$$

where $\mathfrak{x} \in \mathbb{R}^n$, and let it possess r conservation laws

$$h_i(\mathfrak{x}) = c_i, \quad i = 1, \dots, r, \tag{2}$$

where $c_1, \dots, c_r$ are integration constants. The orbit O, or the trajectory of the system motion in the phase space $\mathbb{R}^n$, is parametrically respresentable as

$$\mathfrak{x} = \mathfrak{x}(t), \quad t \in \mathbb{R}.$$

However, this does not mean that the orbit O is a line, i.e. a manifold of dimension 1, or even that O is a closed set. In fact, the closure of the orbit O in $\mathbb{R}^n$ may be greater than O, but it cannot go beyond the integral manifold V defined by the equations (2). A dynamical system (1) is called transitive on the manifold V if for almost any orbit $O \subset V$ we have $\overline{O} = V$. If this equality is satisfied, then all conservation laws have been found.

The justification for the existence of a set of conservation laws (2) that is complete in the sense that the dynamical system is transitive, or even better, ergodic on the manifold (2), essentially depends on the choice of the class to which the functions h_i belong. Back in the 1880s, Bruns [2, 3] developed a method for finding all integrals of motion in the class of algebraic functions for the many-body problem. This method can be extended to some Hamiltonian systems [4] and to the dynamics of a top [5, 6]. However, this class is too narrow to single out a manifold on which the system will be transitive. For example, the predator-prey system has a transcendental integral that prevents the orbit of the system from filling the entire phase plane. Another class considered in the 20th century was the class of integrals that are meromorphic in a certain sense; this formulation goes back to the paper by Poincaré about Bruns' theorem [7].

An issue of not less importance is the development of algorithms for finding manifolds on which the system is transitive. In fact, the problem is to find a complete set of conservation laws. Note that these manifolds can be seen by observing the trajectory over a long period of time. For example, in the framework of numerical experiments with Kahan schemes [8–11], we have repeatedly observed how the points of the orbit found by the reversible difference scheme arrange along lines or surfaces [12]. Thus, it is not difficult to roughly estimate the dimension of the manifold on which the integral curve "winds" and hence determine the number of independent conservation laws. However, the accumulation of errors in numerical methods makes conclusions based on these experiments not entirely convincing.

A natural first step in verifying such judgments is to apply numerical methods to dynamic systems that have been exactly integrated. However, one difficulty arises here: these systems have been integrated because their trajectories are very simple. For example, nonlinear oscillators that are integrated in elliptic functions have elliptic curves as trajectories, so the question of describing the integral manifold is solved immediately, namely, this is the well-known algebraic curve.

Interesting and non-trivial from this point of view are nonlinear dynamic systems integrable in Abelian functions of two variables [13, 14]. In this paper, we consider the Vanhaecke system [15, 16]. This is a completely integrable Hamiltonian system with two degrees of freedom, the general solution of which is expressed by means of Abelian functions of two variables. The algebraic integral manifold of this problem is a surface. Numerical experiments show that integral curves are wound on this surface (Section 2). On the other hand, from theoretical considerations it seems that the system should have integral manifolds of dimension 1, since the system is integrable both in the Liouville sense and in the sense of its solvability in classical transcendental functions [17]. To describe the behavior of orbits on a two-dimensional integral algebraic manifold, we, relying on [15], obtained the main formulas describing the behavior of the Vanhaecke system in the Sage computer algebra system and analyzed the orbits based on known results in the theory of Abelian functions [14, 16, 18].

2. Vanhaecke system and its integral manifolds

The Vanhaecke system is a Hamiltonian system with two degrees of freedom and Hamiltonian

$$H = \frac{p_1^2 + p_2^2}{2} + (q_1^2 + q_2^2)^2 + aq_1^2 + bq_2^2$$

for $a \neq b$, integrated by P. Vanhaecke in 1994 [15]. The trivial case $a = b$ is briefly considered in [16].

The Vanhaecke system has two polynomial integrals:

$$F = (p_2q_1 - p_1q_2)^2 - \left(2\,q_1^4 + 2\,q_1^2q_2^2 + 2\,aq_1^2 + p_1^2\right)(a - b)$$

and

$$G = (p_2q_1 - p_1q_2)^2 + \left(2\,q_1^2q_2^2 + 2\,q_2^4 + 2\,bq_2^2 + p_2^2\right)(a - b).$$

Energy does not give a new integral, since

$$F - G = 2(b - a)H.$$

It is interesting that for $a = b$ both integrals coincide and are the square of the area integral $p_2q_1 - p_1q_2$.

Equations

$$F(p_1, p_2, q_1, q_2) = f, \quad G(p_1, p_2, q_1, q_2) = g \tag{3}$$

define in $\mathbb{R}^4$ a surface S, the integral surface of the Vanhaecke system. We calculated its projection into $q_1q_2p_1$ using the Gröbner basis [19], it is given by the formula

$$s_2p_1^4 + s_1p_1^2 + s_0 = 0, \tag{4}$$

where s_0, s_1, s_2 are polynomials in q_1, q_2, whose coefficients depend on the parameters a, b, f, g. For what follows, it is important to note that these polynomials contain only even powers of q_1 and q_2. Therefore, the surface (4) is symmetric with respect to reflection in the coordinate planes.

Let us see how the integral curve is located on the integral surface. For the numerical experiments given below, we took without any intention the parameters

$$a = 1, \quad b = 2 \tag{5}$$

and the particular solution at

$$p_1(0) = 1, \quad p_2(0) = 0, \quad q_1(0) = 0, \quad q_2(0) = 1. \tag{6}$$

We calculated the solution of the Vanhaecke system in our package fdm for sage [20] for these initial conditions using the Runge–Kutta method for $0 < t < 200$ with a very small step, such that its further decrease did not lead to any noticeable change in the figures discussed below.

Figure 1 shows the projection of the integral surface (4) into the space $q_1q_2p_1$ and the integral curve lying on this surface. The integral curve does indeed lie on this surface and covers it densely enough for us to see the “skeleton” of the integral surface, but not enough to say that the integral curve winds onto this surface everywhere densely. At the same time, the integral curve seems to be closed and therefore the solution is expected to be periodic.

Let us now discuss how the integral manifold and its projection are structured. First of all, we note that for $fg \neq 0$ the integral surface has no singular points. This is easy to verify if we calculate the Gröbner basis of the ideal

$$(F - f, G - g, F_{q_1}, \ldots, G_{p_2}).$$

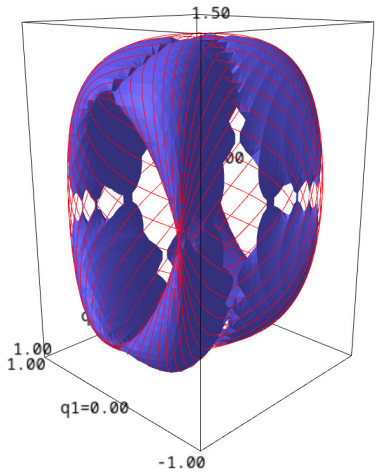

Figure 1. Projection of the integral manifold into the space $q_1q_2p_1$ (blue surface) and one of the integral curves (red line) that winds onto this manifold

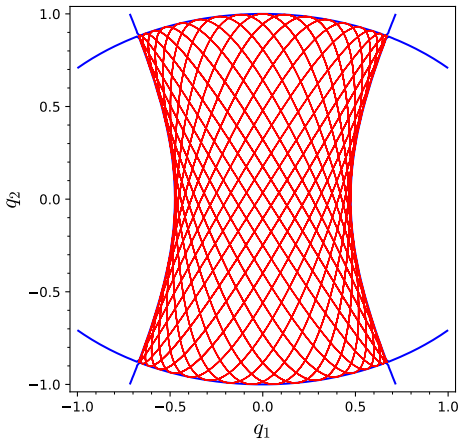

Figure 2. Projection of the integral manifold onto the plane q_1q_2 and one of the integral curves (red line) that winds onto this manifold

The orbits of the dynamical system lie on this manifold, and they cannot intersect or have self-intersection points by virtue of the Cauchy theorem. However, when projecting from four-dimensional space to three-dimensional space, the integral surface acquires double lines on which the orbit has self-intersection points. To find the double lines of the surface (4) we calculated the Gröbner basis of the corresponding ideal and verified that there are three such lines — these are sections of the surface by coordinate planes. For example,

$$s_2p_1^4 + s_1p_1^2 + s_0\big|_{q_1=0} = (p_1^2q_2^2 - ap_1^2 + bp_1^2 - f)^2(a-b)^2.$$

To plot the surface in Figure 1 we used the standard function `implicit_plot3d`, which cannot correctly depict the behavior of the surface near double lines. Therefore, in the Figure 1 you can see how the orbit from the outer surface of the cylinder passes through a double line to the inner side of the tube and returns back. This, in fact, prevents us from describing the situation as the winding of the orbit onto the surface.

Figure 2 shows the projection of this integral manifold onto the plane q_1q_2. The quadratic equation (4) with respect to p_1^2 has real roots if its discriminant

$$D = s_1^2 - 4s_2s_0 \geq 0.$$

Therefore, the line $D = 0$ on the plane q_1q_2 defines the boundaries of the projection of the integral surface onto this plane. This allows depicting the boundaries of the projection of the surface S in Figure 2 as the zeros of the discriminant.

Looking at these illustrations, it is very difficult to answer the question about the dimension of the integral manifold on which the integral curves are wound. Much depends not on the experiment with the numerical solution itself, but on the interpreter. One can cling to the fact that the solution seems periodic and then the integral curve in Figure 1 is closed, and the dimension of the integral manifold is 1. On the contrary, one can assume that the solution is not periodic and over time the

integral curve everywhere densely fills the entire integral surface. In this case, the dimension of the integral manifold is 2 and cannot be reduced.

As noted above, in the theory we are interested in the dimension of the orbit closure as a manifold in $\mathbb{R}^4$ as $t \in \mathbb{R}$ changes. The trouble with this definition and similar-type ones is that it is not constructive. We can observe the integral curve only on a finite segment $0 < t < T$ rather than on the entire interval $0 < t < \infty$. At the same time, increasing T leads to an increase in the error. We see in Figure 1 that for $0 < t < 200$ the trajectory is a closed curve. If we increase the segment several times, the figure will not change, but the integral curve will become thicker. This can be explained both by the accumulation of rounding error and by the fact that the solution is slightly non-periodic and therefore over time the orbit will fill the entire surface.

3. A note on Liouville's theorem

One might think that Liouville's theorem [3] allows finding two more integrals from two polynomial integrals in involution, and therefore the system has two more integrals depending on t. After eliminating t, we should obtain a one-dimensional integral manifold, not a surface. This kind of reasoning leads to the wrong result, since it suggests carrying over local results to the global case. In fact, Liouville's theorem leads to two additional integrals of motion, which are very complicated Abelian quadratures. Vanhaecke obtains this kind of integral of motion in a different way, but, as will be shown below, it does not help at all to reduce the dimension of $\overline{O}$.

4. Integration of the Vanhaecke system in Abelian functions

The general scheme for integrating the Hamilton equations based on Liouville's theorem [3] leads to very cumbersome quadratures. Vanhaecke managed to bypass the stage of writing quadratures from Liouville's theorem by using a convenient change of variables. As a result, it was possible to describe the solution using the Jacobi problem [14], which allows an unambiguous answer to the question posed about the dimension of the minimal integral manifold.

Vanhaecke's passage to the Jacobi problem is based on a guess, which we will describe as follows: we pass from the variables q_1, q_2 to two new variables x_1, x_2, which are roots of the equation

$$x^2 + (q_1^2 + q_2^2 + a + b)x + aq_2^2 + bq_1^2 + ab = 0.$$

In other words, we change variables as

$$\begin{cases} x_1 + x_2 = -q_1^2 - q_2^2 - a - b, \\ x_1 x_2 = bq_1^2 + aq_2^2 + ab. \end{cases} \tag{7}$$

Let us leave aside the discussion of the guess itself, noting only that the appearance of q_1^2 and q_2^2 here significantly simplifies the calculations, since Eq. (4) contains only even powers of q_1 and q_2.

Let us prove in the Sage system that to find x_1, x_2 from the differential equations describing the Vanhaecke system, we will obtain the Jacobi problem [14]:

$$\begin{cases} \dfrac{dx_1}{\sqrt{R(x_1)}} + \dfrac{dx_2}{\sqrt{R(x_2)}} = 0, \\ \dfrac{x_1 dx_1}{\sqrt{R(x_1)}} + \dfrac{x_2 dx_2}{\sqrt{R(x_2)}} = dt. \end{cases} \tag{8}$$

where R is some polynomial of degree 5, which we will write out explicitly below.

For this purpose, we express the coordinates and momenta of the system through x_1, x_2 and their derivatives with respect to t. To do this, we differentiate (7) with respect to t and find the relationship between the momenta and the derivatives of x_1, x_2 with respect to t:

$$\begin{cases} \dfrac{dx_1}{dt} + \dfrac{dx_2}{dt} = -2\,p_1q_1 - 2\,p_2q_2, \\ \dfrac{dx_2}{dt}x_1 + \dfrac{dx_1}{dt}x_2 = 2\,bp_1q_1 + 2\,ap_2q_2. \end{cases}$$

This system is linear with respect to p_1, p_2, so it is not difficult to solve it explicitly:

$$\begin{cases} p_1 = -\dfrac{1}{2}\dfrac{(a+x_2)\dot{x}_1 + (a+x_1)\dot{x}_2}{(a-b)q_1}, \\ p_2 = \dfrac{1}{2}\dfrac{(b+x_2)\dot{x}_1 + (b+x_1)\dot{x}_2}{(a-b)q_2}. \end{cases}$$

Substitute these expressions into Eqs. (3) of the surface S and combine these two equations with Eqs. (7). As a result, we arrive at a system of 4 equations containing $x, \dot{x}$ and q. Eliminating q_1, q_2 and expressing $\dot{x}_2$ through x_1, x_2 and constants using the Gröbner basis method, we obtain two consequences of this system:

$$(x_1 - x_2)^2\dot{x}_1^2 - R(x_1) = 0$$

and similarly

$$(x_1 - x_2)^2\dot{x}_2^2 - R(x_2) = 0.$$

In both equatilies, $R(x)$ is a polynomial of degree 5 in x:

$$R = \frac{4}{a-b}(x+b)(x+a)(2(a-b)x^3 + 2(a^2-b^2)x^2 + 2(a^2b - 2ab^2 + f - g)x + bf - ag).$$

We rewrite the obtained equalities as

$$\frac{\dot{x}_1^2}{R(x_1)} = \frac{1}{(x_2-x_1)^2}, \quad \frac{\dot{x}_2^2}{R(x_2)} = \frac{1}{(x_1-x_2)^2}$$

and extract the root

$$\frac{\dot{x}_1}{\sqrt{R(x_1)}} = \pm\frac{1}{x_1-x_2}, \quad \frac{\dot{x}_2}{\sqrt{R(x_2)}} = \pm\frac{1}{x_2-x_1}.$$

Upon permutation $x_1 \to x_2$, one formula must yield another, so the sign in both formulas is either plus or minus.

But then

$$\frac{dx_1}{\sqrt{R(x_1)}} + \frac{dx_2}{\sqrt{R(x_2)}} = 0$$

and

$$\frac{x_1dx_1}{\sqrt{R(x_1)}} + \frac{x_2dx_2}{\sqrt{R(x_2)}} = \pm\frac{x_1dt}{x_1-x_2} \pm \frac{x_2dt}{x_2-x_1} = \pm dt.$$

Since $dt > 0$, the plus sign should be chosen, which yields system (8).

5. Integration in Abelian functions

In the previous Section we came to the conclusion that the problem of integrating the Vanhaecke system on the integral surface (3) can be written as a Jacobi problem (8), which is convenient to write in quadratures

$$\begin{cases} \int\limits_{x_1^0}^{x_1} \frac{d\xi}{\sqrt{R(\xi)}} + \int\limits_{x_2^0}^{x_2} \frac{d\xi}{\sqrt{R(\xi)}} = 0, \\ \int\limits_{x_1^0}^{x_1} \frac{\xi d\xi}{\sqrt{R(\xi)}} + \int\limits_{x_2^0}^{x_2} \frac{\xi d\xi}{\sqrt{R(\xi)}} = t, \end{cases} \tag{9}$$

where t is the time passed between the initial position of the system (x_1^0, x_2^0) and the final position (x_1, x_2).

The fundamental theorem of the theory of Abelian functions states that every rational symmetric function s of quantities x_1, x_2 satisfying the Jacobi problem

$$\begin{cases} \int\limits_{x_1^0}^{x_1} \frac{d\xi}{\sqrt{R(\xi)}} + \int\limits_{x_2^0}^{x_2} \frac{d\xi}{\sqrt{R(\xi)}} = u, \\ \int\limits_{x_1^0}^{x_1} \frac{\xi d\xi}{\sqrt{R(\xi)}} + \int\limits_{x_2^0}^{x_2} \frac{\xi d\xi}{\sqrt{R(\xi)}} = v, \end{cases}$$

is a meromorphic function of u, v:

$$s = \mathrm{Al}(u, v),$$

referred to as Abelian function. The above is true if the polynomial R has no multiple roots. For parameters a, b, f, g in general position, this is indeed the case. The zeros of R define the periods of the Abelian function: if e_1, e_2 are zeros of R and there are no other zeros in the interval $[e_1, e_2]$, then the numbers

$$\omega_1 = 2\int\limits_{e_1}^{e_2} \frac{d\xi}{\sqrt{R(\xi)}}, \quad \omega_2 = 2\int\limits_{e_1}^{e_2} \frac{\xi d\xi}{\sqrt{R(\xi)}}$$

form a system of common periods of the Abelian function:

$$\mathrm{Al}(u + \omega_1, v + \omega_2) = \mathrm{Al}(u, v).$$

Returning to our dynamical system, we see that

$$x_1 + x_2 = \mathrm{Al}_1(0, t), \quad x_1 x_2 = \mathrm{Al}_2(0, t).$$

By virtue of (7) the squares of q_1 and q_2 are expressed linearly through these functions, which allows expressing the general solution of the Vanhaecke system in terms of two Abelian functions of two arguments.

Note that the polynomial R depends on the parameters a, b and f, g, but does not contain two more constants x_1^0 and x_2^0. Changing these quantities contributes additive constants to the left-hand sides of the equalities (9). Therefore, the formulas

$$x_1 + x_2 = \mathrm{Al}_1(c_1, t + c_2), \quad x_1 x_2 = \mathrm{Al}_2(c_1, t + c_2) \tag{10}$$

determine a two-parameter family of solutions of the Vanhaecke system on a fixed integral surface (3). Let us see how these integral curves are wound onto the algebraic integral surface.

