

ISSN 1812-5093

**ТРУДЫ
ИНСТИТУТА МАТЕМАТИКИ
НАН БЕЛАРУСИ**

Государственное научное учреждение
«Институт математики Национальной
академии наук Беларуси»

Журнал основан в 1998 г.
До 2024 г. выходил под названием
«Труды Института математики».

**PROCEEDINGS OF THE
INSTITUTE OF MATHEMATICS
OF THE NAS OF BELARUS**

State Scientific Institute
«The Institute of Mathematics
of the National Academy of Sciences
of Belarus»

The journal was founded in 1998.
Before 2024, the name of the journal was
«Proceedings of the Institute of Mathematics».

ТОМ 33, № 1
Минск, 2025

ОТ РЕДАКЦИИ

В журнале «Труды Института математики НАН Беларуси» публикуются оригинальные статьи фундаментальной и прикладной математики. Обзорные статьи публикуются по решению редакционной коллегии. Журнал входит в «Перечень научных изданий Республики Беларусь для опубликования результатов диссертационных исследований», утвержденный Высшей аттестационной комиссией Республики Беларусь.

До 2024 г. журнал издавался под названием «Труды Института математики».

Основные разделы журнала:

Алгебра и теория чисел
Вещественный, комплексный и функциональный анализ
Вычислительная математика
Дискретная математика и математическая кибернетика
Дифференциальные уравнения, динамические системы и оптимальное управление
Математическое моделирование и численные методы
Математические методы защиты информации
Теория вероятностей и математическая статистика
Краткие сообщения

Адрес редакции: Address:

ул. Сурганова, 11, к. 45 11, Surganov Str., room 45,
220072, г. Минск, Беларусь Minsk, Belarus, 220072
Тел. +375-17-379-17-84 Phone: +375-17-379-17-84
e-mail: tbusel@im.bas-net.by

EDITOR'S NOTE

The journal "Proceedings of the Institute of Mathematics of the NAS of Belarus" publishes original articles in fundamental and applied mathematics. Publication of review articles requires approval of the editorial board. The journal is approved by the Higher Attestation Commission of the Republic of Belarus for publishing results of dissertation research, and is included in the official list of such publications.

Before 2024, the name of the journal was "Proceedings of the Institute of Mathematics" (Trudy Instituta Matematiki).

Main sections of the journal:

Algebra and number theory
Real, complex and functional analysis
Computational mathematics
Discrete mathematics and mathematical cybernetics
Differential equations, dynamic systems and optimal control
Mathematical modeling and numerical methods
Mathematical methods for information security
Probability theory and mathematical statistics
Brief communications

Главный редактор

Сафонов В. Г. – Институт математики НАН Беларуси, Беларусь

Редакционная коллегия

Гороховик В. В. – Институт математики НАН Беларуси, Беларусь (*заместитель главного редактора*)
Матус П. П. – Институт математики НАН Беларуси, Беларусь (*заместитель главного редактора*)
Бусел Т. С. – Институт математики НАН Беларуси, Беларусь (*ответственный секретарь*)

Антоневич А. Б. – Белорусский государственный университет, Беларусь
Асташкин С. В. – Самарский национальный исследовательский университет, Московский государственный университет имени М. В. Ломоносова, Россия

Асташова И. В. – Московский государственный университет имени М. В. Ломоносова, Россия

Бахтин В. И. – Белорусский государственный университет, Беларусь

Берник В. И. – Институт математики НАН Беларуси, Беларусь

Вабищевич П. Н. – Московский государственный университет имени М. В. Ломоносова, Россия

Васильев Д. В. – Институт математики НАН Беларуси, Беларусь

Галанин М. П. – Федеральный исследовательский центр Институт прикладной математики имени М. В. Келдыша РАН, Россия

Гладков А. Л. – Белорусский государственный университет, Беларусь

Го Веньбинь – Хайнаньский университет, Школа математики и статистики, Китай

Громак В. И. – Белорусский государственный университет, Беларусь

Добровольский Н. М. – Тульский государственный педагогический университет имени Л. Н. Толстого, Россия

Егоров А. Д. – Институт математики НАН Беларуси, Беларусь

Зайцев В. А. – Институт математики, информационных технологий и физики Удмуртского государственного университета, Россия

Изобов Н. А. – Институт математики НАН Беларуси, Беларусь

Ильин А. В. – Московский государственный университет имени М. В. Ломоносова, Россия

Кигурадзе И. Т. – Математический институт имени А. Размадзе, Тбилисский государственный университет имени И. Джавахишвили, Грузия

Ковалев М. Я. – Объединенный институт проблем информатики НАН Беларуси, Беларусь

Корзюк В. И. – Белорусский государственный университет, Институт математики НАН Беларуси, Беларусь

Костюкова О. И. – Институт математики НАН Беларуси, Беларусь

Лебедев А. В. – Белорусский государственный университет, Беларусь

Леваков А. А. – Белорусский государственный университет, Беларусь

Лыков К. В. – Белорусский государственный университет, Институт математики НАН Беларуси, Беларусь

Махнев А. А. – Институт математики и механики имени Н. Н. Красовского УрО РАН, Россия

Ни Мин Кан – Восточно-китайский педагогический университет, Институт математических наук, Китай

Осиновская А. А. – Институт математики НАН Беларуси, Беларусь

Попова С. Н. – Удмуртский государственный университет, Россия

Ровба Е. А. – Гродненский государственный университет имени Янки Купалы, Беларусь

Сарванов В. И. – Институт математики НАН Беларуси, Беларусь

Сергеев И. Н. – Московский государственный университет имени М. В. Ломоносова, Россия

Скиба А. Н. – Гомельский государственный университет имени Ф. Скорины, Беларусь

Старовойтов А. П. – Гомельский государственный университет имени Ф. Скорины, Беларусь

Фомичев В. В. – Московский государственный университет имени М. В. Ломоносова, Россия

Editor-in-Chief

Safonov V. G. – Institute of Mathematics of the National Academy of Sciences of Belarus, Belarus

Editorial Board

Gorokhovich V. V. – Institute of Mathematics of the National Academy of Sciences of Belarus, Belarus (*Associate Editor-in-Chief*)

Matus P. P. – Institute of Mathematics of the National Academy of Sciences of Belarus, Belarus (*Associate Editor-in-Chief*)

Busel T. S. – Institute of Mathematics of the National Academy of Sciences of Belarus, Belarus (*Executive secretary*)

Antonevich A. B. – Belarusian State University, Belarus

Astashkin S. V. – Samara National Research University, Lomonosov Moscow State University, Russia

Astashova I. V. – Lomonosov Moscow State University, Russia

Bakhtin V. I. – Belarusian State University, Belarus

Bernik V. I. – Institute of Mathematics of the National Academy of Sciences of Belarus, Belarus

Dobrovol'skii N. M. – Tula State Lev Tolstoy Pedagogical University, Russia

Egorov A. D. – Institute of Mathematics of the National Academy of Sciences of Belarus, Belarus

Fomichev V. V. – Lomonosov Moscow State University, Russia

Galanin M. P. – Keldysh Institute of Applied Mathematics of Russian Academy of Sciences, Russia

Gladkov A. L. – Belarusian State University, Belarus

Gromak V. I. – Belarusian State University, Belarus

Guo Wenbin – Hainan University, The School of Mathematics and Statistics, China

Il'in A. V. – Lomonosov Moscow State University, Russia

Izobov N. A. – Institute of Mathematics of the National Academy of Sciences of Belarus, Belarus

Kiguradze I. T. – Andrea Razmadze Mathematical Institute of Ivane Javakhishvili, Tbilisi State University, Georgia

Korzyuk V. I. – Belarusian State University, Institute of Mathematics of the National Academy of Sciences of Belarus, Belarus

Kostyukova O. I. – Institute of Mathematics of the National Academy of Sciences of Belarus, Belarus

Kovalev M. Ya. – United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Belarus

Lebedev A. V. – Belarusian State University, Belarus

Levakov A. A. – Belarusian State University, Belarus

Lykov K. V. – Belarusian State University, Institute of Mathematics of the National Academy of Sciences of Belarus, Belarus

Makhnev A. A. – N. N. Krasovskii Institute of Mathematics and Mechanics of the Ural Branch of the Russian Academy of Sciences, Russia

Ni Min Kan – East China Normal University, School of Mathematical Sciences, China

Osinovskaya A. A. – Institute of Mathematics of the National Academy of Sciences of Belarus, Belarus

Popova S. N. – Udmurt State University, Russia

Rovba E. A. – Yanka Kupala State University of Grodno, Belarus

Sarvanov V. I. – Institute of Mathematics of the National Academy of Sciences of Belarus, Belarus

Sergeev I. N. – Lomonosov Moscow State University, Russia

Skiba A. N. – Francisk Skorina Gomel State University, Belarus

Starovoitov A. P. – Francisk Skorina Gomel State University, Belarus

Vabishchevich P. N. – Lomonosov Moscow State University, Russia

Vasilyev D. V. – Institute of Mathematics of the National Academy of Sciences of Belarus, Belarus

Zaitsev V. A. – Institute of Mathematics, Information Technologies and Physics of Udmurt State University, Russia

ТРУДЫ
ИНСТИТУТА МАТЕМАТИКИ НАН БЕЛАРУСИ
Минск. 2025. Т. 33. № 1

СОДЕРЖАНИЕ

АЛГЕБРА И ТЕОРИЯ ЧИСЕЛ

Берник В. И., Калоша Н. И., Васильев Д. В. Закон Бенфорда и приближение логарифмов натуральных чисел рациональными	7
Курочкина Д. Н. О норме Шаттена линейного отображения евклидовых пространств	15
Малинин Д. А. Неразветвленные расширения Галуа и подгруппы GL_n	20
Мурашко В. И., Васильев А. Ф. К теореме К. Дёрка	28
Шведков П. Н., Лыков К. В. Алгебраическое доказательство эквивалентности двух вариантов cut-нормы для многомерных симметричных матриц	34
Ядченко А. А. О разрешимости и факторизации некоторых π -разрешимых неприводимых линейных групп примарной степени. Часть V	44

ВЕЩЕСТВЕННЫЙ, КОМПЛЕКСНЫЙ И ФУНКЦИОНАЛЬНЫЙ АНАЛИЗ

Антоневич А. Б., Ёжикова М. Д. Топологические структуры на градуированных множествах	58
Старовойтов А. П., Кругликов И. В. Существование и явный вид нелинейных аппроксимаций Эрмита–Чебышёва	75

ВЫЧИСЛИТЕЛЬНАЯ МАТЕМАТИКА

Жерело А. В. Приближенная формула для математических ожиданий от решения стохастического дифференциального уравнения с дрейфом	87
Жерело А., Крылов Г., Оболонкин В., Йоргенсен М., Спелман Р. Об одной стохастической модели дойного стада	95

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Мэй Лю, Дудин А. Н. Стационарный анализ многолинейной системы массового обслуживания с неоднородными приборами и распределением времени обслуживания фазового типа	111
---	-----

PROCEEDINGS
OF THE INSTITUTE OF MATHEMATICS OF THE NAS OF BELARUS
Minsk. 2025. Vol. 33. N 1

CONTENTS

ALGEBRA AND NUMBER THEORY

Bernik V. I., Kalosha N. I., Vasilyev D. V. Benford's law and approximation of logarithms of natural numbers by rational numbers	7
Kurochkina D. N. On the Schatten norm of a linear map on Euclidean spaces	15
Malinin D. A. Unramified Galois extensions and subgroups of GL_n	20
Murashka V. I., Vasil'ev A. F. To the theorem of K. Doerk	28
Shvedkov P. N., Lykov K. V. Algebraic proof of the equivalence of two variants of the cut-norm for multidimensional symmetric matrices	34
Yadchenko A. A. On the solvability and factorization of some π -solvable irreducible linear groups of primary degree. Part V	44

REAL, COMPLEX, AND FUNCTIONAL ANALYSIS

Antonevich A. B., Yozhikova M. D. Topological structures on graded sets	58
Starovoitov A. P., Kruglikov I. V. Existence and explicit form of nonlinear Hermite–Chebyshev approximations	75

COMPUTATIONAL MATHEMATICS

Zherelo A. V. Approximate formula for mathematical expectations of a solution of a stochastic differential equation with drift	87
Zherelo A., Krylov G., Obolonkin V., Jorgensen M., Spielman R. About one stochastic model of a dairy herd	95

PROBABILITY THEORY AND MATHEMATICAL STATISTICS

Liu M., Dudin A. N. Steady-state analysis of the multi-server retrial queueing system with heterogeneous servers and phase type distribution of service times	111
--	-----



АЛГЕБРА И ТЕОРИЯ ЧИСЕЛ
ALGEBRA AND NUMBER THEORY



UDC 511.42

BENFORD'S LAW AND APPROXIMATION OF LOGARITHMS OF NATURAL
NUMBERS BY RATIONAL NUMBERS

V. I. Bernik, N. I. Kalosha, D. V. Vasilyev

Institute of Mathematics of the National Academy of Sciences of Belarus, Minsk, Belarus
e-mail: bernik.vasili@mail.ru, kalosha@im.bas-net.by, vasilyev@im.bas-net.by

Received: 02.04.2025

Revised: 19.05.2025

Accepted: 23.05.2025

Keywords: diophantine approximation, Benford's law, first digit distribution, powers of integers.

Abstract. The paper is devoted to studying the frequencies at which first digits occur in series formed by powers of integer numbers. A number of generalizations of this problem are considered, and the relation between the distribution of first digits and Diophantine properties of logarithms is discussed. In conclusion of the article, several interesting problems in modern theory of Diophantine approximation are proposed.

ЗАКОН БЕНФОРДА И ПРИБЛИЖЕНИЕ ЛОГАРИФМОВ НАТУРАЛЬНЫХ ЧИСЕЛ
РАЦИОНАЛЬНЫМИ

В. И. Берник, Н. И. Калоса, Д. В. Васильев

Институт математики НАН Беларуси, Минск, Беларусь
e-mail: bernik.vasili@mail.ru, kalosha@im.bas-net.by, vasilyev@im.bas-net.by

Поступила: 02.04.2025

Исправлена: 19.05.2025

Принята: 23.05.2025

Ключевые слова: диофантовы приближения, закон Бенфорда, распределение первых цифр, степени целых чисел.

Аннотация. В статье исследованы частотные свойства первых цифр в последовательности, образованной степенями целых чисел. Рассматривается ряд обобщений этой проблемы, а также обсуждается связь между распределением первых цифр и диофантовыми свойствами логарифмов. Предлагается ряд актуальных проблем в теории диофантовых приближений.

1. Introduction

If we exclude initial zeros in the decimal representation of any non-zero real number, there will be nine possibilities for the leading digit: $a = 1, 2, \dots, 9$. Surprisingly, in most cases these digits occur with different frequencies, and the frequency of a digit a is roughly equal to

$$\lg \frac{a+1}{a}. \quad (1)$$

The first publication describing this phenomenon is due to astronomer and mathematician Simon Newcomb [1]. Frank Benford [2] provided numerous real-world examples that exhibit this distribution of first digits. Newcomb's observation became known as Benford's law, and the distribution (1) as Benford's distribution.

Benford's law holds for many random and deterministic sequences, and the sequence 2^n , $n = 1, 2, \dots$, is a notable example. Recently, Benford's law for power sequences was studied by Hürliemann [3].

We are going to prove that sequences a^n , where $a = 2, 3$, follow Benford's law. We are also going to estimate the residual term and show how our estimate is related to Diophantine properties of logarithms of natural numbers and their combinations. Several generalizations of these facts will be discussed, and computer simulations will be used to evaluate the strength of the obtained theoretical results.

This study originates from a conversation between Vasili Bernik and academician Yuri Prokhorov at a number-theoretic conference in Bielefeld. Not every result presented in the paper is new; some of them are repeated for completeness and ease of understanding.

2. The main results

Let $\nu(A, Q)$ be the frequency with which a positive integer A occurs as leading digits in the first Q values of the sequence 2^n ($1 \leq n \leq Q$). Let $[\beta]$ and $\{\beta\}$ denote respectively the integer and fractional parts of a real number β .

Theorem 2.1. *Let A be an arbitrary positive integer. Then there are infinitely many positive integers $n = n(A)$ such that the decimal representation of A coincides with the leading digits in the decimal representation of 2^n .*

The proof of the theorem is based on two simple lemmas.

Lemma 2.2. *The number $\lg 2$ is irrational.*

Lemma 2.3. *For any irrational number α , the sequence $\{n\alpha\}$, $n = 1, 2, \dots$, is everywhere dense in $[0, 1)$.*

The proof of Lemma 2.2 is commonly known, and Lemma 2.3 can be proved using Dirichlet's pigeonhole principle. In fact, Weyl's criterion [4] implies a much stronger result.

Lemma 2.4. *The sequence $\{n\alpha\}$, $n = 1, 2, \dots$, is uniformly distributed on $[0, 1)$ if and only if α is an irrational number.*

Now it is easy to prove Theorem 2.1. Let A have k decimal digits, $A = \overline{a_1 a_2 \dots a_k}$, $0 \leq a_j \leq 9$, $a_1 \neq 0$. Let us write 2^n in the form

$$2^n = 10^{n \lg 2} = 10^{[10 \lg 2]} \cdot 10^{\{n \lg 2\}}.$$

The leading digits of 2^n coincide with A if

$$\lg \frac{A}{10^{k-1}} \leq \{n \lg 2\} < \lg \frac{A+1}{10^{k-1}}$$

or

$$\{n \lg 2\} \in [\lg A, \lg(A+1)) \pmod{1}. \quad (2)$$

It follows from Lemmas 2.2–2.4 that the condition (2) holds for infinitely many n .

Let $I \subset [0, 1)$ be an interval or a finite union of intervals, and let $N_I(\alpha, Q)$ be the number of positive integers n , $1 \leq n \leq Q$, such that $\{n\alpha\} \in I$. If $\{n\alpha\}$ is uniformly distributed, we have

$$N_I(Q) = (1 + o(1)) Q |I|.$$

Now taking $\alpha = \lg 2$, $I = [\lg A, \lg(A+1)) \pmod{1}$ yields

$$N_I(\lg 2, Q) = (1 + o(1)) Q \lg \frac{A+1}{A} = Q \lg \frac{A+1}{A} + R(Q), \quad (3)$$

$$\lim_{Q \rightarrow \infty} Q^{-1} R(Q) = 0,$$

i. e., the frequency with which the leading digits of 2^n coincide with the digits of A is asymptotically equal to $\nu(A) = \lg \frac{A+1}{A}$. In particular,

$$\nu(1) = \lg 2 = 0.3010\dots, \quad \nu(2) = \lg \frac{3}{2} = 0.1760\dots,$$

$$\nu(8) = \lg \frac{9}{8} = 0.0511\dots, \quad \nu(9) = \lg \frac{10}{9} = 1 - \lg 9 = 0.0457\dots$$

Thus decimal representations of the numbers 2^n start with the digit 1 more than 6 times more frequently compared to the digit 9.

A natural question arises: how accurately is the value $N_I(\lg 2, Q)$ approximated by the number $Q \lg \frac{A+1}{A}$? To answer this question, we must estimate from above the remainder $R(Q)$ in (3). This estimate, in turn, is determined by the measure of irrationality of $\lg 2$, i. e., by how well $\lg 2$ is approximated by rational numbers.

Lemma 2.5. *Let the following inequality hold for an irrational number β and any integers $p, q \in \mathbb{Z} \times \mathbb{N}$:*

$$\left| \beta - \frac{p}{q} \right| > c(\beta) q^{-\lambda}, \quad \lambda \geq 2. \tag{4}$$

Then for any interval $I \subset [0, 1)$ we have

$$N_I(\beta, Q) = |I|Q + O\left(Q^{1-\frac{1}{\lambda-1}} \ln Q\right), \tag{5}$$

where the implicit constant in the Vinogradov symbol O does not exceed $\frac{2^{2\lambda+4}}{c(\beta)} + 1$.

Lemma 2.5 can be proved using Vinogradov's "little glasses" method [5] by decomposing the characteristic function of the interval I into a Fourier series. Then the residual term can be estimated from the inequality

$$\sum_{\nu=1}^L \|\nu\beta\|^{-1} < \frac{2^{\lambda+1}}{c(\beta)} L^{\lambda-1} \ln L, \quad L \geq 8,$$

where $\|x\|$ is the distance from the real number x to the nearest integer.

The class of numbers $M(\lambda)$ such that the inequality (4) holds is very broad. All real numbers with bounded partial quotients in their continued fraction representations (for example, all quadratic irrationals) lie in the class $M(2)$. For an arbitrary $\lambda > 2$, all real algebraic numbers (Roth, [6]) and almost all real numbers in the sense of Lebesgue measure (Khinchine, [7]) lie in $M(\lambda)$. Today it is known [8] that $\lg 2 \in M(\lambda)$ for $\lambda = \lambda_0 = 2^{42}$, and thus Lemma 2.5 leads to the following quantitative form of Theorem 2.1.

Theorem 2.6. *Let $B(A, Q)$ be the number of positive integers n , $1 \leq n \leq Q$, such that the leading decimal digits of 2^n coincide with the decimal representation of A . Then for any $\varepsilon > 0$ we have*

$$B(A, Q) = Q \lg \frac{A+1}{A} + O\left(Q^{1-1/(\lambda_0-1)+\varepsilon}\right). \tag{6}$$

If we replace the sequence $2^n, n = 1, 2, \dots$, by the sequence $e^n, 1, 2, \dots$, then from results of Masayoshi Hata and Elena Rukhadze [9; 10] we obtain that the residual term in (6) can be replaced by

$$O(Q^{0.66}).$$

It is easy to see that the number 2 in Theorems 2.1 and 2.6 can be replaced by an arbitrary natural number $b \geq 2, b \neq 10^l, l = 0, 1, 2, \dots$. This yields results similar to Theorem 2.6 with the same residual term as in (5) if for all $p, q \in \mathbb{Z} \times \mathbb{N}$ the inequality

$$\left| \lg b - \frac{p}{q} \right| > c(b) q^{-\lambda(b)} \tag{7}$$

is satisfied [8].

Another natural generalization of the problem can be stated as follows. Take two natural numbers A_1 and A_2 . Let $B(A_1, A_2, Q)$ be the number of integers n , $1 \leq n \leq Q$, for which 2^n starts with A_1 , and 3^n starts with A_2 .

Theorem 2.7. *There exists a real number $\mu, 0 < \mu < 1$, such that for any $0 < \varepsilon < 1 - \mu$ we have*

$$B(A_1, A_2, Q) = Q \lg \frac{A_1+1}{A_1} \lg \frac{A_2+1}{A_2} + O_\varepsilon(Q^{\mu+\varepsilon})$$

as $Q \rightarrow \infty$.

Theorem 2.7 can be proved similarly to Theorem 2.6 since the sequence of two-dimensional vectors $\vec{a}_n = (\{n \lg 2\}, \{n \lg 3\})$ is uniformly distributed in the square $[0, 1) \times [0, 1)$. Hence, these vectors infinitely often belong to the rectangle

$$[\lg A, \lg(A+1)) \times [\lg B, \lg(B+1)) \bmod 1.$$

This follows from the fact that the numbers $1, \lg 2, \lg 3$ are linearly independent over the field of rational numbers and the multivariate Weyl criterion [4]. Moreover, there exists a constant $\lambda_1 = \lambda_1(2, 3)$ that provides a quantitative characteristic of this linear independence of the form

$$|a_3 \lg 3 + a_2 \lg 2 + a_1| > c(2, 3) H^{-\lambda_1}, \tag{8}$$

where $a_j \in \mathbb{Z}$, $a_2^2 + a_3^2 \neq 0$, $H = \max_{1 \leq j \leq 3} |a_j|$.

The estimates (8) and more general estimates

$$|a_k \lg p_k + a_{k-1} \lg p_{k-1} + \dots + a_2 \lg 2 + a_1| > c(2, \dots, p_k) H^{-\lambda_2}, \quad (9)$$

$a_j \in \mathbb{Z}$, $\sum_{j=2}^k a_j^2 \neq 0$, where p_j is the j th prime number, were first obtained by A. Baker [11]. In the paper [8], Baker wrote these estimates in the form (8), (9). From these estimates, we can obtain quantitative results about the uniform distribution of fractions in a k -dimensional unit cube, which are similar to (5). Having obtained such results, we can directly prove Theorem 2.7 and the following more general theorem.

Theorem 2.8. *Let p_k be the k th prime number, and for arbitrary natural numbers A_1, A_2, \dots, A_k let $B(A_1, \dots, A_k, Q)$ be the number of positive integers n , $1 \leq n \leq Q$, such that the leading digits of 2^n coincide with the digits A_1 , 3^n has the same property with respect to A_2, \dots, p_k^n – to A_k . Then we can specify μ_1 , $0 < \mu_1 < 1$, such that for any ε_1 , $0 < \varepsilon_1 < 1 - \mu_1$, as $Q \rightarrow \infty$ we have*

$$B(A_1, \dots, A_k, Q) = Q \prod_{s=1}^k \lg \frac{A_s + 1}{A_s} + O_{\varepsilon_1}(Q^{\mu_1 + \varepsilon_1}).$$

Note that estimates from below for linear forms of the type (9) can be much more accurate if we take particular combinations of numbers b_1, \dots, b_k .

Of course, it isn't necessary to restrict Theorem 2.8 only to prime numbers. The result holds for composite numbers, as long as the set b_1, \dots, b_k satisfies the requirement that their logarithms, taken together with the number 1, are linearly independent over the field of rational numbers. For example, we can take $b_1 = 4$ and $b_2 = 9$, but can't take $b_1 = 2$, $b_2 = 4$ or $b_1 = 2$, $b_2 = 3$, $b_3 = 6$.

Let us consider another generalization. Let $B_2^{(s)}(Q)$ be the number of positive integers n , $1 \leq n \leq Q$, such that the decimal digits of the number 2^n , starting from the $(s+1)$ -th position, coincide with the number A_1 . From now on, we allow A_1 to have leading zeros, for example, $A_1 = 002$). Further, let

$$v_s(A_1) = \sum_{t=10^{s-1}}^{10^s-1} \lg \frac{10^s t + A_1 + 1}{10^s t + A_1}.$$

Theorem 2.9. *As $Q \rightarrow \infty$, we have*

$$B_2^{(s)}(Q) = Q v_s(A_1) + O_\varepsilon(Q^{1-1/(\lambda_0-1)+\varepsilon}). \quad (10)$$

The proof of Theorem 2.9 is similar to the proof of Theorem 2.6, since in the decimal representation of 2^n , the digits of A_1 appear starting from the $(s+1)$ -th position if $\{n \lg 2\}$ lies in the interval $[\lg(10^s t + A), \lg(10^s t + A + 1)) \bmod 1$ for some s -digit number t . Theorem 2.6 holds for any such interval. Now it remains to calculate a sum of the right-hand sides of the expressions of the type (6) over all $9 \cdot 10^{s-1}$ possible values of t .

Clearly, this direct approach leads to a significant increase of the residual term in (10) because of the implicit constant in the Vinogradov symbol. Take, for example, $s = 6$, then the residual term in (10) for the sequence e^n becomes at least 10^6 . Therefore, to obtain meaningful estimates of the remainder term of (10), we need to take Q of the order 10^{21} , which is very large.

The results of numerical experiments (see Section 3) suggest that fluctuations of the residual term for the individual intervals cancel each other out, to a degree, when summation is performed. We were able to formally prove that this type of interference does occur for a union of evenly spaced intervals of equal length.