6. Results of computer experiments and discussion

How is the integral curve of the Vanhaecke system structured. — Let the roots of the polynomial R be real, denote them as $e_1, \ldots$ in ascending order. Using them it is easy to compose two real systems of common periods

$$\omega_1 = 2\int\limits_{e_1}^{e_2} \frac{d\xi}{\sqrt{R(\xi)}}, \quad \omega_2 = 2\int\limits_{e_1}^{e_2} \frac{\xi d\xi}{\sqrt{R(\xi)}}$$

and

$$\omega_1' = 2\int\limits_{e_3}^{e_4} \frac{d\xi}{\sqrt{R(\xi)}}, \quad \omega_2' = 2\int\limits_{e_3}^{e_4} \frac{\xi d\xi}{\sqrt{R(\xi)}}$$

of Abelian functions [14, 18]. For any $n, m \in \mathbb{Z}$, it is true that

$$\mathrm{Al}_i(c_1 + n\omega_1 + m\omega_1', t + c_2 + n\omega_2 + m\omega_2') = \mathrm{Al}_i(c_1, t + c_2).$$

For example, for the considered values of constants (5) and initial conditions (6), we have

$$R = 4(2x^2 - 3)(x + 3)(x + 2)(x + 1),$$

therefore there are two real systems of common periods

$$\omega_1 = 2\int\limits_{-3}^{-2} \frac{d\xi}{\sqrt{R(\xi)}}, \quad \omega_2 = 2\int\limits_{-3}^{-2} \frac{\xi d\xi}{\sqrt{R(\xi)}}$$

and

$$\omega_1' = 2\int\limits_{-\sqrt{3/2}}^{-1} \frac{d\xi}{\sqrt{R(\xi)}}, \quad \omega_2' = 2\int\limits_{-\sqrt{3/2}}^{-1} \frac{\xi d\xi}{\sqrt{R(\xi)}}$$

Theorem 1. *The Vanhaecke system has a periodic solution if and only if the periods ω_1, ω_1' are commensurable.*

Proof. (i) If the numbers ω_1, ω_1' are commensurable, then we can choose $n, m \in \mathbb{Z}$ so that

$$n\omega_1 + m\omega_1' = 0$$

and then q_i^2 have period $n\omega_2 + m\omega_2'$. In this case, the integral curve is a closed curve and the dimension of the minimal integral manifold is 1.

(ii) Let the solution have period T, then

$$\mathrm{Al}_i(u, v + T) = \mathrm{Al}_i(u, v).$$

Every pair of common periods of an Abelian function $(0, T)$ can be expressed as a linear combination of four pairs of common periods with integer coefficients of the theorem on the fundamental system of periods proved by Weierstrass [18, ch. 16]. Two of these four pairs are imaginary, so they can be omitted. But then

$$\begin{cases} 0 &= n\omega_1 + m\omega_1', \\ T &= n\omega_2 + m\omega_2', \end{cases}$$

and so the periods are commensurable. □

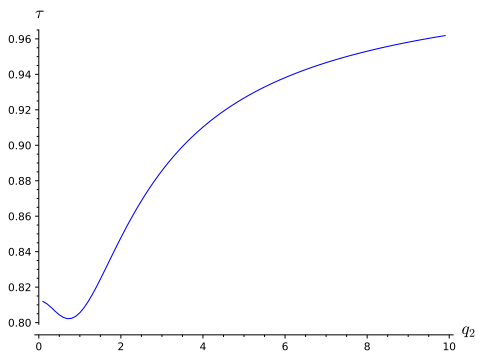

Figure 3. Dependence of τ on the choice of the initial value of q_2 for $q_1 = 0$, $p_1 = 1$, and $p_2 = 0$

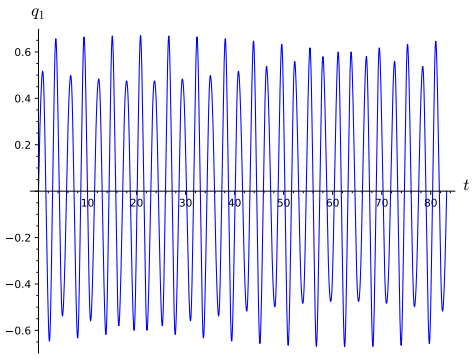

Figure 4. Dependence of q_1 on t in an interval equal to twice the approximate period $29\omega_2' - 72\omega_2$

Note that the polynomial R depends on the parameters a, b and f, g, but not on the choice of initial conditions on the integral surface (3). Therefore, either all integral curves on a fixed surface (3) are closed, or all are open. And this depends on the relationship between the two numbers ω_1 and ω_1'.

The ratio

$$\tau = \frac{\omega_1}{\omega_1'}$$

depends on parameters f and g in a continuous way, see Figure 3. Therefore, by changing these parameters arbitrarily little, we can obtain either a rational number or an irrational one. Thus, in the neighborhood of any point of the four-dimensional space $q_1 q_2 p_1 p_2$ there are infinitely many integral surfaces on which the integral curves are closed, and infinitely many surfaces on which they are not closed.

If for the considered values of the constants f, g the quantity τ is rational, i.e.,

$$\tau(f, g) = -m/n \in \mathbb{Q},$$

then any solution of the Vanhaecke system, the initial data of which are taken on the surface (3), is periodic, and the real number

$$T = n\omega_2 + m\omega_2'$$

is the period of these solutions. The orbits corresponding to periodic solutions are smooth closed curves without self-intersection. Therefore, the dynamical Vanhaecke system is not transitive on the algebraic surface (3) for such values of f, g.

The periods of Abelian functions depend on R, that is, on f, g, but not on the choice of initial conditions on the integral surface (3). Therefore, all solutions lying on the surface (3) for such values of f, g have the same period. This allows revealing the integral surface (3) in numerical experiments if considering all solutions having the same period rather than one individual periodic solution. We proposed to call the manifold formed by the points of all possible solutions having the same period an equiperiodic manifold [21]. In the case of the Vanhaecke system, equiperiodic manifolds are algebraic surfaces belonging to the family of integral surfaces (3) and correspond to those particular values of the parameters f, g for which $\tau \in \mathbb{Q}$.

A small rational change in τ in the topology of $\mathbb{R}$ can lead to a huge change in the period. Thus, the value of the period is not stable to small changes in the parameters of the integral surface. This does not at all contradict the stability of the initial problem solution with respect to a change in the initial data, i.e, a change in f and g. If in the unperturbed state we had a solution with a period of $T = n\omega_2 + m\omega_2'$, then with compensation we obtain a solution that almost returns back in time T.

The behavior of the integral curve for irrational τ can be deduced from the following theorem.

Theorem 2 (Jacobi theorem on infinitesimal periods). *Let ω and ω' be two real numbers. Then either they are commensurable, i.e., there are two integers n and m such that*

$$n\omega + m\omega' = 0,$$

or they are incommensurable, i.e.,

$$n\omega + m\omega' \neq 0 \quad \forall n, m \in \mathbb{Z},$$

but for any $\epsilon > 0$ there are integers n and m such that

$$|n\omega + m\omega'| < \epsilon$$

In the absence of a suitable source, we present a proof. It repeats the Euclidean integer division algorithm.

Proof. Without loss of generality, we can assume that $|\omega| > |\omega'|$. Then we can choose a number $n \in \mathbb{Z}$ such that

$$|\omega - n\omega'| \leq \frac{|\omega'|}{2}$$

We set

$$\omega'' = \omega - n\omega', \quad |\omega''| \leq \frac{|\omega'|}{2}.$$

But then $|\omega'| > |\omega''|$ and we can choose a number $n' \in \mathbb{Z}$ such that

$$|\omega' - n'\omega''| \leq \frac{|\omega''|}{2}$$

Let us put

$$\omega''' = \omega' - n'\omega'' = -n'\omega + m'\omega', \quad |\omega'''| \leq \frac{|\omega''|}{2}.$$

Proceeding in this way, we obtain the number

$$\omega^{(n)} = n\omega + m\omega', \quad n, m \in \mathbb{Z},$$

for which the estimate

$$|\omega^{(n)}| \leq \frac{|\omega^{(n-1)}|}{2} \leq \cdots \leq \frac{|\omega'|}{2^{n-1}}$$

is valid. Thus, we obtain an infinitesimal sequence $\{\omega^{(n)}\}$, which was to be proved. □

Note that for

$$\Delta u = n\omega_1 + m\omega_1', \Delta v = n\omega_2 + m\omega_2'$$

it is true that

$$\mathrm{Al}_i(c_1 + k\Delta u, t + c_2) = \mathrm{Al}_i(c_1, t - k\Delta v + c_2), \quad k \in \mathbb{Z}.$$

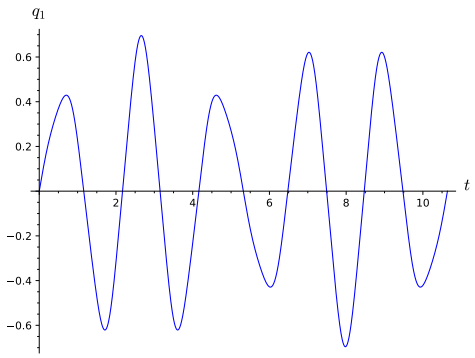

Figure 5. Dependence of q_1 on t in an interval equal to twice the period $-6\omega_2 + 5\omega_2'$

From the Jacobi theorem 2 it follows that Δu can be taken arbitrarily small. Therefore $k\Delta u$ can be take close to any specified number $u \in \mathbb{R}$. But then the integral curve

$$x_1 + x_2 = \mathrm{Al}_1(0,t) = \mathrm{Al}_1(k\Delta u, t - k\Delta v), \quad x_1 x_2 = \mathrm{Al}_2(0,t) = \mathrm{Al}_2(k\Delta u, t - k\Delta v)$$

passes arbitrarily close in the topology of $\mathbb{R}$ to any point on the surface

$$x_1 + x_2 = \mathrm{Al}_1(u,v), \quad x_1 x_2 = \mathrm{Al}_2(u,v), \quad (u,v) \in \mathbb{R}^2.$$

Therefore, the closure of the set of all points of the integral curve (10) is a surface and the dimension of the integral manifold onto which the orbit winds is equal to 2.

Since the set of rational numbers is inferior in power to the set of irrational numbers, it seems surprising that we see a periodic curve in Figure 1. Random initial data should not yield closed curves. This effect is again explained by the Jacobi theorem. Taking n, m such that $n\omega_1 + m\omega_1'$ is arbitrarily small in absolute value, we see, by virtue of the uniform continuity of Abelian functions, that the equality

$$q_i^2(t + n\omega_2 + m\omega_2') \simeq q_i^2(t)$$

is satisfied with arbitrarily high accuracy.

This is exactly the case we observe in our example. The equality

$$\frac{\omega_1}{\omega_1'} \simeq \frac{29}{72}$$

is satisfied with very good accuracy:

$$29\omega_1' - 72\omega_1 = 1 \cdot 10^{-4},$$

i.e., two-digit n, m give a 4th-order zero. Therefore, in numerical experiments it seems that the solution is periodic and the squares q_i have a period

$$29\omega_2' - 72\omega_2 = 41.624490476319636.$$

It is easy to see that q_i themselves are single-valued functions of t, and their period is twice as large, see Figure 4.

The interval of τ variation in Figure 3 is not large, but it contains the value 5/6. It is easy to select the initial values corresponding to such a value of τ from this plot:

$$q_1 = 0, \quad q_2 = 1.688, \quad p_1 = 1, \quad p_2 = 0.$$

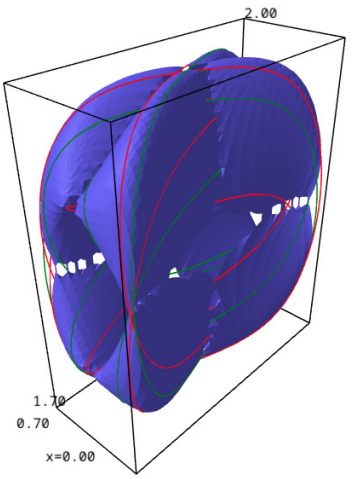

Figure 6. Projection of the integral manifold into the space $q_1q_2p_1$ for initial conditions corresponding to $\tau = 5/6$ and two orbits lying on this surface (red and green)

In this case, the period of oscillations for the squares q_1 and q_2 will be equal to

$$T = -6\omega_2 + 5\omega_2' = 5.31985,$$

which can be easily verified from the plot of the solution found by the Runge–Kutta method (Figure 5).

Figure 6 shows the projection of the integral manifold into the space $q_1q_2p_1$ for this case. As noted above, it is eight bodies linked into a cylinder along double lines. But now, the cavities in these bodies are larger and therefore more noticeable. The orbit covers the surface less densely and is easier to follow. Now it is clearly visible how the trajectories pass onto the inner surface of the pipe through double lines. It is interesting to note that both trajectories shown in Figure 6 have the same period, but this is not at all visible in this figure.

Numerical experiments suggest that the larger in absolute value the integers n, m, the ratio of which gives τ, the more oscillations the periodic solution makes over the period and the more densely it covers the integral surface, and, exactly, more densely, but not everywhere densely.

7. Conclusion

The Vanhaecke system is a completely integrable Hamiltonian system. Its orbits can be divided into two classes, periodic and non-periodic. The points of the orbits corresponding to solutions with the same period form an equiperiodic manifold. In this case, there are infinitely many different equiperiodic manifolds, and all of them are algebraic integral surfaces belonging to the family (2).

For arbitrary initial data, the orbit is not closed and fills at $t \in \mathbb{R}$ some piece of the algebraic surface of the surface everywhere densely. In other words, for almost any initial data, the restriction of the Vanhaecke system to an integral surface is transitive. This does not contradict either the complete integrability of the Vanhaecke system in the sense of Liouville or its solvability in classical transcendental functions. The very intricate course of the orbit on the surface is described analytically by means of Abelian functions of two variables and is determined by four real numbers, a set of

periods of Abelian functions. The key point here is the analytical description of functions of one variable t using periodic functions of two variables.

The behavior of orbits in numerical experiments is determined not so much by the rationality or irrationality of the ratio of two periods τ, but by the possibility of approximating this number with a rational fraction with high accuracy. Therefore, for arbitrary initial data, we see a seemingly periodic solution with a very large period and a closed orbit that runs along the integral surface densely enough for this surface to become visible, but not densely enough to allow saying that it covers the surface everywhere densely.

We believe that the Vanhaecke system is a good test to hone methods for studying integral manifolds of conservative dynamical systems.

Author Contributions: Conceptualization, M. Malykh and L. Sevastianov.; methodology, M. Malykh.; software, Wang Shiwei; validation, M. Malykh and , Wang Shiwei; writing—original draft preparation, Wang Shiwei; writing—review and editing, L. Sevastianov; project administration, A. Zorin. All authors have read and agreed to the published version of the manuscript.

Funding: The work was carried out with the financial support of the Russian Science Foundation (project No. 20-11-20257).

Data Availability Statement: Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Declaration on Generative AI: The authors have not employed any Generative AI tools.

References

[1] J. Moser, *Stable and Random Motions in Dynamical Systems: With Special Emphasis on Celestial Mechanics* (Annals of Mathematics Studies 77). 1973.

[2] H. Bruns, "Über die Integrale der Vielkörper-Problems," *Acta math.*, vol. 11, pp. 25–96, 1887.

[3] E. T. Whittaker, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*. Cambridge: Cambridge University Press, 1988. DOI: 10.1017/CBO9780511608797

[4] P. Painlevé, "Mémore sur les intégrales du problème des n corps," in *Œuvres de Paul Painlevé*. 1975, vol. 2.

[5] P. Y. Polubarinova-Kochina, "On unambiguous solutions and algebraic integrals of a problem about rotation of a gyroscope at a motionless point," in *Dvizhenie tverdogo tela vokrug nepodvizhnoj tochki*, S. A. Chaplygin, Ed., In Russian, Moscow-Leningrad: Academy of Sciences of the USSR, 1940.

[6] V. V. Kozlov, "The nonexistence of an additional analytic integral in the problem of the motion of a nonsymmetric heavy solid around a fixed point," *Vestnik Moskov. Univ. Ser. I Mat. Meh.*, vol. 39, no. 1, pp. 105–110, 1975.

[7] V. V. Kozlov, "Integrability and non-integrability in Hamiltonian mechanics," *Russian Math. Surveys*, vol. 38, no. 1, pp. 1–76, 1983.

[8] J. M. Sanz-Serna, "An unconventional symplectic integrator of W. Kahan," *Applied Numerical Mathematics*, vol. 16, pp. 245–250, 1994.

[9] E. Celledoni, R. I. McLachlan, B. Owren, and G. R. W. Quispel, "Geometric properties of Kahan's method," *J. Phys. A: Math. Theor.*, vol. 46, p. 025 201, 2013. DOI: 10.1088/1751-8113/46/2/025201

[10] M. Petrera and Y. B. Suris, "On the Hamiltonian structure of Hirota-Kimura discretization of the Euler top," *Math. Nachr.*, vol. 283, no. 11, pp. 1654–1663, 2010. DOI: 10.1002/mana.200711162

[11] M. Petrera, J. Smirin, and Y. B. Suris, "Geometry of the Kahan discretizations of planar quadratic Hamiltonian systems," *Proc. R. Soc. A*, vol. 475, p. 20 180 761, 2019. DOI: 10.1098/rspa.2018.0761

[12] M. Malykh, M. Gambaryan, O. Kroytor, and A. Zorin, "Finite Difference Models of Dynamical Systems with Quadratic Right-Hand Side," *Mathematics*, vol. 12, no. 1, p. 167, 2024. DOI: 10.3390/math12010167

[13] V. V. Golubev, *Lectures on integration of the equations of motion of a rigid body about a fixed point*. Jerusalem: Israel Program for Scientific Translations, 1960.

[14] A. I. Markushevich, *Introduction to the Classical Theory of Abelian Functions*. Translations of Mathematical Monographs, 1992. DOI: 10.1090/mmono/096

[15] P. Vanhaecke, "A special case of the Garnier system, (1,4)-polarized abelian surfaces and their moduli," *Compositio Mathematica*, vol. 92, pp. 157–203, 2 1994.

[16] P. Vanhaecke, *Integrable Systems in the Realm of Algebraic Geometry* (Lecture Notes in Mathematics), 2nd. Springer, 2001.

[17] N. J. Hitchin, G. B. Segal, and R. S. Ward, *Integrable Systems: Twistors, Loop Groups, and Riemann Surfaces*. Oxford University Press, 2013.

[18] K. Weierstrass, *Math. Werke*. Berlin: Mayer & Müller, 1902, vol. 4.

[19] D. Cox, J. Little, and D. O'Shea, *Ideals, varieties, and algorithms*, 3rd ed. Springer, 2007.

[20] A. Baddour, M. Gambaryan, L. Gonzalez, and M. D. Malykh, "On Implementation of Numerical Methods for Solving Ordinary Differential Equations in Computer Algebra Systems," *Program. Comput. Soft.*, vol. 49, pp. 412–422, 2023. DOI: 10.1134/S0361768823020044

[21] A. Baddour, M. D. Malykh, and L. A. Sevastianov, "On periodic approximate solutions of dynamical systems with a quadratic right-hand side," *J. Math. Sci.*, vol. 261, no. 5, pp. 698–708, 2022. DOI: 10.1007/s10958-022-05781-4

Information about the authors

Wang, Shiwei—PhD Student of Department of Computational Mathematics and Artificial Intelligence, RUDN University (e-mail: 1995wsw@gmail.com, ORCID: 009-0007-6504-8370)

Malykh, Mikhail D.—DSc., Head of Department of Computational Mathematics and Artificial Intelligence, RUDN University; Senior Researcher of Joint Institute for Nuclear Research (e-mail: malykh-md@rudn.ru, ORCID: 0000-0001-6541-6603, ResearcherID: P-8123-20168, Scopus Author ID: 6602318510)

Sevastianov, Leonid A.—DSc., Professor of Department of Computational Mathematics and Artificial Intelligence, RUDN University; Senior Researcher of Joint Institute for Nuclear Research (e-mail: sevastianov-la@rudn.ru, ORCID: 0000-0002-1856-4643, ResearcherID: B-8497-2016, Scopus Author ID: 8783969400)

Zorin, Aleksander V.—DSc., Professor of Department of Computational Mathematics and Artificial Intelligence of RUDN University (e-mail: zorin-av@rudn.ru, ORCID: 0000-0002-5721-4558)

RUSSIAN METADATA

УДК 519.872, 519.217
DOI: 10.22363/2658-4670-2026-34-2-226-240 EDN: JFRCXF

О поведении орбит системы Ванхекке на интегральных поверхностях

Ш. Ван[1], М. Д. Малых[1,2], Л. А. Севастьянов[1], А. В. Зорин[1]

[1] *Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация*

[2] *Объединённый институт ядерных исследований, ул. Жолио-Кюри, д. 6, Дубна, 141980, Российская Федерация*

Аннотация. В 1990-х годах П. Ванхекке описал гамильтоновую систему с двумя степенями свободы и полиномиальной гамильтоном, которая интегрируема в абелевы функции двух переменных. Эта система доставляет удобный пример вполне интегрируемой системы, в которой интегральные кривые навиты на двумерное многообразие — алгебраическую поверхность в 4-мерном фазовом пространстве. В статье показано, что все необходимые расчёты могут быть выполнены в системе Sage. Представлены результаты численных экспериментов, выполненных в нашем пакете FDM для Sage. Показано, что орбиты системы Ванхекке можно разделить на два класса: периодические и непериодические. Точки периодических орбит, соответствующие решениям с одинаковым периодом, образуют эвкипериодическую поверхность, причём существует бесконечно много различных алгебраических эквипериодических поверхностей. Поведение непериодических орбит в численных экспериментах определяется не столько рациональностью или иррациональностью соотношения двух периодов абелевых функций, сколько возможностью аппроксимации этого числа рациональной дробью с высокой точностью.

Ключевые слова: абелевы функции, динамические системы, вполне интегрируемые гамильтоновы системы

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2026, 34 (2) 241–252
http://journals.rudn.ru/miph

Research article

UDC 517.581, 517.954
PACS 44.05.+e, 45.10.Hj
DOI: 10.22363/2658-4670-2026-34-2-241-252 EDN: JUGPSJ

About one method of approximate solution of the first boundary value problem for the fractional diffusion equation

Ivan I. Zakharov, Temirkhan S. Aleroev

Moscow State University of Civil Engineering, 26 Yaroslavskoye Shosse, Moscow, 129337, Russian Federation

(received: March 9, 2026; revised: March 25, 2026; accepted: April 2, 2026)

Abstract. *Background* This article focuses on a rapidly developing area of fractional calculus that describes anomalous diffusion. Since the correct form of the fractional-order equation describing anomalous diffusion — remains an open question, the authors use examples of models based on fractional diffusion equations to examine the advantages of various fractional-order equations proposed for modelling the advection-diffusion process in media with a fractal structure. *Purpose* This paper examines boundary value problems for the advection-diffusion equation with Caputo and Riemann–Liouville fractional operators. Given that many authors consider in their work models with fractional-order operators obtained simply by replacing ordinary derivatives with fractional ones, the authors of this paper consider it necessary to compare fractional differential equations with various operators. We also note that one of the aims of this work is to construct an effective approximate method for solving the first initial-boundary value problem for a homogeneous fractional differential equation. *Method* A fractional calculus approach is used. The approximate method under consideration is based on the analytical method of separation of variables (the Fourier method). The proposed approximate method and all the necessary calculations are implemented using the Matlab programming language. *Results* Approximate solutions have been obtained for boundary value problems for the advection-diffusion equation with fractional operators of Caputo and Riemann–Liouville. The solutions obtained have been compared both with one another and with the classical model based on ordinary derivatives. *Conclusions* Based on the results obtained, the authors recommend using an equation containing the Riemann–Liouville fractional operator to model anomalous diffusion processes.

Key words and phrases: approximate calculus, fractional calculus, fractional advection-diffusion equation, fractional Riemann–Liouville derivative, fractional Caputo derivative.

For citation: I. I. Zakharov, T. S. Aleroev "About one method of approximate solution of the first boundary value problem for the fractional diffusion equation," *Discrete and Continuous Models and Applied Computational Science*, vol. 34, no. 2, pp. 241–252, 2026. DOI: 10.22363/2658-4670-2026-34-2-241-252. EDN: JUGPSJ.

1. Introduction

The problem of advection-diffusion is considered, whose model is viewed as a differential equation with various fractional differential operators (in the form of Caputo and in the form of Riemann–Liouville). To date, many works have been prepared on solving various problems containing fractional differential operators, but most of them are devoted to numerical methods (finite difference method, spline method, collocation method, etc.), in particular [1, 2]. We would also like to mention the eight-volume 'Handbook of Fractional Calculus with Applications'[3], which contains the latest achievements in the field of fractional differentiation. This work focuses on the development of approximate calculation methods based on analytical methods (the Fourier method). Although numerical methods demonstrate high efficiency in solving the proposed problems, it is analytical methods that allow the development of the necessary set of tools for verifying the solution of problems, for which it is absolutely necessary to develop effective methods for implementing approximate solutions.

In this work, boundary value problems for the advection-diffusion equation with a fractional differential operator are solved using the method of separation of variables (Fourier method). One of the main points in solving a boundary value problem containing a fractional differential operator is the structure of the spectral problem obtained by separation of variables. This is due to the non-self-adjointness of the operators of this spectral problem. The structure of this spectral problem is extremely complex and has not been fully studied. This work fills this gap.

2. The first initial boundary value problem for a homogeneous fractional dispersion equation

Let us consider the first boundary value problem for a homogeneous fractional dispersion equation:

$$\frac{\partial U(x,t)}{\partial t} = -D \cdot {}_0^R D_x^\alpha U(x,t), \tag{1}$$

with boundary and initial conditions

$$U(0,t) = U(1,t) = 0, \tag{2}$$

$$U(x,0) = \phi(x), \tag{3}$$

where ${}_0^R D_x^\alpha = \frac{1}{\Gamma(2-\alpha)} \frac{\partial^2}{\partial x^2} \int_0^x \frac{U(\tau,t)d\tau}{(x-\tau)^{\alpha-1}}$ — fractional derivative (in the sense of Riemann–Liouville), $1 < \alpha < 2$ [4]. And a similar problem with the Caputo operator.

$$\frac{\partial U(x,t)}{\partial t} = -D \cdot {}_0^C D_x^\alpha U(x,t), \tag{4}$$

with boundary and initial conditions

$$U(0,t) = U(1,t) = 0, \tag{5}$$

$$U(x,0) = \phi(x), \tag{6}$$

where ${}_0^C D_x^\alpha = \frac{1}{\Gamma(2-\alpha)} \int_0^x \frac{U(\tau,t)'' d\tau}{(x-\tau)^{\alpha-1}}$ — fractional derivative (in the sense of Caputo), $1 < \alpha < 2$ [5].

Works [4, 5] were devoted to the development of approximate methods for studying mathematical models of heat and mass transfer, and a significant part of this work is devoted to analysing the results obtained in these works. This is primarily due to the fact that the correct form of fractional order equations describing anomalous diffusion remains a subject of debate. This work brings us closer to solving this problem to a certain extent.

It is known that the above tasks (when solved using the Fourier method) generate the following two spectral tasks:

$$\begin{cases} {}_0^R D_x^\alpha X(x) = -\lambda X(x), \\ X(0) = 0, X(1) = 0. \end{cases} \tag{7}$$

$$\begin{cases} {}_0^C D_x^\alpha X(x) = -\mu X(x), \\ X(0) = 0, X(1) = 0. \end{cases} \tag{8}$$

Such Sturm–Liouville problems (7), (8) are studied in detail in [6–9]. It is worth mentioning that problems similar to (7) and (8) also arise when studying the Fokker–Planck equation [10].

The paper [11] shows that the spectral problem with the Caputo fractional derivative (8), when $\alpha \to 1$, has no real roots (i.e., there is no principal eigenvalue). This is strikingly different from the Riemann–Liouville fractional derivative problem (7), where the principal eigenvalue exists for all $1 < \alpha < 2$ (i.e., there is always a fundamental tone). This is a significant difference between the principal eigenvalue of the spectral problem with fractional operators and the classical Sturm–Liouville problem. This phenomenon has a significant impact on the adequacy of the modelled diffusion process.

The works [4], [5] demonstrate that the numbers λ_n, μ_n are eigenvalues of problems (7), (8) only when they are zeros of the corresponding Mittag–Leffler functions:

$$E_{\alpha,\alpha}(-\lambda) = \sum_{j=0}^{\infty} \frac{(-\lambda)^j}{\Gamma(\alpha j + \alpha)}, \tag{9}$$

$$E_{\alpha,2}(-\mu) = \sum_{j=0}^{\infty} \frac{(-\mu)^j}{\Gamma(\alpha j + 2)}. \tag{10}$$

With corresponding eigenfunctions (11), (12)

$${}^R W_n(-\lambda_n, x) = x^{\alpha-1} E_{\alpha,\alpha}(-\lambda_n x^\alpha), \tag{11}$$

$${}^C W_n(-\mu_n, x) = x E_{\alpha,2}(-\mu_n x^\alpha). \tag{12}$$

It should be noted that, when studying models with a fractional operator, we (and not only we) heavily rely on the method for constructing eigenfunctions and adjoint functions proposed in the works of M. M. Jerbashyan and A. B. Nersesyan [12], [13].

Similarly, how demonstrated in the work Y. Luchko [14], problems (1) and (4) generate problems that are conjugate to problems (7) and (8).

$$\begin{cases} {}_0^R D_{1-x}^\alpha Y(x) = -\lambda Y(x), \\ Y(0) = 0, Y(1) = 0. \end{cases}$$

$$\begin{cases} {}_0^C D_{1-x}^{\alpha} Y(x) = -\mu Y(x), \\ Y(0) = 0, Y(1) = 0. \end{cases}$$

The eigenvalues of the conjugate problems coincide with the eigenvalues of problems (7), (8). And the eigenfunctions have the form

$${}^R Z_n(-\lambda_n, x) = (1-x)^{\alpha-1} E_{\alpha,\alpha}(-\lambda_n (1-x)^{\alpha}),$$

$${}^C Z_n(-\mu_n, x) = (1-x) E_{\alpha,2}(-\mu_n (1-x)^{\alpha}).$$

As noted in works [4, 5], the solution to problems (1), (2), (3) and (4), (5), (6) is as follows:

$${}^R U^{\alpha}(x,t) = \sum_{n=1}^{\infty} \frac{(\phi, {}^R Z_n)}{({}^R W_n, {}^R Z_n)} e^{-D\lambda_n t} x^{\alpha-1} E_{\alpha,\alpha}(-\lambda_n x^{\alpha}), \tag{13}$$

$${}^C U^{\alpha}(x,t) = \sum_{n=1}^{\infty} \frac{(\phi, {}^C Z_n)}{({}^C W_n, {}^C Z_n)} e^{-D\mu_n t} x E_{\alpha,2}(-\mu_n x^{\alpha}). \tag{14}$$

To analyse these solutions, we will need the following theorems, which were discussed in detail in [15].

Theorem 1. *If*

$$2 - \left(\frac{32\pi^2}{9} + \frac{2}{3}\right)^{-1} < \alpha < 2,$$

then the first eigenvalue of the problem (4)–(6) *is positive and simple (the multiplicity of this eigenvalue is equal to 1), and the fundamental (main) tone has no nodes (i.e., the eigenvalue corresponding to the first eigenvalue does not vanish in (1,2)).*

Theorem 2. *The first eigenvalue* λ_1 *of the problem* (1)–(3) *is positive, simple, and satisfies the condition*

$$0 < \lambda_1^{-1} < \frac{\Gamma(2+2\alpha)}{\Gamma(1+\alpha)},$$

and the principal (main) tone has no nodes for all $1 < \alpha < 2$.

Since the proofs are described in great detail in [15] and are rather cumbersome, we will not reproduce them here.

Considering that for different $\alpha \neq 2$, the solutions to problems (1)–(3), (4)–(6) differ significantly from each other (as can be seen in Figure 1–3), the following recommendations are given: which of these two models should be used when modelling a particular process.

From [15] it is known that problem (7) is equivalent to the integral equation (15)

$$U(x) - \frac{\lambda}{\Gamma(2-\alpha)} \left[\int_0^1 x^{1-\alpha}(1-t)^{1-\alpha} U(t)dt - \int_0^x (x-t)^{1-\alpha} U(t)dt \right] = \\ = U(x) - \frac{\lambda}{\Gamma(2-\alpha)} \int_0^x G_1(x,t) U(t)dt = 0.$$

And problem (8) is in turn equivalent to integral equation (16)

$$U(x) - \frac{\mu}{\Gamma(2-\alpha)}\left[\int_0^1 x(1-t)^{1-\alpha}U(t)dt - \int_0^x (x-t)^{1-\alpha}U(t)dt\right] = \\ = U(x) - \frac{\mu}{\Gamma(2-\alpha)}\int_0^x G_0(x,t)U(t)dt = 0.$$

As can be seen from [15]: the kernel $G_1(x,t)$ (also called the influence function in mechanics) is sign-definite and persymmetric, while the kernel $G_0(x,t)$ is not sign-definite and not persymmetric. From the above and from Theorems 1-2, it follows that when dealing with processes with oscillatory properties, it is necessary to use problems (1)–(3). It turns out that equation (4), which is referred to as an oscillatory equation [16], [17], can hardly be considered as such.

These two theorems also show that for some α, the difference between U_R^α and U_C^α will be significant. It is known that when $\alpha \to 2$, the following relations hold: $\lambda_n \to (\pi n)^2$ and $\mu_n \to (\pi n)^2$, $n = 0, 1, 2, \ldots$. This statement was first obtained in 1998 [18], then in 2005 it was established by other methods [19]. Finally, a generalisation of the result was obtained [15].

Further, when $\alpha \to 2$, the functions of the problem have the form:

$$\lim_{\alpha\to 2} {}^R W_n(-\lambda_n, x) = \lim_{\alpha\to 2} x^{\alpha-1}E_{\alpha,\alpha}(-\lambda_n, x^\alpha) = xE_{2,2}(-\lambda_n, x^2),$$

$$\lim_{\alpha\to 2} {}^C W_n(-\mu_n, x) = \lim_{\alpha\to 2} xE_{\alpha,2}(-\mu_n, x^\alpha) = xE_{2,2}(-\mu_n, x^2).$$

It is known [19] that there is such a formula

$$E_{2,2}(z^2) = \frac{sh(z)}{z}.$$

From this formula it follows that

$$E_{2,2}(-z^2) = \frac{sin(z)}{z}.$$

From these relationships, it follows that when $\alpha = 2$, the solutions to problems (1), (2), (3), (4), (5), and (6) coincide with the solution to the corresponding problem for the differential equation (17):

$$\frac{\partial u(x,t)}{\partial t} = -D \cdot \frac{\partial^2 u(x,t)}{\partial x^2}. \tag{17}$$

That is, the following equality will hold:

$$\lim_{\alpha\to 2} {}^R U^\alpha(x,t) = \lim_{\alpha\to 2} {}^C U^\alpha(x,t) = u(x,t),$$

where

$$u(x,t) = \sum_{n=1}^{\infty} B_n \sin(\pi n x)e^{-D(\pi n)^2 t},$$

in which

$$B_n = 2\int_0^1 \phi(x)\sin(\pi n x)dx.$$

Table 1

The first few eigenvalues λ_n for $1 < \alpha < 2$

α	**1.1**	**1.2**	**1.3**	**1.4**	
λ_n	5.101	4.567	4.518	4.708	
	$10.795 \pm 7.989i$	$14.139 \pm 6.9214i$	$17.920 \pm 3.920i$	16.726	
	$14.338 \pm 16.876i$	$21.916 \pm 18.035i$	$32.274 \pm 17.144i$	26.220	
	$17.273 \pm 26.049i$	$28.880 \pm 30.344i$	$45.830 \pm 32.677i$	$45.424 \pm 12.811i$	
	$19.920 \pm 35.456i$	$35.547 \pm 43.497i$	$59.425 \pm 50.079i$	$69.176 \pm 30.809i$	
	$22.400 \pm 45.061i$	$42.078 \pm 57.328i$	$73.206 \pm 68.999i$	$94.023 \pm 52.268i$	
α	**1.5**	**1.6**	**1.7**	**1.8**	**1.9**
λ_n	5.075	5.610	6.322	7.240	8.404
	17.472	19.513	22.570	26.737	32.253
	32.129	38.158	45.945	56.345	70.247
	55.834	62.410	76.129	95.440	121.869
	64.586	88.220	111.768	143.271	186.675
	$99.718 \pm 21.304i$	123.879	153.696	199.672	264.408

3. Approximate solution of the first initial-boundary value problem

To find an approximate solution, we first find several initial eigenvalues λ_n and μ_n from the corresponding equations:

$$E_{\alpha,\alpha}(-\lambda) = \sum_{j=0}^{\infty} \frac{(-\lambda)^j}{\Gamma(\alpha j + \alpha)} = 0,$$

$$E_{\alpha,2}(-\mu) = \sum_{j=0}^{\infty} \frac{(-\mu)^j}{\Gamma(\alpha j + 2)} = 0.$$

Table 1–2 shows the results of calculations of several first eigenvalues λ_n and μ_n for different values of $1 < \alpha < 2$.