Lemma 2.10. *Let the inequality (7) hold for an irrational β and all $(p, q) \in \mathbb{Z} \times \mathbb{N}$. Then for all integers $M > 1$, $S \leq M$, real numbers a and b , $0 < b - a < M^{-1}$, and*

$$V = \cup_{j=0}^S [a + jM^{-1}, b + jM^{-1}),$$

the asymptotic equality

$$N_V(\beta, Q) = |V|Q + O_\varepsilon(Q^{1-\frac{1}{\lambda-1}} \ln Q),$$

holds, where the implicit constant in the Vinogradov symbol O does not exceed

$$2^{2\lambda+10} \ln M \left(\frac{1}{c(M\beta)M^{\lambda-2}} + \frac{1}{c(\beta)} \right).$$

The result of Lemma 2.10 depends on Diophantine properties of both β and $M\beta$. The relation between $c(\beta)$ and $c(M\beta)$ is not simple, but for most numbers (in the sense of Lebesgue measure) these quantities can be considered to be equal.

Recall that the integer base 2 can be replaced with an arbitrary real base $a > 1$. Take

$$\lg a_1 = \frac{\sqrt{5}-1}{2}, \quad \lg a_2 = \sqrt{2}, \quad \lg a_3 = e, \quad \lg a_4 = \pi$$

and let $B_{a_j}(Q)$, $1 \leq j \leq 4$, be the number of integers n , $1 \leq n \leq Q$, such that the leading digits of the number a_j^n coincide with the digits of a positive integer A . Using well-known results about rational approximation of $\lg a_j$, we can retrace the proof of Theorem 2.6 to obtain the following result.

Theorem 2.11. *As $Q \rightarrow \infty$, we have*

$$B_{a_1}(Q) = Q \lg \frac{A+1}{A} + O(\ln Q); \tag{11}$$

$$B_{a_2}(Q) = Q \lg \frac{A+1}{A} + O(\ln Q); \tag{12}$$

$$B_{a_3}(Q) = Q \lg \frac{A+1}{A} + O(\ln^2 Q); \tag{13}$$

$$B_{a_4}(Q) = Q \lg \frac{A+1}{A} + O\left(Q^{\frac{5}{6}}\right). \tag{14}$$

For almost all a and an arbitrary positive constant $\varepsilon > 0$, we have

$$B_a(A, Q) = Q \lg \frac{A+1}{A} + O(\ln^{2+\varepsilon} Q). \tag{15}$$

There is every reason to expect that the estimate of the residual term in (11) will be the tightest among (11)–(15) since the golden ratio $\frac{\sqrt{5}-1}{2}$ has the continued fraction representation $[1, 1, \dots]$, leading to the estimate (4) with $c(a_1) = (\sqrt{5})^{-1} - \delta$, $\lambda = 2$, for any $\delta > 0$ and $q > q_0(\delta)$.

Let us consider another generalization of the original problem. Take $m_1 = 10^x, \dots, m_k = 10^k$. For these sequences, we can define $B'(A_1, \dots, A_k, Q)$ similarly to $B(A_1, \dots, A_k, Q)$ in Theorem 2.8 by replacing p_j with m_j . Then the well-known metric lower bound on values of polynomials with integer coefficients [12] can be used to obtain the following theorem.

Theorem 2.12. *For almost all x and any $\delta > 0$, we have*

$$B'(A_1, \dots, A_k, Q) = Q \prod_{s=1}^k \lg \frac{A_s+1}{A_s} + O_\delta(\ln^{k+\delta} Q).$$

3. Results of computational experiments

In order to evaluate the accuracy of our theoretical bounds, we have performed a number of computations for large numbers Q , obtaining the following results.

Calculation of $B(A, Q)$ for $b = 2$ and all $A = 1, 2, \dots, 9$ shows that the deviation of $B(A, Q)$ from the asymptotic estimate $Q \lg \frac{A+1}{A}$ does not exceed 7 for $1 \leq Q \leq 10^6$. This suggests that $\lg 2$ is like most real numbers, i. e., that it lies in the class $M(2)$.

Looking at the sequences 2^n and 3^n simultaneously, for the quantity $B(A_1, A_2, Q)$ we have obtained that for all combination of the numbers (A_1, A_2) , $1 \leq A_1 \leq 9$, $1 \leq A_2 \leq 9$, $Q \leq 10^6$, we have

$$\left| B(A_1, A_2, Q) - Q \lg \frac{A_1+1}{A_1} \lg \frac{A_2+1}{A_2} \right| \leq 12.$$

Finally, considering the second, third, fourth, fifth and sixth digits for bases 2 and 3, and for all possible digits $0 \leq A_1 \leq 9$, we obtain that the maximum deviation of $B_2^{(s)}(A_1, Q)$ and $B_3^{(s)}(A_1, Q)$ from $Q \nu_s(A_1)$ for $0 \leq Q \leq 10^6$ does not exceed 22 ($s = 2$), 65 ($s = 3$), 122 ($s = 4$), 405 ($s = 5$), 921 ($s = 6$).

These results are visualized in the Figures 1 and 2 below.

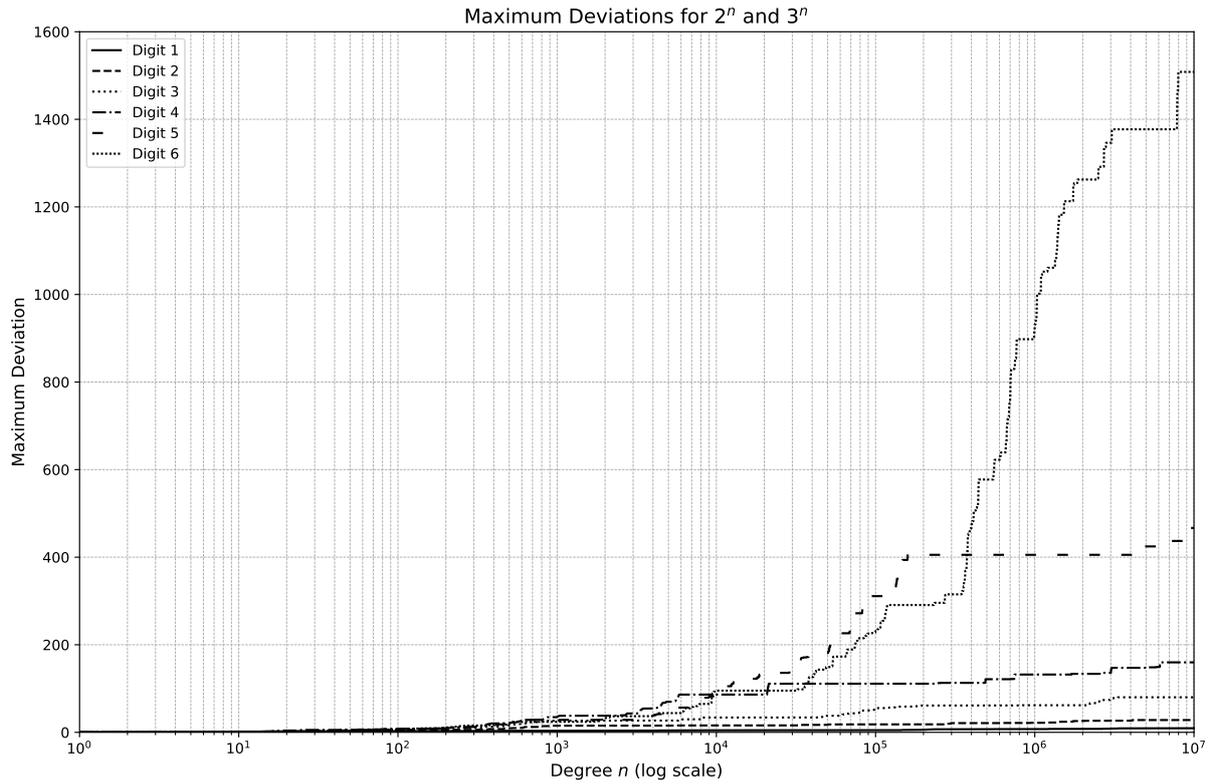


Fig. 1. Maximum deviations of $B_2^{(s)}(A_1, Q)$ and $B_3^{(s)}(A_1, Q)$ from the respective asymptotic values

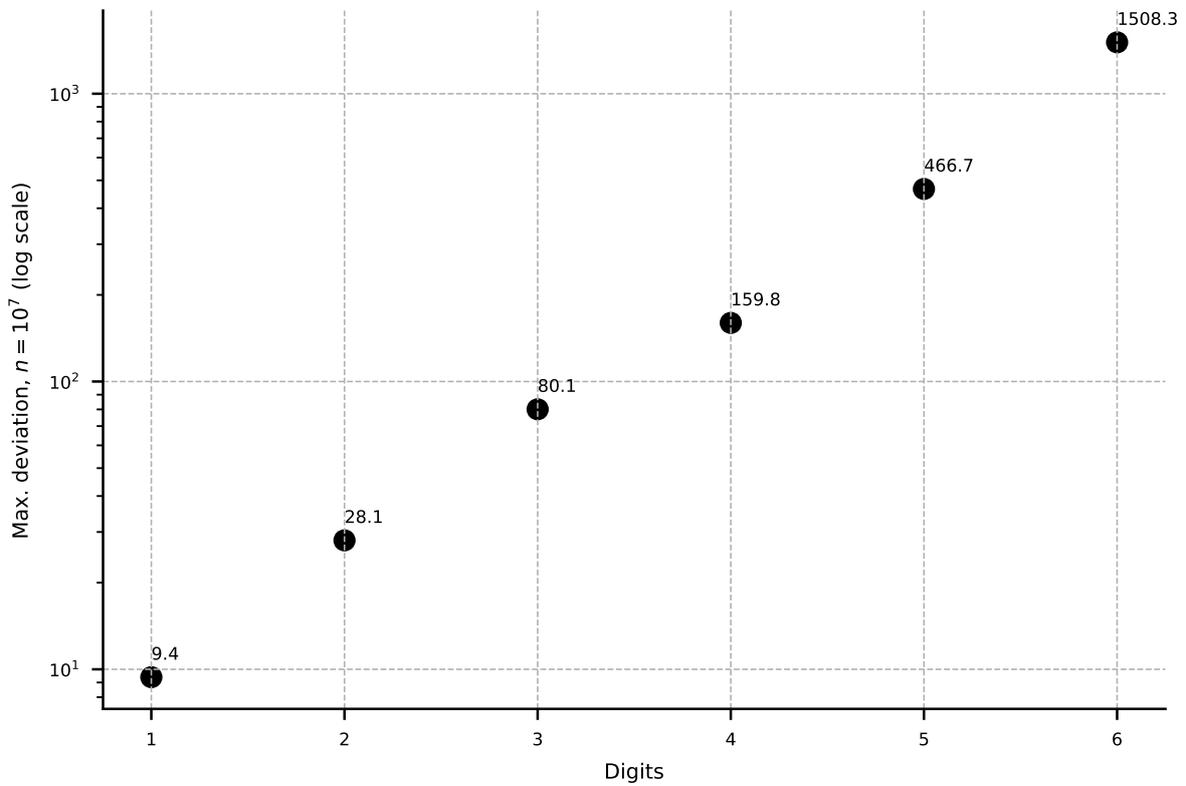


Fig. 2. Maximum deviations of $B_2^{(s)}(A_1, Q)$ and $B_3^{(s)}(A_1, Q)$ for $Q = 10^7$ as s changes from 1 to 6

Fig. 2 suggests that the asymptotic growth of the respective deviation is proportional to e^s and not 10^{s-1} , as expected from considering the 10^{s-1} intervals.

4. Conclusion

The problem of quantitative characterization of statistical properties of the first digits of integer powers has a long history. We have established the connection between this problem and Diophantine properties of logarithms, and obtained estimates for the remainder term in the asymptotic expression for the number of integer powers with specific first digits. We have also proposed numerous generalizations of this problem and provided state of the art solutions. Our results rely on known theorems establishing Diophantine properties of the respective real logarithms. However, these theorems are often very inexact, and improving them often requires solving difficult classical problems of Diophantine approximation. In certain cases, metric approach allows researchers to circumvent this obstacle by proving the desired Diophantine properties for a subset of a box $T \subset \mathbb{R}^n$ of sufficiently large Lebesgue measure [12].

5. Topical problems in the theory of Diophantine approximation

The results presented in the paper show how Diophantine properties of numbers can have surprising consequences in other areas of mathematics. To close out the article, let us formulate several topical problems in Diophantine approximation.

Problem 1. Consider the well-known Dirichlet-type theorem on solutions of the inequality

$$|P(x)| < Q^{-n} \quad (16)$$

in polynomials $P(x)$, $\deg P \leq n$, $H(P) \leq Q$. Obtain bounds on the measure of sets $\sigma(P)$ such that inequality (16) holds for points in these sets.

Problem 2. Study inequality (16) for a) irreducible polynomials and b) reducible polynomials.

Problem 3. Generalize inequality (16) to the fields of a) complex numbers and b) p -adic numbers.

Problem 4. Consider the inequality

$$|D(P)| < 2^{2n-2-2v}, \quad v \geq 0, \quad (17)$$

where $D(P)$ is the discriminant of a polynomial P with integer coefficients of degree n and height $H(P) \leq Q$. Find upper and lower bounds for the number of such polynomials satisfying the inequality (17) in the fields of real and p -adic numbers.

The authors would like to thank Y. V. Prokhorov, Y. V. Nesterenko, A. Dubickas, F. Goetze and V. G. Safonov for a number of useful constructive comments.

This work was supported by the Institute of Mathematics of the National Academy of Sciences of Belarus within the framework of the state programme ‘‘Convergence–2020’’.

References

1. Newcomb S. On the frequency of use of the different digits in natural numbers. *American Journal of Mathematics*, 1881, vol. 4, pp. 39–40.
2. Benford F. The law of anomalous numbers. *Proceedings of the American Philosophical Society*, 1938, vol. 78, pp. 551–572.
3. Hürlimann W. Generalizing Benford's Law Using Power Laws: Application to Integer Sequences. *International Journal of Mathematics and Mathematical Sciences*, 2009, art. 970284, 10 pp. <https://doi.org/10.1155/2009/970284>.
4. Kuipers L., Niederreiter H. *Uniform Distribution of Sequences*. New York, Wiley, 1974, 390 p. (Pure and Applied Mathematics).
5. Vinogradov I. M. *Selected Papers*. Berlin, Springer-Verlag, 1985, 401 p.
6. Roth K. Rational approximations to algebraic numbers. *Mathematica*, 1955, vol. 2, pp. 1–20.
7. Khinchine A. Zur metrischen Theorie der diophantischen Approximationen. *Mathematische Annalen*, 1924, vol. 92, pp. 115–125.
8. Baker A., Wüstholtz G. Linear forms in logarithms of algebraic numbers. *Journal für die reine und angewandte Mathematik*, 1993, vol. 442, pp. 19–62.

9. Hata M. Legendre type polynomials and irrationality measures. *Journal für die reine und angewandte Mathematik*, 1990, vol. 407, pp. 99–125.
10. Rukhadze E. A. A problem on the distribution of first digits. *Moscow University Mathematics Bulletin*, 1987, vol. 42, no. 1, pp. 25–29 (in Russian).
11. Baker A. *Transcendental Number Theory*. Cambridge, Cambridge University Press, 1990, 155 p.
12. Bernik V., Kleinbock D., Margulis G. A. Khinchine-type theorems on manifolds: convergence case for standard and multiplicative versions. *International Mathematics Research Notices*, 2001, vol. 9, pp. 453–486.

УДК 512.64

О НОРМЕ ШАТТЕНА ЛИНЕЙНОГО ОТОБРАЖЕНИЯ ЕВКЛИДОВЫХ ПРОСТРАНСТВ

Д. Н. Курочкина

Белорусский государственный университет, Минск, Беларусь
e-mail: fpm.kurochkiDN@bsu.by

Поступила: 22.01.2025

Исправлена: 23.03.2025

Принята: 23.05.2025

Ключевые слова: линейное отображение, евклидово пространство, сингулярные числа, ортогональная матрица, неравенство Иенсена, средние степени, неравенство Адамара.

Аннотация. Установлена связь нормы Шаттена линейного отображения $\mathcal{A} : V \rightarrow W$ евклидовых пространств размерностей n и m соответственно со средними степенными, порожденными длинами образов векторов ортонормированного базиса V .

ON THE SCHATTEN NORM OF A LINEAR MAP ON EUCLIDEAN SPACES

D. N. Kurochkina

Belarusian State University, Minsk, Belarus
e-mail: fpm.kurochkiDN@bsu.by

Received: 22.01.2025

Revised: 23.03.2025

Accepted: 23.05.2025

Keywords: linear map, Euclidean space, singular values, orthogonal matrix, Jensen's inequality, power means, Hadamard's inequality.

Abstract. We find a connection between the Schatten norm of the linear mapping $\mathcal{A} : V \rightarrow W$ of Euclidean spaces of dimensions n and m , respectively, and the power means generated by the lengths of the images of vectors of the orthonormal basis V .

1. Введение

Пусть V, W – евклидовы пространства размерностей n и m соответственно, тогда $\mathcal{A} : V \rightarrow W$ есть линейное отображение ранга $k \leq \min\{n, m\}$, $\mathcal{A}^* : W \rightarrow V$ – сопряженное к \mathcal{A} отображение. Известно, что $\mathcal{A}^* \circ \mathcal{A} : V \rightarrow V$ есть самосопряженный оператор, матрицей которого в некотором ортонормированном базисе $\mathcal{E}_V = \langle \varepsilon_1, \dots, \varepsilon_n \rangle$ пространства V является диагональная матрица $\text{diag}[\rho_1^2, \dots, \rho_n^2]$, где $\rho_1 \geq \dots \geq \rho_k > 0 = \rho_{k+1} = \dots = \rho_n$ есть так называемые сингулярные числа оператора \mathcal{A} (все используемые в статье сведения о матрицах и линейных отображениях можно найти в [1] и [2, Гл. 4, § 16]). Теорема о сингулярном разложении утверждает, что существует такой ортонормированный базис $\mathcal{F} = \langle f_1, \dots, f_m \rangle$ пространства W , что

$$\mathcal{A}(\varepsilon_i) = \rho_i f_i \text{ при } i \leq n \text{ и } \mathcal{A}^*(f_j) = \rho_j \varepsilon_j \text{ при } j \leq m.$$

Обозначим через $\mathcal{B} = (\mathcal{A}^* \circ \mathcal{A})^{1/2} : V \rightarrow V$ корень из самосопряженного оператора $\mathcal{A}^* \circ \mathcal{A}$, а через $\mathcal{B}^p : V \rightarrow V$ – p -ю степень оператора \mathcal{B} . В базисе $\mathcal{E}_V = \langle \varepsilon_1, \dots, \varepsilon_n \rangle$ матрицей оператора $\mathcal{B}^p : V \rightarrow V$ является диагональная матрица $\text{diag}[\rho_1^p, \dots, \rho_n^p]$.

Введем в рассмотрение число

$$|\mathcal{A}|_p = \text{tr}(\mathcal{B}^p)^{\frac{1}{p}} = (\rho_1^p + \dots + \rho_n^p)^{\frac{1}{p}}, \text{ где } p \geq 1.$$

Ясно, что $|\mathcal{A}|_2 = \sqrt{\rho_1^2 + \dots + \rho_n^2}$, а $|\mathcal{A}|_1 = \rho_1 + \dots + \rho_n$. Известно [1, р. 464–465], что $|\mathcal{A}|_p$ является нормой в векторном пространстве $\text{Lin}(V, W)$ всех линейных отображений, которая называется p -нормой Шаттена отображения \mathcal{A} . Норма Шаттена является матричной, т. е. $|\mathcal{A}_2 \circ \mathcal{A}_1|_p \leq |\mathcal{A}_2|_p \cdot |\mathcal{A}_1|_p$ для композиции $V \xrightarrow{\mathcal{A}_1} W \xrightarrow{\mathcal{A}_2} U$ линейных отображений.

Всюду далее V и W будут евклидовы пространства размерностей n и m соответственно.

Определение 1.1. Обозначим для линейного отображения $\mathcal{A} : V \rightarrow W$ число

$$\|\mathcal{A}\|_p = \sup_{\mathcal{E}} \left(\sum_{i \leq n} \|\mathcal{A}(e_i)\|^p \right)^{1/p},$$

где супремум берется по всем ортонормированным базисам $\mathcal{E} = \langle e_1, \dots, e_n \rangle$ пространства V (через $\|\mathcal{A}(e_i)\|$ обозначена длина вектора $\mathcal{A}(e_i)$).

Впервые такого рода характеристика линейного отображения была рассмотрена в работе [3] для ортогональной проекции $\mathcal{A} = pr : \mathbb{R}^n \rightarrow W$ на подпространство $W \subset \mathbb{R}^n$ в связи с уточнением оценки дисперсии случайной ошибки в теории Гаусса метода наименьших квадратов.

Используя стандартные рассуждения и неравенство Минковского, несложно показать, что $\|\mathcal{A}\|_p$ при $p \geq 1$ является нормой линейного отображения $\mathcal{A} : V \rightarrow W$, где V и W – пространства размерностей m, n соответственно.

Целью данной заметки является сравнение нормы $\|\mathcal{A}\|_p$ и p -нормы Шаттена $|\mathcal{A}|_p$. С одной стороны они похожи по форме и кроме того $\|\mathcal{A}\|_2 = |\mathcal{A}|_2$ (что следует из цепочки равенств

$$\text{tr}(\mathcal{A}^* \circ \mathcal{A}) = \sum_{i \leq n} (\mathcal{A}^* \circ \mathcal{A}(e_i), e_i) = \sum_{i \leq n} (\mathcal{A}(e_i), \mathcal{A}(e_i)) = \sum_{i \leq n} \|\mathcal{A}(e_i)\|^2 \quad (1)$$

для любого ортонормированного базиса $\mathcal{E} = \langle e_1, \dots, e_n \rangle$).

С другой стороны, проверка неравенства треугольника для p -нормы Шаттена $|\mathcal{A}|_p$ вызывает большие сложности (что не так для нормы $\|\mathcal{A}\|_p$). Теорема 1.2 еще сильнее подчеркивает нетривиальность во взаимоотношении между ними.

Теорема 1.2. Для линейного отображения $\mathcal{A} : V \rightarrow W$ справедливо

$$\|\mathcal{A}\|_p = \begin{cases} |\mathcal{A}|_p, & \text{если } p \geq 2 \\ c_p \cdot |\mathcal{A}|_2, & \text{если } 1 \leq p < 2 \end{cases}$$

(здесь константа $c_p = n^{1/p-1/2}$).

Доказательство теоремы 1.2 основывается на точных оценках для средних k -степенных, порождаемых линейным отображением (см. теорему 1.3). Напомним, что средним k -степенным неотрицательных чисел β_1, \dots, β_n называется число

$$A_k(\beta_1, \dots, \beta_n) = \begin{cases} \left(\frac{\beta_1^k + \dots + \beta_n^k}{n} \right)^{\frac{1}{k}} & \text{при } k \neq 0 \\ (\beta_1 \cdot \dots \cdot \beta_n)^{\frac{1}{k}} & \text{при } k = 0. \end{cases}$$

Известно, что

$$A_k(\beta_1, \dots, \beta_n) \geq A_2(\beta_1, \dots, \beta_n) \geq A_l(\beta_1, \dots, \beta_n) \text{ при } k \geq 2 \geq l. \quad (2)$$

Пусть $\mathcal{A} : V \rightarrow W$ есть линейное отображение и $\mathcal{E} = \langle e_1, \dots, e_n \rangle$ – ортонормированный базис V . Через $A_k(\mathcal{E})$ обозначим $A_k(\beta_1, \dots, \beta_n)$, где $\beta_i = \|\mathcal{A}(e_i)\| \geq 0$ при $i \leq n$.

Теорема 1.3. Пусть ρ_1, \dots, ρ_n есть сингулярные числа линейного отображения $\mathcal{A} : V \rightarrow W$. Тогда для любого ортонормированного базиса $\mathcal{E} = \langle e_1, \dots, e_n \rangle$ пространства V среднее степенное $A_k(\mathcal{E})$ удовлетворяет двойному неравенству

1. $A_k(\rho_1, \dots, \rho_n) \geq A_k(\mathcal{E}) \geq A_2(\rho_1, \dots, \rho_n)$ при $k \geq 2$;
2. $A_k(\rho_1, \dots, \rho_n) \leq A_k(\mathcal{E}) \leq A_2(\rho_1, \dots, \rho_n)$ при $k < 2$.

Замечание 1.4. В [3] для ортогональной проекции $\mathcal{A} = pr : \mathbb{R}^n \rightarrow W$ для $k = 4$ было установлено двойное неравенство:

$$(*) \left(\frac{m}{n} \right)^{\frac{1}{4}} \geq A_4(\mathcal{E}) \geq \left(\frac{m}{n} \right)^{\frac{1}{2}}. \quad (3)$$

которое согласуется с теоремой 1.3, так как сингулярные числа ортогональной проекции равны $\rho_1 = \dots = \rho_m = 1$, а $\rho_{m+1} = \dots = \rho_n = 0$. В работе [4] для ортогональной проекции была установлена неулучшаемость оценки (3) снизу, а в [5] были получены необходимые и достаточные условия для существования ортогонального базиса, проектирующегося в векторы равной длины.

2. Редукция теоремы 1.2 к теореме 1.3 и ортогональным матрицам

Ясно, что для сингулярного базиса \mathcal{E}_V $A_p(\mathcal{E}_V)$ совпадает с $A_p(\rho_1, \dots, \rho_n)$. Этот факт и теорема 1.3 совместно влекут:

1. $\max_{\mathcal{E}} A_p(\mathcal{E}) = A_p(\rho_1, \dots, \rho_n)$ при $p \geq 2$;
2. $\min_{\mathcal{E}} A_p(\mathcal{E}) = A_p(\rho_1, \dots, \rho_n)$ при $p < 2$.

Поэтому $\|\mathcal{A}\|_p = \max_{\mathcal{E}} (\sum_{i \leq n} \|\mathcal{A}(e_i)\|^p)^{1/p} = n^{1/p} \max_{\mathcal{E}} A_p(\mathcal{E})$ при $p \geq 2$ совпадает с

$$n^{1/p} A_p(\rho_1, \dots, \rho_n) = \left(\sum_{i \leq n} \rho_i^p \right)^{1/p} = |\mathcal{A}|_p.$$

В учебнике С. М. Агеева «Курс линейной алгебры» содержится, что $A_p(\mathcal{E}) = A_2(\rho_1, \dots, \rho_n)$ для так называемого *равнобедренного базиса* $\mathcal{E} = \mathcal{E}_{eq}$, т. е. такого ортонормированного базиса, у которого длины β_i всех векторов $\mathcal{A}(e_i)$ равны между собой (и, следовательно, $\beta_i = A_2(\beta_1, \dots, \beta_n) = A_2(\rho_1, \dots, \rho_n)$). Следовательно,

3. $\max_{\mathcal{E}} A_p(\mathcal{E}) = A_2(\rho_1, \dots, \rho_n)$ при $p < 2$;
4. $\min_{\mathcal{E}} A_p(\mathcal{E}) = A_2(\rho_1, \dots, \rho_n)$ при $p \geq 2$.

Поэтому $\|\mathcal{A}\|_p = \max_{\mathcal{E}} (\sum_{i \leq n} \|\mathcal{A}(e_i)\|^p)^{1/p} = n^{1/p} \max_{\mathcal{E}} A_p(\mathcal{E})$ при $1 \leq p < 2$ совпадает с

$$n^{1/p} A_2(\rho_1, \dots, \rho_n) = n^{1/p-1/2} \left(\sum_{i \leq n} \rho_i^2 \right)^{1/2} = c_p |\mathcal{A}|_2.$$

Покажем, что доказательство теоремы 1.3 сводится к аналогичному утверждению для ортогональных матриц.

Предложение 2.1. Пусть \mathcal{E}_V есть сингулярный, а \mathcal{E} – ортонормированный базисы пространства V . Пусть также T есть ортогональная матрица $T(\mathcal{E}_V | \mathcal{E})$ перехода от сингулярного базиса \mathcal{E}_V к базису \mathcal{E} , а сингулярные числа отображения \mathcal{A} стоят на диагонали матрицы $R = \text{diag}[\rho_1, \rho_2, \dots, \rho_n]$, где $\rho_1 \geq \rho_2 \geq \dots \geq \rho_n$. Тогда $\|\mathcal{A}(e_i)\| = \beta_i$ для любого $i \leq n$ совпадает с длиной s_i i -го столбца матрицы $T_R = RT$.

Доказательство. Достаточно показать, что T_R есть матрица $T(\mathcal{E}_W | \mathcal{A}(\mathcal{E}))$ перехода от сингулярного базиса \mathcal{E}_W к образу $\mathcal{A}(\mathcal{E})$. Действительно, $e_j = \sum_{i=1}^n t_{ij} \varepsilon_i$ и поэтому

$$A(e_j) = \sum_{i=1}^n t_{ij} A(\varepsilon_i) = \sum_{i=1}^n (t_{ij} \rho_i) f_i.$$

Следовательно, $s_j = \sum_{i=1}^n (t_{ij} \rho_i)^2 = \|A(e_j)\|^2 = \beta_j$. □

Предложение 2.2. Если \mathcal{E} – ортонормированный базис пространства V , T – матрица перехода от сингулярного базиса \mathcal{E}_V к \mathcal{E} , то $A_k(\mathcal{E}) = \left(\frac{\sum_{i=1}^n s_i^k}{n} \right)^{1/k}$, где s_i есть длина i -го столбца матрицы T_R .

Отсюда следует, что при доказательстве теоремы 1.3 (1)–(2) длины векторов $\|\mathcal{A}(e_i)\|$ можно заменить на длины s_i столбцов матрицы T_R .

3. Доказательство теоремы 1.3 (1)

Из неравенства (2) и цепочки равенств (1) следует, что $A_2(\mathcal{E}) = A_2(\rho_1, \dots, \rho_n)$ для любого ортонормированного базиса $\mathcal{E} = \langle e_1, \dots, e_n \rangle$ пространства V , а также

$$A_k(\mathcal{E}) \geq A_2(\rho_1, \dots, \rho_n) \text{ при } k \geq 2 \text{ и } A_k(\mathcal{E}) \leq A_2(\rho_1, \dots, \rho_n) \text{ при } k \leq 2.$$

Отсюда и из предложения 2.2 следует, что доказательство неравенства (1) из теоремы 1.3 сводится к неравенству для ортогональных матриц при $k \geq 2$:

$$\left(\frac{\rho_1^k + \dots + \rho_n^k}{n} \right)^{1/k} \geq \left(\frac{\sum_{i=1}^n s_i^k}{n} \right)^{1/k}.$$

Для его установления воспользуемся неравенством Иенсена: если гладкая числовая функция $y = f(x)$ является выпуклой, то выполняется неравенство:

$$f(\alpha_1 x_1 + \dots + \alpha_n x_n) \leq \alpha_1 f(x_1) + \dots + \alpha_n f(x_n),$$

где числа $\alpha_1, \alpha_2, \dots, \alpha_n \geq 0$ таковы, что их сумма равна единице (если же функция $f(x)$ вогнута, то

$$f(\alpha_1 x_1 + \dots + \alpha_n x_n) \geq \alpha_1 f(x_1) + \dots + \alpha_n f(x_n)).$$

Ясно, что (*) функция $f(x) = x^{\frac{k}{2}}$, имеющая 2-ю произвольную $f^{(2)}(x) = \frac{k}{2}(\frac{k}{2} - 1)x^{\frac{k}{2}-2}$, является выпуклой при $k \geq 2$ и $k < 0$ (является вогнутой при $0 < k < 2$).

Пусть $k \geq 2$. Так как $\sum_{i \leq n} t_{ij}^2 = 1$, то для j -го столбца матрицы T_R имеет место

$$(\rho_1^2 t_{1j}^2 + \dots + \rho_n^2 t_{nj}^2)^{\frac{k}{2}} \leq t_{1j}^2 \rho_1^k + \dots + t_{nj}^2 \rho_n^k.$$

Применив неравенство Иенсена для каждого столбца, получим:

$$(\rho_1^2 t_{11}^2 + \dots + \rho_n^2 t_{n1}^2)^{\frac{k}{2}} + \dots + (\rho_1^2 t_{1n}^2 + \dots + \rho_n^2 t_{nn}^2)^{\frac{k}{2}} \leq A,$$

где $A = (t_{11}^2 \rho_1^k + \dots + t_{n1}^2 \rho_n^k) + \dots + (t_{1n}^2 \rho_1^k + \dots + t_{nn}^2 \rho_n^k) = (t_{11}^2 + \dots + t_{1n}^2) \rho_1^k + \dots + (t_{n1}^2 + \dots + t_{nn}^2) \rho_n^k$. Поскольку матрица T ортогональна, то $(t_{i1}^2 + \dots + t_{in}^2) \rho_i^k = \rho_i^k$ для всех $i \leq n$, и поэтому $A = \rho_1^k + \dots + \rho_n^k$.

Поскольку функция $f(x) = x^{\frac{k}{2}}$ монотонно возрастает при $k \geq 2$, то

$$\left(\frac{\sum_{i=1}^n s_i^k}{n} \right)^{\frac{1}{k}} \leq \left(\frac{\rho_1^k + \dots + \rho_n^k}{n} \right)^{\frac{1}{k}}.$$

Отсюда и из предложения 2.2 получаем

$$A_k(\mathcal{E}) = \left(\frac{\sum_{i=1}^n s_i^k}{n} \right)^{\frac{1}{k}} \leq \left(\frac{\rho_1^k + \dots + \rho_n^k}{n} \right)^{\frac{1}{k}} = A_k(\mathcal{E}_V).$$

4. Доказательство теоремы 1.3 (2)

Доказательство неравенства (2) из теоремы 1.3 сводится к неравенству для ортогональных матриц:

$$\left(\frac{\sum_{i=1}^n s_i^k}{n} \right)^{\frac{1}{k}} \geq \left(\frac{\rho_1^k + \dots + \rho_n^k}{n} \right)^{\frac{1}{k}} \text{ при } k \leq 2, k \neq 0; \quad (4)$$

и

$$\left(\prod_{i=1}^n s_i \right)^{\frac{1}{n}} \geq \left(\prod_{i=1}^n \rho_i \right)^{\frac{1}{n}} \text{ при } k = 0. \quad (5)$$

Сначала докажем неравенство (5), которое можно переписать в виде

$$\left(\prod_{j=1}^n \left(\sum_{i=1}^n t_{ij}^2 \rho_i^2 \right)^{\frac{1}{2}} \right)^{\frac{1}{n}} \geq \left(\prod_{i=1}^n \rho_i \right)^{\frac{1}{n}}$$

или в равносильном виде

$$\prod_{j=1}^n \left(\sum_{i=1}^n t_{ij}^2 \rho_i^2 \right)^{\frac{1}{2}} \geq \prod_{i=1}^n \rho_i.$$

Для доказательства последнего неравенства воспользуемся неравенством Адамара, примененным к матрице T_R . Учитывая, что $\det(T_R) = \det(T) \prod_{i=1}^n \rho_i$, а матрица T ортогональна, получим требуемое:

$$\prod_{i=1}^n \rho_i = |\det(T)| \prod_{i=1}^n \rho_i \leq \prod_{i=1}^n s_i = \prod_{j=1}^n \left(\sum_{i=1}^n t_{ij}^2 \rho_i^2 \right)^{\frac{1}{2}}.$$

Наконец, проводя рассуждения, аналогичные тем, что были приведены в предыдущем параграфе, докажем неравенство (4). Из (*) следует, что $s_i^k = (\rho_1^2 t_{1i}^2 + \dots + \rho_n^2 t_{ni}^2)^{\frac{k}{2}}$ не превосходит $t_{1i}^2 \rho_1^k + \dots + t_{ni}^2 \rho_n^k$ при $k < 0$; и не меньше $t_{1i}^2 \rho_1^k + \dots + t_{ni}^2 \rho_n^k$ при $0 < k < 2$. Отсюда и из того, что

матрица T ортогональна, следует, что

$$\sum_{i=1}^n s_i^k = (\rho_1^2 t_{11}^2 + \dots + \rho_n^2 t_{n1}^2)^{\frac{k}{2}} + \dots + (\rho_1^2 t_{1n}^2 + \dots + \rho_n^2 t_{nn}^2)^{\frac{k}{2}}$$

при $k < 0$ не превосходит

$$(t_{11}^2 + \dots + t_{1n}^2)\rho_1^k + \dots + (t_{n1}^2 + \dots + t_{nn}^2)\rho_n^k = \rho_1^k + \dots + \rho_n^k$$

(и при $0 < k < 2$ не меньше $\rho_1^k + \dots + \rho_n^k$).

Поскольку функция $f(x) = x^{\frac{k}{2}}$ монотонно возрастает при $0 < k < 2$ и монотонно убывает при $k < 0$, то

$$\left(\frac{\sum_{i=1}^n s_i^k}{n}\right)^{\frac{1}{k}} \geq \left(\frac{\rho_1^k + \dots + \rho_n^k}{n}\right)^{\frac{1}{k}} \text{ при } 0 < k < 2 \text{ или } k < 0 - \text{неравенство (4) установлено.}$$

В заключение, автор выражает благодарность профессору С. М. Агееву, под руководством которого выполнена эта работа.

Литература

1. Horn R., Johnson Ch. *Matrix Analysis*. Cambridge: Cambridge University Press, 2014.
2. Мальцев А. И. Основы линейной алгебры. Москва: Лань, 2024.
3. Колмогоров А. Н., Петров А. А., Смирнов Ю. М. Одна формула Гаусса из теории метода наименьших квадратов // Изв. АН СССР. Сер. матем. 1947. Т. 11, № 6. С. 561–566.
4. Мальцев А. И. Замечание к работе А. Н. Колмогорова, А. А. Петрова и Ю. М. Смирнова // Изв. АН СССР. Сер. матем. 1947. Т. 11, № 6. С. 567–568.
5. Нисневич Л. Б., Брызгалов В. И. Об одной задаче n -мерной геометрии // УМН. 1953. Т. 8, № 4(56). С. 169–172.

References

1. Horn R., Johnson Ch. *Matrix Analysis*. Cambridge, Cambridge University Press, 2014.
2. Maltsev A. I. *Fundamentals of Linear Algebra*. Moscow, Lan, 2024 (in Russian).
3. Kolmogorov A. N., Petrov A. A., Smirnov Yu. M. One Gauss Formula from Least Squares Theory. *Izvestiya Akad. Nauk SSSR. Ser. Mat.*, 1947, vol. 11, no. 6, pp. 561–566 (in Russian).
4. Maltsev A. I. A Remark on the Paper of A. N. Kolmogorov, A. A. Petrov, and Yu. M. Smirnov. *Izvestiya Akad. Nauk SSSR. Ser. Mat.*, 1947, vol. 11, no. 6, pp. 567–568 (in Russian).
5. Nisnevich L. B., Bryzgalov V. I. On a Problem of N-Dimensional Geometry. *UMN*, 1953, vol. 8, no. 4(56), pp. 169–172 (in Russian).

УДК 511.2

НЕРАЗВЕТВЛЕННЫЕ РАСШИРЕНИЯ ГАЛУА И ПОДГРУППЫ GL_n

Д. А. Малинин

Институт математики НАН Беларуси, Минск, Беларусь
e-mail: dmalinin@gmail.com

Поступила: 14.01.2025

Исправлена: 27.01.2025

Принята: 23.05.2025

Ключевые слова: неразветвленные расширения Галуа, действие Галуа, конечные устойчивые подгруппы GL_n .

Аннотация. Мы рассматриваем естественное действие групп Галуа неразветвленных расширений Галуа числовых полей на конечных устойчивых при действии группы Галуа подгруппах GL_n .

UNRAMIFIED GALOIS EXTENSIONS AND SUBGROUPS OF GL_n

D. A. Malinin

Institute of Mathematics of the National Academy of Sciences of Belarus, Minsk, Belarus
e-mail: dmalinin@gmail.com

Received: 14.01.2025

Revised: 27.01.2025

Accepted: 23.05.2025

Keywords: unramified Galois extensions, Galois action, Galois stable finite subgroups of GL_n .

Abstract. We consider the natural action of Galois groups of unramified Galois extensions of number fields on finite Galois stable subgroups of GL_n .

1. Введение

Существование глобальных полей с заданной группой Галуа для разных классов групп, а также с предписанными локальными свойствами для ветвления возникает в связи с различными вопросами теории чисел. В связи с этим можно отметить, к примеру, работы [1] и [2].

В этой статье мы рассматриваем некоторое неразветвленное расширение Галуа E/F конечной степени d с группой Галуа Γ для числовых полей E и F , и конечной абелевой подгруппой $G \subset GL_n(E)$ заданной экспоненты t , где мы предполагаем, что G устойчива относительно естественного коэффицициентного Γ -действия.

В данной работе O_K обозначает максимальный порядок числового поля K , а $F(G)$ обозначает поле, которое получается присоединением к F всех матричных коэффицициентов всех матриц $g \in G$.

Основная цель данной статьи – доказать существование абелевых Γ -устойчивых подгрупп $G \subset GL_n(R)$ для заданного n и заданного показателя t группы G (R обозначает некоторые дедекиндовы подкольца E , но наиболее интересным случаем является $R = O_E$) таких, что $F(G) = E$ обеспечивает некоторые разумные ограничения для фиксированного нормального расширения E/F и целые числа n, t, d остаются верными. Установлена нижняя граница для возможных степеней n представлений $G: n \geq C = C(E, F, d, t, h)$ такая, что $\phi_E(t)d \leq C \leq \phi_E(t)dh$, где h – показатель группы классов поля F ; $\phi_E(t) = [E(\zeta_t) : E]$ – обобщенная функция Эйлера для поля E (ζ_t обозначает примитивный t -корень из 1). Также доказано, что в некоторых случаях верхняя граница улучшаема (теорема 1, части 1), 3), 4)), хотя нижняя граница $C = \phi_E(t)d \leq n$ не может быть улучшена (предложение 2).

Эти результаты имеют некоторые приложения к конечным арифметическим группам, их когомологиям и положительно определенным квадратичным решеткам над кольцами целых чисел в полях вполне вещественных чисел (см. [3–5]), например, обобщенный «принцип Хассе»: для арифметических групп G определенного типа ($G_{\mathbb{R}}$ является компактной), вполне вещественного K/\mathbb{Q} и $Gal(K/\mathbb{Q})$ -устойчивой подгруппы G_{O_K} из $GL_n(O_K)$ ядро естественного отображения когомологий $H^1(Gal(K/\mathbb{Q}), G_{O_K}) \rightarrow \prod_v H^1(Gal(K_v/\mathbb{Q}_v), G_v)$ тривиально. Некоторые результаты, связанные с устойчивостью Галуа для порядков в конечномерных алгебрах, можно найти в [6]. Явная конструкция вполне вещественных неразветвленных расширений полей полезна в этой ситуации. Некоторые интересные конструкции неразветвленных и вполне вещественных (а также мнимых)

числовых полей получены в статьях [7–10], см. также [11]; некоторые компьютерные вычисления с использованием KANT и PARI могут быть полезны для этой цели (см., например, [8; 10]). Другая конструкция вполне вещественных неразветвленных расширений, имеющих заданную группу Галуа, приведена в теореме 3.

Статья организована следующим образом. Формулировки результатов приведены в разделе 2, разделы 3, 4, 5 посвящены их доказательствам.

Большинство символов и обозначений, которые мы используем в этой статье, являются традиционными. \mathbb{Q} и \mathbb{Q}_p обозначают поле рациональных чисел и p -адических рациональных чисел. \mathbb{Z} и \mathbb{N} обозначают кольцо рациональных целых чисел и натуральных чисел, \mathbb{R} и \mathbb{C} обозначают поля действительных и комплексных чисел. $GL_n(R)$ обозначает общую линейную группу над R . Мы пишем $[E : F]$ для степени расширения поля E/F . Максимальный порядок числового поля K обозначается O_K . В этой статье мы пишем Γ для групп Галуа, $\sigma, \gamma \in \Gamma$ для элементов Γ . Мы пишем ζ_t для примитивного t -корня из 1, $\phi_K(t) = [K(\zeta_t) : K]$ обозначает обобщенную функцию Эйлера для поля K , I_m обозначает единичную $(m \times m)$ -матрицу, $\det M$ является определителем матрицы M . Если G – конечная линейная группа, $F(G)$ обозначает поле, полученное присоединением к F всех матричных коэффициентов всех матриц $g \in G$. Для Γ , действующего на G , и любых $\sigma \in \Gamma$ и $g \in G$ мы пишем g^σ для образа g под действием σ . K^Γ обозначает подполе Γ -устойчивых элементов поля K , $\dim_K A$ обозначает размерность K -алгебры A над полем K , $M_n(R)$ – полную матричную алгебру над R .

2. Результаты

Пусть E обозначает конечное расширение поля алгебраических чисел F , отличное от F . Пусть O'_E обозначает пересечение колец нормирования всех разветвленных простых идеалов в кольце O_E , и пусть $O'_F = F \cap O'_E$.

Поскольку кольца O'_E и O'_F полулокальны, известно, что они являются областями главных идеалов.

Теорема 1. Пусть $d > 1, t > 1$ будут заданы рациональные целые числа, и пусть E/F будет нормальным неразветвленным расширением полей алгебраических чисел степени d с группой Галуа Γ .

1) Если $n \geq \phi_E(t)d$, существует конечная абелева Γ -устойчивая подгруппа $G \subset GL_n(O'_E)$ экспоненты t такая, что $E = F(G)$.

2) Если $n \geq \phi_E(t)dh$ и h – показатель группы классов группы F , существует конечная абелева Γ -устойчивая подгруппа $G \subset GL_n(O_E)$ показателя t такая, что $E = F(G)$.

3) Если $n \geq \phi_E(t)d$ и h взаимно просты с n , то G , указанная в 1), сопряжена в $GL_n(F)$ с подгруппой группы $GL_n(O_E)$.

4) Если d нечетно, то G , заданная в 1), сопряжена в $GL_n(F)$ с подгруппой в $GL_n(O_E)$.

Во всех приведенных выше случаях G может быть построена как группа, порожденная матрицами $g^\gamma, \gamma \in \Gamma$ для некоторого $g \in GL_n(E)$.

Результаты, связанные с устойчивостью Галуа конечных групп в ситуациях, подобных нашей, возникают в теории определенных квадратичных решеток, арифметических групп и когомологий Галуа. Точнее, пусть E – вполне вещественное числовое поле, H – алгебраическая подгруппа в $GL_n(\mathbb{C})$, определенная над подполем F в E . Если H определена в следующем смысле: вещественная группа Ли $H(\mathbb{R})$, подгруппа \mathbb{R} -точек, компактна, то подгруппа $H(O_E)$ O_E -точек H является конечной Γ -устойчивой подгруппой, и последнее условие имеет некоторые интересные следствия ([3; 4; 12], см. также [13]). Эти результаты также тесно связаны с некоторыми аспектами целочисленных представлений конечных групп, см. [4; 14; 15]. В нашем контексте мы изучаем, можно ли реализовать заданное поле E , нормальное над F , как поле $E = F(G)$ в обоих случаях $G \subset GL_n(O'_E)$ и $G \subset GL_n(O_E)$, и если это так, каковы возможные порядки n матричных реализаций и структура G .

Теорема 1 дает положительный ответ на вопрос: можно ли реализовать любое нормальное неразветвленное расширение числового поля E/F как $E = F(G)$ для некоторого $G \subset GL_n(O_E)$ при условии $n \geq \phi_E(t)dh$. Мы доказываем, что любое конечное нормальное расширение поля E/F может быть получено как $F(G)/F$, если $n \geq \phi_E(t)d$ для некоторого $G \subset GL_n(E)$. Фактически, мы строим некоторые алгебры Галуа в смысле [16] и устанавливаем нижние границы для их возможных размерностей n . В предложении 2 доказано, что ограничения для заданных целых чисел n, t и d в теореме 1 не могут быть улучшены.

Предложение 2. Пусть E/F – заданное нормальное расширение полей алгебраических чисел с группой Галуа Γ , $[E : F] = d$, и пусть $G \subset GL_n(E)$ – конечная абелева Γ -устойчивая подгруппа экспоненты t такая, что $E = F(G)$ и n – минимально возможное. Тогда $n = d\phi_E(t)$ и G неприводима относительно сопряжения в $GL_n(F)$. Более того, если G имеет минимально возможный порядок, то G – группа типа (t, t, \dots, t) и порядка t^m для некоторого положительного целого числа $t \leq d$.

Условия следующей теоремы рассматривались Л. Море-Байи [7] в более общей ситуации. В общем случае существование глобальных полей с заданной группой Галуа и предписанными локальными свойствами для ветвления является довольно тонким вопросом. Л. Море-Байи доказал существование относительных расширений числовых полей с заданной локальной структурой ветвления над заданным множеством простых делителей и неразветвленных в других местах. Однако наша конструкция в теореме 3 дает вполне вещественные неразветвленные расширения более явным и простым способом.

Теорема 3. Для данной конечной группы Γ существует бесконечно много нормальных неразветвленных расширений вполне вещественных полей E/F , имеющих группу Галуа Γ .

3. Доказательство теоремы 1

Доказательство теоремы 1. Для любого расширения числового поля L/L_1 оба кольца O'_L и O'_{L_1} являются полулокальными, поэтому они являются кольцами главных идеалов, и O'_L имеет базис над O'_{L_1} . Начнем с доказательства 1). Для заданного базиса w_1, w_2, \dots, w_n из O'_E над O'_F мы намерены построить матрицу $g = [g_{ij}]_{i,j} = \sum_{i=1}^d B_i w_i$ и попарно коммутирующие матрицы B_i таким образом, чтобы нормальное замыкание поля $F(g_{11}, g_{12}, \dots, g_{nn})$ над F совпадало с E и, таким образом, группа G , порожденная g^σ , $\sigma \in \Gamma$, является абелевой Γ -устойчивой группой экспоненты t . Во-первых, мы определяем собственные значения, которые должны иметь матрицы B_i , если g имеет заданный набор собственных значений. Собирая заданные собственные значения попарно коммутирующих полупростых матриц и используя регулярное представление, мы строим Γ -устойчивую абелеву группу G для целых параметров, указанных в предложении.

Доказательство 1) используется в доказательстве остальной части теоремы. Фактически, некоторые результаты из теории представлений порядков в полупростых алгебрах (см. [17, § 75]) применяются к порядку $D = O_F[B_1, B_2, \dots, B_d] \subset A$ внутри F -алгебры $A = F[B_1, B_2, \dots, B_d]$. Утверждения 2), 3) и 4) заключаются в том, что конструкция представления G , заданная в 1), может быть реализована над O_E без использования целочисленного базиса O'_E над O'_F (в общем случае кольцо O_E не требует базис над O_F). Этого можно достичь, используя теорему Штейница–Шевалле для модулей над кольцами Дедекинда, которая применяется к порядку D или прямой сумме его копий, а также один результат Шура для 3) и результат, доказанный Фрелихом [18] для 4).

Доказательство 1). В нашем доказательстве мы рассматриваем два разных случая.

Случай 1. Мы предполагаем, что $F(\zeta_t)$ и E линейно не пересекаются над F и $[E : F] = d$. В этом случае $\phi_E(t) = \phi_F(t)$. Пусть $w_1 = 1, w_2, \dots, w_d$ будет базисом O'_E над O'_F , и пусть Γ будет группой Галуа $E(\zeta_t)$ над $F(\zeta_t)$. Пусть g будет полупростой $(d \times d)$ -матрицей с собственными значениями $\zeta_t, 1, \dots, 1$. Используя разложение $g = B_1 + w_2 B_2 + \dots + w_d B_d$, мы можем построить матрицы $B_i, i = 1, 2, \dots, d$, и мы можем доказать, что группа G , порожденная $g^\gamma, \gamma \in \Gamma$, является абелевой Γ -устойчивой группой экспоненты t . Рассмотрим матрицу $W = [w_i^{\sigma_j}]_{i,j}$ для $\{\sigma_1 = 1, \sigma_2, \dots, \sigma_d\} = \Gamma$. Поскольку E/F неразветвлен, $\det W$ является единицей O'_E . Обозначим через W_i матрицу W , i -й столбец которой заменен на d выбранных собственных значений $\zeta_t, 1, \dots, 1$. Мы можем вычислить

$$\lambda_i = \frac{\det W_i}{\det W}$$

и построить матрицы B_i как регулярное представление $B_i = R(\lambda_i)$ $\lambda_i \in O'_E[\zeta_t]$ в базисе w_1, w_2, \dots, w_d расширения кольца $O'_E[\zeta_t] \supset O'_F[\zeta_t]$, которое получается путем присоединения ζ_t к основному кольцу. Пусть α_{ij} будут коэффициентами обратной матрицы $W^{-1} = [\alpha_{ij}]_{i,j}$. Тогда $\alpha_{i1}^{\sigma_j} = \alpha_{ij}$ и $\lambda_i = (\zeta_t - 1)\alpha_{i1}$ для $i \neq 1$, и $\lambda_1 = 1 + (\zeta_t - 1)\alpha_{11}$. Так что $\lambda_i^{\sigma_j} = (\zeta_t - 1)\alpha_{i1}^{\sigma_j} = (\zeta_t - 1)\alpha_{ij}$ для $i \neq 1$, и $\lambda_1^{\sigma_j} = (\zeta_t - 1)\alpha_{11}^{\sigma_j} + 1 = (\zeta_t - 1)\alpha_{1j} + 1$. Поскольку любое линейное отношение

$$k_1(\lambda_1 - 1) + \sum_{i=2}^d k_i \lambda_i = 0, \quad k_i \in F(\zeta_t), \quad i = 1, 2, \dots, d$$

подразумевает линейное отношение

$$k_1(\lambda_1^{\sigma_j} - 1) + \sum_{i=2}^d k_i \lambda_i^{\sigma_j} = 0, \quad k_i \in F(\zeta_t), \quad i = 1, 2, \dots, d$$

для всех $\sigma_j \in \Gamma$, это также означало бы $\det W^{-1} = 0$, что невозможно. Следовательно, $\lambda_1 - 1, \lambda_2, \dots, \lambda_d$ генерируют поле $E(\zeta_t)$ над $F(\zeta_t)$, и поэтому $B_1 - I_d, B_2, \dots, B_d$ генерируют $F(\zeta_t)$ -диапазон $F(\zeta_t)[B_1, \dots, B_d]$ над $F(\zeta_t)$.

Обратите внимание, что B_i можно выразить как линейную комбинацию $g^{\sigma_i}, i = 1, 2, \dots, d$, с коэффициентами в E : $B_i = \sum_{j=1}^d \alpha_{ij} g^{\sigma_j}$. Это можно получить из системы матричных уравнений

$$g^{\sigma_j} = \sum_{i=1}^d w_i^{\sigma_j} B_i, \quad j = 1, 2, \dots, d,$$

если рассматривать B_i как неопределенности. Поскольку G имеет показатель t , $F(\zeta_t)$ является полем расщепления для G , группы, порожденной всеми $g^\sigma, \sigma \in \Gamma$. Следовательно, размерность $E(\zeta_t)$ -пространства $E(\zeta_t)G = E(\zeta_t) \otimes_{F(\zeta_t)} F(\zeta_t)G$ над $E(\zeta_t)$ равна d , и поэтому размерность $F(\zeta_t)$ -пространства $F(\zeta_t)G$ также равна d .