As shown in works [4, 5], for an approximate solution to problems (1)–(3) and (4)–(6), it is sufficient to take the first few terms of the series (13), (14):

$${}^{R}U^{\alpha}(x,t) \approx \sum_{n=1}^{N} \frac{(\phi, {}^{R}Z_n)}{({}^{R}W_n, {}^{R}Z_n)} e^{-D\lambda_n t} x^{\alpha-1} E_{\alpha,\alpha}(-\lambda_n x^{\alpha}),$$

$${}^{C}U^{\alpha}(x,t) \approx \sum_{n=1}^{N} \frac{(\phi, {}^{C}Z_n)}{({}^{C}W_n, {}^{C}Z_n)} e^{-D\mu_n t} x E_{\alpha,2}(-\mu_n x^{\alpha}).$$

Let us denote the approximate solution as

$${}^{R}U_0^{\alpha}(x,t) = \sum_{n=1}^{N} \frac{(\phi, {}^{R}Z_n)}{({}^{R}W_n, {}^{R}Z_n)} e^{-D\lambda_n t} x^{\alpha-1} E_{\alpha,\alpha}(-\lambda_n x^{\alpha}),$$

Table 2

The first few eigenvalues μ_n for $1 < \alpha < 2$

α	**1.1**	**1.2**	**1.3**	**1.4**	
μ_n	$1.353 \pm 7.250i$	$3.173 \pm 7.985i$	$5.473 \pm 8.273i$	$8.190 \pm 7.812i$	
	$2.810 \pm 15.731i$	$7.132 \pm 18.959i$	$13.393 \pm 21.811i$	$22.012 \pm 23.568i$	
	$4.303 \pm 24.691i$	$11.399 \pm 31.246i$	$22.371 \pm 37.953i$	$38.507 \pm 43.723i$	
	$5.829 \pm 33.969i$	$15.902 \pm 44.438i$	$32.169 \pm 55.962i$	$57.132 \pm 67.167i$	
	$7.3825 \pm 43.488i$	$20.600 \pm 58.338i$	$42.647 \pm 75.471i$	$77.563 \pm 93.329i$	
	$8.960 \pm 53.204i$	$25.466 \pm 72.825i$	$53.715 \pm 96.250i$	$99.581 \pm 121.845i$	
α	**1.5**	**1.6**	**1.7**	**1.8**	**1.9**
μ_n	$11.147 \pm 6.123i$	13.420	9.933	9.457	9.514
	$33.314 \pm 23.130i$	14.645	23.2510	28.477	33.598
	$61.196 \pm 46.651i$	$47.293 \pm 18.851i$	$63.101 \pm 5.930i$	62.200	73.039
	$93.818 \pm 75.367i$	$91.705 \pm 43.625i$	$130.587 \pm 29.601i$	97.063	124.419
	$130.545 \pm 108.499i$	$145.569 \pm 75.805i$	$215.458 \pm 60.453i$	155.450	191.146
	$170.946 \pm 145.542i$	$207.859 \pm 114.486i$	$316.291 \pm 99.750i$	196.596	267.945

$$ {}^{C}U_0^{\alpha}(x,t) = \sum_{n=1}^{N} \frac{(\phi, {}^{C}Z_n)}{({}^{C}W_n, {}^{C}Z_n)} e^{-D\mu_n t} x E_{\alpha,2}(-\mu_n x^{\alpha}). $$

It is worth noting that, unfortunately, as has been pointed out on more than one occasion, it is very difficult to find good experimental data (whether from a computational experiment or an experiment conducted using laboratory equipment). However, based on, for example, [20], [21], it can be said that the calculations presented in this paper have been carried out quite well.

4. Results of the approximate solution of the first initial-boundary value problem

The solutions to the first initial-boundary value problem are calculated for various values of the fractional derivative α. In order to compare the solution to problem (1)–(3), with a fractional derivative in the form of Riemann–Liouville, and the solution to problem (4)–(6), with the fractional derivative in the form of Caputo, with the solution of the classical diffusion equation (17), we perform the calculation at $\alpha = 2$.

As can be seen in Figure 1, the surface of solutions to problems (1)–(3), (4)–(6) at $\alpha = 2$ coincides with the solution to the classical problem (17).

When considering $\alpha \neq 2$, for example $\alpha = 1.9$, the surfaces constructed for problems with fractional derivatives in the form of Caputo and Riemann–Liouville not only differ significantly from the classical heat conduction problem (17), but, as can be seen from Figure 2, also differ significantly from each other.

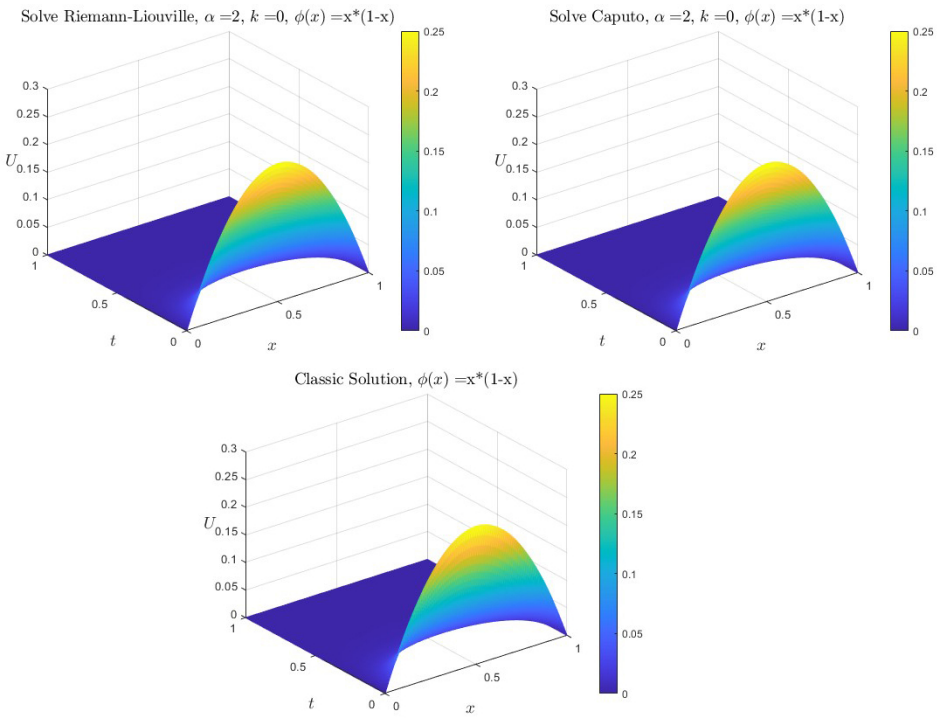

Figure 1. An approximate solution to the classical diffusion problem and problems with derivatives in the sense of Caputo and Riemann–Liouville for $\alpha = 2$

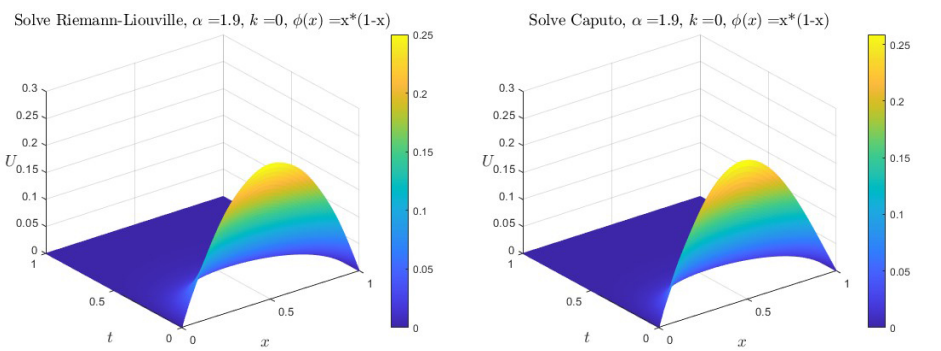

Figure 2. Approximate solution of the boundary value problem with operators in the sense of Caputo and Riemann–Liouville for $\alpha = 1.9$

As noted by the authors earlier, these differences arise due to the significant difference in the generated spectral problems (7), (8), which generate completely different structures of eigenvalues ((9), (10)) and eigenfunctions ((11), (12)).

When $\alpha \to 1$, the differences become more noticeable (Figure 3).

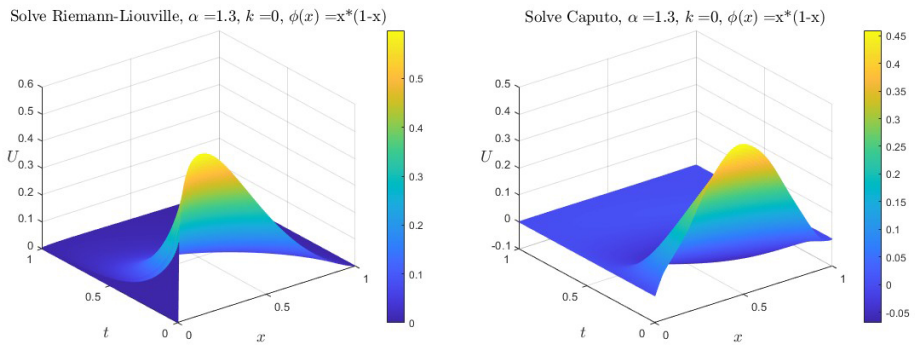

Figure 3. Approximate solution of the boundary value problem with operators in the sense of Caputo and Riemann–Liouville for $\alpha = 1.3$

5. Discussion

Presented in this paper results demonstrate how different the models can be, depending on whether the equation involves the Riemann–Liouville fractional derivative or the Caputo fractional derivative. As mentioned earlier, a significant advantage of models based on the Riemann-Liouville operator is that its first eigenvalue remains real for all $1 < \alpha < 2$. In the case of models based on the Caputo fractional derivative, their first eigenvalue ceases to be real, which creates certain difficulties in their application for modelling advection-diffusion processes.

The principal eigenvalue is of fundamental meaning for modelling advection-diffusion processes. For example, the paper [22] investigates the influence of various diffusion and advection coefficients (tending to zero or infinity) on the first eigenvalue for an elliptic operator with the Neumann zero boundary condition. Without going into the physical details, we note that although the occurrence of complex eigenvalues is possible (in the case where the operator is non-self-adjoint), such models cannot be used to describe diffusion processes.

6. Conclusion

Since the correct form of the fractional order equation describing anomalous diffusion is still a matter of debate, the authors, using examples of specific fractional diffusion equations of fractional diffusion, consider the advantages of one or another fractional order equation presented to describe the process of heat and mass transfer in media with a fractal structure. Approximate methods for solving initial boundary value problems for these equations have been developed.

The presented results demonstrate the possibility of the developed method for the approximate solution of the boundary value problem for the fractional diffusion equation. A significant difference between spectral problems arising with fractional derivatives in the sense of Caputo and Riemann–Liouville is demonstrated. What differences arise when using different operators to solve the advection-diffusion problem. Based on the calculations presented (Figure 1–3) and the analysis of the influence functions $G_0(x,t)$, $G_1(x,t)$, it can be concluded that the model with the Riemann-Liouville fractional derivative (problem (1)–(3)) is recommended for modelling processes of highly anomalous diffusion.

Author Contributions: Conceptualisation, writing, review and editing: Temirkhan Sultanovich Aleroev; methodology, writing and preparation of the first draft: Ivan Ivanovich Zakharov. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Declaration on Generative AI: The author has not employed any generative AI tools.

References

[1] R. Gorenflo and F. Mainardi, "Fractional Calculus: Integral and Differential Equations of Fractional Order," in *Fractals and Fractional Calculus in Continuum Mechanics*, A. Carpinteri and F. Mainardi, Eds., New York: Springer-Verlag, 1997, pp. 223–276.

[2] F. Mainardi, "The fundamental solutions for the fractional diffusion-wave equation," *Applied Mathematics Letters*, vol. 9, no. 6, pp. 23–28, 1996.

[3] K. Anatoly, L. Yuri, E. K. George, E. T. Vasily, P. Ivo, B. Dumitru, M. L. Antonio, and o. many others, *The Handbook of Fractional Calculus with Applications*. Berlin: De Gruyter Brill, 2019.

[4] I. I. Zakharov and T. S. Aleroev, "About one method of approximate solution of the first boundary value problem for the fractional diffusion equation," *Proceeding of MIPT*, no. 61, pp. 60–67, 2024, (in Russian).

[5] I. Zakharov and T. Aleroev, "About one method of approximate solution of the first boundary value problem for the fractional diffusion equation used in gas dynamics," *Trudy MAI*, no. 136, 2024, (in Russian).

[6] T. S. Aleroev, "Solving the Boundary Value Problems for Differential Equations with Fractional Derivatives by the Method of Separation of Variables," *Mathematics*, vol. 8, no. 11, 1877 2020.

[7] A. Tfayli, "Sur quelques equations aux derivees partielles fractionnaires, theorie et applications," ffNNT : 2020LAROS031ff. fftel-03549142f, (in French), Ph.D. dissertation, Universite de La Rochelle, 2020.

[8] W. Wyss, "The fractional diffusion equation," *Journal of Mathematical Physics*, vol. 27, no. 11, pp. 2782–2785, 1986.

[9] O. Agrawal, "Solution for a fractional diffusion-wave equation defined in a bounded domain," *Nonlinear Dynamics*, vol. 29, pp. 145–155, 2002.

[10] T. Aleroev, H. Aleroeva, J. Huang, M. Tamm, Y. Tang, and Y. Zhao, "Boundary value problems of fractional Fokker-Planck equations," *Computers and Mathematics with Applications*, vol. 73, no. 6, pp. 959–969, 2017.

[11] T. S. Aleroev and Y. Li, *Fractional Sturm-Liouville problem with Caputo derivative loses the principal eigenvalue*, 2025.

[12] M. Dzhrbashyan and A. Nersesyan, "Decompositions on some biorthogonal systems and boundary value problems for fractional order differential equations," *Trudy Moskovskogo Matematicheskogo Obshchestva*, vol. 10, pp. 89–179, 1961, (in Russian).

[13] M. Dzhrbashyan and A. Nersesyan, "On the construction of some special biorthogonal systems," *Izvestiya Akademii nauk Armyanskoi SSR*, vol. 12, no. 5, pp. 17–42, 1959, (in Russian).

[14] Y. Luchko, "Some uniqueness and existence results for the initial-boundary value problems for the generalized time-fractional diffusion equation," *Computers and Mathematics with Applications*, vol. 59, no. 5, pp. 1766–1772, 2010.

[15] T. Aleroev and H. Aleroeva, "Problems of Sturm-Liouville type for differential equations with fractional derivatives," in *Handbook of Fractional Calculus with Applications. Volume 2: Fractional Differential Equations*, A. Kochubei and Y. Luchko, Eds., Berlin: De Gruyter Brill, 2019.

[16] A. Nakhushev, *Fractional calculus and its applications*. Moscow: Fizmatlit, 2003, 271 pp., (in Russian).

[17] F. Mainardi, "Fractional Relaxation-Occilation and Fractional Diffusion-Wave Phenomena Chaos," *Solitions and Fractals*, vol. 7, no. 9, pp. 1461–1477, 1996.

[18] T. S. Aleroev, "On a boundary value problem for a fractional-order differential operator," *Differentsial'nye Uravneniya*, vol. 34, no. 1, p. 123, 1998, (in Russian).

[19] A. Y. Popov, "On the number of real eigenvalues of a certain boundary-value problem for a second-order equation with fractional derivative," *Fundamentalnaya i Prikladnaya Matematika*, vol. 12, no. 6, pp. 137–155, 2006, (in Russian).

[20] J. Bangti, L. Raytcho, P. Joseph, and R. William, "A finite element method for the fractional Sturm-Liouville problem," *Mathematics of Computation*, vol. 84, no. 296, pp. 2561–2583, 2015.

[21] A. Nakhushev, "Sturm-Liouville problem for a second-order ordinary differential equation with fractional derivatives in lower terms," *Doklady Akademii Nauk SSSR*, vol. 234, no. 2, pp. 308–311, 1977, (in Russian).

[22] C. Xinfu and L. Yuan, "Effects of Diffusion and Advection on the Smallest Eigenvalue of an Elliptic Operator and Their Applications," *Indiana University Mathematics Journal*, vol. 61, no. 1, pp. 45–80, 2012.

Information about the authors

Zakharov, Ivan I.—Lecturer of the Department of Higher Mathematics of Moscow State University of Civil Engineering (e-mail: kroshvanya@yandex.ru, ORCID: 0009-0004-7182-838X)

Aleroev, Temirkhan S.—Doctor of Physics and Mathematics, Professor of the Department of Higher Mathematics of Moscow State University of Civil Engineering (e-mail: aleroev@mail.ru, ORCID: 0000-0003-3096-7784)

RUSSIAN METADATA

УДК 517.581, 517.954
PACS 44.05.+e, 45.10.Hj
DOI: 10.22363/2658-4670-2026-34-2-241-252 EDN: JUGPSJ

Об одном методе приближённого решения первой начально-краевой задачи для дробного уравнения диффузии

И. И. Захаров, Т. С. Алероев

Национальный исследовательский Московский государственный строительный университет, Ярославское шоссе, д. 26, 129337, Москва, Российская Федерация

Аннотация. *Предпосылки* Статья посвящена активно развивающемуся в настоящее время разделу дробного исчисления, описывающего аномальную диффузию. Так как корректная форма уравнения дробного порядка, описывающего аномальную диффузию, — вопрос до сих пор открытый, то авторы на примерах моделей, основанных на уравнениях дробной диффузии, рассматривают преимущества того или иного уравнения дробного порядка, представленных для моделирования процесса адвекции-диффузии в средах с фрактальной структурой. *Цель* В работе рассматриваются краевые задачи для уравнения адвекции-диффузии с дробными операторами Капуто и Римана–Лиувилля. Учитывая, что многие авторы рассматривают в своих работах модели с операторами дробного порядка, полученные простой заменой обыкновенных производных на дробные, то авторы работы считают необходимым сравнение дробно дифференциальных уравнений с различными операторами. Так же отметим, что одной из целей работы является построение эффективного приближённого метода решения первой начально-краевой задачи для однородного дробно-дифференциального уравнения. *Методы* Используется аппарат дробного исчисления. В основе рассматриваемого приближённого метода лежит аналитический метод разделения переменных (Метод Фурье). Реализация предложенного приближённого метода и все необходимые вычисления выполняются на языке программирования Matlab. *Результаты* Получены приближённые решения краевых задач для уравнения адвекции-диффузии с дробными операторами Капуто и Римана–Лиувилля. Произведено сравнение полученных решений как между собой, так и с классической моделью, основанной на обыкновенных производных. *Выводы* На основе полученных результатов для моделирования процессов аномальной диффузии, авторами рекомендуется использовать уравнение, содержащее дробный оператор Римана–Лиувилля.

Ключевые слова: приближённые вычисления, дробное исчисление, дробное уравнение адвекции-диффузии, дробная производная Римана–Лиувилля, дробная производная Капуто.