Обозначим через E' образ $E(\zeta_t)$ при регулярном представлении $E(\zeta_t)$ над $F(\zeta_t)$ в базисе w_1, \dots, w_d . Тогда $A = E(\zeta_t)G = E(\zeta_t) \otimes_{F(\zeta_t)} F(\zeta_t)G$, $E(\zeta_t)$ -оболочка G является E' -алгеброй Галуа в смысле [16], т. е. это ассоциативная и коммутативная отделимая E' -алгебра, имеющая нормальный базис. Мы можем выбрать идемпотенты

$$\varepsilon_j = \frac{1}{\zeta_t - 1} (g^{\sigma_j} - I_d), \quad j = 1, 2, \dots, d,$$

как нормальный базис A над E' , так что $\varepsilon_j = \varepsilon_1^{\sigma_j}$.

У нас есть $F(\zeta_t)G = F(\zeta_t)[\langle g^{\sigma_1}, \dots, g^{\sigma_d} \rangle] = F(\zeta_t)[(g - I_d)^{\sigma_1}, \dots, (g - I_d)^{\sigma_d}]$, и $\dim_{F(\zeta_t)} F(\zeta_t)G = d$. Поскольку длина орбиты $M = [m_{ij}] = (g - I_d)$ под действием Γ равна d , мы можем использовать коэффициенты матриц $M^{\sigma_i}, i = 1, 2, \dots, d$, чтобы построить элемент $\theta = \sum_{i,j} k_{ij} m_{ij}, k_{ij} \in F(\zeta_t)$, который генерирует нормальный базис $E(\zeta_t)/F(\zeta_t)$. Следовательно, для любого заданного $\alpha \in E(\zeta_t)$ мы имеем $\alpha = \sum_i k_i \theta^{\sigma_i}$ для некоторого $k_i \in F(\zeta_t)$.

Следовательно, наш выбор собственных значений подразумевает, что $F(\zeta_t)(G) = E(\zeta_t)$.

Теперь мы можем применить регулярное представление R_F матрицы $O'_F[\zeta_t]$ над O'_F к матрицам $M = [m_{ij}]_{i,j}, m_{ij} \in O'_F[\zeta_t]$ следующим образом: $R_F(M) = [R_F(m_{ij})]_{i,j}$. Таким образом, используя R_F для всех компонентов матриц $B_i \in M_n(F(\zeta_t))$, мы можем получить абелеву подгруппу $G \subset GL_{n_1}(E), n_1 = [F(\zeta_t) : F]d$ экспоненты t , которая Γ -устойчива, если мы определим изоморфные группы Галуа расширений E/F и $E(\zeta_t)/F(\zeta_t)$. Мы снова имеем $\dim_F FG = \dim_E EG$, E снова является алгеброй Галуа, и $F(G) = E$. Теперь, используя естественное вложение G в $GL_n(E), n \geq n_1$, мы завершаем доказательство теоремы 1 в случае 1).

Случай 2. В силу случая 1 мы можем рассмотреть случай, когда пересечение $F_0 = E \cap F(\zeta_t) \neq F$. Мы можем использовать регулярное представление $R_{O'_E}$ над O'_F . Пусть $\Gamma_0 = \{\sigma'_1, \sigma'_2, \dots, \sigma'_d\}$ будет множеством некоторых расширений элементов $\Gamma = \{\sigma_1, \sigma_2, \dots, \sigma_d\}$ до $E(\zeta_t)/F$, и пусть $w_1 = 1, w_2, \dots, w_d$ будет базисом O'_E над O'_F . Итак, мы можем использовать наши предыдущие обозначения и применить аналогичный аргумент, как в случае 1 доказательства, для построения $g = \sum_{i=1}^d B_i w_i$ и матриц B_i как регулярных представлений R_0 собственных значений

$$\lambda_i = \frac{\det W_i}{\det W} = \sum_{j=1}^{\Phi_E(t)} \lambda_{ij} \zeta^j, \quad i = 1, 2, \dots, d,$$

следующим образом: мы рассматриваем

$$B_i = R_0(\lambda_i) = \sum_{j=1}^{\Phi_E(t)} R(\lambda_{ij}) \zeta^j,$$

где R – регулярное представление O'_E над O'_F . Мы также имеем $\lambda_1^{\sigma_j} = \alpha_{1j} + 1, \lambda_i^{\sigma_j} = \alpha_{ij}$ для $j = 2, \dots, d$. Теперь, если у нас есть какая-либо линейная связь между строками матрицы $[\alpha_{ij}(\zeta_t^{\sigma_j} - 1)]_{i,j}$, это

будет означать линейную связь между ее столбцами, и поэтому столбцы $W^{-1} = [\alpha_{ij}]$ линейно зависимы, и $\det W^{-1} = 0$, что является противоречием. Итак, снова получаем, что $\lambda_1 - 1, \lambda_2, \dots, \lambda_d$ линейно независимы над F , поэтому $\dim_F F G' = \dim_F F[B_1 - I_d, B_2, \dots, B_d] = \dim_E E G' = d$ для G' , порожденного $g^{\sigma_i}, i = 1, 2, \dots, d$. Как и ранее, мы можем рассмотреть регулярное представление $R_E(B_i)$ для коэффициентов матриц B_i в расширении кольца $O'_E[\zeta_t] \supset O'_E$. Итак, получаем $g_0 = \sum_{i=1}^d R_E(B_i)w_i$, и можем взять группу G , порожденную всеми $g_0^{\sigma_i}, i = 1, 2, \dots, d$. Так как $[E(\zeta_t) : F] = [E(\zeta_t) : E][E : F] = \phi_E(t)d$, то порядок $n = \phi_E(t)d$ совпадает с целым числом, требуемым в формулировке теоремы 1. Таким образом, мы можем построить Γ -стабильную группу G , удовлетворяющую условиям 1) в теореме 1.

Доказательство 2). Рассмотрим O_F -порядок $D = O_F[B_1, B_2, \dots, B_d] \subset A$ в полупростой F -алгебре $A = F[B_1, B_2, \dots, B_d]$, где B_i – это $(n' \times n')$ -матрицы, взятые из 1). Используя нашу конструкцию B_i , мы можем предположить $n' = \phi_E(t)d$. Пусть M будет соответствующим модулем представления в n' -мерном F -векторном пространстве V . Мы утверждаем, что матрицы B_i из 1) могут быть реализованы над O_F путем взятия прямой суммы h копий O_F -модуля M . Мы можем использовать теорему Штейница–Шевалле (см., например, [17]) для M , чтобы получить разложение: $M = v_1 \mathfrak{O}_{\mathfrak{F}} + v_2 \mathfrak{O}_{\mathfrak{F}} + \dots + v_{n'} \mathfrak{O}_{\mathfrak{F}} + v_{n'} \mathfrak{a} O_F$ для некоторых элементов $v_1, v_2, \dots, v_{n'} \in V$ и некоторого дробного идеала \mathfrak{a} множества O_F . Взяв прямую сумму $M_1 = \bigoplus h$ копий M , мы заключаем, что класс Штейница M_1 равен \mathfrak{a}^h , поэтому он тривиален, и M_1 становится свободным O_F -модулем: $M_1 = c_1 O_F + c_2 O_F + \dots + c_{hn'} O_F$ для некоторых элементов $c_1, c_2, \dots, c_{hn'} \in FM_1$. Следовательно, матрицы $B'_1 = \bigoplus^h B_1, B'_2 = \bigoplus^h B_2, \dots, B'_d = \bigoplus^h B_d$ (h копий B_i) $GL_n(F)$ -сопряжены к матрицам, содержащимся в $GL_{hn'}(O_F)$ (мы можем рассмотреть $n = hd\phi_E(t)$ на мгновение, а затем распространить результат на любые $n \geq hd\phi_E(t)$, взяв прямые суммы). Мы можем заключить, что все матрицы $g^\sigma, \sigma \in \Gamma$ также сопряжены в $GL_n(F)$ с матрицами, содержащимися в $GL_n(O_E)$. Поскольку G порождается этими матрицами, мы получаем утверждение 3) для матриц порядка $hd\phi_E(t)$. Для распространения этого результата на произвольные $n \geq hd\phi_E(t)$ мы можем зафиксировать положительные целые числа k и r с $n = khd\phi_E(t) + r, r < hd\phi_E(t)$ и взять прямую сумму k копий построенной реализации G и r копий единичных представлений. Это завершает доказательство 2).

Доказательство 3). Это следует из утверждения (75.5) в [17], примененного к порядку $D = O_F[B_1, B_2, \dots, B_d] \subset A$ в F -алгебре $A = F[B_1, B_2, \dots, B_d]$. Так как все матрицы $B_i, i = 1, 2, \dots, d$, сопряжены в $GL_n(F)$ с матрицами, содержащимися в $GL_n(O_F)$ (здесь можно рассмотреть $n = d\phi_E(t)$), то мы заключаем, что все матрицы $g^\sigma, \sigma \in \Gamma$, также сопряжены в $GL_n(F)$ с матрицами из $GL_n(O_E)$. Так как G порождается этими матрицами, то получаем утверждение 3).

Доказательство 4). По [18], теореме 4.3, в любом нормальном неразветвленном расширении числовых полей нечетной степени кольцо целых чисел имеет свободный базис. В нашем случае $O_E = w_1 O_F + w_2 O_F + \dots + w_d O_F$ для некоторых w_1, w_2, \dots, w_d . Поэтому матрицы $B_i, i = 1, 2, \dots, d$, сопряжены в $GL_n(F)$ матрицам из $GL_n(O_F)$, и наш аргумент 1) может быть непосредственно применен к кольцам O_E и O_F вместо O'_E и O'_F . Это означает, что G сопряжена в $GL_n(F)$ подгруппе $GL_n(O_E)$, как и утверждалось.

Это завершает доказательство теоремы 1.

4. Доказательство предложения 2

Доказательство предложения 2. Мы можем использовать доказательство теоремы 1.

Пусть $G \subset GL_n(E)$ – группа, заданная в формулировке предложения 2, и пусть n – минимально возможное. Тогда мы имеем следующее разложение E -оболочки $A = EG$:

$$A = \varepsilon_1 A + \varepsilon_2 A + \dots + \varepsilon_k A$$

для некоторых примитивных идемпотентов $\varepsilon_1, \dots, \varepsilon_k$ группы A . ε_i сопряжены под действием группы Галуа $\Gamma = \{\sigma_1, \dots, \sigma_d\}$. Ибо если сумма $\varepsilon_i^{\sigma_j}, j = 1, 2, \dots, d$, не равна I_n , то $I_n = e_1 + e_2$ для $e_1 = \varepsilon_1^{\sigma_1} + \dots + \varepsilon_1^{\sigma_d}$ и $e_2 = I_n - e_1$, и e_1, e_2 фиксируются Γ , так что e_1, e_2 сопряжены в $GL_n(F)$ до диагональной формы. Поскольку любой из 2 компонентов $e_i G$ имеет ранг меньше n , существует матричная группа, удовлетворяющая условиям предложения 2 степени меньше n .

Следовательно, $\varepsilon_i = \varepsilon_1^{\sigma_i}, k = d$ и идемпотенты $\varepsilon_1, \dots, \varepsilon_d$ образуют нормальный базис A . Но ранг матрицы ε_i не меньше, чем $\phi_E(t)$. Действительно, $\varepsilon_i G$ содержит элемент $\varepsilon_i g$, для некоторого

$g \in G$ порядка t такого, что $(\varepsilon_i g)^t = \varepsilon_i$, но $(\varepsilon_i g)^k \neq \varepsilon_i$ для $k < t$. Мы можем найти $g \in G$ следующим образом. Так как $I_n = \varepsilon_1 + \dots + \varepsilon_k$ для любого $h \in G$ порядка t существует ε_j такой, что $(\varepsilon_j h)^t = \varepsilon_j$, но $(\varepsilon_j h)^k \neq \varepsilon_j$ для $k < t$, и то же свойство выполняется для $\varepsilon_j h$ с любым $\sigma \in \Gamma$. Тогда, используя свойство нормального базиса $\varepsilon_k = \varepsilon_1^{\sigma_k}$, мы можем взять $g = h^{\sigma_j^{-1} \sigma_i}$.

Таким образом, неприводимая компонента $\varepsilon_i G$ определяет точное неприводимое представление циклической группы, порожденной g . Но если $T : C \rightarrow GL_r(E)$ является точным неприводимым представлением циклической группы C , порожденной элементом g порядка t , то его степень r равна $\phi_E(t)$. Из этого следует, что ранг матриц ε_i равен $\phi_E(t)$. Таким образом, размерность A над E равна $\phi_E(t)d$.

Если G порождается g^γ , $\gamma \in \Gamma$, и его порядок минимален, Γ -устойчивость подразумевает, что g имеет d сопряженных относительно Γ -действия, и поэтому G является абелевой группой экспоненты t и порядка t^m для некоторого положительного целого числа $m \leq d$. Это завершает доказательство предложения 2.

5. Доказательство теоремы 3

Доказательство теоремы 3. Во-первых, будет построено вполне вещественное расширение L/\mathbb{Q} степени n . Для этой цели мы можем зафиксировать простые числа q_1, q_2, q_3 таким образом, что q_1 и a являются взаимно простыми, и выбрать многочлен $H(x) = (x - a_1 q_1)(x - a_2 q_2) \dots (x - a_n q_n) + a q_1$, группа которого имеет транспозицию и один или 2 множителя нечетной степени по модулю q_2 , $(n - 1)$ -цикл по модулю q_3 для целых чисел a_i , достаточно больших по сравнению с $|q_1 a|$ и малых по сравнению с $|a_i - a_j|$, $i \neq j$, таких, что все корни $H(x)$ являются действительными. Поле расщепления $H(x)$ вполне действительно, а его группа Галуа является симметрической группой S_n . Следовательно, L имеет подполе L степени n над \mathbb{Q} .

Зафиксируем множество R всех простых чисел, разветвленных в L/\mathbb{Q} .

Рассмотрим следующие условия:

- 1) $F(x) = (x - b_1 p_1)(x - b_2 p_1) \dots (x - b_n p_1) + p_1 b$;
- 2) $b_1, b_2, \dots, b_n \in \mathbb{Z}$ различны, и p_1 не делит b ; $2bn < (\prod_{j=1(j \neq i)}^n |b_i - b_j|) p_1^{n-2}$ для $i = 1, \dots, n$;
- 3) $F(x)$ имеет транспозицию и 1 или 2 множителя нечетной степени по модулю p_2 ;
- 4) $F(x)$ имеет $(n - 1)$ -цикл по модулю p_3 ;
- 5) p_1, p_2, p_3 – простые числа, не содержащиеся в $R \cup q$, а $q \notin R$ – простое число, сравнимое с 1 по модулю n .

Условия 1) и 2) гарантируют неприводимость $F(x)$, поскольку $F(x)$ – многочлен Эйзенштейна, а также все корни $F(x)$ действительны. Действительно, коэффициент при x^{n-1} в $x^n F(1/x + b_i p_1)$ равен $\prod_{j=1(j \neq i)}^n (b_j p_1 - b_i p_1)$. Следовательно, для корней x_1, x_2, \dots, x_n уравнения $F(x)$ справедливо следующее равенство:

$$\left| \frac{1}{x_1 - b_i p_1} + \frac{1}{x_2 - b_i p_1} + \dots + \frac{1}{x_n - b_i p_1} \right| = \left| \frac{\prod_{j=1(j \neq i)}^n |b_j p_1 - b_i p_1|}{b p_1} \right|,$$

и поэтому

$$|x_{k_i} - b_i p_1| \leq \frac{nb p_1}{\prod_{j=1(j \neq i)}^n |b_j p_1 - b_i p_1|}$$

при условии $|x_{k_i} - b_i p_1| \leq |x_j - b_i p_1|$ для всех $j \neq k_i$.

Теперь, если $2nb p_1 < \prod_{j=1(j \neq i)}^n |b_j p_1 - b_i p_1|$, то $|x_{k_i} - b_j p_1| \leq \frac{1}{2}$, и все корни x_{k_i} содержатся в окружностях радиуса $\frac{1}{2}$ с различными центрами $b_j p_1$, поэтому среди x_{k_i} нет комплексно сопряженных.

Условия 3) и 4) подразумевают совпадение группы Галуа $F(x)$ и симметрической группы S_n .

Из теоремы плотности Фробениуса ([19], см. также [20], теорема 42) следует, что заданный многочлен имеет ту же факторизацию, соответствующую перестановке заданного типа цикла по модулю бесконечного числа простых чисел. Следовательно, существует целое число M , не делящееся на p_1, p_2, p_3, q и простые числа из R , такое, что сравнение

- 6) $f(x) \equiv F(x) \pmod{M}$ влечет, что группа Галуа $f(x)$ равна S_n .

Пусть K будет полем разложения $f(x)$. Тогда q -круговое поле $\mathbb{Q}(\zeta_q)$ имеет подполе степени n над \mathbb{Q} , которое может быть определено как поле разложения целочисленного многочлена $k(x)$. По лемме Краснера существует $t_1 \in \mathbb{N}$ такое, что сравнения

7) $f(x) \equiv H(x) \pmod{p^{t_1}}$ для всех $p \in R$ подразумевают совпадение локализаций: $L\mathbb{Q}_p = K\mathbb{Q}_p$ для $p \in R$. Если максимальное абелево подполе K_{ab} поля K не является \mathbb{Q} , то $K_{ab} = \mathbb{Q}(\sqrt{r})$ для некоторого $r \in \mathbb{Z}$, и для достаточно большого целого числа t_2 сравнение

8) $f(x) \equiv k(x) \pmod{q^{t_2}}$ подразумевает $K \cap L = \mathbb{Q}$, рассматривая разветвление в q . Но составной KL не разветвлен над K , потому что $L\mathbb{Q}_p = K\mathbb{Q}_p$ для всех простых чисел p , разветвленных в L .

Мы можем найти многочлен $f(x)$, удовлетворяющий условиям 1)–8), так что его корни действительны, согласно 2). Действительно, используя теорему о слабой аппроксимации (или китайскую теорему о наложении), мы можем удовлетворить 1) и 3)–8), и в факторизации $f(x) = (x - c_1)(x - c_2) \dots (x - c_n) + c_0$ можно увеличить, добавив несколько кратных модулей сравнений 3)–8), чтобы сделать $c_i, i = 1, \dots, n$, достаточно большим по сравнению с c_0 и малым по сравнению с $|c_i - c_j|, i \neq j (i, j \neq 0)$. Следовательно, поля $E = LK$ и K вполне вещественны, а расширение E/K неразветвлено, нормально, и его группа Галуа равна S_n . По теории Галуа, для заданной конечной группы $\Gamma \subset S_n$ (для подходящего n) существует нормальное подрасширение E/F , где $F = E^\Gamma$ является подполем Γ -неподвижных элементов E , которое также неразветвлено и имеет Γ в качестве группы Галуа. Это замечание завершает доказательство теоремы 3.

Автор благодарен рецензенту за полезные замечания, которые способствовали улучшению работы.

Работа поддержана Институтом математики НАН Беларуси в рамках задания 1.1.01 государственной программы научных исследований «Конвергенция–2025».

Литература

1. Harbater D. Galois groups with prescribed ramification // Contemporary Math. 1994. Vol. 174. P. 35–60.
2. Ozaki Manabu. Construction of maximal unramified p -extensions with prescribed Galois groups // Invent math. 2011. Vol. 183. P. 649–680.
3. Bartels H.-J. Zur Galoiskohomologie definiter arithmetischer Gruppen // J. reine angew. Math. 1978. Vol. 298. P. 89–97.
4. Малинин Д. А. Целочисленные представления конечных групп, устойчивые при действии группы Галуа // Алгебра и анализ. 2000. Т. 12, № 3. С. 106–145.
5. Малинин Д. А. Целочисленные представления конечных групп с действием Галуа // Докл. РАН. 1996. Т. 349, № 3. С. 303–305.
6. Ritter J., Weiss A. Galois action on integral representations // J. London Math. Soc. (2). 1992. Vol. 46. P. 411–431.
7. Moret-Bailly L. Extensions de corps globaux a ramification et groupe de Galois donnees // C. R. Acad. Sci. Paris, Serie I. 1990. Vol. 311. P. 273–276.
8. Maire Ch. On infinite unramified extensions // Pacific J. Math. 2000. Vol. 192, N 1. P. 135–142.
9. Kondo T. Algebraic number fields with the discriminant equal to that of quadratic number field // J. Math. Soc. Japan. 1995. Vol. 47, N 1. P. 31–36.
10. Yamamura K. Maximal unramified extensions of imaginary quadratic fields of small conductors // Journal de Theorie des Nombres de Bordeaux. 1997. Vol. 9. P. 405–448.
11. Pohst M. Berechnung kleiner Diskriminanten total reeller algebraischer Zahlkörper // J. Reine angew. Math. 1975. Vol. 278/279. P. 278–300.
12. Bartels H.-J., Kitaoka Y. Endliche arithmetische Untergruppen der GL_n // J. reine angew. Math. 1980. Vol. 313. P. 151–156.
13. Rohlfes J. Arithmetische definierte Gruppen mit Galois-operation // Invent. Math. 1978. Vol. 48. P. 185–205.
14. Малинин Д. А. О целочисленных представлениях p -групп над локальными полями // Докл. АН СССР. 1989. Т. 309, № 5. С. 1060–1063.
15. Малинин Д. А. Целочисленные представления p -групп заданного класса нильпотентности над локальными полями // Алгебра и анализ. 1998. Т. 10, № 1. С. 58–67.

16. Ишханов В. В., Лурье Б. Б., Фаддеев Д. К. Задача погружения в теории Галуа. М.: Наука, Гл. ред. физ.-мат. лит., 1990.
17. Curtis C. W., Reiner I. Representation theory of finite groups and associative algebras. New York: Interscience, 1962.
18. Fröhlich A. Discriminants of algebraic number fields // Math Zeitschr. 1960. Vol. 74. P. 18–28.
19. Frobenius G. Über Beziehungen zwischen den Primidealen eines algebraischen Zahlkörpers und den Substitutionen seiner Gruppe. Sitzber: Preussen Akad. Wiss., 1896. S. 689–705.
20. Чеботарев Н. Г. Основы теории Галуа. Ч. 2. ОГИЗ: Ленинград–Москва, 1937.

References

1. Harbater D. Galois groups with prescribed ramification. *Contemporary Math.*, 1994, vol. 174, pp. 35–60.
2. Ozaki Manabu. Construction of maximal unramified p -extensions with prescribed Galois groups. *Invent math.*, 2011, vol. 183, pp. 649–680.
3. Bartels H.-J. Zur Galoiskohomologie definiter arithmetischer Gruppen. *J. reine angew. Math.*, 1978, vol. 298, pp. 89–97.
4. Malinin D. A. Galois stability for integral representations of finite groups *St. Petersburg Math. J.*, 2000, vol. 12, no. 3, pp. 106–145.
5. Malinin D. A. Integral representations of finite groups with Galois action. *Dokl. Russ. Akad. Nauk*, 1996, vol. 349, no. 3, pp. 303–305.
6. Ritter J., Weiss A. Galois action on integral representations. *J. London Math. Soc. (2)*, 1992, vol. 46, pp. 411–431.
7. Moret-Bailly L. Extensions de corps globaux a ramification et groupe de Galois donnees. *C. R. Acad. Sci. Paris, Serie I*, 1990, vol. 311, pp. 273–276.
8. Maire Ch. On infinite unramified extensions. *Pacific J. Math.*, 2000, vol. 192, no. 1, pp. 135–142.
9. Kondo T. Algebraic number fields with the discriminant equal to that of quadratic number field. *J. Math. Soc. Japan*, 1995, vol. 47, no. 1, pp. 31–36.
10. Yamamura K. Maximal unramified extensions of imaginary quadratic fields of small conductors. *Journal de Theorie des Nombres de Bordeaux*, 1997, vol. 9, pp. 405–448.
11. Pohst M. Berechnung kleiner Diskriminanten total reeller algebraischer Zahlkörper. *J. Reine angew. Math.*, 1975, vol. 278/279, pp. 278–300.
12. Bartels H.-J., Kitaoka Y. Endliche arithmetische Untergruppen der GL_n . *J. reine angew. Math.*, 1980, vol. 313, pp. 151–156.
13. Rohlf J. Arithmetische definierte Gruppen mit Galois-operation. *Invent. Math.*, 1978, vol. 48, pp. 185–205.
14. Malinin D. A. Integral representations of p -groups over local fields. *Sov. Math. Dokl.*, 1990, vol. 40, no. 3, pp. 619–622.
15. Malinin D. A. Integral representations over local fields for p -groups of a given class of nilpotency. *St. Petersburg Math. J.*, 1998, vol. 10, no. 1, pp. 58–67.
16. Ishkhanov V. V., Lur'e B. B., Faddeev D. K. *The embedding problem in Galois theory*. Moscow, Nauka, 1990.
17. Curtis C. W., Reiner I. *Representation theory of finite groups and associative algebras*. New York, Interscience, 1962.
18. Fröhlich A. Discriminants of algebraic number fields. *Math Zeitschr.*, 1960, vol. 74, pp. 18–28.
19. Frobenius G. *Über Beziehungen zwischen den Primidealen eines algebraischen Zahlkörpers und den Substitutionen seiner Gruppe*. Sitzber, Preussen Akad. Wiss., 1896, s. 689–705.
20. Chebotarev N. G. *Foundations of Galois theory, Part II*. Noordhoff, 1950 (in German).

UDC 512.542

TO THE THEOREM OF K. DOERK

V. I. Murashka, A. F. Vasil'ev

Francisk Skorina Gomel State University, Gomel, Belarus
e-mail: mvimath@yandex.ru, formation56@mail.ru

Received: 05.05.2025

Revised: 21.05.2025

Accepted: 23.05.2025

Keywords: finite group; the generalized Fitting subgroup; the generalized Fitting height; the non- p -soluble length; hereditary Plotkin radical; σ -nilpotent group.

Abstract. For a finite group G and its maximal subgroup M we proved that the generalized Fitting height of G minus the generalized Fitting height of M is not greater than 2 and the non- p -soluble length of G minus the non- p -soluble length of M is not greater than 1. We constructed a hereditary saturated formation \mathfrak{F} such that $\{n_\sigma(G, \mathfrak{F}) - n_\sigma(M, \mathfrak{F}) \mid G \text{ is finite } \sigma\text{-soluble and } M \text{ is a maximal subgroup of } G\} = \mathbb{N} \cup \{0\}$ where $n_\sigma(G, \mathfrak{F})$ denotes the σ -nilpotent length of the \mathfrak{F} -residual of G . This construction shows the results about the generalized lengths of maximal subgroups published in Math. Nachr. (1994) and Mathematics (2020) are not correct.

К ТЕОРЕМЕ К. ДЁРКА

В. И. Мурашко, А. Ф. Васильев

Гомельский государственный университет им. Ф. Скорины, Гомель, Беларусь
e-mail: mvimath@yandex.ru, formation56@mail.ru

Поступила: 05.05.2025

Исправлена: 21.05.2025

Принята: 23.05.2025

Ключевые слова: конечная группа; обобщенная подгруппа Фиттинга; обобщенная высота Фиттинга; не- p -разрешимая длина; наследственный радикал Плоткина; σ -нильпотентная группа.

Аннотация. Для конечной группы G и ее максимальной подгруппы M мы доказали, что обобщенная высота Фиттинга группы G минус обобщенная высота Фиттинга подгруппы M не превосходит 2, а не- p -разрешимая длина группы G минус не- p -разрешимая длина подгруппы M не превосходит 1. Мы построили наследственную насыщенную формацию \mathfrak{F} так, что $\{n_\sigma(G, \mathfrak{F}) - n_\sigma(M, \mathfrak{F}) \mid G \text{ конечна } \sigma\text{-разрешима и } M \text{ является максимальной подгруппой группы } G\} = \mathbb{N} \cup \{0\}$, где $n_\sigma(G, \mathfrak{F})$ обозначает σ -нильпотентную длину \mathfrak{F} -корадикала группы G . Эта конструкция показывает, что результаты об обобщенных длинах максимальных подгрупп, опубликованные в Math. Nachr. (1994) и Mathematics (2020), являются некорректными.

1. Introduction and the Main results

All groups considered here are finite. One way to study the structure of finite groups is to study their given normal series. An important parameter of such series is their length. For example the derived length, the nilpotent length and the p -length encode information about the structure of a group. Note that the series defining the nilpotent length were used for computations in soluble (polycyclic) groups [1]. One of the main disadvantages of the above mentioned lengths is that they are not defined for all groups. Khukhro and Shumyatsky [2; 3] introduced the following lengths associated with every group.

Definition 1.1 (Khukhro, Shumyatsky). (1) The generalized Fitting height $h^*(G)$ of a finite group G is the least number h such that $F_h^*(G) = G$, where $F_{(0)}^*(G) = 1$, and $F_{(i+1)}^*(G)$ is the inverse image of the generalized Fitting subgroup $F^*(G/F_{(i)}^*(G))$.

(2) Let p be a prime, $1 = G_0 \leq G_1 \leq \dots \leq G_{2h+1} = G$ be the shortest normal series in which for i odd the factor G_{i+1}/G_i is p -soluble (possibly trivial), and for i even the factor G_{i+1}/G_i is a (non-empty) direct product of nonabelian simple groups. Then $h = \lambda_p(G)$ is called the non- p -soluble length of a group G .

(3) $\lambda_2(G) = \lambda(G)$ is the nonsoluble length of a group G .

For the properties and applications of these lengths see [2–7].

Note that if G is a soluble group, then $h^*(G) = h(G)$ is the nilpotent length of G . K. Doerk [8, Satz 1] proved that the difference of the nilpotent lengths of a soluble group and its maximal subgroup can be only 0, 1 or 2. The analogues of this result were obtained for the π -length of a π -soluble group [9] and the σ -nilpotent length of a σ -soluble group [10]. From [4, Theorem 5.6] it follows that for a group G

and its subgroup H the differences $h^*(G) - h^*(H)$ and $\lambda_p(G) - \lambda_p(H)$ are not bounded from below by a constant. Here we prove

Theorem 1.2. *Let M be a maximal subgroup of a group G and p be a prime. Then*

$$h^*(G) - h^*(M) \leq 2, \lambda(G) - \lambda(M) \leq 1 \text{ and } \lambda_p(G) - \lambda_p(M) \leq 1.$$

This theorem is the consequence of two general results obtained via the functorial method. According to Plotkin [11] a functorial is a function γ which assigns to each group G its subgroup $\gamma(G)$ satisfying $f(\gamma(G)) = \gamma(f(G))$ for any isomorphism $f : G \rightarrow G^*$. From [12, p. 27 and Proposition 3.2.3] follows the following definition:

Definition 1.3. A functorial γ is called a hereditary Plotkin radical if it satisfies:

(P1) $f(\gamma(G)) \subseteq \gamma(f(G))$ for every epimorphism $f : G \rightarrow G^*$.

(P2) $\gamma(G) \cap N = \gamma(N)$ for every $N \trianglelefteq G$.

Note that the \mathfrak{F} -radical for a Fitting formation is a hereditary Plotkin radical. Recall [11] that for functorials γ_1 and γ_2 the upper product $\gamma_2 \star \gamma_1$ is defined by $(\gamma_2 \star \gamma_1)(G)/\gamma_2(G) = \gamma_1(G/\gamma_2(G))$. This operation is an associative one. With every functorial one can associate the following length.

Definition 1.4 [7, Definition 2.4]. Let γ be a functorial. Then the γ -series of G is defined starting from $\gamma_{(0)}(G) = 1$, and then by induction $\gamma_{(i+1)}(G) = (\gamma_{(i)} \star \gamma)(G)$. The least number h such that $\gamma_{(h)}(G) = G$ is defined to be the γ -length $h_\gamma(G)$ of G . If there is no such number, then $h_\gamma(G) = \infty$.

If $\gamma = F$ assigns to every group its Fitting subgroup, then the γ -length is just the nilpotent length (height) and for a group G is denoted by $l(G)$ or $h(G)$. For $\gamma = F^*$ we get the generalized Fitting height. One of our main results is

Theorem 1.5. *Let γ be a hereditary Plotkin radical which satisfies $F^*(G) \subseteq \gamma(G)$ for any group G with $h_\gamma(G) < \infty$. If M is a maximal subgroup of a group G and $h_\gamma(G), h_\gamma(M) < \infty$, then $h_\gamma(G) - h_\gamma(M) \leq 2$.*

From [13, Theorem 3.1 and Corollary 3.4(A)] it follows that if γ is a hereditary Plotkin radical iff $\mathfrak{F} = (G \mid \gamma(G) = G)$ is a Q -closed Fitting class and γ is the \mathfrak{F} -radical. The example of a Q -closed Fitting class of soluble groups which is not a formation follows from [14, IX, Examples 2.21(b)]. Theorem 1.5 gives the analogues of Doerk's result for any Q -closed Fitting class of soluble groups.

Corollary 1.6. *Let \mathfrak{F} be a Q -closed Fitting class of soluble groups, γ assigns to every group its \mathfrak{F} -radical and $\pi = \pi(\mathfrak{F})$. If M is a maximal subgroup of a soluble π -group G , then $h_\gamma(G) - h_\gamma(M) \leq 2$.*

Note that $h(H) \leq h(G)$ holds for any soluble group G and its subgroup H . Hence from Theorem 1.5 for $\gamma = F$ follows

Corollary 1.7 [8]. *Let M be a maximal subgroup of a soluble group G . Then $h(G) - h(M) \in \{0, 1, 2\}$.*

Let $\sigma = \{\sigma_i \mid i \in I\}$ be a partition of the set of all primes \mathbb{P} . Recall [15] that a group G is called σ -soluble if for every its chief factor H/K there exists $\sigma_i \in \sigma$ such that H/K is a σ_i -group (i. e. all prime divisors of $|H/K|$ belong to σ_i); σ -nilpotent if it has a normal Hall σ_i -subgroup for every $\sigma_i \in \sigma$. The greatest normal σ -nilpotent subgroup of G is denoted by $F_\sigma(G)$. The γ -length of G for $\gamma = F_\sigma$ is denoted by $l_\sigma(G)$. Note that a group is σ -soluble iff $l_\sigma(G) < \infty$.

Corollary 1.8 [10]. *Let σ be a partition of \mathbb{P} and M be a maximal subgroup of a σ -soluble group G . Then $l_\sigma(G) - l_\sigma(M) \in \{0, 1, 2\}$.*

According to [12, p. 27 and Proposition 3.2.3] a hereditary Kurosh–Amitsur radical can be defined in the following way:

Definition 1.9. A hereditary Plotkin radical γ is called a hereditary Kurosh–Amitsur radical if it satisfies (P3): $\gamma(G/\gamma(G)) \simeq 1$ for every group G .

For a class of simple groups \mathfrak{J} the greatest normal subgroup $O_{\mathfrak{J}}(G)$ of G all whose composition factors belong to \mathfrak{J} is the example of hereditary Kurosh–Amitsur radical. Kurosh–Amitsur radicals (of groups) were studied in [16].

Theorem 1.10. *Let ρ be a hereditary Kurosh–Amitsur radical which contains the soluble radical in every group and $\gamma = \rho \star F^* \star \rho$. If M is a maximal subgroup of a group G and $h_\gamma(G), h_\gamma(M) < \infty$, then $h_\gamma(G) - h_\gamma(M) \leq 1$.*

Recall that for a formation \mathfrak{F} and a group G the \mathfrak{F} -residual of G is denoted by $G^{\mathfrak{F}}$. The nilpotent and σ -nilpotent lengths of the \mathfrak{F} -residual are denoted by $n_{\mathfrak{F}}(G)$ [17] and $n_\sigma(G, \mathfrak{F})$ [10] respectively. Let \mathfrak{F} be a hereditary saturated formation. In [17] it was claimed that $n_{\mathfrak{F}}(G) - n_{\mathfrak{F}}(M) \in \{0, 1, 2\}$ for any soluble group G and its maximal subgroup M . For a partition σ of \mathbb{P} in the paper [10] it was proved

that $n_\sigma(G, \mathfrak{F}) - n_\sigma(M, \mathfrak{F}) \in \{0, 1, 2\}$ for any σ -soluble group G and its maximal subgroup M . Our next result shows that the two above mentioned facts are **wrong**.

Theorem 1.11. *Let σ be a partition of \mathbb{P} with $|\sigma| > 1$. Then there exists a hereditary saturated formation $\mathfrak{F} = \mathfrak{F}(\sigma)$ of soluble groups such that*

$$\{n_\sigma(G, \mathfrak{F}) - n_\sigma(M, \mathfrak{F}) \mid G \text{ is } \sigma\text{-soluble and } M \text{ is a maximal subgroup of } G\} = \mathbb{N} \cup \{0\}.$$

In particular, there exists a hereditary saturated formation \mathfrak{F} such that

$$\{n_{\mathfrak{F}}(G) - n_{\mathfrak{F}}(M) \mid G \text{ is soluble and } M \text{ is a maximal subgroup of } G\} = \mathbb{N} \cup \{0\}.$$

2. Preliminaries

All unexplained notations and terminologies are standard. The reader is referred to [14] if necessary. Recall that \mathbb{N} and \mathbb{P} denote the sets of all natural and prime numbers respectively.

Recall that a *class of groups* is a collection \mathfrak{F} of groups with the property that if $G \in \mathfrak{F}$ and if $H \simeq G$, then $H \in \mathfrak{F}$; a *formation* is a class of groups \mathfrak{F} which is closed under taking epimorphic images (i. e. from $G \in \mathfrak{F}$ and $N \trianglelefteq G$ it follows that $G/N \in \mathfrak{F}$) and subdirect products (i. e. from $G/N_1 \in \mathfrak{F}$ and $G/N_2 \in \mathfrak{F}$ it follows that $G/(N_1 \cap N_2) \in \mathfrak{F}$). A formation \mathfrak{F} is called: hereditary if $H \in \mathfrak{F}$ whenever $H \leq G \in \mathfrak{F}$; saturated if $G \in \mathfrak{F}$ whenever $G/\Phi(G) \in \mathfrak{F}$. The smallest normal subgroup of G with quotient in \mathfrak{F} is called the \mathfrak{F} -residual of G . A group G is called p -closed if it has a normal Sylow p -subgroup. The class of all soluble p -closed groups is the example of a hereditary saturated formation.

From [7, Proposition 2.3 and Lemma 2.6] the next result follows.

Lemma 2.1. *If γ is a hereditary Plotkin radical, then $\gamma_{(n)}$ is a hereditary Plotkin radical for any $n \in \mathbb{N}$ and $\max\{h_\gamma(N), h_\gamma(G/N)\} \leq h_\gamma(G) \leq h_\gamma(G/N) + h_\gamma(N)$ for any $N \trianglelefteq G$.*

One of the characteristic properties of Kurosh–Amitsur radicals is the following

Lemma 2.2. *Let γ be a hereditary Kurosh–Amitsur radical. Then $\gamma(G/N) = \gamma(G)/N$ for any $N \trianglelefteq G$ with $N \subseteq \gamma(G)$.*

Proof. Assume that $\gamma(G)/N < \gamma(G/N) = H/N$. Then $1 \not\leq H/\gamma(G) = \gamma(H/\gamma(G)) \subseteq \gamma(G/\gamma(G)) \simeq \simeq 1$, a contradiction. \square

3. Proves of the Main Results

3.1. Proof of Theorem 1.5

Assume the contrary. Let a group G be a minimal order counterexample. Hence G has a maximal subgroup M with $h_\gamma(G) - h_\gamma(M) > 2$. It is clear that $h_\gamma(G) \geq 3$. Let $M_i = \gamma_{(i)}(M)$ and $G_i = \gamma_{(i)}(G)$. If $MG_1 = G$, then $h_\gamma(G) - 1 = h_\gamma(G/G_1) = h_\gamma(MG_1/G_1) = h_\gamma(M/(M \cap G_1)) \leq h_\gamma(M)$ by Lemma 2.1. It means that $h_\gamma(G) - h_\gamma(M) \leq 1$, a contradiction. Therefore $G_1 \subseteq M$.

If $M_0 = M$, then G is a cyclic group of prime order and $h_\gamma(G) - h_\gamma(M) = 1$, a contradiction. So $M_0 \neq M$. Suppose that $M_i \subseteq G_{i+1} \subseteq M$ and $M_i \neq M$ for some $i \geq 0$. At least it is true for $i = 0$. Let prove that $M_{i+1} \subseteq G_{i+2} \subseteq M$ and $M_{i+1} \neq M$.

Note that $h_\gamma(G) > i + 1$ and $h_\gamma(M) > i$. From $M_i \subseteq G_{i+1} \subseteq G_{i+2}$ it follows that $M_i \subseteq M \cap G_{i+2}$. If $G_{i+2} \not\subseteq M$, then by Definition 1.4 and Lemma 2.1

$$h_\gamma(G) - (i + 2) = h_\gamma(G/G_{i+2}) = h_\gamma(MG_{i+2}/G_{i+2}) = h_\gamma(M/(M \cap G_{i+2})) \leq h_\gamma(M) - i.$$

Therefore $h_\gamma(G) - h_\gamma(M) \leq (i + 2) - i = 2$, a contradiction. Thus $G_{i+2} \subseteq M$.

Now $G_{i+2}, M_{i+1} \trianglelefteq M$. Let $I = G_{i+2} \cap M_{i+1} \trianglelefteq M$. From $I \trianglelefteq M_{i+1}$ it follows that $\gamma_{(i+1)}(I) = I$ by Lemma 2.1. From the other hand $I \trianglelefteq G_{i+2}$ and $G_{i+1} = \gamma_{(i+1)}(G) = \gamma_{(i+1)}(G) \cap G_{i+2} = \gamma_{(i+1)}(G_{i+2})$ by (P2) and Lemma 2.1. Thus $I \leq G_{i+1}$ by (P2).

Let $F/G_{i+1} = F^*(G/G_{i+1})$. From $h_\gamma(G) < \infty$ it follows that $h_\gamma(G/G_{i+1}) < \infty$. Therefore $F/G_{i+1} \subseteq \subseteq \gamma(G/G_{i+1}) = G_{i+2}/G_{i+1}$. Hence $F \leq G_{i+2}$. Now

$$(M_{i+1}G_{i+1}/G_{i+1}) \cap G_{i+2}/G_{i+1} = (M_{i+1} \cap G_{i+2})G_{i+1}/G_{i+1} = G_{i+1}/G_{i+1} \simeq 1.$$

From [18, X, Theorem 13.12] it follows that

$$M_{i+1}G_{i+1}/G_{i+1} \subseteq C_{G/G_{i+1}}(G_{i+2}/G_{i+1}) \subseteq C_{G/G_{i+1}}(F/G_{i+1}) \subseteq F/G_{i+1} \subseteq G_{i+2}/G_{i+1}.$$

Thus $M_{i+1} \subseteq G_{i+2}$. If $M_{i+1} = M$, then $G_{i+2} = M < G$. By our assumption $M_i \neq M$. Hence $h_\gamma(M) = i + 1$ and $h_\gamma(G) = i + 3$. Therefore $h_\gamma(G) - h_\gamma(M) = 2$, a contradiction. Thus $M_{i+1} \neq M$.

It means $M_i \subseteq G_{i+1} \subseteq M$ and $M_i \neq M$ for every natural i . Thus $h_\gamma(G) = \infty$, the contradiction.

3.2. Proof of Corollary 1.6

Since \mathfrak{F} is a Q -closed Fitting class of soluble groups, we see that γ is a hereditary Plotkin radical by [13, Theorem 3.1 and Corollary 3.4(A)] and \mathfrak{F} contains a group of order p for any $p \in \pi$. It means that $h_\gamma(G) < \infty$ iff G is a soluble π -group and \mathfrak{F} contains all nilpotent π -groups by [14, IX, Theorem 1.9]. So $F^*(G) = F(G) \subseteq \gamma(G)$ for any group G with $h_\gamma(G) < \infty$. Thus Corollary 1.6 directly follows from Theorem 1.5.

3.3. Proof of Theorem 1.10

Note that γ is a hereditary Plotkin radical by [7, Proposition 2.3].

Assume the contrary. Let a group G be a minimal order counterexample. Hence G has a maximal subgroup M with $h_\gamma(G) - h_\gamma(M) > 1$. It is clear that $h_\gamma(G) > 1$.

If $M\gamma(G) = G$, then by Definition 1.4 and Lemma 2.1

$$h_\gamma(G) - 1 = h_\gamma(G/\gamma(G)) = h_\gamma(M/(M \cap \gamma(G))) \leq h_\gamma(M).$$

Therefore $h_\gamma(G) - h_\gamma(M) \leq 1$, a contradiction. Hence $\gamma(G) \subseteq M$. Since ρ satisfies (P2), we see that $\rho(G) \subseteq \rho(M)$. Note that $(M/\rho(G))/(\rho(M/\rho(G))) = (M/\rho(G))/(\rho(M)/\rho(G)) \simeq M/\rho(M)$ and $\rho(G/\rho(G)) \simeq 1$ by Lemma 2.2 and (P3). From the definition of γ it follows that $\gamma(G)/\rho(G) = \gamma(G/\rho(G))$ and $\gamma(M/\rho(G)) = \gamma(M)/\rho(G)$. If $\rho(G) = M$, then $h_\gamma(G) - h_\gamma(M) = 1 - 1 = 0$, a contradiction. Hence $h_\gamma(G/\rho(G)) = h_\gamma(G)$ and $h_\gamma(M/\rho(G)) = h_\gamma(M)$. From our assumption it follows that $\rho(G) = 1$. So $\rho(\gamma(G)) = 1$. From $\gamma(G), \rho(M) \trianglelefteq M$ it follows that $\rho(M) \cap \gamma(G) \trianglelefteq \gamma(G)$. Hence $\rho(M) \cap \gamma(G) = \rho(\gamma(G)) = 1$. Now from [18, X, Theorem 13.12] it follows that

$$\rho(M) \subseteq C_G(\gamma(G)) \subseteq C_G(F^*(G)) \subseteq F^*(G) \subseteq \gamma(G).$$

It means that $\rho(M) = 1$. Now $M_{\mathfrak{E}} = 1$. Therefore $F^*(M)$ is the direct products of minimal normal non-abelian subgroups of M by [18, X, Definition 13.14 and Lemma 13.16]. Let M_1 be one of them. If $M_1 \not\subseteq F^*(G)$, then $M_1 \cap F^*(G) = 1$. So $M_1 \subseteq C_G(F^*(G)) \subseteq F^*(G)$, a contradiction. Hence $F^*(M) \subseteq F^*(G) \subseteq \gamma(G) \subseteq M$. Thus $F^*(M) = F^*(G)$.

Since ρ is a Kurosh-Amitsur radical and $h_\gamma(G) > 1$, we see that $h_\gamma(G/F^*(G)) = h_\gamma(G) - 1$. If $h_\gamma(M) > 1$, then $h_\gamma(M/F^*(G)) = h_\gamma(M/F^*(M)) = h_\gamma(M) - 1$ and we get the contradiction with the initial assumption. Thus $h_\gamma(M) = 1$. It means that $M/F^*(G) = \rho(M/F^*(G))$. Therefore $\gamma_2(G) \not\subseteq M$. Now $G/\gamma_2(G) = M/(M \cap \gamma_2(G))$. So $1 \simeq \rho(G/\gamma_2(G)) = \rho(M/(M \cap \gamma_2(G)))$ by Lemma 2.2 and definition of γ . From $\gamma(G) \subseteq \gamma_2(G) \cap M$ and $M/\gamma(G) = \rho(M/\gamma(G))$ it follows that $M/(M \cap \gamma_2(G)) = \rho(M/(M \cap \gamma_2(G))) \simeq 1$. Thus $h_\gamma(G) = 2$ and $h_\gamma(M) = 1$, the final contradiction.

3.4. Proof of Theorem 1.2

If $\gamma = F^*$, then from Theorem 1.5 it follows that $h^*(G) - h^*(M) \leq 2$ for any group G and its maximal subgroup M .

Assume that ρ is the p -soluble radical and $\gamma = \rho \star F^* \star \rho$. Then γ satisfies the assumptions of Theorem 1.10. Hence if H is not a p -soluble group, then $h_\gamma(H) = \lambda_p(H)$ by [7, Lemma 2.7]. Let a group G be a minimal order group with a maximal subgroup M such that $\lambda_p(G) - \lambda_p(M) > 1$. It means that G is a non- p -soluble group, $\lambda_p(G) > 1$ and M is p -soluble. If $M\gamma(G) = G$, then from $\lambda_p(G/\gamma(G)) \geq 1$ and $G/\gamma(G) \simeq M/(M \cap \gamma(G))$ it follows that a p -soluble group M has a non- p -soluble composition factor, a contradiction. Thus $\gamma(G) \trianglelefteq M$. Hence a p -soluble group M has a non- p -soluble composition factor, the final contradiction. It means that $\lambda_p(G) - \lambda_p(M) \leq 1$ and $\lambda(G) - \lambda(M) \leq 1$ ($\lambda = \lambda_2$) for any group G and its maximal subgroup M .

3.5. Proof of Theorem 1.11

From $|\sigma| > 1$ it follows that there exists $p \in \mathbb{P}$ such that $|\sigma \cap (\mathbb{P} \setminus \{p\})| > 1$. Let \mathfrak{F} be a formation of all p -closed soluble groups. Then \mathfrak{F} is a hereditary saturated formation. Note that $n_\sigma(G, \mathfrak{F}) - n_\sigma(M, \mathfrak{F}) = 0$ for every soluble p -closed group G and its maximal subgroup M .

For every $n > 0$ there exists a sequence of not necessary different primes $p = p_0, p_1, p_2, \dots, p_n$ such that every two of its consecutive elements belong to different elements of σ and $p_i \neq p$ for all $i > 0$. Let G_1 be a cyclic group of order p . Define a sequence of subgroups G_i inductively. Note that for G_i there exists a faithful irreducible module V_i over \mathbb{F}_{p_i} [14, B, Theorem 10.3]. Let G_{i+1} be the semidirect product of V_i with G_i corresponding to the action of G_i on V_i as an $\mathbb{F}_p G_i$ -module. Since p_i and p_{i-1} belong to different elements of σ and V_i is the unique minimal normal subgroup of G_{i+1} , we see that $F_\sigma(G_{i+1}) = V_i$.

Let $G = G_{n+1}$ and $M_i = V_i V_{i-1} \dots V_1$. Then M_n is a maximal subgroup of G and a p' -group. Hence $l_\sigma(M_n^{\mathfrak{F}}) = l_\sigma(1) = 0$. Note that G has the unique chief series and $G_2 \simeq G/(V_2 V_3 \dots V_n)$ is not p -closed. It means that $G^{\mathfrak{F}} = M_n$. Note that $M_i \trianglelefteq G_{i+1}$. Now $F_\sigma(M_i) = F_\sigma(G_{i+1}) \cap M_i = V_i \cap M_i = V_i$. It means that $l_\sigma(M_n) = n$. Therefore $n_\sigma(G, \mathfrak{F}) - n_\sigma(M, \mathfrak{F}) = n$. Since every soluble group is σ -soluble, we see that

$$\{n_\sigma(G, \mathfrak{F}) - n_\sigma(M, \mathfrak{F}) \mid G \text{ is } \sigma\text{-soluble and } M \text{ is a maximal subgroup of } G\} = \mathbb{N} \cup \{0\}.$$

In particular if $|\sigma_i| = 1$ for every $\sigma_i \in \sigma$, then

$$\{n_{\mathfrak{F}}(G) - n_{\mathfrak{F}}(M) \mid G \text{ is soluble and } M \text{ is a maximal subgroup of } G\} = \mathbb{N} \cup \{0\}.$$

4. Final Remarks and Open Questions

Note that $h(H^{\mathfrak{N}}) = h(H) - 1$ for any non-unit soluble group H and $h(H^{\mathfrak{N}}) = h(H)$ for a unit group H . If a unit group is a maximal subgroup M of G , then G is cyclic and $n_{\mathfrak{F}}(G) - n_{\mathfrak{F}}(M) = 0$. If M is a non-unit subgroup of a soluble group G , then $h(M^{\mathfrak{N}}) = h(M) - 1$ and $h(G^{\mathfrak{N}}) = h(G) - 1$. Hence $n_{\mathfrak{N}}(G) - n_{\mathfrak{N}}(M) \in \{0, 1, 2\}$ for any soluble group G and its maximal subgroup M by Corollary 1.7. That is why the main result of [17] is wrong not for all hereditary saturated formations. Therefore the following question seems natural:

Question 4.1. Describe all hereditary saturated formations \mathfrak{F} such that $n_{\mathfrak{F}}(G) - n_{\mathfrak{F}}(M) \in \{0, 1, 2\}$ for any soluble group G and its maximal subgroup M .

Proposition 4.2. Let \mathfrak{F} be a hereditary saturated formation containing all nilpotent groups. Assume that there exists a constant n such that $h(G) \leq n$ for any soluble \mathfrak{F} -group G . Then $n_{\mathfrak{F}}(G) - n_{\mathfrak{F}}(M) \leq n + 1$ for any soluble group G and its maximal subgroup M .

Proof. Note that $H^{\mathfrak{N}} \subseteq H^{\mathfrak{F}}$ for any group H . It means that $h(H) - h(H^{\mathfrak{F}}) \leq n$ for any group H by Lemma 2.1. If $h(G) = h(G^{\mathfrak{F}})$, then $G \simeq 1$ and has no maximal subgroups. Assume that $G \not\simeq 1$. Then $1 \leq h(G) - h(G^{\mathfrak{F}}) \leq n$, $h(M) - h(M^{\mathfrak{F}}) \leq n$ and $h(G) - h(M) \leq 2$. So $(h(G) - h(G^{\mathfrak{F}})) - (h(M) - h(M^{\mathfrak{F}})) \geq 1 - n$ or $n + 1 \geq h(G) - h(M) + n - 1 \geq h(G^{\mathfrak{F}}) - h(M^{\mathfrak{F}})$. Thus $n_{\mathfrak{F}}(G) - n_{\mathfrak{F}}(M) \leq n + 1$. \square

Example 4.3. There exist formations \mathfrak{F} for which the value $n_{\mathfrak{F}}(G) - n_{\mathfrak{F}}(M)$ is bounded but not by 2. Let p be a prime and \mathfrak{F} be a class of all p -closed soluble groups of nilpotent length at most 3. Then \mathfrak{F} is a hereditary saturated formation and $n_{\mathfrak{F}}(G) - n_{\mathfrak{F}}(M) \leq 4$ by Proposition 4.2 for any soluble group G and its maximal subgroup M .