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2026, 34 (2) 253–259
http://journals.rudn.ru/miph

Research article

UDC 539.12
DOI: 10.22363/2658-4670-2026-34-2-253-259 EDN: JKNJJK

Stochastic representation of quantum mechanics and entangled solitons

Yuri P. Rybakov

RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation

(received: September 3, 2025; revised: October 10, 2025; accepted: October 20, 2025)

Abstract. *Background* Using the Einstein's idea on particles-solitons, the stochastic representation of the wave function as a large sum of solitons with random phases is suggested. *Purpose* To illustrate the main idea of the stochastic representation, let us recall the lectures by N. Wiener on nonlinear problems in random theory. *Method* In these lectures Wiener used the so-called α-representation of the wave function for a nonrelativistic particle in three-dimensional space, the wave function being considered as an element of the random Hilbert space with the Gaussian dispersion. *Results* This representation appears to be equivalent, in accordance with the central limiting theorem, to the sum of many complex solitons with random phases. *Conclusions* On the basis of this representation it is possible to substantiate the Born's rule for measuring physical observables and also to explain the "spin–statistics" correlation in quantum mechanics.

Key words and phrases: Stochastic representation, Brioschi identity, "spin–statistics" correlation

For citation: Y. P. Rybakov "Stochastic representation of quantum mechanics and entangled solitons," *Discrete and Continuous Models and Applied Computational Science*, vol. 34, no. 2, pp. 253–259, 2026. DOI: 10.22363/2658-4670-2026-34-2-253-259. EDN: JKNJJK.

1. Introduction

It is worth-while to stress that in the period 1912–1913 the outstanding German physicist Gustav Mie published a series of works devoted to the field theory of matter [1–3]. The main idea of Mie was to dispense with the point-like charges as sources in electrodynamics, the linear Maxwell equations for the electromagnetic field being replaced with some new nonlinear ones, which would admit only regular solutions, without singularities.

This revolutionary approach by Mie promoted Einstein to discover adequate equations of General Relativity for describing gravitational fields. As a result, Einstein formulated a tremendous program of geometrizing physics, where particles were considered as pulsating extended objects, so-called solitons, that is clots of some universal "unitary field" [4–6]. This fundamental approach to creating the new physical picture of the world became popular especially within the scope of the quantum theory [7, 8]. However, the physical origin of that universal "unitary field" was unknown. It should be remarked that the important contributions to solving this enigma was made by the outstanding mathematicians Leonard Euler, Francesco Brioschi, and Norbert Wiener.

2. Euler problem on n squares and Brioschi identity

Euler formulated the following problem on n squares:

Given n real numbers $a_i,\ i = \overline{1,n}$, find n new real numbers $c_i,\ i = \overline{1,n}$, such that: 1) c_i is bilinear in a_k; 2) the following relation is valid:

$$\left(\sum_{i=1}^{n} a_i^2\right)^2 = \sum_{k=1}^{n} c_k^2.$$

In 1748 Euler knew solutions for this problem for $n = 2, 3, 4$ [9, 10]. However, the remarkable German mathematician A. Hurwitz in 1898 showed that for $n > 4$ the solution existed only for $n = 8$. That special solution was found before Hurwitz by the outstanding Italian geometrician Francesco Brioschi (1824–1897) [11], who used the complex projective coordinates (16 spinors Ψ) to study the geometry of the $8D$-space. In particular, Brioschi proved the remarkable identity satisfied by any 8-spinor ψ (semispinor) [12]:

$$j^\mu j_\mu - \tilde{j}^\mu \tilde{j}_\mu = s^2 + p^2 + \vec{v}^2 + \vec{a}^2,$$

where the standard bilinear spinor quantities are used:

$$s = \bar{\psi}\psi, \quad p = \imath\bar{\psi}\gamma_5\psi, \quad \vec{v} = \bar{\psi}\vec{\lambda}\psi,$$

$$\vec{a} = \imath\bar{\psi}\gamma_5\vec{\lambda}\psi, \quad j_\mu = \bar{\psi}\gamma_\mu\psi, \quad \tilde{j}_\mu = \bar{\psi}\gamma_\mu\gamma_5\psi,$$

with Dirac matrices γ_μ, γ_5 and Pauli isotopic ones $\vec{\lambda}$ being included.

Suppose that one constructs the Lagrangian for some ψ-field model, which contains the so-called Higgs potential, the latter one monotonically depending on the invariant (2). Therefore, if one searches for stationary states in our model, the vacuum boundary condition at the space infinity should be valid:

$$\lim_{|\vec{x}|\to\infty} \psi = \psi_0.$$

Thus, fixing the vacuum value of the invariant (2), one gets, in view of (2), the field manifold

$$s^2 + p^2 + \vec{v}^2 + \vec{a}^2 = \text{const},$$

which is homeomorphic to the sphere S^7. On the other hand, the following inclusions take place:

$$S^7 \supset S^3 \supset S^2.$$

Taking into account the nontriviality of the homotopy groups

$$\pi_3(S^3) = \pi_3(S^2) = \mathbb{Z},$$

one can expect due to (2) the existence in our field model of topological solitons endowed with the corresponding topological charges of two kinds. The first one is the winding number $\deg(S^3 \to S^3)$, which can be interpreted as the baryon charge $\mathbb{B}$. The second one refers to the so-called Hopf invariant Q_{H} and can be interpreted as the lepton charge $\mathbb{L}$. The corresponding physical models proved their effectiveness in nuclear physics (Skyrme model [13–15]) and in the theory of lepton interactions (Faddeev model [16–18]). Topological analysis for the stability of solitons appears to be important in many physical models [19–21].

3. Stochastic representation of quantum mechanics

Wiener found a special stochastic representation of the one-particle wave function considered as an element of the random Hilbert space with the Gaussian dispersion [22–26]. Nowadays this representation is known as the stochastic integral by Wiener–Ito–Stratonowich. To show its relation to the Einstein's idea on particles-solitons, one intends to represent the wave function as a large sum of complex solitons with random phases. In accordance with the central limiting theorem [27] this sum behaves as a Gaussian random variable.

As an illustration of this effect let us consider the classical T. Young's 2-slit diffraction experiment, the photons being replaced with the Brioschi solitons $\phi(t,\vec{x}) = \phi(t,\vec{x})^*$ in the real representation. Introducing the Lagrangian density $\mathcal{L}(\phi, \partial_\mu\phi)$ and calculating the canonical momentum

$$\pi(t,\vec{x}) = \frac{\partial\mathcal{L}}{\partial(\partial_t\phi)},$$

consider the auxiliary complex function

$$\varphi(t,\vec{x}) = 2^{-1/2}\,(\nu\phi + \imath\pi/\nu)\,,$$

where the number ν can be found from the normalization:

$$\hbar = \int dV\,|\varphi|^2 \approx \int\limits_{V_0} dV\,|\varphi|^2,$$

with $\hbar$ being the Planck constant and $V_0 = \ell^3$ denoting the proper volume of the soliton. Here dV stands for the elementary 3-volume.

Let us now consider an impenetrable screen with 2 slits, for which the width $d \gg \ell$ is supposed to be large with respect to the proper size ℓ of the particles-solitons. During Young's experiment the particles are pushed to the screen one by one with some fixed initial velocity. Our aim is to find the probability for observing a center $\vec{r}$ of a particle-soliton in some small volume $\Delta V(\vec{X}_0) \gg V_0$ located behind the screen at large distance from it, with $\vec{X}_0$ denoting the central point of this small volume.

Assuming now the Young's experiment with the soliton's scattering to be repeated $N \gg 1$ times, one can define the wave function as follows:

$$\psi(t,\vec{x}) = (\hbar N)^{-1/2}\sum_{j=1}^{N}\varphi_j(t,\vec{x}),$$

where $\varphi_j(t,\vec{x})$ denotes the function (3) for the j-th trial. With the intention to show that $|\psi(t,\vec{x})|^2$ plays the role of the probability density $\rho(t,\vec{x})$ let us calculate the following quantity:

$$\Delta V\,\rho = \int\limits_{\Delta V} dV\,|\psi|^2 = \frac{1}{\hbar N}\Big(\sum_{i=1}^{N} a_{ii} + s\Big),$$

where the denotations are used:

$$s = \sum_{i\neq k=1}^{N} a_{ik}, \quad a_{ik} = \frac{1}{2}\int\limits_{\Delta V} dV\,\left(\varphi_i^*\varphi_k + \varphi_k^*\varphi_i\right).$$

Taking into account the normalization (3), one gets

$$\sum_{i=1}^{N} a_{ii} = \hbar\Delta N, \quad \rho\Delta V = \frac{1}{\hbar N}(\hbar\Delta N + s),$$

where ΔN is the number of trials, for which the center of our soliton can be observed in the domain ΔV. Considering φ_j as independent random variables with zero mean values, one can estimate the probability for $|s|$ to surpass $\hbar\Delta N$ through applying the Chebyshev's inequality [27]:

$$P(|s| > \hbar\Delta N) \leq (\hbar\Delta N)^{-2} < s^2 >,$$

where the symbol <> denotes calculating the mean value.

Taking into account that the profiles φ_i and φ_k for $i \neq k$ can overlap only in the domain of the volume V_0, one obtains the estimate

$$< s^2 > \leq \alpha(\hbar)^2 V_0 \Delta N \frac{\Delta N}{\Delta V},$$

where $\alpha \sim 1$ stands for the effective "packing" coefficient. Thus, unifying (3) and (3), one finds the estimate:

$$P(|s| > \hbar\Delta N) < \alpha \frac{V_0}{\Delta V} \ll 1,$$

whence one concludes that with the probability close to unity the quantity $\rho = |\psi|^2$ can be interpreted as the probability density for measuring the coordinates of the soliton's center behind the screen. It means that the stochastic wave function (3) plays the role of the probability amplitude in accordance with the Born's interpretation.

4. Measuring physical observables

Let us now discuss measuring some physical observable A, which can be expressed within the scope of the Lagrangian formalism, if one follows the E. Nther's theorem in application to soliton configurations considered as island-like systems with the asymptotically flat space-time [20]. In particular, if the generator $\hat{M}_A$ of the corresponding asymptotic symmetry group is given, our physical observable takes the form:

$$A = \int dV(\pi\imath\hat{M}_A\phi) = \int dV(\varphi^*\hat{M}_A\varphi).$$

Applying (4) to the j-th trial, one can express the mean value in question as follows:

$$< A > = \frac{1}{N}\sum_{j=1}^{N} A_j = \frac{1}{N}\int dV \sum_{j=1}^{N}(\varphi_j^*\hat{M}_A\varphi_j).$$

Using the definition (3) of the stochastic wave function and taking into account the estimate (3), one can rewrite (4) with the probability close to unity in the standard quantum mechanical form (the Born's rule):

$$< A > = \int dV(\psi^*\hat{A}\psi),$$

where the Hermitian operator $\hat{A} = \hbar\hat{M}_A$ is introduced.

5. Stochastic representation for the case of several particles. "Spin–statistics" correlation

Let us consider the case of n particles-solitons. Here a new definition of the wave function in the configuration space of the dimension $3n$ should be given as the generalization of (3):

$$\psi_n\left(t, \vec{x}_1, \ldots, \vec{x}_n\right) = (\hbar^n N)^{-1/2} \sum_{j=1}^{N}\prod_{k=1}^{n}\varphi_j^{(k)}\left(t, \vec{x}_k\right),$$

where $\varphi_j^{(k)}$ stands for the soliton configuration of the form (3) for the k-th particle in the j-th trial. In this case the mean value expression for some observable A_n is similar to (4) and ensues from (5):

$$< A_n > = \int dV_n \left(\psi_n^* \hat{A}_n \psi_n\right),$$

where dV_n denotes the elementary volume in the new configuration space and the operator $\hat{A}_n$ in (5) is the superposition of the one-particle generators:

$$\hat{A}_n = \sum_{k=1}^{n} \hbar \hat{M}_A^{(k)}.$$

As can be seen from (5), the wave function has the structure of the so-called entangled solitons [28].

Now it is worth-while to apply the definition (5) to explaining the well-known "spin–statistics" correlation in quantum mechanics. To this end, suppose our particles to be identical, with the soliton profiles $\varphi_j^{(k)}$ belonging to the irreducible representation D^J of the rotation group SO(3), where the weight J is the spin of our particles-solitons. Since for identical particles the profiles $\varphi_j^{(k)}$ should be equal, one can write:

$$\varphi_j^{(k)} = \varphi_j \left(\vec{x}_k - \vec{r}_k\right),$$

with $\vec{r}_k$ denoting the center of the soliton in question.

Let us now choose two arbitrary particles with the centers $\vec{r}_1$, $\vec{r}_2$ and perform their transposition. It means that our two-particle system is rotated over angle π around the bisector $\vec{b}$ of the central line $\vec{r}_1 - \vec{r}_2$. However, in this position the extended character of our particles-solitons comes into play. In fact, to restore the previous state of the system, it would be inevitable to perform proper rotations of our particles-solitons over angle π around their central axes parallel to $\vec{b}$. As can be easily seen, this latter operation appears to be equivalent to the relative rotation of our two particles over angle 2π. Due to the properties of the representation D^J the 2π-rotation implies multiplying the wave function (5) by $(-1)^{2J}$. Therefore, one concludes that the wave function should be symmetrical under transpositions of particles for the integer spin J and antisymmetrical — for the semi-integer J. This property of the wave function corresponds to the well-known "spin–statistics" correlation in quantum mechanics.

6. Results and Discussions

Motivated by the works of Gustav Mie on field theory of matter, Einstein formulated a tremendous program of geometrizing physics, where particles are represented as solitons, that is clots of some fundamental "unitary field". Basing on the central limiting theorem, one can construct the stochastic representation of the wave function as a large sum of solitons with random phases, the wave function being considered as a Gaussian random variable. This fact permits one to interpret the wave function as the probability amplitude for observing center coordinates of particles-solitons. Moreover, it's possible to substantiate the Born's rule for measuring physical observables as bilinear functionals in the wave function and also to explain the well-known "spin–statistics" correlation in quantum mechanics.

7. Conclusions

Taking into account the idea by G. Mie and A. Einstein on particles-solitons, the stochastic representation of the wave function is suggested. Using this new picture, it's possible to explain the M. Born's quantum rule for measuring physical observables and also the "spin–statistics" correlation.

Author Contributions: Conceptualization, Rybakov Yu. P.; methodology, Rybakov Yu. P.; writing—review and editing Rybakov Yu. P.; supervision, Rybakov Yu. P.; project administration, Rybakov Yu. P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data sharing is not applicable.

Acknowledgments: Non profit acknowledgments.

Conflicts of Interest: The authors declare no conflict of interest.

Declaration on Generative AI: The authors have not employed any Generative AI tools.

References

[1] G. Mie, "Grundlagen einer Theorie der Materie, I," *Ann. der Physik*, vol. 37, pp. 511–534, 1912.
[2] G. Mie, "Grundlagen einer Theorie der Materie, II," *Ann. der Physik*, vol. 39, pp. 1–40, 1912.
[3] G. Mie, "Grundlagen einer Theorie der Materie, III," *Ann. der Physik*, vol. 40, pp. 1–66, 1913.
[4] A. Einstein, "On the generalized theory of gravitation," *Sci. Amer.*, vol. 182, pp. 13–70, 1950.
[5] A. Einstein, "Physics and reality," *J. of Franklin Inst.*, vol. 221, pp. 349–382, 1936.
[6] A. Einstein, *The meaning of relativity*. 1955.
[7] D. Bohm, *Causality and chance in modern physics*. Hurrisburg: Univ. of Pennsilvania Press, 1957.
[8] L. de Broglie, *Une tentative d'interprétation causale et nonlinéaire de la mécanique ondulatoire: la téorie de la double solution*. Paris: Gauthier–Villars, 1956.
[9] J. H. Conway and D. A. Smith, *On quaternions and octonions: their geometry, arithmetic, and symmetry*. Natick, Massachusets: A. K. Peters, Ltd, 2003.
[10] D. A. Grave, *Treatise on algebraic analysis, Vol. 1, Science foundations*. Kiev: Publishing House of Ukr. Acad. Sci., 1938.
[11] M. Noether, "Francesco Brioschi," *Math. Ann.*, vol. 50, pp. 477–491, 1898. DOI: 10.1007/BF01444296
[12] E. Cartan, *The theory of spinors*. Paris: Hermann, 1966.
[13] T. H. R. Skyrme, "A nonlinear theory of strong interactions," *Proc. Roy. Soc. Lond., ser. A*, vol. 247, pp. 260–278, 1958.
[14] T. H. R. Skyrme, "A nonlinear field theory," *Proc. Roy. Soc. Lond., ser. A*, vol. 260, pp. 127–138, 1961.
[15] T. H. R. Skyrme, "A unified field theory of mesons and baryons," *Nucl. Phys.*, vol. 31, pp. 556–569, 1962.
[16] L. D. Faddeev, "Some comments on many-dimensional solitons," *Lett. Math. Phys.*, vol. 1, pp. 283–293, 1976.
[17] L. D. Faddeev, "Gauge invariant model of electromagnetic and weak interactions of leptons," *Rep. Acad. Sci. USSR*, vol. 210, pp. 807–810, 1973.
[18] L. D. Faddeev, "Hadrons from leptons," *JETP Lett.*, vol. 21, pp. 141–144, 1975.
[19] R. Palais, "The principle of symmetric criticality," *Comm. Math. Phys.*, vol. 69, pp. 19–30, 1979.
[20] B. Felsager, *Geometry, particles and fields*. Odense: Odense Univ. Press, 1981.
[21] D. Finkelstein, "Kinks," *J. Math. Phys.*, vol. 7, pp. 1218–1221, 1966. DOI: 10.1063/1.1705025

[22] N. Wiener, *Nonlinear problems in random theory*. New York: John Wiley & Sons, inc., 1958.
[23] T. Hida, *Brownian motion*. New York, Heidelberg, Berlin: Springer–Verlag, 1980.
[24] K. Itô, "Multiple Wiener integral," *J. Math. Soc. Japan*, vol. 3, pp. 157–169, 1951.
[25] R. L. Stratonowitch, *Selected items of fluctuations theory in radiotechnology*. Moscow: Radio USSR, 1961.
[26] H. H. Kuo, *Gaussian measures in Banach spaces*. Berlin, Heidelberg, New York: Springer–Verlag, 1975.
[27] M. Loève, *Probability theory*. D. van Nostrand comp., inc., 1960.
[28] Y. P. Rybakov, "On the causal interpretation of quantum mechanics," *Found. Phys.*, vol. 4, pp. 149–161, 1974.