Let G_4 and M_3 be the same as in the proof of Theorem 1.11. Note that $h(M_3) \leq 3$ and hence $M_3^{\mathfrak{F}} = 1$. Therefore $n_{\mathfrak{F}}(G_4) - n_{\mathfrak{F}}(M_3) = 3 > 2$.

In the view of this example the following question seems interesting:

Question 4.4. Describe all hereditary saturated formations \mathfrak{F} such that there exists a constant n with $n_{\mathfrak{F}}(G) - n_{\mathfrak{F}}(M) \leq n$ for any soluble group G and its maximal subgroup M . For such formation \mathfrak{F} do there exists a constant m with $h(G) \leq m$ for every soluble \mathfrak{F} -group G .

Recall that \mathfrak{N}_σ denotes the formation of all σ -nilpotent groups. With the help of Corollary 3.3 one can prove that $n_\sigma(G, \mathfrak{N}_\sigma) - n_\sigma(M, \mathfrak{N}_\sigma) \in \{0, 1, 2\}$ for any σ -soluble group G and its maximal subgroup M .

Question 4.5. Consider analogues of Questions 4.1 and 4.4 for $n_\sigma(G, \mathfrak{F})$.

The work was supported by BRFFR grant no. $\Phi 23\text{PH}\Phi$ -237.

References

1. Cannon J. J., Eick B., Leedham-Green C. R. Special polycyclic generating sequences for finite soluble groups. *J. Symb. Comput.*, 2004, vol. 38, iss. 5, pp. 1445–1460.

2. Khukhro E. I., Shumyatsky P. Nonsoluble and non- p -soluble length of finite groups. *Isr. J. Math.*, 2015, vol. 207, iss. 2, pp. 507–525.
3. Khukhro E. I., Shumyatsky P. On the length of finite factorized groups. *Ann. Mat. Pura Appl.*, 2015, vol. 194, iss. 6, pp. 1775–1780.
4. Fumagalli F., Leinen F., Puglisi O. A reduction theorem for nonsolvable finite groups. *Isr. J. Math.*, 2019, vol. 232, iss. 1, pp. 231–260.
5. Guralnick R. M., Tracey G. On the generalized Fitting height and insoluble length of finite groups. *Bull. London Math. Soc.*, 2020, vol. 52, iss. 5, pp. 924–931.
6. Khukhro E. I., Shumyatsky P. On the length of finite groups and of fixed points. *Proc. Amer. Math. Soc.*, 2015, vol. 143, iss. 9, pp. 3781–3790.
7. Murashka V. I., Vasil'ev A. F. On the lengths of mutually permutable products of finite groups. *Acta Math. Hungar.*, 2023, vol. 170, iss. 1, pp. 412–429.
8. Doerk K. Über die nilpotente Länge maximaler Untergruppen bei endlichen auflösbaren Gruppen. *Rend. Sem. Mat. Univ. Padova*, 1994, vol. 91, pp. 20–21.
9. Monakhov V. S., Shpyrko O. A. The nilpotent π -length of maximum subgroups in finite π -soluble groups. *Vestnik Moskov. Univ. Ser. I. Mat. Mekh.*, 2009, iss. 6, pp. 3–8.
10. Heliel A., Al-Shomrani M., Ballester-Bolinches A. On the σ -Length of Maximal Subgroups of Finite σ -Soluble Groups. *Mathematics*, 2020, vol. 8, iss. 12.
11. Plotkin B. I. Radicals in groups, operations on group classes and radical classes. *Selected Questions of Algebra and Logic*. Novosibirsk, Nauka, 1973, pp. 205–244 (in Russian).
12. Gardner B. J., Wiegandt R. *Radical Theory of Rings*. Marcel Dekker, New York, 2003.
13. Baer R. Group theoretical properties and functions. *Colloq. Math.*, 1966, vol. 14, pp. 285–327.
14. Doerk K., Hawkes T. O. *Finite Soluble Groups*. De Gruyter Exp. Math., De Gruyter, Berlin, New York, 1992, vol. 4.
15. Skiba A. N. On σ -subnormal and σ -permutable subgroups of finite groups. *J. Algebra*, 2015, iss. 436, pp. 1–16.
16. Krempa J., Malinowska I. A. On Kurosh-Amitsur radicals of finite groups. *An. Stiint. Univ. "Ovidius" Constanta Ser. Mat.*, 2011, vol. 19, iss. 1, pp. 175–190.
17. Ballester-Bolinches A., Pérez-Ramos M. D. A Note on the \mathfrak{F} -length of Maximal Subgroups in Finite Soluble Groups. *Math. Nachr.*, 1994, vol. 166, iss. 1, pp. 67–70.
18. Huppert B., Blackburn N. *Finite Groups III*. Grundlehren Math. Wiss., Springer-Verlag, Berlin, Heidelberg, 1982, vol. 243.

УДК 519.16, 512.64:(004.85+538.9)

АЛГЕБРАИЧЕСКОЕ ДОКАЗАТЕЛЬСТВО ЭКВИВАЛЕНТНОСТИ ДВУХ ВАРИАНТОВ CUT-НОРМЫ ДЛЯ МНОГОМЕРНЫХ СИММЕТРИЧНЫХ МАТРИЦ

П. Н. Шведков¹, К. В. Лыков²

¹Белорусский государственный университет, Минск, Беларусь

²Институт математики НАН Беларуси, Минск, Беларусь

e-mail: shvedkovpavel@gmail.com, alkv@list.ru

Поступила: 09.04.2025

Исправлена: 23.05.2025

Принята: 23.05.2025

Ключевые слова: cut-норма, матричная норма, полилинейные формы, эквивалентность норм, теория графов, комбинаторная оптимизация, квантовые вычисления, ограниченная машина Больцмана, многомерный массив, тензор.

Аннотация. В работе доказана эквивалентность двух специальных матричных норм. Обе нормы возникают в моделях, формулируемых в терминах взаимодействия бинарных переменных. При этом одна норма связана со взаимодействием этих переменных внутри одной группы, а другая – со взаимодействием переменных из разных групп. Утверждение позволяет легко переносить содержательные результаты со второго (более простого) случая на первый.

ALGEBRAIC PROOF OF THE EQUIVALENCE OF TWO VARIANTS OF THE CUT-NORM FOR MULTIDIMENSIONAL SYMMETRIC MATRICES

P. N. Shvedkov¹, K. V. Lykov²

¹Belarusian State University, Minsk, Belarus

²Institute of Mathematics of the National Academy of Sciences of Belarus, Minsk, Belarus

e-mail: shvedkovpavel@gmail.com, alkv@list.ru

Received: 09.04.2025

Revised: 23.05.2025

Accepted: 23.05.2025

Keywords: cut-norm, matrix norm, multilinear forms, equivalence of norms, graph theory, combinatorial optimization, quantum computing, restricted Boltzmann machine, multidimensional array, tensor.

Abstract. The paper proves the equivalence of two special matrix norms. Both norms arise in models formulated in terms of interactions between binary variables. One norm is associated with the interaction of these variables within a single group, while the other is related to the interaction of variables from different groups. The statement allows for an easy transfer of meaningful results from the second (simpler) case to the first.

1. Введение

Известно, что некоторые NP-трудные задачи (например, задачи коммивояжера, о раскраске графа, о покрытии множества и др.) могут быть эффективно решены с помощью квантовых компьютеров. Один из вариантов квантовых компьютеров основан на принципе квантового отжига [1; 2]. Такой квантовый компьютер называется адиабатическим, а наиболее известными вариантами исполнения являются компьютеры канадской компании D-Wave. Опуская тонкости, отметим, что в итоге решение вышеупомянутых задач сводится к минимизации гамильтониана Изинга [3]

$$-\sum_{i<j} J_{ij}s_i s_j - \sum_j h_j s_j,$$

где $s_{i,j} \in \{-1, 1\}$ – значения проекций спинов частиц (кубитов); J_{ij} – энергия взаимодействия частицы i с частицей j , а h_j – воздействие внешнего поля на частицу j , и нахождению минимизирующих этот гамильтониан значений бинарных переменных s_i . Возникающее взаимодействие и динамика соответствующих процессов частично описываются математическими моделями спиновых стекол [4; 5]. Несколько более простая ситуация возникает, когда частицы разделены на две группы, и

взаимодействий внутри групп нет. В таком случае и при отсутствии внешнего поля минимизируется билинейная форма (вместо квадратичной)

$$-\sum_{i,j} J_{ij} s_i s'_j,$$

где переменные s_i и s'_j независимы, т. е. представляют спины из разных групп частиц.

Аналогично, в нейросетевых моделях ассоциативной памяти, основанных идейно на теории Хебба [6], сети Хопфилда [7] и ограниченной машине Больцмана [8], минимизируются, соответственно, бинарная квадратичная форма

$$Q(x) := \sum_{i<j} w_{ij} x_i x_j, \quad x_i \in \{0, 1\}, \quad x = (x_1, x_2, \dots, x_n),$$

и бинарная билинейная форма

$$B(y, y') := \sum_{i,j} w_{ij} y_i y'_j, \quad y_i, y'_j \in \{0, 1\}, \quad y = (y_1, y_2, \dots, y_n), \quad y' = (y'_1, y'_2, \dots, y'_m).$$

В других задачах (например, в задачах из геометрии банаховых пространств [9]) оцениваются распределения абсолютных значений $|S|$ бинарных квадратичных и/или билинейных форм:

$$S(t) := \sum_{i<j} w_{ij} r_i(t) r_j(t) \quad \text{или} \quad S(t, s) := \sum_{i,j} w_{ij} r_i(t) r_j(s), \quad t \in [0, 1],$$

где

$$r_i(t) := (-1)^{\lfloor 2^i t \rfloor}, \quad i = 1, 2, \dots, n,$$

– функции Радемахера (конкретная реализация последовательности независимых симметричных бернуллиевских случайных величин в виде функций на отрезке $[0, 1]$ с мерой Лебега в роли вероятности). При этом для многих задач распределения S и $|S|$ достаточно уметь оценивать лишь с точностью до фиксированных растяжений области значений. При такой постановке задачу минимизации последней формы $S(t, s)$ можно рассматривать как задачу нахождения так называемой cut-нормы матрицы $W = (w_{ij})_{i \in \overline{[1, n]}, j \in \overline{[1, m]}}$, используемой в аппроксимационных алгоритмах комбинаторной оптимизации [10]:

$$\|W\|_{\text{cut}} := \max \left\{ \left| \sum_{i \in I, j \in J} w_{ij} \right| : I \subset \overline{[1, n]}, J \subset \overline{[1, m]} \right\},$$

где через $\overline{[1, n]}$ здесь и далее обозначен отрезок $\{1, 2, \dots, n\}$ натурального ряда. Как отмечено в [10],

$$\|W\|_{\text{cut}} \leq - \min_{t, s \in [0, 1]} S(t, s) = \max_{t, s \in [0, 1]} |S(t, s)| \leq 4 \|W\|_{\text{cut}}.$$

Легко видеть также, что для введенной выше формы $B(y, y')$, отличающейся от $S(t, s)$ условием $y_i, y'_j \in \{0, 1\}$, имеет место точное равенство

$$\max_{y \in \{0, 1\}^n, y' \in \{0, 1\}^m} |B(y, y')| = \|W\|_{\text{cut}}.$$

Хотя задача точного вычисления cut-нормы является NP-трудной, равно как и задача хорошей аппроксимации [11], существуют эффективные алгоритмы вычисления этой нормы с точностью до некоторого фиксированного множителя [12].

Что касается квадратичных форм $S(t)$ и $Q(x)$, то для них, как легко видеть, справедливы соотношения

$$\|W^s\|_{\square} = \max_{x \in \{0, 1\}^n} |Q(x)| \leq \max_{t \in [0, 1]} |S(t)| \leq 5 \|W^s\|_{\square},$$

где матрица W^s определяется своими элементами

$$(W^s)_{i,j} := \begin{cases} w_{ij}/2, & \text{при } i < j \\ 0, & \text{при } i = j \\ w_{ji}/2, & \text{при } i > j \end{cases},$$

а

$$\|W\|_{\square} := \max \left\{ \left| \sum_{i,j \in I} w_{ij} \right| : I \subset \overline{[1, n]} \right\}$$

– симметричный вариант cut-нормы. Несложно показать также, что

$$\|W^s\|_{\square} \leq \|W^s\|_{\text{cut}} \leq 4\|W^s\|_{\square}.$$

Соотношения, аналогичные приведенным, имеют место и в многомерном случае. В частности, для многомерной симметричной внедиагональной матрицы A справедливы неравенства (точные определения норм и класса рассматриваемых матриц см. далее)

$$\|A\|_{\square} \leq \|A\|_{\text{cut}} \leq C_d \|A\|_{\square}, \quad (1)$$

где константа C_d зависит от порядка (кратности) d матрицы A , но не зависит ни от размерности этой матрицы (длины одномерных строк, столбцов и т. д.), ни от значений элементов матрицы. Это соотношение не только связывает оценки максимумов абсолютных значений полиномиальных и полилинейных форм, но позволяет решать и другие задачи, связанные с разделением переменных. В работе [13] с помощью соотношения (1) осуществлен переход от многодольных гиперграфов к полным гиперграфам в задаче о разбросе. Однако доказательство правого неравенства в (1) для многомерного случая, по-видимому, не сводится к простым алгебраическим преобразованиям, как это имеет место в двумерном случае, и требует более тонкого анализа. В [13] это неравенство доказывается с помощью обращения к сильному инструменту – известной теореме о декаплинге случайных величин (см. [14, теорема 3.1.1]). В настоящей работе мы даем прямое доказательство неравенства (1), без обращения к теореме о декаплинге.

2. Определения и формулировка основного результата

Через $\overline{[1, n]}$, $n \in \mathbb{N}$, будем обозначать множество $\{1, 2, \dots, n\}$, состоящее из первых n натуральных чисел. Через $\pi(d)$ будем обозначать множество всех перестановок множества $\overline{[1, d]} = \{1, 2, \dots, d\}$, $d \in \mathbb{N}$.

Рассмотрим кубическую матрицу $A = (a_{i_1 \dots i_d})_{i_1 \in I \dots i_d \in I}$, $I = \overline{[1, n]}$, порядка d с вещественными или комплексными элементами, и определим для нее две полунормы:

$$\|A\|_{\text{cut}} := \max \left\{ \left| \sum_{i_1 \in I_1} \sum_{i_2 \in I_2} \dots \sum_{i_d \in I_d} a_{i_1 i_2 \dots i_d} \right| : I_k \subset \overline{[1, n]}, k = 1, 2, \dots, d \right\}$$

и

$$\|A\|_{\square} := \max \left\{ \left| \sum_{i_1 \in I} \sum_{i_2 \in I} \dots \sum_{i_d \in I} a_{i_1 i_2 \dots i_d} \right| : I \subset \overline{[1, n]} \right\}.$$

Будем называть кубическую матрицу A порядка d *симметричной*, если для ее элементов выполняются следующие условия:

$$a_{i_1 i_2 \dots i_d} = a_{i_{\sigma(1)} i_{\sigma(2)} \dots i_{\sigma(d)}}$$

для произвольной перестановки $\sigma \in \pi(d)$, и, кроме этого, $a_{i_1 i_2 \dots i_d} = 0$ при совпадении хотя бы двух индексов $i_k = i_l$, $1 \leq k < l \leq d$.

Оказывается, что для таких матриц справедлива следующая теорема, являющаяся основным результатом настоящей работы.

Теорема 2.1. *Для каждого натурального $d \geq 2$ существует константа C_d такая, что для любой кубической симметричной матрицы A порядка d выполняются неравенства*

$$\|A\|_{\square} \leq \|A\|_{\text{cut}} \leq C_d \|A\|_{\square}.$$

Перед доказательством теоремы сформулируем и докажем два следствия. Как отмечалось во введении, для квадратичной формы $Q(x)$ и билинейной формы $B(y, y')$ выполняются следующие равенства

$$\max_{x \in \{0, 1\}^n} |Q(x)| = \|W^s\|_{\square}$$

и

$$\max_{y \in \{0,1\}^n, y' \in \{0,1\}^m} |B(y, y')| = \|W\|_{\text{cut}}.$$

Аналогично для полилинейной формы $T(x_1, \dots, x_d)$ и полиномиальной формы $P(x) = T(x, \dots, x)$ с кубической матрицей A порядка d верно

$$\max_{x_j \in \{0,1\}^n, j=1, \dots, d} |T(x_1, \dots, x_d)| = \|A\|_{\text{cut}}$$

и

$$\max_{x \in \{0,1\}^n} |P(x)| = \|A\|_{\square}.$$

Тогда из теоремы 2.1 прямо вытекают следующие утверждения (с той же константой C_d , что и в теореме 2.1).

Следствие 2.2. Для полиномиальных форм $P(x)$ и полилинейных форм $T(x_1, \dots, x_d)$ с матрицами, как в формулировке теоремы 2.1, выполняются неравенства

$$\max_{x \in \{0,1\}^n} |P(x)| \leq \max_{x_j \in \{0,1\}^n, j=1, \dots, d} |T(x_1, \dots, x_d)| \leq C_d \max_{x \in \{0,1\}^n} |P(x)|.$$

Доказательство. Прямо следует из теоремы, если принять во внимание приведенные выше равенства для норм. \square

Следствие 2.3. Для полиномиальных форм $P(x)$ и полилинейных форм $T(x_1, \dots, x_d)$ с матрицами, как в формулировке теоремы 2.1, выполняются неравенства

$$\max_{x \in \{-1,1\}^n} |P(x)| \leq \max_{x_j \in \{-1,1\}^n, j=1, \dots, d} |T(x_1, \dots, x_d)| \leq 2^d C_d \max_{x \in \{-1,1\}^n} |P(x)|.$$

Доказательство. Первое неравенство очевидно, докажем второе.

Пусть $\max_{x_j \in \{-1,1\}^n, j \in \overline{1, d}} |T(x_1, \dots, x_d)|$ достигается на наборе x'_1, \dots, x'_d . Тогда

$$\begin{aligned} \max_{x_j \in \{-1,1\}^n, j \in \overline{1, d}} |T(x_1, \dots, x_d)| &= \left| \sum_{i_1=0}^n \dots \sum_{i_d=0}^n a_{i_1 \dots i_d} x'_{1i_1} \dots x'_{di_d} \right| \leq \\ &\leq \sum_{k=0}^d \sum_{S \subset \overline{1, d}, |S|=k} \left| \sum_{\substack{i_s : x'_{is} = +1, \\ i_t : x'_{it} = -1 \\ s \in S, t \in \overline{1, d} \setminus S}} a_{i_1 \dots i_d} \right| \leq 2^d \|A\|_{\text{cut}}. \end{aligned}$$

Из теоремы 2.1 $\|A\|_{\text{cut}} \leq C_d \|A\|_{\square}$. Остается показать, что

$$\|A\|_{\square} = \max_{x \in \{0,1\}^n} |P(x)| \leq \max_{x \in \{-1,1\}^n} |P(x)|.$$

Здесь справедлива следующая цепочка рассуждений. Полиномиальная форма $P(x)$ линейна по каждому аргументу, так как элементы матрицы A с совпадающими индексами равны нулю. Тогда, если рассматривать ее на кубе $[-1, 1]^n$, то максимум ее модуля достигается в крайних точках. Отсюда и получаем, что $\max_{x \in \{0,1\}^n} |P(x)| \leq \max_{x \in \{-1,1\}^n} |P(x)|$. Окончательно имеем

$$\max_{x_j \in \{-1,1\}^n, j \in \overline{1, d}} |T(x_1, \dots, x_d)| \leq 2^d \|A\|_{\text{cut}} \leq 2^d C_d \|A\|_{\square} \leq 2^d C_d \max_{x \in \{-1,1\}^n} |P(x)|.$$

\square

Вернемся к обсуждению теоремы 2.1. Константа в правом неравенстве теоремы 2.1, конечно, зависит от *порядка (кратности) d* матрицы A . Важно, что эта константа не зависит от *размерности n* этой матрицы, а также от значений $a_{i_1 i_2 \dots i_d}$ элементов этой матрицы. Например, $C_2 = 4$, как это уже отмечалось во введении, и эта константа подходит для обычной (в наших обозначениях, имеющей порядок 2) симметричной внедиагональной матрицы с произвольными значениями элементов a_{ij} при $i > j$. Утверждение теоремы 2.1 для $d = 2$ легко следует из алгебраической формулы

$$2XY = (X + Y)^2 - X^2 - Y^2,$$

примененной специальным (но достаточно очевидным) способом к алгебре всех формальных сумм декартовых произведений вида

$$A_1 \times B_1 + A_2 \times B_2 + \dots + A_k \times B_k,$$

где A_i и B_i – подмножества множества $I = \overline{[1, n]}$. Мы называем такую систему алгебр в теоретико-множественном смысле *алгебры подмножеств* множества $I \times I$, так как эту систему можно рассматривать как часть системы $2^{I \times I}$, состоящей из всех подмножеств $I \times I$. Однако прямое применение подобного подхода к случаю $d = 3$ приводит к формуле

$$3(X^2Y + Y^2X) = (X + Y)^3 - X^3 - Y^3,$$

в которой слагаемые X^2Y и Y^2X , различные даже при условии коммутативности $XY = YX$, не разделены. Тем не менее, в настоящей работе мы приводим элементарное доказательство теоремы 2.1. В следующем разделе проводится подготовительная работа перед доказательством основного утверждения.

3. Вспомогательные утверждения

В этом разделе мы докажем несколько утверждений, которые будем использовать при доказательстве теоремы 2.1. Первый результат можно назвать комбинаторно-алгебраическим. Мы предполагаем, что он известен и следует из более общих конструкций алгебры или комбинаторики. Однако, не найдя этого утверждения в известных и доступных нам учебниках, монографиях и справочниках, мы решили привести здесь его с полным доказательством.

Лемма 3.1. *В любом коммутативном кольце справедливо тождество*

$$f(x_1, x_2, \dots, x_d) := \sum_{k=1}^d (-1)^{d-k} \sum_{S \subset \overline{[1, d]}; |S|=k} \left(\sum_{i \in S} x_i \right)^d = d! \cdot x_1 x_2 \dots x_d.$$

Доказательство.

Для доказательства заметим, что

$$\sum_{S \subset \overline{[1, d]}; |S|=k} \left(\sum_{i \in S} x_i \right)^d \Big|_{x_1=0} = \sum_{S \subset \overline{[2, d]}; |S|=k-1} \left(\sum_{i \in S} x_i \right)^d + \sum_{S \subset \overline{[2, d]}; |S|=k} \left(\sum_{i \in S} x_i \right)^d.$$

Поэтому

$$\begin{aligned} f(0, x_2, \dots, x_d) &= \sum_{k=1}^d (-1)^{d-k} \left(\sum_{S \subset \overline{[2, d]}; |S|=k-1} \left(\sum_{i \in S} x_i \right)^d + \sum_{S \subset \overline{[2, d]}; |S|=k} \left(\sum_{i \in S} x_i \right)^d \right) = \\ &= \sum_{k=0}^{d-1} (-1)^{d-k-1} \sum_{S \subset \overline{[2, d]}; |S|=k} \left(\sum_{i \in S} x_i \right)^d + \sum_{k=1}^d (-1)^{d-k} \sum_{S \subset \overline{[2, d]}; |S|=k} \left(\sum_{i \in S} x_i \right)^d = 0. \end{aligned}$$

Здесь мы учли, что

$$\sum_{S \subset \overline{[2, d]}; |S|=d} \left(\sum_{i \in S} x_i \right)^d = 0 \quad \text{и} \quad \sum_{S \subset \overline{[2, d]}; |S|=0} \left(\sum_{i \in S} x_i \right)^d = 0.$$

Ясно, что доказываемое тождество достаточно проверить только в кольце целых чисел (так как в этом кольце различные полиномы являются разными функциями). Если в разложении однородного полинома $f(x_1, x_2, \dots, x_d)$ степени d в сумму мономов $x_1^{d_1} x_2^{d_2} \dots x_d^{d_d}$, $d_1 + d_2 + \dots + d_d = d$, найдется моном $m = x_2^{d_2} \dots x_d^{d_d}$ с $d_1 = 0$, то при подстановке $x_1 = 0$ в равенство

$$f(x_1, x_2, \dots, x_d) = cx_2^{d_2} \dots x_d^{d_d} + \dots, \quad c \neq 0,$$

получится противоречие. Значит, таких мономов в разложении $f(x_1, x_2, \dots, x_d)$ нет. Аналогично, не будет и мономов без переменной x_2 , без переменной x_3 и т. д. Получается, что

$$f(x_1, x_2, \dots, x_d) = ax_1 x_2 \dots x_d.$$

Так как слагаемые $x_1 x_2 \dots x_d$ могут появиться в выражении

$$\left(\sum_{i \in S} x_i\right)^d$$

лишь при $|S| = d$, коэффициент a совпадает с аналогичным коэффициентом у многочлена $(x_1 + x_2 + \dots + x_d)^d$, и тогда $a = d!$ \square

Приведем пояснение, как применить лемму 3.1 к подмножествам $X_k \subset \overline{[1, n]}$ индексов симметричной матрицы вместо элементов x_k коммутативного кольца.

Напомним, что *мультимножеством* \mathcal{A} с основанием Ω называется подмножество $A \subset \Omega$, в котором каждый элемент a представлен с некоторой кратностью $c(a) \in \mathbb{N}$. На мультимножествах естественным образом определена операция сложения $+$, при которой кратности элементов складываются. Разрешая кратности элементов принимать произвольные целые значения, мы очевидным образом превращаем систему всех мультимножеств с основанием Ω в модуль над кольцом целых чисел, который мы обозначим через $Mod(\Omega)$. Применяя стандартную конструкцию тензорного произведения модулей, превратим все мультимножества с основанием

$$\tilde{\Omega} := \Omega \cup \Omega^2 \cup \dots \cup \Omega^k \cup \dots$$

в кольцо с операциями $+$ (сложение мультимножеств) и \times (тензорное произведение \otimes мультимножеств как элементов модуля $Mod(\tilde{\Omega})$) в качестве сложения и умножения. Мы считаем, что если мультимножества \mathcal{A} и \mathcal{B} имеют основаниями Ω^k и Ω^m соответственно, то $\mathcal{A} \otimes \mathcal{B}$ – это мультимножество с основанием Ω^{k+m} , поэтому введенная операция умножения действует в пределах $Mod(\tilde{\Omega})$. Если мы теперь применим к элементам из Ω^k и к полученному кольцу функтор забывания порядка умножения (декартова для Ω^k и тензорного для кольца), в частности, отождествим мультимножества $\mathcal{A} \otimes \mathcal{B}$ и $\mathcal{B} \otimes \mathcal{A}$, то новое кольцо, которое мы обозначим через K , будет коммутативным. По смыслу элементы кольца K – это мультимножества с основанием из множества всех неупорядоченных конечных наборов $[x_1, x_2, \dots, x_k]$ разной длины, $x_i \in \Omega$, а операции $+$ и \times соответствуют объединению (с учетом кратности) и коммутативному варианту тензорного произведения. При таком подходе $[x] + [x] = 2[x] \neq [x, x]$, $\{[x], [y]\} = [x] + [y] = [y] + [x] \neq [x, y] = [y, x]$ (мы опускаем фигурные скобки для мультимножеств с одноточечным основанием). Для большей ясности рассмотрим пример. Пусть

$$\mathcal{A} = \{[x, y], [x, y], [x]\} \quad \text{и} \quad \mathcal{B} = \{[x, x, y], [x, y], [x, y, x], [y]\}, \quad x, y \in \Omega.$$

Тогда в кольце K будут справедливы равенства

$$\mathcal{A} = 2[x, y] + [x], \quad \mathcal{B} = 2[x, x, y] + [x, y] + [y], \quad \mathcal{A} + \mathcal{B} = \{[x], [y], 3[x, y], 2[x, x, y]\}$$

и

$$\mathcal{A} \times \mathcal{B} = \{[x, y], 2[x, y, y], [x, x, y], 2[x, x, y, y], 2[x, x, x, y], 4[x, x, x, y, y]\}.$$

Следствие 3.2. *Предположим, что множества $X_1, X_2, \dots, X_d \subset \Omega$ попарно не пересекаются. Тогда в кольце K справедливо равенство*

$$d! \cdot (X_1 \times X_2 \times \dots \times X_d) = \sum_{k=1}^d (-1)^{d-k} \sum_{S \subset \overline{[1, d]}: |S|=k} \left(\bigsqcup_{i \in S} X_i\right)^{\times d}.$$

В следствии 3.2, доказательство которого является прямым и очевидным применением леммы 3.1 к кольцу K , мы использовали знак \sqcup для объединения попарно непересекающихся множеств, как это принято, например, в теории меры.