Information about the authors

Rybakov, Yuri P. (Russian Federation)—Professor, Doctor of Sciences in Physics and Mathematics, Professor at the Institute of Physical Research and Technologies of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: rybakov-yup@rudn.ru, phone: +7(916)262-55-36, ORCID: 0000-0002-7744-9725, ResearcherID: S-4813-2018, Scopus Author ID: 16454766600)

RUSSIAN METADATA

УДК 539.12
DOI: 10.22363/2658-4670-2026-34-2-253-259 EDN: JKNJJK

Стохастическое представление квантовой механики и запутанные солитоны

Ю. П. Рыбаков

Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация

Аннотация. *Актуальность* Используя идею Эйнштейна о частицах-солитонах, предлагается стохастическое представление волновой функции в виде большой суммы солитонов со случайными фазами. *Цель* Для иллюстрации основной идеи стохастического представления вспомним лекции Н. Винера по нелинейным задачам в теории случайных величин. *Метод* В этих лекциях Винер использовал так называемое α-представление волновой функции для нерелятивистской частицы в трёхмерном пространстве, рассматривая волновую функцию как элемент случайного гильбертова пространства с гауссовой дисперсией. *Результаты* Это представление, согласно центральной предельной теореме, оказывается эквивалентным сумме многих комплексных солитонов со случайными фазами. *Выводы* На основе этого представления можно обосновать правило Борна для измерения физических наблюдаемых величин, а также объяснить корреляцию «спин–статистика» в квантовой механике.

Ключевые слова: стохастическое представление, тождество Бриоски, корреляция спин–статистика

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2026, 34 (2) 260–273
http://journals.rudn.ru/miph

Research article

UDC 004.89; 004.93
PACS 07.05.Mh 07.05.Tp 42.62.Be
DOI: 10.22363/2658-4670-2026-34-2-260-273 EDN: JMNMUP

Efficient Polyp Segmentation and Classification with YOLO11–SAM3 and LoRA

Eugene Yu. Shchetinin[1], Leonid A. Sevastianov[2], Anastasiia A. Tiutiunnik[2]

[1] *Sevastopol State University, 33 Universitetskaya St, Sevastopol, 299053, Russian Federation*

[2] *RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation*

(received: February 22, 2026; revised: March 12, 2026; accepted: March 20, 2026)

Abstract. *Purpose* To develop and experimentally evaluate an autonomous computer–aided diagnosis (CADx) system for colonoscopy that integrates polyp segmentation with binary histological classification (adenoma/non–adenoma) while maintaining computational efficiency for near real–time clinical deployment. *Materials and Methods* We propose a cascaded YOLO11–SAM3–LoRA architecture wherein a YOLO11–medium detector generates region proposals (bounding boxes), followed by SAM 3 performing box–prompted segmentation with parameter–efficient domain adaptation via Low–Rank Adaptation (LoRA) and a semantic anchoring mechanism based on text prompt ensembling. Histological classification is implemented through a dual–stream classifier that fuses region–of–interest (ROI) features extracted by EfficientNet–B0 with morphological mask descriptors. Training was conducted on the combined Kvasir–SEG and CVC–ClinicDB datasets (detection/segmentation) and CVC–HDClassif (classification). Generalization was evaluated on the external ETIS–Larib dataset under an external zero–shot protocol. *Results* On the primary test set, the system achieved a mean Dice coefficient (mDice) of 0.962 for segmentation, adenoma sensitivity of 92.3%, and area under the receiver operating characteristic curve (AUROC) of 0.942 for binary classification. Under external zero–shot evaluation on ETIS–Larib, the system maintained segmentation quality (mDice = 0.884), indicating robustness to domain shift. End–to–end processing throughput reached 38.2 FPS on an NVIDIA A100 GPU. *Conclusions* The results demonstrate that the combination of cascaded localization, foundation model segmentation with LoRA adaptation, and text prompt ensembling achieves high segmentation quality and competitive CADx accuracy while preserving computational efficiency. Clinical deployment requires further prospective multicenter validation and error analysis stratified by lesion subtypes.

Key words and phrases: polyp segmentation, polyp classification, computer–aided diagnosis, CADx, colonoscopy, YOLO11, SAM 3, LoRA, prompt ensembling, deep learning

For citation: E. Y. Shchetinin, L. A. Sevastianov, A. A. Tiutiunnik "Efficient Polyp Segmentation and Classification with YOLO11–SAM3 and LoRA," *Discrete and Continuous Models and Applied Computational Science*, vol. 34, no. 2, pp. 260–273, 2026. DOI: 10.22363/2658-4670-2026-34-2-260-273. EDN: JMNMUP.

1. Introduction

The substantial mortality burden associated with colorectal cancer (CRC) necessitates improvements in early detection quality. Although colonoscopy remains the primary screening modality, its effectiveness is constrained by inter-operator variability: systematic reviews indicate that the adenoma miss rate (AMR) ranges from approximately 14 to 30% [1]. This observation motivates the development of computer–aided systems capable of enhancing lesion detection reliability (CADe) and supporting optical stratification (CADx) under real–time constraints. When appropriately validated according to clinical criteria, CADx approaches may potentially support resect–and–discard strategies, thereby reducing unnecessary histopathological examinations and associated costs [2].

The evolution of automated endoscopic analysis has progressed from specialized convolutional architectures to the adaptation of foundation models. Early segmentation solutions based on U-Net [3] and its variants (UNet++ [4], ResUNet [5]) achieved high accuracy on images with well-defined boundaries; however, their limited effective receptive field and sensitivity to local textures could lead to errors on typical endoscopic artifacts (mucosal folds, specular reflections, air bubbles). Subsequent developments incorporated attention mechanisms and multi-scale contextual aggregation (e.g., PraNet [6]) alongside transformer-based architectures (e.g., Polyp–PVT [7]), which better capture global dependencies but frequently require trade-offs between computational complexity, spatial resolution, and video-rate inference speed.

At the current stage, two complementary research directions can be identified, each addressing only part of the clinical requirements. First, object–centric detectors of the YOLO family [8, 9] provide high throughput suitable for video streams; however, their bounding box outputs are insufficient for precise boundary delineation and morphological assessment–both critical for measurement and subsequent texture–based or morphological classification. Second, promptable foundation segmentation models (SAM/MedSAM/SAM2/SAM3) can produce detailed masks and demonstrate transferability across diverse visual conditions, yet their application typically requires external prompts (points, boxes, or text queries) and may exhibit sensitivity to domain shift in the specialized endoscopic environment [10–13]. This complicates the construction of fully autonomous pipelines without additional localization mechanisms and domain adaptation, particularly for low–contrast and morphologically ambiguous lesions including flat and sessile serrated polyps.

Existing attempts to combine detectors with foundation segmentation models (e.g., YOLO–SAM hybrids) are often limited to naive cascading, which in certain implementations leads to: (i) redundant computation when multiple localization mechanisms operate simultaneously, and (ii) underutilization of semantic prompts and textual concepts. Accordingly, the scientific and applied objective of this study is to develop an autonomous and computationally efficient mechanism for localization and semantic adaptation that transfers the advantages of foundation segmentation to the specialized colonoscopy domain while minimizing manual prompting requirements and preserving suitability for near real–time deployment.

1.1. Scientific Novelty and Contributions

The principal contributions of this work are as follows:

1. **Autonomous cascade architecture for colonoscopy detection and foundation–based segmentation.** We propose a computationally efficient YOLO11–SAM3–LoRA architecture that combines rapid candidate localization with subsequent semantic boundary refinement. This design jointly satisfies video-stream requirements (throughput) and clinically meaningful segmentation (precise masks) within a unified automatic pipeline.

2. **Semantic stabilization via text prompt ensembling.** We develop a semantic anchoring mechanism based on an ensemble of clinical terminology and aggregated textual representations, which enhances prompting robustness to formulation variability and domain shift. Ablation analysis demonstrates a measurable contribution of this component to mDice compared with single–prompt baselines.
3. **Integration of segmentation with CADx classification and external generalization assessment.** We implement a dual–stream classification scheme that fuses ROI features with morphological mask descriptors and conduct external zero-shot evaluation on ETIS–Larib. The obtained results (mDice = 0.884 on ETIS-Larib; AUROC = 0.942 for binary histological classification) indicate practical transferability of the proposed configuration to conditions differing from the training distribution.

2. Materials and Methods

2.1. Datasets

The following datasets were employed for training, evaluation, and generalization assessment.

2.1.1. Training and Validation (Detection and Segmentation)

Kvasir–SEG [14] and CVC–ClinicDB [15] were used for training detection and segmentation tasks. The combined volume comprises 1,612 images with binary polyp masks.

2.1.2. Training (Classification)

CVC–HDClassif [16] was employed for training the binary classifier, as it contains histologically verified labels (adenoma/non–adenoma). The class ratio (adenomas/non–adenomas) was 1.32:1; class imbalance was addressed through weighted loss functions.

2.1.3. In-Domain Testing

For in–domain evaluation, a held–out test partition comprising 20% of the training datasets was utilized, allocated at the image level with a fixed seed (seed = 5) and stratified by dataset to maintain proportional representation. Images from the external ETIS–Larib dataset were excluded from the test set.

2.1.4. External Zero–Shot Evaluation

ETIS-Larib Polyp DB [17] contains 196 images and exhibits pronounced domain shift relative to the training data (lighting conditions, contrast, equipment, and scene characteristics). External evaluation followed a zero–shot protocol: ETIS data were not used for training or hyperparameter tuning. All images were resized to an input resolution of 512×512 pixels and normalized identically to the training pipeline. Augmentations were applied exclusively to the training partition and were not employed during metric computation on ETIS.

2.2. Model Architecture

The system implements a three–stage processing pipeline as illustrated in Figure 1.

Table 1

Configuration of trainable parameters in the YOLO11–SAM3–LoRA model

Component	**Status**	**Method**
YOLO11 Backbone	Frozen	COCO pretraining
SAM 3 Image Encoder	Frozen	SA–Co pretraining
SAM 3 Attention Projections	Trainable	LoRA ($r = 16$)
SAM 3 Text Prompt Encoder	Frozen	Built–in SAM 3 encoder
Classification Module (EfficientNet–B0 + MLP)	Trainable	Full training

2.2.1. Coarse Detection (Region Proposal Generation)

YOLO11–medium [18] serves as the region proposal network. It performs initial background filtering and generates bounding boxes $B_i = [x, y, w, h]$ with associated confidence scores C_d. Unlike standard SAM deployment, the built–in DETR detector in SAM 3 is not utilized to avoid computational redundancy and prediction conflicts.

2.2.2. Semantic Refinement with Prompt Ensembling

For each candidate B_i, the SAM 3 model [19] is applied in box-prompted mode. To enhance semantic selectivity, instead of a single prompt, we employ an ensemble of clinical concepts T_{ens}:

$$T_{\text{ens}} = \{\text{“colorectal polyp”}, \text{“sessile lesion”}, \text{“adenoma”}, \text{“mucosal anomaly”}\}$$

For semantic prompting, we utilize the built–in text encoder of SAM 3. We construct an ensemble of K textual class descriptions (prompt ensemble), obtain corresponding embeddings, normalize them, and aggregate via averaging (with subsequent normalization), which reduces sensitivity to single-prompt formulation. Notably, this approach enables confident detection of low–contrast flat neoplasms that are systematically missed by the model when using a unitary “polyp” prompt.

2.2.3. Low–Rank Adaptation (LoRA)

For domain adaptation to medical imaging, trainable low-rank matrices ($r = 16$, $\alpha = 32$) are injected into the attention layers of the SAM 3 encoder. Weights are updated according to:

$$W' = W_{\text{frozen}} + \Delta W = W_{\text{frozen}} + BA$$

where $A \in \mathbb{R}^{r\times k}$, $B \in \mathbb{R}^{d\times r}$, and W_{frozen} denotes the frozen pretrained weights [20]. At rank $r = 16$, only 2.03% of parameters are trainable, providing a balance between plasticity (adaptation to intestinal textures) and stability (preservation of boundary detection capabilities). The parameter distribution is summarized in Table 1.

2.2.4. Dual–Path Classifier

Traditional approaches utilize mask features for classification; however, histological type is determined by surface texture (pit pattern) and vascularization, which may be lost in binary masks. The YOLO11–SAM3–LoRA classifier integrates two streams:

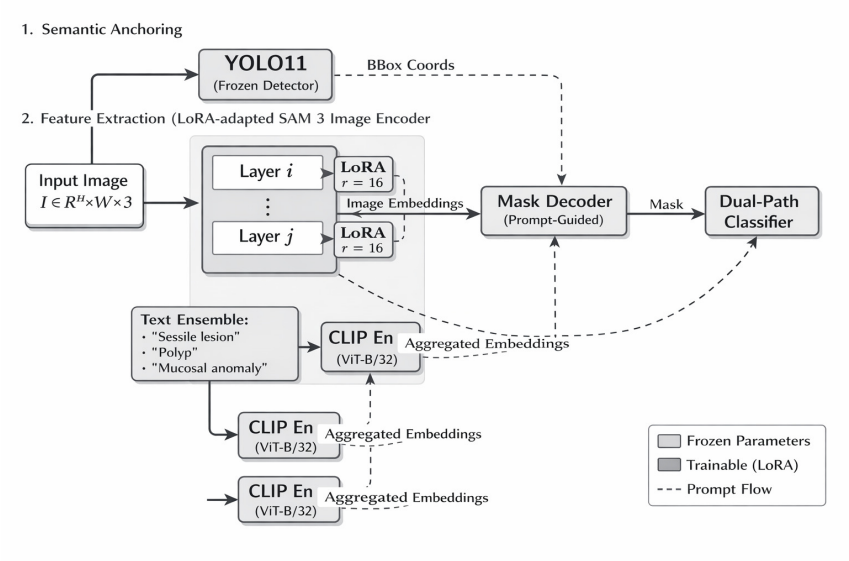

Figure 1. Block diagram of the proposed CADx system architecture: candidate generation using YOLO11, semantic refinement based on SAM 3, and dual–path classification

Visual Path ($V_{\mathbf{vis}}$). Feature extraction within the image ROI is performed using the lightweight convolutional network EfficientNet–B0, enabling texture and vascular pattern analysis.

Morphological Path ($V_{\mathbf{morph}}$). Geometric mask descriptors are analyzed via a multilayer perceptron (MLP), including area, perimeter, shape factor, and boundary irregularity.

The final class probability P_{cls} is computed by applying the softmax function to the output of a fully connected layer over the concatenated feature vectors V_{vis} and V_{morph}:

$$P_{\text{cls}} = \text{Softmax}\left(\text{MLP}\left(V_{\text{vis}} \oplus V_{\text{morph}}\right)\right).$$

3. Results

3.1. Comparative Segmentation Performance Analysis

Table 2 presents comparative analysis results on the primary test set derived from Kvasir–SEG and CVC–ClinicDB. These datasets are characterized by high image quality and standardized lighting conditions, enabling assessment of maximal architectural potential. Evaluation on standard datasets (Kvasir/CVC) was conducted using 5–fold cross–validation.

The proposed YOLO11–SAM3–LoRA architecture demonstrates superior performance across all metrics on the combined Kvasir–SEG and CVC–ClinicDB test set, achieving mDice of 0.962 and mIoU of 0.927. This confirms the effectiveness of the hierarchical approach integrating YOLO11 for candidate generation with SAM 3 featuring LoRA adaptation and prompt ensembling, providing a balance between segmentation precision and recall.

Table 2
Comparative segmentation performance of the proposed architecture against state–of–the–art methods on the combined Kvasir–SEG and CVC–ClinicDB test set

Model	**mDice**	**mIoU**	**FPS (A100)**
U–Net [3]	0.678	0.530	55.1
PraNet [6]	0.899	0.845	45.0
ColonSegNet [21]	0.820	0.756	140
UNet++ [4]	0.821	0.710	35
ResUNet++ [22]	0.842	0.740	32.0
YOLOv8–SAM2 [23]	0.908	0.830	44.0
Polyp–PVT [7]	0.917	0.846	28
MedSAM [11]	0.951	0.907	31
YOLO11–SAM3 (Ours)	**0.962**	**0.927**	38.2

Compared to U–Net (mDice = 0.678, mIoU = 0.530), the improvement in mDice is 28.4 percentage points (pp) and in mIoU is 39.7 pp. ColonSegNet (mDice = 0.820, mIoU = 0.756) achieves high throughput (140 FPS), surpassing U–Net by 14.2 pp in mDice, but underperforming the proposed system by 14.2 pp in mDice and 17.1 pp in mIoU. PraNet (mDice = 0.899, mIoU = 0.845) exceeds U-Net by 22.1 pp in mDice but falls short of the proposed system by 6.3 pp in mDice and 8.2 pp in mIoU at 45.0 FPS.

UNet++ variants, including the original (mDice = 0.821, mIoU = 0.710) and ResUNet++ (mDice = 0.842, mIoU = 0.740), demonstrate improvements through nested structures and residual connections, enhancing recall and F1-score relative to baseline U-Net; however, they still underperform the proposed system by 12.0–14.1 pp in mDice and 18.7–21.7 pp in mIoU at lower processing speeds (32–35 FPS versus 38.2 FPS). The hybrid YOLOv8-SAM2 model (mDice = 0.908, mIoU = 0.830) represents a strong competitor, demonstrating good performance on CVC–ClinicDB but with notable degradation on Kvasir–SEG, indicating reduced robustness to inter–dataset variation; the proposed model surpasses it by 9.7 pp in mIoU, although YOLOv8–SAM2 achieves higher throughput (44.0 FPS).

Transformer–based and SAM–derived approaches, including Polyp-PVT (mDice = 0.917, mIoU = 0.846) and MedSAM (mDice = 0.951, mIoU = 0.907), achieve high performance through pyramidal structures and semantic refinement but underperform the proposed system by 1.1–4.5 pp in mDice and 2.0–8.1 pp in mIoU at lower throughput (28–31 FPS). The overall improvement underscores the system's potential for minimizing polyp miss rates in real–time "optical biopsy" scenarios; however, validation on prospective clinical data and additional datasets is necessary to confirm generalizability and reduce false–positive rates.

Note: "pp" (percentage points) denotes the absolute difference between metric values: $\Delta_{\text{pp}} = (x_{\text{new}} - x_{\text{base}}) \times 100$.

3.2. Generalization Capability Assessment

For objective evaluation of the proposed architecture's effectiveness, comparison was conducted against a broad spectrum of methods, including classical CNNs (UNet, UNet++, ResUNet++),

Table 3

Comparative zero–shot generalization performance on the independent ETIS–Larib dataset

Model	mDice	mIoU	FPS	Notes
U–Net [3]	0.398	0.248	55.1	Severe degradation on small objects and artifacts
ColonSegNet [21]	0.658	0.583	140	Moderate robustness, but misses flat lesions
PraNet [6]	0.709	0.628	45.0	Good boundary accuracy, but degradation under complex lighting
UNet++ [4]	0.762	0.616	35	Significant drop on small objects
ResUNet++ [22]	0.793	0.666	32.0	Improved boundaries, but limited generalization to flat morphologies
YOLOv8–SAM2 [23]	0.840	0.724	44.0	Unstable boundaries
Polyp–PVT [7]	0.865	0.762	28	Good generalization, but misses on artifacts
MedSAM [11]	0.815	0.688	31	Misses flat lesions
YOLO11–SAM3 (Ours)	**0.884**	**0.792**	38.2	Robust to artifacts and flat morphologies

specialized attention-based models (PraNet, Polyp–PVT), and contemporary hybrid approaches (MedSAM, YOLOv8–SAM2). To assess system robustness to challenging cases (small and flat polyps), testing was performed on the ETIS–Larib dataset without additional fine–tuning (zero–shot). Neither the YOLO11 detector nor the SAM 3 segmentor underwent fine-tuning or hyperparameter optimization on ETIS images. For each ETIS image, the detector generated bounding boxes; the box with maximum confidence was passed to SAM 3. If no box exceeded the confidence threshold τ, the algorithm returned an empty mask, and metrics for that image were recorded as zero (e.g., Dice = 0). Segmentation metrics were computed at the pixel level per image and subsequently averaged (macro–average). For robust uncertainty estimation, 95% confidence intervals were additionally calculated via bootstrap resampling across images.

Table 3 presents generalization capability (zero–shot) evaluation results on the independent ETIS–Larib dataset. This dataset was selected as a stress test due to its high proportion of small (<6 mm) and flat (sessile) polyps, as well as challenging lighting conditions, substantially differentiating it from the training distribution (Kvasir/CVC).

Analysis of generalization metrics on the independent ETIS–Larib dataset (Table 3) yields the following observations:

Performance under domain shift. Despite the expected decline in absolute metrics compared to in-domain evaluation, the proposed YOLO11–SAM3–LoRA system demonstrated the highest segmentation quality on ETIS–Larib: mDice = 0.884. The advantage over MedSAM (mDice = 0.815) was +6.9 pp, and over the nearest hybrid competitor YOLOv8–SAM2 (mDice = 0.840) was +4.4 pp. The observed superiority may be attributed to the combination of cascaded detection and domain adaptation of the segmentor (LoRA), which potentially enhances robustness to low-contrast

boundaries under varying illumination and endoscopic equipment characteristics. Relative to classical CNN models U–Net (mDice = 0.398) and UNet++ (mDice = 0.762), the advantage is +48.6 pp and +12.2 pp respectively, underscoring the limited transferability of certain CNN architectures in zero–shot scenarios.

Robustness to quality degradation. Classical convolutional networks (e.g., ResUNet++: mDice = 0.793; PraNet: mDice = 0.709) exhibit notable quality degradation on the external dataset. In contrast, the proposed system maintains metrics above 0.88, indicating more robust target class representation under domain shift. For confirmation of statistical significance of these differences, confidence intervals and/or significance tests across ETIS images should additionally be reported.

Role of prompts and clinical interpretation. Results on ETIS–Larib are consistent with the hypothesis that text prompt ensembling may reduce sensitivity to formulation variability and help focus the model on subtle mucosal changes. In particular, inclusion of formulations such as "sessile lesion" and "mucosal anomaly" in the ensemble may facilitate better localization of flat/low–contrast lesions; however, claims regarding specific effectiveness on subtypes (e.g., SSL) require separate quantitative analysis by subgroups with corresponding annotations.

Overall, the mDice = 0.884 result on ETIS–Larib demonstrates the competitiveness of the proposed system under domain shift conditions. Nevertheless, further validation on additional independent datasets and, where possible, prospective clinical data is necessary to confirm generalizability and robustness to false-positive detections.

3.3. Classification Results

Table 4 presents quantitative diagnostic performance metrics for the dual–stream classifier integrating visual (ROI) and morphological segmentation features. On the test set, the classifier achieved Accuracy = 0.891 (95% CI: 0.862–0.920). For CADx/optical diagnosis applications, the key safety metric is sensitivity to adenomatous (neoplastic) lesions, which reached Sensitivity = 0.923 (95% CI: 0.884–0.962), indicating a low false-negative rate within this sample. Specificity was 0.847 (95% CI: 0.812–0.882), reflecting a moderate tendency of the algorithm to classify borderline cases as adenomas. Such a bias toward enhanced sensitivity is often considered preferable from the perspective of minimizing clinically significant lesion misses; however, it is accompanied by increased false-positive decisions and may potentially lead to additional interventions and costs.

The integrated metric AUROC = 0.942 (95% CI: 0.923–0.961) indicates high discriminative ability across the full range of classification thresholds. The relatively narrow confidence intervals suggest estimation stability on this test set (under the chosen CI computation method). Confusion matrix analysis revealed that a substantial proportion of false-positive detections within the binary labeling scheme corresponded to sessile serrated lesions (SSL), which may morphologically mimic adenomas and represent a known diagnostic challenge even for expert endoscopists [24]. Integration of the morphological stream partially reduces the frequency of such errors; further improvement may be achieved through extension to multiclass classification with SSL as a separate class, potentially improving specificity without substantial sensitivity loss.

3.4. Ablation Study

The conducted ablation analysis, results of which are presented in Table 5, enabled quantitative assessment of each architectural component's contribution to overall system performance. Comparisons were performed across 5 mDice values (one per fold); paired Wilcoxon tests ($\alpha = 0.05$) were employed. Differences between the full model and ablation configurations were evaluated at the fold level of 5–fold cross–validation. For each configuration, mDice was computed separately per fold,

Table 4

Performance metrics of the dual–stream classifier for binary histological polyp classification. $N_{\text{total}} = 470$, $N_{\text{adenoma}} = 272$, $N_{\text{non-adenoma}} = 198$, TP = 251, FN = 21, Sens = 251/272 = 0.923, TN = 168, FP = 30, Spec = 168/198 = 0.848. Confidence intervals were estimated via clustered bootstrap by polyps/patients (2,000 iterations): at each iteration, cluster resampling with replacement was performed, followed by metric recalculation across all frames of selected clusters; 95% CI defined as 2.5th and 97.5th percentiles.

Metric	**Value**	**95% CI**
Accuracy	0.891	[0.862; 0.920]
Sensitivity (Adenoma)	0.923	[0.884; 0.962]
Specificity (Non–adenoma)	0.847	[0.812; 0.882]
AUROC	0.942	[0.923; 0.961]

after which pairwise differences (Full − Ablation) were compared using the Wilcoxon signed–rank test (two–sided). p-values were corrected via the Holm method for multiple comparisons. Effect size is reported as rank–biserial correlation (r_{rb}). Additionally, 95% confidence intervals for mean ΔmDice were calculated via bootstrap across folds. Sequential component exclusion revealed the following significance hierarchy:

Role of the specialized detector (YOLO11). The greatest quality reduction was observed when replacing the external YOLO11 detector with the built–in SAM 3 localization mechanism (DETR): mDice decreased from 0.962 to 0.933 (ΔmDice = −0.029). This indicates that under endoscopic conditions, candidate localization quality is a critical factor for stable segmentation: the specialized detector better rejects typical artifacts (mucosal folds, specular reflections, air bubbles) and provides the segmentor with more accurate spatial prompts.

Significance of semantic ensembling. Replacing the prompt ensemble with a single "polyp" prompt led to mDice reduction of 0.009 (0.962 → 0.953). Despite the small absolute magnitude, this component enhances robustness to formulation variability and, according to qualitative error analysis observations, particularly affects "difficult" cases (flat/low–contrast lesions). For quantitative confirmation of this effect, stratified error analysis by subgroups (flat/SSL, etc.) and/or external dataset results should additionally be reported.

Comparison of LoRA and full fine–tuning. Replacing LoRA adaptation with full fine–tuning yielded similar mDice values (0.962 vs. 0.959; ΔmDice = −0.003; NS), indicating comparable effectiveness of parameter–efficient adaptation within this protocol. Practically, this means LoRA can achieve comparable quality at substantially reduced training computational costs. For claims regarding domain shift robustness, additional comparison of LoRA, full fine–tuning, and no adaptation on an external dataset (e.g., ETIS–Larib) with confidence intervals is recommended.

Thus, the proposed YOLO11 + SAM 3 (LoRA) + Prompt Ensemble configuration represents an optimal balance wherein each component addresses a specific challenge: YOLO ensures localization, LoRA provides domain adaptation, and the prompt ensemble ensures semantic stability.

The ablation analysis (Table 5) demonstrates that the specialized localization stage contributes most substantially to final segmentation accuracy: replacing the external YOLO11 detector with the built–in SAM 3 localization mechanism (DETR) leads to mDice reduction of 0.029. Abandoning the prompt ensemble in favor of a single "polyp" prompt reduces mDice by 0.009, consistent with the role of semantic stabilization in challenging/low–contrast cases. Finally, replacing LoRA adaptation with full fine–tuning yields similar mDice values ($\Delta = -0.003$), indicating comparable quality with more

Table 5
Ablation study results: quantitative assessment of architectural component contributions and statistical significance of changes. * indicates values to be computed; NS = not significant

Configuration	**mDice** (mean)	**ΔmDice** vs Full	p (Wilcoxon)	$p_{\mathbf{adj}}$ (Holm)	**95% CI** (ΔmDice)	$r_{\mathbf{rb}}$
Full Model (YOLO11 + SAM3 + LoRA + Prompt Ensemble)	0.962	—	—	—	—	—
SAM3 native localization (DETR) instead of YOLO11	0.933	−0.029	*	*	*	*
Single prompt (“polyp”) instead of ensemble	0.953	−0.009	*	*	*	*
Full fine-tuning instead of LoRA	0.959	−0.003	NS	NS	*	*
No adaptation (SAM3 frozen)	0.951	−0.011	*	*	*	*

parameter–efficient tuning; statistical conclusions are confirmed by fold–level tests with multiple comparison correction (p_{adj}, Holm) and effect sizes (r_{rb}).

4. Discussion

In this study, we presented a cascaded YOLO11–SAM3–LoRA system for endoscopic CAD that combines a rapid candidate localization stage with subsequent boundary refinement using a foundation approach augmented by parameter–efficient domain adaptation. Experimental evaluation on open datasets demonstrates high segmentation metrics on in-domain testing and quality preservation under external assessment on ETIS–Larib, as well as high discriminative ability in binary histological classification. Collectively, results indicate that the combination of specialized localization, LoRA adaptation, and semantic prompts can enhance system robustness to endoscopic condition variability while preserving computational efficiency–critical for near real–time decision support scenarios.

Comparison with baseline architectures (U–Net, U–Net++, ResU–Net++, PraNet) and hybrid competitors (YOLOv8–SAM2), presented in Tables 2–3, demonstrates the advantage of the proposed approach. Specifically, the difference relative to YOLOv8–SAM2 was +1.8 pp in mDice on combined in–domain evaluation (Kvasir–SEG + CVC–ClinicDB) and +4.4 pp in mDice on ETIS–Larib (external zero–shot). These differences may be attributed to cascaded localization characteristics and subsequent mask refinement at the individual frame level, as well as parameter–efficient domain adaptation of the segmentation component (LoRA). For rigorous causal inference, additional comparison of SAM2 and SAM3 under identical localization and inference protocols is required. These findings align with recent comprehensive reviews of intelligent colonoscopy systems [25].

4.1. Generalization Capability and Semantic Anchoring

A key indicator of system transferability is external evaluation results on ETIS–Larib (Table 3), where under domain shift conditions the model maintained high values (mDice = 0.884; mIoU = 0.792). Ablation analysis (Table 5) demonstrates that prompt ensemble usage provides measurable contribution to the final metric (ΔmDice relative to single prompt). The observed advantage is consistent with the hypothesis that prompt ensembling reduces sensitivity to formulation variability

and may be particularly useful for low–contrast/flat lesions. Compared to baseline models, the advantage reaches +6.9–48.6 pp in mDice (depending on baseline), underscoring the limited capability of certain CNN architectures in zero-shot scenarios and the potential role of prompt engineering in enhancing foundation approach transferability.

4.2. Error Analysis and Clinical Interpretation

Confusion matrix analysis reveals that a substantial proportion of false-positive detections (FP) in the current binary formulation is associated with sessile serrated lesions (SSL), which may morphologically mimic adenomas and represent a challenging differential diagnosis even for expert endoscopists [24]. The proportion of SSL among FP should be specified as X% (Y/Z) to avoid formulation ambiguity. High adenoma sensitivity (Sensitivity = 92.3%; Table 4) reduces false-negative risk within this test sample, while specificity of 84.7% reflects bias toward "cautious" classification. However, compliance with clinical criteria for resect–and–discard strategies (e.g., PIVI–oriented NPV/high–confidence metrics and surveillance interval agreement) was not evaluated in this work; therefore, clinical implications should be interpreted as preliminary. For comparisons with "typical AUROC values," references to relevant CADx studies should be provided.

4.3. Computational Efficiency and Deployment

Inference throughput of 38.2 FPS on NVIDIA A100 indicates potential for integration into high-performance endoscopic systems under comparable settings (input resolution, batch = 1, FP16/FP32 mode, and inclusion/exclusion of post-processing). Preliminary acceleration assessment using TensorRT INT8 for Jetson Orin indicates potential performance of 12–15 FPS under specified conditions; final throughput depends on specific device variant, input resolution, and whether timing is end-to-end or neural network forward-pass only.

4.4. Limitations and Regulatory Considerations

This study is retrospective in nature and based on open de–identified data; limitations include absence of prospective clinical validation in real workflow, potential frame correlation, and binary CADx formulation without SSL designation as a separate class. Future work involves multicenter validation and assessment of impact on clinically meaningful endpoints (e.g., ADR), as well as preparation for SaMD regulatory evaluation (FDA/CE), including reproducibility, calibration, interpretability, and risk management considerations.

5. Conclusions

In this work, we developed and experimentally evaluated a hybrid CAD architecture for endoscopy that integrates polyp localization, segmentation, and binary histological stratification. The proposed approach is based on hierarchical integration of the YOLO11 detector with SAM 3 foundation segmentation, augmented by parameter–efficient LoRA domain adaptation, achieving high segmentation metrics on in–domain testing and quality preservation under external evaluation on ETIS–Larib (mDice = 0.884).

Ablation analysis demonstrates measurable contribution of semantic text prompts: prompt ensembling reduces model sensitivity to formulation variability and promotes more robust behavior under domain shift. Collectively, results indicate that the combination of specialized localization, LoRA adaptation, and semantic prompts can enhance practical applicability of foundation approaches for endoscopic analysis, including processing of low–contrast and morphologically ambiguous cases.

From an applied perspective, achieved adenoma sensitivity (92.3%) and computational throughput (38.2 FPS on NVIDIA A100) demonstrate potential for integration into near real–time decision support workflows. However, clinical conclusions (including resect–and–discard scenarios) require separate prospective validation on multicenter data with reporting of metrics directly related to clinical endpoints (e.g., NPV/high–confidence mode and surveillance interval agreement), as well as analysis of impact on detection rates in real workflow (including ADR) and error profiles stratified by subtypes (including SSL).

Author Contributions: Conceptualization, methodology, software, Shchetinin Eu. Yu..; validation, formal analysis, supervision, Sevastianov L. A.; writing—original draft preparation, review and editing, Tiutiunnik A. A. All authors have read and agreed to the published version of the manuscript.

Funding: The work was carried out with the financial support of the Sevastopol State University, project 42-01-09/319/2025-1.

Data Availability Statement: Data sharing is not applicable.

Acknowledgments: The authors thank the creators of the open Kvasir, CVC, and ETIS–Larib datasets.

Conflicts of Interest: The authors declare no conflict of interest.

Declaration on Generative AI: The authors have not employed any Generative AI tools.

References

[1] J. C. van Rijn, J. B. Reitsma, J. Stoker, P. M. Bossuyt, S. J. van Deventer, and E. Dekker, "Polyp miss rate determined by tandem colonoscopy: a systematic review.," *The American journal of gastroenterology*, vol. 101, no. 2, pp. 343–350, 2006. DOI: 10.1111/j.1572-0241.2006.00390.x

[2] D. K. Rex, C. Kahi, M. O'Brien, et al., "The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on real-time endoscopic assessment of the histology of diminutive colorectal polyps.," *Gastrointestinal endoscopy*, vol. 73, no. 3, pp. 419–422, 2011. DOI: 10.1016/j.gie.2011.01.023

[3] O. Ronneberger, P. Fischer, and T. Brox, "U–Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, ser. Lecture Notes in Computer Science, vol. 9351, Springer, Cham, 2015, pp. 234–241. DOI: 10.1007/978-3-319-24574-4_28

[4] Z. Zhou, M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A Nested U-Net Architecture for Medical Image Segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMIA ML-CDS 2018*, ser. Lecture Notes in Computer Science, vol. 11045, Springer, Cham, 2018, pp. 3–11. DOI: 10.1007/978-3-030-00889-5_1

[5] F. Diakogiannis, F. Waldner, P. Caccetta, and C. Wu, "ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 162, pp. 94–114, 2020. DOI: 10.1016/j.isprsjprs.2020.01.013

[6] D.-P. Fan and et al., "PraNet: Parallel Reverse Attention Network for Polyp Segmentation," in *Proceedings of the 2020 International Conference on Medical Image Computing and Computer-Assisted Intervention*, ser. Lecture Notes in Computer Science, vol. 12266, Springer, Cham, 2020, pp. 263–273. DOI: 10.1007/978-3-030-59725-2_26

[7] B. Dong and et al., *Polyp-PVT: Polyp segmentation with pyramid vision transformers*, 2021.

[8] C. Wang, I. Yeh, and H.-Y. Liao, *YOLOv9: Learning what you want to learn using programmable gradient information.* 2024.

[9] A. Wang, H. Chen, L. Liu, K. Chen, Z. Lin, J. Han, and G. Ding, *YOLOv10: Real-time end-to-end object detection.* 2024.

[10] A. Kirillov and et al., *Segment Anything*, 2023.

[11] J. Ma, Y. He, F. Li, and et al., "Segment anything in medical images.," *Nature Communications*, vol. 15, pp. 1–9, 2024. DOI: 10.1038/s41467-024-44824-z

[12] N. Ravi, V. Gabeur, Y. Hu, R. Hu, C. Ryali, T. Ma, and et al., *SAM 2: Segment anything in images and videos.* 2024.

[13] E. Shchetinin and A. Tiutiunnik, "On Segmentation of Polyps Using Segment Anything Model.," *Research result. Information technologies*, vol. 10, no. 3, pp. 55–63, 2025. DOI: 10.18413/2518-1092-2025-10-3-0-5

[14] D. Jha, P. Smedsrud, M. Riegler, P. Halvorsen, T. de Lange, D. Johansen, and H. Johansen, "Kvasir-SEG: A segmented polyp dataset.," in *MultiMedia Modeling*, ser. Lecture Notes in Computer Science, vol. 11962, Springer International Publishing, 2020, pp. 451–462. DOI: 10.1007/978-3-030-37734-2_37

[15] J. Bernal, F. Sánchez, G. Fernández-Esparrach, D. Gil, C. Rodríguez, and F. Vilariño, "WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians.," *Computerized Medical Imaging and Graphics*, vol. 43, pp. 99–111, 2015. DOI: 10.1016/j.compmedimag.2015.02.007

[16] Y. Tudela, M. Majó, N. de la Fuente, and et al., "A complete benchmark for polyp detection, segmentation and classification in colonoscopy images.," *Frontiers in Oncology*, vol. 14, pp. 1–19, 2024. DOI: doi.org/10.3389/fonc.2024.1417862

[17] J. Silva, A. Histace, O. Romain, X. Dray, and B. Granado, "Toward embedded detection of polyps in WCE images for early diagnosis of colorectal cancer.," *International Journal of Computer Assisted Radiology and Surgery*, vol. 9, pp. 283–293, 2014. DOI: doi.org/10.1007/s11548-013-0926-3

[18] G. Jocher and J. Qiu, *Ultralytics YOLO11 [Internet]. Version 11.0.0. GitHub*, 2024.