Далее рассмотрим кубическую симметричную матрицу $A = (a_{i_1, \dots, i_d})_{i_1 \dots i_d \in I}$, $I = \overline{[1, n]}$, порядка d , размерности $\underbrace{n \times \dots \times n}_d$. Для произвольного мультимножества \mathcal{M} , состоящего из элементов

$i = (i_1, i_2 \dots i_d)$, $i_j \in I$ с кратностями $c(i)$ определим «интеграл»

$$A(\mathcal{M}) := \sum_{i \in \mathcal{M}} c(i) a_{i_1, \dots, i_d}.$$

Этот «интеграл» обладает на множествах из $I^{\times d}$ свойством аддитивности по отношению к операции $+$ в кольце K . Поэтому из следствия 3.2 вытекает

Следствие 3.3. Для произвольного набора попарно непересекающихся множеств $X_1, X_2, \dots, X_d \subset I = \overline{[1, n]}$ и произвольной симметричной матрицы $A = (a_{i_1, \dots, i_d})_{i_1, \dots, i_d \in I}$ порядка d справедливо равенство

$$\sum_{i_1 \in X_1} \cdots \sum_{i_d \in X_d} a_{i_1, \dots, i_d} = \frac{1}{d!} \sum_{k=1}^d (-1)^{d-k} \sum_{S \subset \overline{[1, d]}; |S|=k} \left(\sum_{i_1 \in \bigsqcup_{j \in S} X_j} \cdots \sum_{i_d \in \bigsqcup_{j \in S} X_j} a_{i_1, \dots, i_d} \right).$$

Еще одно утверждение, которое нам понадобится, по смыслу носит аналитический характер, но доказывается снова обращением к алгебраическому тождеству.

Будем называть матрицу $B = (b_{i_1, \dots, i_k})_{i_1, \dots, i_k \in \overline{[1, m]}}$ порядка k *внедиагональной*, если $b_{i_1, i_2, \dots, i_k} = 0$ при совпадении хотя бы двух индексов $i_s = i_t$, $1 \leq s < t \leq k$. В частности, любая симметричная матрица порядка k в наших определениях будет и внедиагональной.

Лемма 3.4. Предположим, что внедиагональная матрица $B = (b_{i_1, \dots, i_k})_{i_1, \dots, i_k \in \overline{[1, m]}}$ порядка k и размерности $m \times \dots \times m$ обладает следующим свойством: для любого разбиения $X_1 \sqcup \dots \sqcup X_k = \overline{[1, m]}$ верно

$$\left| \sum_{i_1 \in X_1} \cdots \sum_{i_k \in X_k} b_{i_1, \dots, i_k} \right| \leq c.$$

Тогда

$$\left| \sum_{i_1, \dots, i_k \in \overline{[1, m]}} b_{i_1, \dots, i_k} \right| \leq k^k c.$$

Доказательство. Из легко проверяемого (простым подсчетом количества вхождений элементов в суммы) равенства

$$\sum_{i_1, \dots, i_k \in \overline{[1, m]}} b_{i_1, \dots, i_k} = k^{k-m} \sum_{\substack{X_1, X_2, \dots, X_k: \\ X_1 \sqcup X_2 \sqcup \dots \sqcup X_k = \overline{[1, m]}}} \sum_{i_1 \in X_1} \cdots \sum_{i_k \in X_k} b_{i_1, \dots, i_k},$$

получим

$$\left| \sum_{i_1, \dots, i_k \in \overline{[1, m]}} b_{i_1, \dots, i_k} \right| \leq k^{k-m} \sum_{\substack{X_1, X_2, \dots, X_k: \\ X_1 \sqcup X_2 \sqcup \dots \sqcup X_k = \overline{[1, m]}}} \left| \sum_{i_1 \in X_1} \cdots \sum_{i_k \in X_k} b_{i_1, \dots, i_k} \right| \leq k^{k-m} k^m c = k^k c. \quad \square$$

4. Доказательство основного результата

Доказательство теоремы 2.1. Пусть матрица A порядка d имеет размерность $\underbrace{n \times \dots \times n}_d$ и

$\|A\|_{\square} = M$. В таком случае для любого множества $X \subset \overline{[1, n]}$

$$\left| \sum_{i_1 \in X} \sum_{i_2 \in X} \cdots \sum_{i_d \in X} a_{i_1, i_2, \dots, i_d} \right| \leq M,$$

и далее мы будем этим пользоваться без специальных оговорок и ссылок.

Левое неравенство в теореме очевидно. В доказательстве нуждается только правое неравенство, для доказательства которого нам нужно оценить абсолютные значения сумм вида

$$\sum_{i_1 \in Y_1} \cdots \sum_{i_d \in Y_d} a_{i_1, \dots, i_d}, \quad Y_1, Y_2, \dots, Y_d \subset I = \overline{[1, n]}, \quad (2)$$

сверху через M .

Обозначим

$$Y_j^0 := \overline{[1, n]} \setminus Y_j, \quad Y_j^1 := Y_j, \quad j \in \overline{[1, n]},$$

и для каждого двоичного слова $b = b_1 \dots b_d$, где $b_1, \dots, b_d \in \{0, 1\}$,

$$I_b = I_{b_1 \dots b_d} := \bigcap_{j=1}^d Y_j^{b_j}.$$

Тогда все возможные 2^d множеств $I_{b_1 \dots b_d}$, $b_1, \dots, b_d \in \{0, 1\}$ (или, что равносильно, $I_b, b \in \{0, \dots, 2^d - 1\}$), попарно не пересекаются, и для каждого j множество Y_j является объединением

2^{d-1} подмножеств такого вида:

$$Y_j = \bigcup_{\substack{b_1, \dots, b_d \in \{0,1\} \\ b_j = 1}} I_{b_1 \dots b_d}.$$

Значит сумму

$$\sum_{i_1 \in Y_1} \cdots \sum_{i_d \in Y_d} a_{i_1 \dots i_d}$$

можно разбить на $(2^{d-1})^d$ слагаемых вида

$$\sum_{i_1 \in I_{b_1}} \cdots \sum_{i_d \in I_{b_d}} a_{i_1 \dots i_d}, \quad (3)$$

где $b^j = b_1^j \dots b_d^j$, $b_1^j, \dots, b_d^j \in \{0, 1\}$. В каждом таком слагаемом любые два множества из соответствующей этому слагаемому системы множеств $\{I_{b_1}, \dots, I_{b_d}\}$ либо совпадают, либо не пересекаются.

Если любые два множества из системы $\{I_{b_1}, \dots, I_{b_d}\}$ не пересекаются, то, в силу симметричности матрицы A ,

$$\sum_{i_1 \in I_{b_1}} \cdots \sum_{i_d \in I_{b_d}} a_{i_1 \dots i_d} = \frac{1}{d!} \left(\sum_{\sigma \in \pi(d)} \sum_{i_1 \in I_{b_{\sigma(1)}}} \sum_{i_2 \in I_{b_{\sigma(2)}}} \cdots \sum_{i_d \in I_{b_{\sigma(d)}}} a_{i_1 i_2 \dots i_d} \right),$$

и, согласно следствию 3.3,

$$\sum_{i_1 \in I_{b_1}} \cdots \sum_{i_d \in I_{b_d}} a_{i_1 \dots i_d} = \frac{1}{d!} \sum_{k=1}^d (-1)^{d-k} \sum_{S \subset [1, d]: |S|=k} \left(\sum_{i_1 \in \bigcup_{j \in S} I_{b_j}} \cdots \sum_{i_d \in \bigcup_{j \in S} I_{b_j}} a_{i_1 \dots i_d} \right).$$

Значит, для попарно непересекающихся множеств $I_{b_1} \dots I_{b_d}$

$$\begin{aligned} \left| \sum_{i_1 \in I_{b_1}} \cdots \sum_{i_d \in I_{b_d}} a_{i_1 \dots i_d} \right| &= \frac{1}{d!} \left| \sum_{k=1}^d (-1)^{d-k} \sum_{S \subset [1, d]: |S|=k} \left(\sum_{i_1 \in \bigcup_{j \in S} I_{b_j}} \cdots \sum_{i_d \in \bigcup_{j \in S} I_{b_j}} a_{i_1 \dots i_d} \right) \right| \leq \\ &\leq \frac{1}{d!} \sum_{k=1}^d \sum_{S \subset [1, d]: |S|=k} \left| \sum_{i_1 \in \bigcup_{j \in S} I_{b_j}} \cdots \sum_{i_d \in \bigcup_{j \in S} I_{b_j}} a_{i_1 \dots i_d} \right| \leq \frac{1}{d!} \sum_{k=1}^d \sum_{S \subset [1, d]: |S|=k} M = \frac{2^d - 1}{d!} M. \quad (4) \end{aligned}$$

Осталось рассмотреть случай, когда среди множеств I_{b_1}, \dots, I_{b_d} есть повторяющиеся. В силу симметричности матрицы можно перегруппировать слагаемые так, чтобы повторяющиеся множества шли подряд:

$$\sum_{i_1 \in I_{b_1}} \cdots \sum_{i_d \in I_{b_d}} a_{i_1 \dots i_d} = \sum_{i_1 \in J_1} \cdots \sum_{i_{d_1} \in J_1} \cdots \sum_{i_{d_1+1} \in J_1} \cdots \sum_{i_{d_1} \in J_1} a_{i_1 \dots i_{d_1}}.$$

Здесь J_k попарно не пересекаются и $1 \leq d_1 < d_2 < \dots < d_l = d$. Рассмотрим теперь матрицу B порядка d_1 и размерности $\underbrace{|J_1| \times \dots \times |J_1|}_{d_1}$ с элементами

$$b_{i_1 \dots i_{d_1}} = \sum_{\substack{i_{d_r+t} \in J_{r+1} \\ r=1, \dots, l-1 \\ t=1, \dots, d_{r+1}-d_r}} a_{i_1 \dots i_{d_1} i_{d_1+1} \dots i_d}.$$

По лемме 3.4

$$\left| \sum_{i_1 \in I_{b_1}} \cdots \sum_{i_d \in I_{b_d}} a_{i_1 \dots i_d} \right| = \left| \sum_{i_1, \dots, i_{d_1} \in J_1} b_{i_1 \dots i_{d_1}} \right| \leq d_1^{d_1} \max_{F_1, \dots, F_{d_1}} \left| \sum_{i_k \in F_k, k=1, \dots, d_1} b_{i_1 \dots i_{d_1}} \right|,$$

где максимум взят по всем попарно непересекающимся системам подмножеств F_k множества J_1 . Аналогично, разбивая множества J_2, \dots, J_l , получим

$$\left| \sum_{i_1 \in I_{b_1}} \cdots \sum_{i_d \in I_{b_d}} a_{i_1 \dots i_d} \right| \leq d_1^{d_1} (d_2 - d_1)^{d_2 - d_1} \dots (d_l - d_{l-1})^{d_l - d_{l-1}} \max_{X_1, \dots, X_d} \left| \sum_{i_k \in X_k, k=1, \dots, d} a_{i_1 \dots i_d} \right|,$$

где максимум взят по всем попарно непересекающимся системам подмножеств X_k множества $\overline{[1, n]}$.

В силу оценки для попарно непересекающихся множеств (4) и неравенства

$$\begin{aligned} & d_1^{d_1} (d_2 - d_1)^{d_2 - d_1} \dots (d_l - d_{l-1})^{d_l - d_{l-1}} \leq \\ & \leq (d_1 + (d_2 - d_1) + \dots + (d_l - d_{l-1}))^{d_1 + (d_2 - d_1) + \dots + (d_l - d_{l-1})} = d^d \end{aligned}$$

имеем

$$\left| \sum_{i_1 \in I_{b_1}} \dots \sum_{i_d \in I_{b_d}} a_{i_1 \dots i_d} \right| \leq d^d \frac{2^d - 1}{d!} M, \quad (5)$$

где $I_{b_i} \cap I_{b_j} = \emptyset$ или $I_{b_i} = I_{b_j}$ для всех пар i, j . Применяя теперь разложение суммы (2) на слагаемые вида (3) и доказанное неравенство (5), получим

$$\left| \sum_{i_1 \in Y_1} \dots \sum_{i_d \in Y_d} a_{i_1 \dots i_d} \right| \leq 2^{d(d-1)} d^d \frac{2^d - 1}{d!} M \leq \frac{(d2^d)^d}{d!} M \leq (2^d e)^d M = (2^d e)^d \|A\|_{\square},$$

что завершает доказательство. \square

Работа К. В. Лыкова поддержана Институтом математики НАН Беларуси в рамках государственной программы «Конвергенция–2025» (задание 1.3.05).

Литература

1. *McGeoch C. C.* Adiabatic Quantum Computation and Quantum Annealing. Springer Nature Switzerland AG, 2014.
2. What is Quantum Annealing? [Electronic resource]. – Mode of access: https://docs.dwavesys.com/docs/latest/c_gs_2.html.
3. *Lucas A.* Ising formulations of many NP problems // *Front. in Phys.* 2014. Vol. 2. Art. 5. doi.org/10.3389/fphy.2014.00005
4. *Talagrand M.* Mean field models for spin glasses. Berlin, Heidelberg: Springer, 2011.
5. *Talagrand M.* Spin Glasses: A Challenge for Mathematicians. Cavity and Mean Field Models. Berlin, Heidelberg: Springer, 2003.
6. *Hebb D. O.* The Organization of Behavior. New York: Wiley & Sons, 1949.
7. *Hopfield J. J.* Neural networks and physical systems with emergent collective computational abilities // *Proc. Natl. Acad. Sci. USA.* 1982. Vol. 79, iss. 8. P. 2554–2558. DOI: 10.1073/pnas.79.8.2554
8. *Salakhutdinov R., Hinton G.* Deep Boltzmann Machine // *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics.* PMLR. 2009. Vol. 5. P. 448–455.
9. *Асташкин С. В., Лыков К. В.* О безусловности дробного хаоса Радемахера в симметричных пространствах // *Изв. РАН. Сер. матем.* 2024. Т. 88, № 1. С. 3–20. doi.org/10.4213/im9406
10. *Frieze A., Kannan R.* Quick Approximation to Matrices and Applications // *Combinatorica.* 1999. Vol. 19. P. 175–220. DOI: <https://doi.org/10.1007/s004930050052>
11. *Håstad J.* Some Optimal Inapproximability Results // *J. Assoc. Comput. Machinery.* 2001. P. 798–859. DOI: 10.1145/502090.502098
12. *Alon N., Naor.* A Approximating the cut-norm via grothendieck’s inequality // *STOC’04: Proceedings of the thirty-sixth annual ACM symposium on Theory of computing.* June 13–16, 2004. New York, USA, 2004. P. 72–80. doi.org/10.1145/1007352.1007371
13. *Astashkin S., Lykov K.* Random unconditional convergence of Rademacher chaos in L_∞ and sharp estimates for discrepancy of weighted graphs and hypergraphs, Ver. 1 released 28.12.2024 [Electronic resource]. – Mode of access: <https://arxiv.org/abs/2412.20107>.
14. *De la Peña V. H. and Giné E.* Decoupling: from dependence to independence. New York, Berlin, Heidelberg: Springer-Verlag, 1999.

References

1. McGeoch C. C. *Adiabatic Quantum Computation and Quantum Annealing*. Springer Nature Switzerland AG, 2014.
2. *What is Quantum Annealing?* [Electronic resource]. – Mode of access: https://docs.dwavesys.com/docs/latest/c_gs_2.html.
3. Lucas A. Ising formulations of many NP problems. *Front. in Phys.*, 2014, vol. 2, art. 5. doi.org/10.3389/fphy.2014.00005
4. Talagrand M. *Mean field models for spin glasses*. Berlin, Heidelberg, Springer, 2011.
5. Talagrand M. *Spin Glasses: A Challenge for Mathematicians. Cavity and Mean Field Models*. Berlin, Heidelberg, Springer, 2003.
6. Hebb D. O. *The Organization of Behavior*. New York, Wiley & Sons, 1949.
7. Hopfield J. J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA*, 1982, vol. 79, iss. 8, pp. 2554–2558. DOI: 10.1073/pnas.79.8.2554
8. Salakhutdinov R., Hinton G. Deep Boltzmann Machine. *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics*, PMLR, 2009, vol. 5, pp. 448–455.
9. Astashkin S., Lykov K. On the unconditionality of Rademacher’s fractional chaos in symmetric spaces. *Izvestiya: Mathematics*, 2024, vol. 88, no. 1, pp. 3–20 (in Russian). doi.org/10.4213/im9406
10. Frieze A., Kannan R. Quick Approximation to Matrices and Applications. *Combinatorica*, 1999. vol. 19, pp. 175–220. DOI: <https://doi.org/10.1007/s004930050052>
11. Håstad J. Some Optimal Inapproximability Results. *J. Assoc. Comput. Machinery*, 2001, pp. 798–859. DOI: 10.1145/502090.502098
12. Alon N., Naor. A Approximating the cut-norm via grothendieck’s inequality. *STOC’04: Proceedings of the thirty-sixth annual ACM symposium on Theory of computing (June 13–16, 2004, New York)*. New York, 2004, pp. 72–80. doi.org/10.1145/1007352.1007371
13. Astashkin S., Lykov K. *Random unconditional convergence of Rademacher chaos in L_∞ and sharp estimates for discrepancy of weighted graphs and hypergraphs*, Ver. 1 released 28.12.2024 [Electronic resource]. – Mode of access: <https://arxiv.org/abs/2412.20107>
14. De la Peña V. H., Giné E. *Decoupling: from dependence to independence*. New York, Berlin, Heidelberg, Springer-Verlag, 1999.

УДК 511.2

О РАЗРЕШИМОСТИ И ФАКТОРИЗАЦИИ НЕКОТОРЫХ π -РАЗРЕШИМЫХ НЕПРИВОДИМЫХ ЛИНЕЙНЫХ ГРУПП ПРИМАРНОЙ СТЕПЕНИ. ЧАСТЬ V

А. А. Ядченко

Институт математики НАН Беларуси, Минск, Беларусь
e-mail: yadchenko_56@mail.ru

Поступила: 23.04.2025

Исправлена: 08.05.2025

Принята: 23.05.2025

Ключевые слова: конечные группы, характеры, факторизация групп.

Аннотация. Работа является пятой и заключительной из серии статей, где для множества π , состоящего из нечетных простых чисел, исследуются конечные π -разрешимые неприводимые комплексные линейные группы степени $2|H| + 1$, у которых холловы π -подгруппы H являются TI -подгруппами и не являются нормальными в группах. Цель серии – доказать разрешимость и определить условия факторизации таких групп.

ON THE SOLVABILITY AND FACTORIZATION OF SOME π -SOLVABLE IRREDUCIBLE LINEAR GROUPS OF PRIMARY DEGREE. PART V

A. A. Yadchenko

Institute of Mathematics of the National Academy of Sciences of Belarus, Minsk, Belarus
e-mail: yadchenko_56@mail.ru

Received: 23.04.2025

Revised: 08.05.2025

Accepted: 23.05.2025

Keywords: finite groups, characters, factorizations of groups.

Abstract. The work is the fifth and final one in a series of articles, where for a set π consisting of odd primes, finite π -solvable irreducible complex linear groups of degree $2|H| + 1$ are investigated, for which Hall π -subgroups are TI -subgroups and are not normal in groups. The purpose of the series is to prove solvability and to determine the conditions for factorization of such groups.

1. Введение

Исследуются конечные π -разрешимые неприводимые комплексные линейные группы степени $n = 2|H| + 1$, у которых холлова π -подгруппа H является TI -подгруппой.

В первой части работы [1] были доказаны некоторые предварительные результаты, а также получены некоторые свойства минимального контрпримера Γ к теореме (*), которая является основой доказательства главной теоремы. Во второй [2], третьей [3] и четвертой [4] ее частях продолжено изучение свойств минимального контрпримера к теореме (*).

Условие (*). Скажем, что для Γ , A , B , C , χ и n выполнено условие (*), если $\Gamma = BA$, где B – нормальная в Γ подгруппа, $(|B|, |A|) = 1$, A – группа нечетного порядка, большего 3, которая не является нормальной в группе Γ , $C_B(a) = C_B(A) = C$ для каждого элемента $a \in A^\#$, и B имеет точный неприводимый характер χ степени n , который является a -инвариантным хотя бы для одного элемента $a \in A^\#$.

Теорема (*). Пусть для Γ , A , B , C , χ и $n = 2|A| + 1$ выполнено условие (*). Тогда группа Γ разрешима, n является степенью простого числа q , подгруппа C_2 абелева и, если подгруппа C не абелева, то в обозначениях леммы 2.7 [1] характер $\chi_C = \beta + |A|\beta_1$, где $\beta(1) = |A| + 1$ и $\beta_1(1) = 1$ либо $\beta(1) = 1$ и $\beta_1(1) = 2$. Также, если $|\pi| > 1$ и при $|\pi| = 1$ характер $\hat{\chi}$ примитивный, то $B = O_q(B)C$.

Условие 1. Пусть π – множество нечетных простых чисел и G – конечная не π -замкнутая π -разрешимая группа с π -холловой TI -подгруппой H , имеющая точный неприводимый характер степени n .

Теорема. Пусть группа G удовлетворяет условию 1 и $n = 2|H| + 1 > 7$. Тогда n – степень простого числа q , группа G разрешима и, если при $|\pi| = 1$ характер χ примитивный, то $G = N_G(H)O_q(G)$.

2. Некоторые определения, обозначения и предварительные результаты

\mathbf{Z}_+ – множество целых неотрицательных чисел; $i = \overline{1, n} = 1, 2, \dots, n$; если ψ – характер некоторой подгруппы $X \subseteq G$, то $\text{Irr}(\psi)$ обозначает множество всех неприводимых компонент характера ψ ; $\pi = \pi(H)$; $\pi' = \pi(X) \setminus \pi$; $X_{\pi'}$ – холлова π' -подгруппа группы X . Если $X \triangleleft G$ и φ – неприводимый характер подгруппы X , то условие, что φ – g -инвариантен для некоторого элемента $g \in G \setminus X$, запишем для краткости в виде $I_G(\varphi) \neq X$. Все остальные обозначения и определения обычны и их можно найти, например, в [5] или [6]. Всюду под характером группы будем понимать комплексный характер, а под группой – конечную группу.

Пусть $\Gamma = AB$ – группа с подгруппами A и B , где $B \triangleleft \Gamma$, $(|A|, |B|) = 1$ и $|A|$ нечетен (A – группа простых автоморфизмов группы B). Тогда она удовлетворяет условию теоремы 13.1 [6]. Согласно этой теореме существует взаимно-однозначное соответствие $\pi(B, A) : \text{Irr}_A(B) \longrightarrow \text{Irr}(C_B(A))$ между множеством всех A -инвариантных неприводимых характеров группы B и множеством всех неприводимых характеров подгруппы $C_B(A)$, которое обладает рядом свойств, зависящих, в частности, от свойств подгруппы A . Пусть $\chi \in \text{Irr}_A(B)$. Тогда, по лемме 13.3 [6], существует такой единственный неприводимый характер $\hat{\chi}$ группы Γ , что $\hat{\chi}_B = \chi$ и $A \subseteq \ker(\det \hat{\chi})$. Он называется каноническим продолжением характера χ на группу Γ . В дальнейшем под $\hat{\chi}$ будем понимать именно такой характер.

3. Основная часть

Продолжим нумерацию формулировок лемм, начатую в первой части [1] работы. В ней, напомним, мы получили некоторые предварительные результаты, показали, что неприводимый характер $\hat{\chi}$ точный и что $\hat{\chi}(1) = q^\alpha$ для некоторого нечетного простого числа q и $\alpha \in \mathbf{N}$, а также начали изучать свойства минимального контрпримера к теореме (*). Во второй части [2] мы продолжили изучение минимального контрпримера к теореме (*) и доказали свойства подгруппы C и ее характера χ_C .

Напомним, что $N = N_\Gamma(B_q)$, где $\hat{\chi}(1) = q^\alpha$ и B_q – A -инвариантная силовская q -подгруппа, т. е. $A \subseteq N$. В третьей части [3] мы установили, что $N \neq \Gamma$ (лемма 3.13) и что $N^{(q_1)} = N_\Gamma(B_{q_1}) \neq \Gamma$ для $q \neq q_1 \in \pi'$ (лемма 3.14). В лемме 3.15 [4] показано, что характер $\hat{\chi}_N$ неприводим. С применением теоремы Клиффорда легко убедиться в том, что характер $\hat{\chi}_{B_q}$ тоже неприводим. Следовательно, $1 \neq Z(B_q) \subseteq Z(\Gamma)$ и ограничение характера $\hat{\chi}$ на любую подгруппу, содержащую B_q , также является неприводимым характером.

Пусть $B \supset Y^{(q)}$ – такая собственная максимальная A -инвариантная подгруппа, что $N_{\pi'} \subseteq Y^{(q)}$. Так как $B_q \not\subseteq C$, что вытекает из леммы 3.3 [2], то $Y^{(q)} \not\subseteq C$. Обозначим $\Gamma_1 = AY^{(q)}$. Очевидно, $Z(B_q) \subseteq \Gamma_1$.

Дальше мы предполагаем, что $q \neq q_1 \in \pi(|B : C|)$. Пусть для таких простых чисел q_1 подгруппа $Y^{(q_1)} \subseteq B$ является собственной A -инвариантной и такой, что $B_{q_1} \subseteq Y^{(q_1)}$. Так как подгруппа $(N^{(q_1)})_{\pi'}$ также является A -инвариантной, то мы также при необходимости можем считать, что $(N^{(q_1)})_{\pi'} \subseteq Y^{(q_1)}$, и что подгруппа $Y^{(q_1)}$ является максимальной A -инвариантной в группе B . Мы замечаем, что $Y^{(q_1)} \not\subseteq C$. Обозначим $\Gamma^{(q_1)} = AY^{(q_1)}$. Также очевидно, что $Z(B_q) \subseteq \Gamma^{(q_1)}$.

Лемма 3.16. (1) Характер $\hat{\chi}_{\Gamma_1}$ неприводим и $Y^{(q)} = O_q(Y^{(q)})C_{Y^{(q)}}(A)$;

(2) если подгруппа $C_{Y^{(q)}}(A)$ не является абелевой, то

$$\chi_{C_{Y^{(q)}}(A)} = (\chi_C)_{C_{Y^{(q)}}(A)} = \beta_{C_{Y^{(q)}}(A)} + |A|(\beta_1)_{C_{Y^{(q)}}(A)}$$

с неприводимыми характерами $\beta_{C_{Y^{(q)}}(A)}$ и $(\beta_1)_{C_{Y^{(q)}}(A)}$, где $\beta(1) = |A| + 1$ и $\beta_1(1) = 1$ или $\beta(1) = 1$ и $\beta_1(1) = 2$;

(3) если подгруппа C не является абелевой, то и подгруппа $C_{Y^{(q)}}(A)$ также не является абелевой.