[19] N. Carion, L. Gustafson, Y. Hu, and et al., *SAM 3: Segment anything with concepts.* 2025.

[20] E. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, and et al., *LoRA: Low-rank adaptation of large language models.* 2022.

[21] D. Jha, S. Ali, N. K. Tomar, H. D. Johansen, D. Johansen, J. Rittscher, M. A. Riegler, and P. Halvorsen, "Real-time polyp detection, localization and segmentation in colonoscopy using deep learning.," *IEEE Access*, vol. 9, pp. 40 496–40 510, 2021. DOI: 10.1109/ACCESS.2021.3063716

[22] D. Jha, P. Smedsrud, M. Riegler, D. Johansen, T. de Lange, P. Halvorsen, and H. Johansen, "ResUNet++: An advanced architecture for medical image segmentation.," in *Proceedings of the IEEE International Symposium on Multimedia (ISM)*, ser. Lecture Notes in Computer Science, IEEE, 2019, pp. 225–230. DOI: 10.1109/ISM46123.2019.00049

[23] M. Mansoori, S. Shahabodini, J. Abouei, K. Plataniotis, and A. Mohammadi, *Self-prompting polyp segmentation in colonoscopy using hybrid YOLO-SAM 2 model.* 2024.

[24] D. Rex, D. Ahnen, J. Baron, and et al., "Serrated lesions of the colorectum: Review and recommendations from an expert panel.," *American Journal of Gastroenterology*, vol. 107, no. 9, pp. 1315–1329, 2012. DOI: 10.1038/ajg.2012.161

[25] G. Ji, J. Liu, P. Xu, N. Barnes, F. Khan, S. Khan, and D. Fan, *Frontiers in intelligent colonoscopy.* 2025.

Information about the authors

Shchetinin, Eugene Yu.—Doctor of Physical and Mathematical Sciences, Professor of the Department of Information Technology and Systems of the Sevastopol State University (e-mail: riviera-molto@mail.ru, ORCID: 0000-0003-3651-7629, ResearcherID: O-8287-2017, Scopus Author ID: 16408533100)

Sevastianov, Leonid A.—Doctor of Physical and Mathematical Sciences, Professor of the Department of Mathematical Modeling and Artificial Intelligence of the RUDN University (e-mail: sevastianov-la@rudn.ru, ORCID: 0000-0002-1856-4643, ResearcherID: B-8497-2016, Scopus Author ID: 8783969400)

Tiutiunnik, Anastasia A.—PhD in Physics and Mathematics, Associate Professor of the Department of Mathematical Modeling and Artificial Intelligence of the RUDN University (e-mail: tyutyunnik-aa@rudn.ru, ORCID: 0000-0002-4643-327X, Scopus Author ID: 55978487200)

RUSSIAN METADATA

УДК 004.89; 004.93
PACS 07.05.Mh 07.05.Tp 42.62.Be
DOI: 10.22363/2658-4670-2026-34-2-260-273 EDN: JMNMUP

Эффективная сегментация и классификация полипов с помощью YOLO11–SAM3 и LoRA

Е. Ю. Щетинин[1], Л. А. Севастьянов[2], А. А. Тютюнник[2]

[1] *Севастопольский государственный университет, ул. Университетская, д. 33, Севастополь, 299053, Российская Федерация*

[2] *Российский университет дружбы народов, ул. Миклухо–Маклая, д. 6, Москва, 117198, Российская Федерация*

Аннотация. *Цель* Разработать и экспериментально оценить автономную систему компьютерной поддержки для колоноскопии, объединяющую сегментацию полипов и бинарную гистологическую классификацию (аденома/не–аденома) с сохранением вычислительной эффективности для near real–time применения. *Материал и методы* Предложена каскадная архитектура YOLO11–-SAM3-–LoRA: детектор YOLO11–medium формирует кандидаты (bounding boxes), после чего SAM3 выполняет box–prompted сегментацию с параметро–эффективной доменной адаптацией (LoRA) и механизмом семантического якорения на основе ансамбля текстовых промптов. Гистологическая классификация реализована двухпоточным классификатором, объединяющим признаки ROI (EfficientNet–B0) и морфологические дескрипторы маски. Обучение проведено на объединённом наборе Kvasir–SEG и CVC–ClinicDB (детекция/сегментация) и CVC–HDClassif (классификация). Переносимость оценивалась на внешнем датасете ETIS–Larib в режиме external zero–shot. *Результаты* На основной тестовой выборке достигнуты mDice = 0,962 для сегментации, чувствительность к аденомам (Sensitivity) = 92,3% и AUROC = 0,942 в задаче бинарной классификации. Во внешнем тестировании на ETIS–Larib (zero-shot) система сохраняет качество сегментации (mDice = 0,884), что указывает на устойчивость к доменному сдвигу. Скорость end–to–end обработки составляет 38,2 FPS на NVIDIA A100. *Заключение* Результаты показывают, что сочетание каскадной локализации, foundation–сегментации с LoRA–адаптацией и ансамблирования текстовых промптов обеспечивает высокое качество сегментации и конкурентоспособную точность CADx при сохранении вычислительной эффективности. Для клинического внедрения необходима дальнейшая проспективная многоцентровая валидация и анализ ошибок по подтипам поражений.

Ключевые слова: сегментация полипов, классификация полипов, YOLO11, SAM 3, LoRA, ансамбль промптов, CAD, CADx, колоноскопия, глубокое обучение

Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2026, 34 (2) 274–278
http://journals.rudn.ru/miph

Research article

UDC 519.872, 519.217
PACS 47.10.Tp, 47.30.Pn,
DOI: 10.22363/2658-4670-2026-34-2-274-278 EDN: JOTBVJ

On the classical Volterra spectral equation

Dashdondog Tsedenbayar[1,2], Bayaraa Ser-Od[2], Wuren Qiqige[3]

[1] *Institute of Mathematics and Digital Technology, Mongolian Academy of Sciences, Ulaanbaatar, 13330, Mongolia*

[2] *Mongolian University of Science and Technology, Ulaanbaatar, 14191, Mongolia*

[3] *College of Mathematics and Computer Science, Chifeng University, No. 1, Yingbin Road, Hongshan District, Chinfeng city, Chifeng, 024000, Inner Mongolia, China*

(received: March 11, 2026; revised: April 12, 2026; accepted: April 20, 2026)

Abstract. We study the characterizations of the classical Volterra operator based on a simple relation between its singular numbers and the eigenvalues of its imaginary part on $L^2[0,1]$.

Key words and phrases: Volterra operator, trace operator, operator norm, eigenvalue

For citation: D. Tsedenbayar, B. Ser-Od, W. Qiqige "On the classical Volterra spectral equation," *Discrete and Continuous Models and Applied Computational Science*, vol. 34, no. 2, pp. 274–278, 2026. DOI: 10.22363/2658-4670-2026-34-2-274-278. EDN: JOTBVJ.

1. Introduction

Let H be a complex Hilbert space equipped with the inner product $(\cdot,\cdot)$ and the induced norm $\|\cdot\|$. Denote by $B(H)$ the Banach algebra of bounded linear operators acting on H with the operator norm

$$\|A\| = \sup_{\|x\|=1}\{\|Ax\| : x \in H\}, A \in B(H).$$

The integral operator A on $L^2[0,1]$ defined as

$$(Af)(x) = \int_0^1 k(x,s)f(s)ds$$

with some kernel $k(t,s) \in L^2([0,1]^2)$. Then A is a Hilbert–Schmidt operator.

Recall that for an operator A the spectrum

$$\sigma(A) = \{\lambda \in \mathbb{C} : A - \lambda I \text{is not invertible}\}$$

is a nonempty compact subset of the complex plane.

For any compact linear operator A in a Hilbert space H the singular numbers $s_k(A)$ are the distances from A to the set of all operators of rank less than or equal to $k-1$ for $k \geqslant 1$. Their squares are the eigenvalues of the compact self-adjoint nonnegative operator A^*A counted according to their multiplicities. (see e.g. [1]) In particular, $s_1(A) = \|A\|$.

Denote by V the Volterra operator

$$(Vf)(x) = \int_0^x f(s)ds$$

and its adjoint

$$(V^*f)(x) = \int_x^1 f(s)ds$$

on $L^2[0,1]$.

The Halmos (see [2]) calculation yields (see [3]):

$$s_k(V) = \frac{2}{(2k-1)\pi}$$

for all $k \geqslant 1$, in particular

$$\|V\| = \frac{2}{\pi}.$$

Note that

$$\sum_{k=0}^{\infty} s_k^2 = \frac{1}{2}.$$

If $A \in B(H)$ and $\{\lambda_k\}$ are the eigenvalues of A, then

$$tr(A) = \sum_k \lambda_k.$$

It is well known that V is quasi-nilpotent, compact and even a Hilbert–Schmidt operator. Due to its excellent properties, many scholars have studied Volterra operator pencils on Hilbert spaces and formulated many interesting results (see e.g. [4–13]).

The set $M_\alpha = \{f \in L^2[0,1] : f(x) = 0, 0 \leqslant x \leqslant \alpha\}$ is an invariant subspace for V, for $0 \leqslant \alpha \leqslant 1$ (see [14, 15]). Recall that

$$\operatorname{Re} V = \frac{1}{2}(V + V^*) = \frac{1}{2}P,$$

where P is the orthogonal projection onto the constant functions.

It is easy to see that the nonzero eigenvalues of the imaginary part

$$\operatorname{Im} V = \frac{1}{2i}(V - V^*)$$

are

$$\pm\frac{1}{(2k-1)\pi}$$

for $k \in \mathbb{N}$, each having multiplicity 1 (see [15]). Then we have the spectral equation

$$2\sigma(\operatorname{Im} V) = \sigma((V^*V)^{\frac{1}{2}}) \cup (-\sigma((V^*V)^{\frac{1}{2}})),$$

where σ denotes the spectrum (see [16]).

Recall that for VV^* the orthonormal basis of $L^2[0,1]$ consisting of eigenfunctions

$$f_k(x) = \sqrt{2}\cos\left(\frac{2k+1}{2}\pi x\right)$$

corresponding to eigenvalues

$$\lambda_k = \frac{4}{(2k+1)^2\pi^2}$$

for $n \geqslant 0$ (see [17]).

We obtain the characterizations of the Volterra operator on $L^2[0,1]$.

2. The Results

Theorem 3. *Suppose that A is a Hilbert–Schmidt operator on $L^2[0,1]$ such that*

$$A + A^* = Q$$

is a rank-one projection, quasinilpotent and

$$2\sigma(Im\, A) = \sigma((A^*A)^{\frac{1}{2}}) \cup (-\sigma((A^*A)^{\frac{1}{2}})).$$

Then $A = V$.

Proof. If A is quasinilpotent and rank-one projection (3) and the spectra on both sides in (3) are countable set, then A is a compact operator (see [18, 19]). On the other hand, we observe that 0 is not an eigenvalue of A, and that A and A^* have no proper common invariant subspace. Also A is rank-one projection and Im A $\geqslant 0$ then A is unitary equivalent to the Volterra operator (see [20]).

Suppose that the eigenvalues of Im A and the singular numbers $\{s_k\}_{k\in\mathbb{N}}$ of A have multiplicity 1. From (3), we have

$$\sum_{k=1}^{\infty} s_k^2 = tr(A^*A) = tr\left((\frac{1}{2}Q - i\text{Im A})(\frac{1}{2}Q + i\text{Im A})\right) =$$

$$= \frac{1}{4}tr(Q) + \frac{i}{2}tr([Q, \text{Im A}]) + tr((\text{Im A})^2) =$$

$$= \frac{1}{4} + 2\sum_{k=1}^{\infty}\left(\frac{s_k}{2}\right)^2 = \frac{1}{4} + \frac{1}{2}\sum_{k=1}^{\infty} s_k^2.$$

Hence,

$$\sum_{k=0}^{\infty} s_k^2 = \frac{1}{2}.$$

Therefore $A = V$. This completes the proof. □

Conclusions

We have seen that characterizations technique for the Volterra operator is based on some relation between its singular numbers and the eigenvalues of its imaginary part on $L^2[0,1]$.

Author Contributions: Development, methodology and original draft preparation, D. Tsedenbayar; writing-original draft preparation, B. Ser-Od; manuscript review and editing, Wuren Qiqige. All authors have reviewed and approved the published version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: Wuren Qiqige's research was supported by a grant from the Doctoral Research and Innovation Development Fund of Chifeng University (Project Number: BSJJ077).

Data Availability Statement: Data sharing is not applicable.

Acknowledgments: The authors wish to thank the editor and the anonymous referees for their valuable suggestions and comments on the first version of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

Declaration on Generative AI: The authors have not employed any Generative AI tools.

References

[1] I. C. Gohberg and M. G. Krein, *Introduction to Theory of Linear Nonselfadjoint Operators*. Translations of Mathematical Monographs, 1978, vol. 18, 378 pp.

[2] P. R. Halmos, *A Hilbert Space Problem Book*. Springer New York, NY, 1982, 373 pp. DOI: 10.1007/978-1-4684-9330-6

[3] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. Rensselaer Polytechnic Institute, USA, 2014, 1220 pp.

[4] M. R. Dostanic, "Estimates for the norm of powers of Volterra's operator through its singular values," *Mathematische Zeitschrift*, vol. 236, pp. 473–459, 2001. DOI: 10.1007/PL00004837

[5] L. Khadkhuu and D. Tsedenbayar, "A note about Volterra operator," *Mathematica Slovaca*, vol. 5, pp. 1117–1120, 2018. DOI: 10.1515/ms-2017-0173

[6] Y. Lyubich and D. Tsedenbayar, "The norms and singular numbers of polynomials of the classical Volterra operator in $L^2(0,1)$," *Studia Mathematica*, vol. 199, pp. 171–184, 2010. DOI: 10.4064/sm199-2-3

[7] L. Khadkhuu and D. Tsedenbayar, "On the numerical range and numerical radius of the Volterra operator," *Bulletin of Irkutsk State University, Series Mathematics*, vol. 24, pp. 102–108, 2018. DOI: 10.26516/1997-7670.2018.24.102

[8] T. Ransford, "Norms of polynomials of the Volterra operator," *Journal of Mathematical Analysis and Applications*, vol. 517, no. 2, pp. 1–15, 2023. DOI: 10.1016/j.jmaa.2022.126626

[9] T. Ransford, "Volterra operator norms: a brief survey," *Moroccan Journal of Pure and Applied Analysis*, vol. 9, no. 2, pp. 276–290, 2023. DOI: 10.2478/mjpaa-2023-0018

[10] T. Ransford and D. Tsedenbayar, "On the real and imaginary parts of powers of the Volterra operatorl," *Integral Equations and Operator Theory*, vol. 96, no. 4, pp. 1–12, 2024. DOI: 10.1007/s00020-024-02755-w

[11] D. Sarason, "A remark on the Volterra operator," *Journal of Mathematical Analysis and Applications*, vol. 12, pp. 244–246, 1965. DOI: 10.1016/0022-247X(65)90035-1

[12] D. Tsedenbayar, "On the power boundedness of certain Volterra operator pencil," *Studia Mathematica*, vol. 156, pp. 59–66, 2003. DOI: 10.4064/sm156-1-4

[13] D. Tsedenbayar, *Some properties of the Volterra and Cesáro operators*. Institute of Mathematics of the Polish Academy of Sciences, 2002, 33 pp.

[14] N. Aronszajn and K. T. Smith, "Invariant subspaces of completely continuous operators," *Annals of Mathematics*, vol. 60, pp. 345–350, 1954. DOI: http://mi.mathnet.ru/mat37

[15] E. B. Davies, *Linear Operators and Their Spectra*. Cambridge University Press, Cambridge, 2007, vol. 106, 465 pp.

[16] N. Dunford and J. T. Schwartz, *Linear Operators, Part-1*. John Wiley and Sons, 1988, 872 pp.

[17] S. M. Garcia, G. Mashreghi, and W. T. Ross, *Operator Theory by Example*. Oxford University Press, Oxford University Press, Oxford Graduate Texts in Mathematics, 60., 2023, 373 pp. DOI: 10.1093/oso/9780192863867.001.0001

[18] J. Anderson, "Commutators of Compact Operators," *Journal Für Die Reine und Angewandte Mathematik*, vol. 291, pp. 128–132, 1977. DOI: 10.1515/crll.1977.291.128

[19] H. P. R. Brown A. and P. C., "Commutators of Operators on Hilbert Space," *Canadian Journal of Mathematics*, vol. 17, pp. 695–708, 1965. DOI: 10.4153/CJM-1965-070-7

[20] I. C. Gohberg and M. G. Krein, *Theory and Applications of Volterra Operators in Hilbert Space*. Translations of Mathematical Monographs, 1970, vol. 24, 430 pp.

Information about the authors

Tsedenbayar, Dashdondog—Academician, Professor, Doctor of Sciences in Physics and Mathematics, (e-mail: tsdnbr@must.edu.mn, ORCID: 0000-0002-4053-7983)

Ser-Od, Bayaraa—Associate professor of Department of Mathematics at School of Applied Sciences, Mongolian University of Science and Technology (e-mail: serod_b@must.edu.mn, Scopus Author ID: 36661503100)

Wuren, Qiqige—College of Mathematics and Computer Science, Chifeng University, No. 1, Yingbin Road, Hongshan District, Chinfeng city, Chifeng-024000, Inner Mongolia, China (e-mail: wurenqiqige@126.com)

RUSSIAN METADATA

УДК 519.872, 519.217

PACS 47.10.Tp, 47.30.Pn,

DOI: 10.22363/2658-4670-2026-34-2-274-278 EDN: JOTBVJ

О классическом спектральном уравнении Вольтерры

Д. Цэдэнбаяр[1,2], Б. Сэр-Од[2], У. Цэцэг[3]

[1] *Институт математики и цифровой технологии, Монгольская Академия Наук, Улан-Батор, 13330, Монголия*

[2] *Монгольский Государственный Университет Науки и Технологии, Улан-Батор, 14191, Монголия*

[3] *Колледж математики и информатики Университета Чифэн, No 1, ул. Инбинь, район Хуншань, город Чинфэн, Чифэн, 024000, Китай*

Аннотация. Мы изучаем характеристики классического оператора Вольтерры на основе простого соотношения между его сингулярными числами и собственными значениями его мнимой части на $L^2[0,1]$.

Ключевые слова: Оператор Вольтерры, оператор следа, норма оператора, собственное значение