Доказательство. По лемме 3.15 [4] характер $\hat{\chi}_N$ неприводим. Так как $N \subseteq \Gamma_1$, то характер $\hat{\chi}_{\Gamma_1}$ также неприводим, и легко убеждаемся в том, что для Γ_1 , A , $Y^{(q)}$, $C_{Y^{(q)}}(A)$, $\chi_{Y^{(q)}}$ и n выполнено условие (*). Поскольку $|\Gamma_1| < |\Gamma|$, то по индукции $Y^{(q)} = O_q(Y^{(q)})C_{Y^{(q)}}(A)$ и подгруппа $C_{Y^{(q)}}(A)$ удовлетворяет заключениям теоремы (*) – утверждение (1).

Предположим, что подгруппа $C_{Y(q)}(A)$ не является абелевой. Тогда подгруппа C не является абелевой и по утверждению (1) леммы 3.5 [2] может выполняться только одно из двух утверждений леммы 2.7 [1]: $\chi_C = \beta + |A|\beta_1$, где $\beta(1) = |A| + 1$ и $\beta_1(1) = 1$ – утверждение (1) или $\beta(1) = 1$ и $\beta_1(1) = 2$ – утверждение (2). Здесь β и β_1 – неприводимые характеры подгруппы C . Поскольку характер $\widehat{\chi}_{\Gamma_1}$ неприводим, то из утверждения (2) леммы 3.5 [2] вытекает, что, если подгруппа $C_{Y(q)}(A)$ не является абелевой, то

$$\chi_{C_{Y(q)}(A)} = (\chi_C)_{C_{Y(q)}(A)} = \beta_{C_{Y(q)}(A)} + |A|(\beta_1)_{C_{Y(q)}(A)}$$

с неприводимыми характерами $\beta_{C_{Y(q)}(A)}$ и $(\beta_1)_{C_{Y(q)}(A)}$, для которых выполняется одно из указанных свойств – утверждение (2).

Из утверждения (2) вытекает, что, если подгруппа C неабелева, то и подгруппа $C_{Y(q)}(A)$ также неабелева – утверждение (3).

Лемма 3.17. *Выполняется: (1) $Z(\Gamma) = Z(B) = Z(B_q)$ и факторгруппа $\bar{B} = B/Z(B)$ простая; (2) если в лемме 3.16 $\beta(1) \neq |A| + 1$, то $\pi((C_{Y(q)}(A))_{q'}) \subseteq \{2; 3\}$.*

Доказательство. Покажем вначале, что $Z(\Gamma) = Z(B) = Z(B_q)$. Поскольку $A \not\triangleleft \Gamma$ и $C_B(a) = C_B(A)$ для каждого элемента $a \in A^\#$ по условию (*), то $Z(\Gamma) \cap A = 1$. Значит, $Z(\Gamma) \subseteq B$. Так как $Z(B_q) \subseteq Z(\Gamma)$, то нам достаточно показать, что $Z(\Gamma) \subseteq B_q$.

Предположим, что существует простое число $q \neq q' \in \pi(Z(\Gamma))$. Так как q' не делит $\widehat{\chi}(1)$ и q' не делит $|\ker \widehat{\chi} \cap \Gamma'|$, ибо $\ker \widehat{\chi} = 1$, то по лемме 9 [7] число q' не делит $|Z(\widehat{\chi}) \cap \Gamma'|$. Поскольку $Z(\widehat{\chi}) = Z(\Gamma)$, то q' не делит $|Z(\Gamma) \cap \Gamma'|$. Поэтому q' делит $|\Gamma/\Gamma'|$. Следовательно, $A \subseteq O^{q'}(\Gamma) \neq \Gamma$. Так как $B_q \subseteq O^{q'}(\Gamma)$, то характер $\widehat{\chi}_{O^{q'}(\Gamma)}$ неприводим. Поскольку $A \not\triangleleft O^{q'}(\Gamma)$, то по индукции

$$(O^{q'}(\Gamma))_{\pi'} = O_q((O^{q'}(\Gamma))_{\pi'})_{C_{(O^{q'}(\Gamma))_{\pi'}}(A)}.$$

Но $\Gamma = O^{q'}(\Gamma)N_\Gamma(A)$ и $N_\Gamma(A) = AN_B(A) = AC$. Поэтому

$$B = \Gamma_{\pi'} = O_q((O^{q'}(\Gamma))_{\pi'})C = O_q(B)C,$$

что доказывает теорему. Это противоречит выбору группы Γ .

Итак, $\pi(Z(\Gamma)) = \{q\}$, т. е. $Z(\Gamma) \subseteq B_q$. Тогда, очевидно, $Z(\Gamma) \subseteq Z(B_q)$. Так как справедливо обратное утверждение, то $Z(\Gamma) = Z(B_q)$.

Убедимся теперь в том, что $Z(B_q)$ является максимальной нормальной в Γ подгруппой.

Пусть $B \triangleright X$ – собственная максимальная A -инвариантная подгруппа. Покажем, что $X \subseteq Z(B_q)$.

Очевидно, что $X \triangleleft \Gamma$ и что B/X – главный фактор группы Γ . С применением теоремы Клиффорда рассмотрим характер

$$\widehat{\chi}_X = e \sum_{t \in T} \varphi^t,$$

где φ – неприводимый характер подгруппы X ; T – множество представителей всех смежных классов группы Γ по подгруппе $I = I_\Gamma(\varphi)$ – инерции характера φ в группе Γ , взятых по одному из каждого класса и

$$\widehat{\chi}(1) = e|T|\varphi(1).$$

Так как $\widehat{\chi}(1) = q^\alpha$, то e , $|T|$ и $\varphi(1)$ также являются степенями простого числа q . Также мы замечаем, что $A \subseteq I$ и по теореме 6.11 [6] $\widehat{\chi} = (\varphi')^\Gamma$ для такого неприводимого характера φ' подгруппы I , что $(\varphi')_X = e\varphi$.

Допустим, что $X \subseteq C$. Мы замечаем, что, с одной стороны, ограничение характера χ_X подчиняется закону Клиффорда выше, а с другой – для характера $\chi_X = (\chi_C)_X$ выполняется одно из утверждений леммы 2.7 [1]. И мы видим, что такие утверждения, как

$$\chi_X = e \sum_{t \in T} \varphi^t = \beta_X + |A|(\beta_1)_X; \chi_X = e \sum_{t \in T} \varphi^t = \beta_X + |A|(\beta_1 + \beta_2)_X,$$

где β , β_1 и β_2 – различные неприводимые характеры группы C , неверны, потому что

$$(\chi_X, \varphi^t)_X = e \neq (\chi_X, (\beta_1))_X = |A|.$$

Нетрудно заметить, что может выполняться только утверждение (12) этой леммы –

$$\chi_X = (\chi_C)_X = (2|A| + 1)\beta_X, \quad \beta(1) = 1$$

или только утверждение (15) –

$$\chi_X = (\chi_C)_X = \beta_X, \quad \beta(1) = \chi(1) = 2|A| + 1.$$

Предположим, что выполняется утверждение (15) указанной леммы, т. е. $\chi_X = \beta_X$, $\beta(1) = \chi(1)$. Тогда из леммы 2.6 [1] следует, что $\widehat{\chi}_A = \beta(1)1_A$. Отсюда вытекает, что $A \subseteq \ker \widehat{\chi}$. Это противоречит лемме 3.1 [1].

Следовательно, выполняется утверждение (12) леммы 2.7 [1], т. е. $\chi_X = (2|A| + 1)\beta_X$, $\beta(1) = 1$. Тогда

$$X \subseteq Z(\chi) = Z(\widehat{\chi}) = Z(\Gamma) = Z(B_q).$$

При $X \subseteq C$ мы получили требуемое утверждение.

Пусть $X \not\subseteq C$.

Предположим, что $|T| = 1$. Тогда $\widehat{\chi}(1) = e\varphi(1)$.

Допустим при этом, что $e = 1$. Значит, характер $\widehat{\chi}_X = \varphi$ неприводим. Следовательно, неприводим и характер $\widehat{\varphi} = \widehat{\chi}_{AX}$, который существует по лемме 2.5 [1]. Легко убедиться в том, что для $AX, A, X, C_X(A)$, φ и n выполнено условие (*). Так как $|AX| < |\Gamma|$, то по индукции

$$X = O_q(X)C_X(A).$$

Понятно, что

$$O_q(X) \subseteq O_q(B) \subseteq B_q$$

и, так как $O_q(B) \subseteq X$, ввиду максимальности последней подгруппы, то $O_q(X) = O_q(B)$. Так как $X \not\subseteq C$, то $O_q(B) \not\subseteq C$. Поэтому $A \not\triangleleft O_q(B)$ и характер $\widehat{\chi}_{AO_q(B)}$ неприводим.

В самом деле. Из предпоследней выделенной формулы вытекает, что

$$AO_q(B) \triangleleft AX$$

и, значит, характер

$$\widehat{\chi}_{AO_q(B)} = \widehat{\varphi}_{AO_q(B)} = e_1 \sum_{t_1 \in T_1} \psi^{t_1}$$

подчиняется формуле Клиффорда. Здесь обозначения e_1, t_1, T_1 и ψ соответствуют обозначениям из приведенной выше формулы Клиффорда.

Предположим, что $\psi(1) \neq \widehat{\varphi}(1)$. Тогда

$$\psi^{t_1}(1) \leq (2|A| + 1)/q < |A| - 1$$

для каждого элемента $t_1 \in T_1$, и из леммы 2.10 [6] вытекает, что

$$A/\ker \psi^{t_1} \triangleleft AO_q(B)/\ker \psi^{t_1}.$$

Поскольку $\bigcap_{t_1 \in T_1} \ker \psi^{t_1} = 1$, то подгруппа $AO_q(B)$ изоморфна некоторой подгруппе из прямого произведения

$$\prod_{t_1 \in T_1} AO_q(B)/\ker \psi^{t_1}$$

π -замкнутых факторгрупп. Следовательно, $A \triangleleft AO_q(B)$, т. е. $O_q(B) \subseteq C$. Из выделенной ранее формулы вытекает, что $X \subseteq C$, что не так.

Итак, характер $\widehat{\chi}_{AO_q(B)}$ неприводим. Следовательно, неприводим и характер $\widehat{\chi}_{AB_{q_1}O_q(B)}$ для каждого простого числа $q_1 \in \pi(|B : C|)$.

Допустим, что $AB_{q_1}O_q(B) \neq \Gamma$. Тогда по индукции $B_{q_1} \subseteq C$, что не так.

Поэтому $A(B_{q_1}O_q(B) = B) = \Gamma$. Следовательно, $B_q = O_q(B) \triangleleft \Gamma$. Мы получили противоречие с тем, что $\Gamma^{(q)} \neq \Gamma$.

Пусть теперь $e \neq 1$. Поскольку $\varphi \in \text{Irr}_A(X)$, то для характера $\widehat{\varphi}$ справедливо выражение

$$\widehat{\varphi}(1) = \widehat{\chi}(1)/e < |A| - 1.$$

Так как характер φ точный в рассматриваемом случае, то $\ker \widehat{\varphi} \subseteq A$. Если $\ker \widehat{\varphi} \neq 1$, то поскольку $\ker \widehat{\varphi} \triangleleft AX$, мы получаем, что $X \subseteq C_B(\ker \widehat{\varphi})$. Тогда по условию (*) $X \subseteq C$. Если же $\ker \widehat{\varphi} = 1$, то из леммы 2.10 [1] вытекает, что $A \triangleleft AX$. Мы вновь получаем, что $X \subseteq C$. Этот случай рассмотрен.

Пусть теперь $|T| = |\Gamma : I| \neq 1$. Тогда характер $\widehat{\chi}$ непримитивный и по лемме 3.4 [2] мы можем считать, что $\varphi'(1) = 1$ и что

$$|\Gamma : I| = \widehat{\chi}(1) = q^\alpha = 2|A| + 1.$$

Следовательно, $\varphi(1) = (\varphi')_X(1) = 1$. Мы видим, что все неприводимые компоненты φ^t , $t \in T$, характера $\widehat{\chi}_X$ линейные. Поэтому $X' \subseteq \ker \widehat{\chi}_X = 1$ и, значит, подгруппа X абелева.

Допустим, что группа $\Gamma^* = AB_qX = \Gamma$. Тогда подгруппа $B = B_qX$ разрешима.

Предположим при этом, что $|\pi| > 1$. По лемме 2.9 [1] $B = F(B)C$. Так как $X \subseteq F(B)$, то мы видим, что

$$F(B) = (((F(B))_q = F(B) \cap B_q))X,$$

т. е. $(F(B))_q \subseteq C_\Gamma(X) \triangleleft \Gamma$. Поскольку подгруппа X максимальная нормальная в Γ , то мы получаем, что $C_\Gamma(X) = X$, т. е. $(F(B))_q = 1$ или $C_\Gamma(X) = \Gamma$, т. е. $X \subseteq Z(\Gamma) = Z(B_q)$.

Поскольку второе утверждение доказывает лемму, то мы полагаем, что $(F(B))_q = 1$. Получаем, что $B_q \subseteq C$, что противоречит лемме 3.3 [2].

Пусть теперь $|\pi| = 1$. Тогда по условию (*) характер $\widehat{\chi}$ примитивный. Следовательно, $X \subseteq Z(\Gamma) = Z(B_q)$.

Желаемое утверждение установлено.

Допустим теперь, что группа $\Gamma^* \neq \Gamma$. Поскольку характер $\widehat{\chi}_{\Gamma^*}$ неприводим, то мы легко замечаем, что к нему и группе Γ^* мы можем применить индукцию. По ней

$$B_qX = O_q(B_qX)C_{B_qX}(A).$$

Отсюда мы видим, что $X \subseteq C_{B_qX}(A) \subseteq C$. Это противоречит рассматриваемому случаю. Мы доказали, что $X = Z(B_q)$.

Так как X максимальная нормальная в B и A -инвариантная подгруппа и $X = Z(B_q)$, то факторгруппа $B/Z(B_q)$ простая – утверждение (1).

Докажем утверждение (2).

Пусть элемент $c \in C_{Y(q)}(A)$ имеет простой порядок $o(c) = q_1 \neq q$. Согласно утверждению (1) доказываемой леммы $c \notin Z(\Gamma)$. Тогда $c \notin C_\Gamma(B_q)$ и поэтому $\varphi = \widehat{\chi}_{B_q} \in \text{Irr}_{<c>}(B_q)$. Обозначим $\widehat{\varphi} = \widehat{\chi}_{<c>B_q}$. По следствию 13.4 [6] $\widehat{\varphi}(c) \in \mathbf{Q}$. Так как $c \notin Z(\Gamma)$ и $c \in C$, то характер $\chi_{<c>} = \widehat{\varphi}_{<c>}$ содержит не менее двух различных неприводимых компонент. Поскольку $\beta(1) \neq |A| + 1$, то согласно лемме 2.7 [1] это компоненты $\beta_{<c>}$, $\beta(1) = 1$ и $\beta_1(<c>)$, $\beta_1(1) = 2$. Третьей компонентой характера $\chi_{<c>}$ может быть характер $\beta_2(<c>)$ степени 1. При этом, если $\beta_1(1) = 2$, то

$$\chi_{<c>} = \beta(<c>) + |A|\beta_1(<c>)$$

– утверждение (2) леммы 2.7 [1].

Применим лемму 1 [8] в обозначениях которой $\text{Spec}(\chi_{<c>})$ – это множество всех различных неприводимых компонент характера $\chi_{<c>}$ без учета их кратности. Это множество называется спектром характера χ на элементе c . Понятно, что

$$|\text{Spec}(\chi_{<c>})| \leq o(c) = q_1,$$

а из предпоследнего выделенного выражения также следует, что

$$|\text{Spec}(\chi_{<c>})| \leq \beta(1) + \beta_1(1).$$

Так как $q_1 \neq q$ не делит $\chi(1) = q^\alpha$, то по утверждению 2) указанной леммы справедливо одно из следующих утверждений:

а) $q_1 = 2$ и $c \in Z(\chi)$;

б) $|Spec(\chi_{\langle c \rangle})| = q_1 - 1$ и $q_1 - 1$ делит $\chi(1)$;

в) $|Spec(\chi_{\langle c \rangle})| = q_1$.

Поскольку $Z(\chi) = Z(B_q)$ и $c \notin B_q$, то не может выполняться утверждение 1), т. е. $q_1 \neq 2$. Тогда $q_1 > 2$ и $q_1 - 1$ – четное число и оно не делит $\chi(1)$. Значит, может выполняться только утверждение в), т. е. $|Spec(\chi_{\langle c \rangle})| = q_1$. Так как $\beta_1(1) = 2$,

$$Spec(\chi_{\langle c \rangle}) = Spec(\beta_{\langle c \rangle}) \cup Spec((\beta_1)_{\langle c \rangle})$$

и $\beta(1) + \beta_1(1) = 3$, то $|Spec(\chi_{\langle c \rangle})| = o(c) \leq 3$.

Итак, $o(c) \in \{2; 3\}$.

Лемма 3.17 доказана.

Так как $A \not\triangleleft \Gamma^{(q_1)}$, то к подгруппе $\Gamma^{(q_1)}$ и к ее характеру $\widehat{\chi}_{\Gamma^{(q_1)}}$ применима лемма 3.2 [1]. По ней $\widehat{\chi}_{\Gamma^{(q_1)}} = \eta + \theta$, где η – такой характер наибольшей степени, что $A \ker \eta / \ker \eta \triangleleft \Gamma^{(q_1)} / \ker \eta$ и θ – такой характер наибольшей степени, что $A \ker \theta / \ker \theta \not\triangleleft \Gamma^{(q_1)} / \ker \theta$.

Лемма 3.18. *Выполняется: характер $\eta \neq 0$ и характер θ неприводим одной из степеней: $\theta(1) = |A|$ или $\theta(1) = 2|A|$.*

Доказательство. Предположим, что $\eta = 0$. Тогда $\theta(1) = 2|A| + 1$. Допустим также, что характер θ неприводим. Поскольку $\Gamma^{(q_1)} \neq \Gamma$, то мы можем применить индукцию, по которой

$$Y^{(q_1)} = O_q(Y^{(q_1)})C_{Y^{(q_1)}}(A).$$

Отсюда вытекает, что $B_{q_1} \subseteq C_{Y^{(q_1)}}(A) \subseteq C$. Это противоречит выбору простого числа q_1 .

Итак, $\eta = 0$ и характер θ приводим. Тогда из леммы 2.10 [1] вытекает, что он является суммой неприводимых компонент θ_1 и θ_2 степеней $|A|$ и $|A| + 1$ соответственно и $|A| + 1 = 2^{\alpha_1}$, $\alpha_1 \in \mathbf{N}$. По лемме 3.2 [1]

$$Y^{(q_1)} = O_2(Y^{(q_1)})C_{Y^{(q_1)}}(A).$$

Отсюда вытекает, что $B_{q_1} \subseteq C_{Y^{(q_1)}}(A)$, если $q_1 \neq 2$. Противоречие с выбором простого числа q_1 .

Поэтому мы делаем вывод, что $q_1 = 2$.

Продолжим исследовать этот случай. Очевидно,

$$q^\alpha = 2|A| + 1 = 2(|A| + 1) - 1 = 2 \cdot 2^{\alpha_1} - 1 = 2^{\alpha_1 + 1} - 1.$$

Значит, $\alpha = 1$ и $\widehat{\chi}(1) = q$ – простое число Мерсенна. Следовательно, $\alpha_1 + 1$ – нечетное число. Поэтому α_1 – четное число. Значит,

$$|A| = (2^{\alpha_1} - 1) = (2^{\alpha_1/2} - 1)(2^{\alpha_1/2} + 1)$$

делится на 3. Следовательно, \overline{B} – простая $3'$ -группа. Такой может быть только группа Судзуки $Sz(d)$, $d = 2^{2m+1}$, $m \in \mathbf{N}$, порядка

$$d^2(d-1)(d^2+1) = d^2(d-1)(d+2r+1)(d-2r+1), \quad r = 2^m.$$

Как в доказательстве утверждения 3.15.2 [4] мы устанавливаем, что в этом случае $|\overline{B} : \overline{C}|$ – степень 2. Так как $\overline{B} = B/(Z(B_q) = Z(\Gamma))$ и $\overline{C} = C/Z(B_q)$, то мы видим, что $|B : C|$ степень 2. По лемме 2.9 [1] $B = O_2(B)C$. Это означает, что $B_q \subseteq C$ и противоречит лемме 3.3 [2]. Значит, характер θ не может быть и приводимым в случае, когда $\theta(1) = 2|A| + 1$.

Остался случай, когда $\theta(1) \neq 2|A| + 1$. Следовательно, $\eta \neq 0$.

По лемме 3.2 [1]

$$\theta(1) \in \{|A| - 1; |A|; |A| + 1; 2(|A| - 1); 2|A| - 1; 2|A|\}.$$

При этом

$$Y^{(q_1)} = (\ker \eta)_{\pi'} C_{Y^{(q_1)}}(A); \quad A \ker \eta \triangleleft \Gamma^{(q_1)}; \quad [Y^{(q_1)}, A] \subseteq (\ker \eta)_{\pi'},$$

характер $\theta_{A \ker \eta}$ точный и, если $\theta(1) \neq 2(|A| - 1)$ и $\theta(1) \neq 2|A|$, то неприводимый. Кроме всего этого, если $\theta(1) \neq |A|$ и $\theta(1) \neq 2|A|$, то $\theta(1)$ – степень некоторого простого числа r .

Исследуем эти значения.

Допустим, что $\theta(1) = |A| - 1$. По лемме 2.11 [1]

$$|A| - 1 = 2^k, \quad |A \ker \eta : Z(\theta_{A \ker \eta})| = |A| 2^{2k}, \quad k \in \mathbf{N},$$

и

$$(\ker \eta)_{\pi'} = O_2(A \ker \eta) Z(\theta_{A \ker \eta}), \quad (O_2(A \ker \eta))' \subseteq Z(\theta_{A \ker \eta}).$$

Очевидно, что $Z(\theta_{A \ker \eta}) \subseteq Z(\theta)$ и что $Z(\theta_{A \ker \eta}) \subseteq \ker \eta \subseteq Z(\eta)$. Стало быть, по лемме 5 [9]

$$Z(\theta_{A \ker \eta}) \subseteq Z(\theta) \cap Z(\eta) = Z(\Gamma^{(q_1)}).$$

Это означает, что

$$Z(\theta_{A \ker \eta}) \subseteq C_{Y^{(q_1)}}(A).$$

Поэтому

$$Y^{(q_1)} = O_2(A \ker \eta) C_{Y^{(q_1)}}(A)$$

и, значит, $|Y^{(q_1)} : C_{Y^{(q_1)}}(A)|$ является степенью 2. Поэтому простое число $r = 2$ может быть одним из выбранных, т. е. $r = q_1 = 2$. Из последнего выделенного равенства следует, что

$$AO_2(A \ker \eta) \triangleleft \Gamma^{(q_1)}.$$

Предположим, что $B_2 \not\subseteq C$. Отсюда, из теоремы Клиффорда, из того, что $A \not\subseteq AO_2(A \ker \eta)$ и из леммы 2.10 [1] вытекает, что характер $\theta_{AO_2(A \ker \eta)}$ неприводим. Тогда и характер $\theta_{O_2(A \ker \eta)}$ также неприводим. Следовательно, $(O_2(A \ker \eta))' \neq 1$.

Пусть инволюция $z \in (O_2(A \ker \eta))'$. Мы видим, что

$$\widehat{\chi}(z) = \eta(z) + \theta(z) = (|A| + 2) + (-(|A| - 1)) = 3,$$

ибо $z \in \ker \eta$,

$$\eta(1) = \widehat{\chi}(1) - \theta(1) = (2|A| + 1) - (|A| - 1) = |A| + 2$$

и характер $\widehat{\chi}$ точный. Однако, с другой стороны, $z \in Z(\Gamma^{(q_1)}) \subseteq C_{Y^{(q_1)}}(A) \subseteq C$. Поэтому

$$\widehat{\chi}(z) = (\chi_C)_{C_{Y^{(q_1)}}(A)}(z)$$

и мы можем воспользоваться леммой 2.7 [1]. Учитывая, что $\chi(1) = 2|A| + 1$, мы легко проверяем, что ни одно из утверждений леммы 2.7 не подходит.

Например, пусть выполняется утверждение (1) леммы 2.7, т. е. в ее обозначениях $\chi_C = \beta + |A|\beta_1$, $\beta(1) = |A| + 1$ и $\beta_1(1) = 1$. Тогда

$$\chi(z) = \beta(z) + |A|\beta_1(z) = \pm(|A| + 1) \mp |A| \neq 3.$$

Подобным образом мы проверяем оставшиеся утверждения леммы 2.7 [1]. Полученное противоречие указывает на то, что и $B_2 \subseteq C$ при $\theta(1) = |A| - 1$. Делаем вывод, что $\theta(1) \neq |A| - 1$.

Пусть $\theta(1) = |A| + 1$. Этот случай мы рассматривали чуть ранее.

Рассмотрим теперь случай, когда $\theta(1) = 2(|A| - 1)$. По утверждению (1_i) леммы 3.2 [1] $2(|A| - 1) = 2^{\alpha_1}$, $\alpha_1 \in \mathbf{N}$, и

$$(\ker \eta)_{\pi'} = O_2((\ker \eta)_{\pi'}) C_{(\ker \eta)_{\pi'}}(A).$$

Тогда $q_1 = 2$ и из выделенного выражения выше вытекает, что

$$Y^{(q_1)} = O_2((\ker \eta)_{\pi'}) C_{Y^{(q_1)}}(A).$$

Отсюда следует, что

$$AO_2((\ker \eta)_{\pi'}) \triangleleft \Gamma^{(q_1)}.$$

Поэтому из теоремы Клиффорда вытекает, что характер $\theta_{AO_2((\ker \eta)_{\pi'})}$ неприводим или имеет одну неприводимую компоненту ψ степени $|A| - 1$ или две сопряженные ψ и ψ^g , $g \in \Gamma^{(2)} \setminus AO_2((\ker \eta)_{\pi'})$, компоненты той же степени.

Если характер $\theta_{AO_2((\ker \eta)_{\pi'})}$ не является неприводимым, то как и в случае, когда $\theta(1) = |A| - 1$, мы придем к противоречию со значением $\chi(z)$ для инволюции $z \in Z(AO_2((\ker \eta)_{\pi'}))$.

Предположим, что характер $\theta_{AO_2((\ker\eta)_{\pi'})}$ неприводим. Тогда неприводим и характер $\theta_{O_2((\ker\eta)_{\pi'})}$. Поэтому

$$1 \neq Z(O_2((\ker\eta)_{\pi'})) = Z(\theta_{AO_2((\ker\eta)_{\pi'})}),$$

и существует инволюция $z \in Z(AO_2((\ker\eta)_{\pi'}))$. Тогда

$$\widehat{\chi}(z) = \eta(z) + \theta(z) = 3 + (-(2|A| - 2)) = -2|A| + 5.$$

Поскольку $z \in C$, то, с другой стороны,

$$\chi(z) = \beta(z) + |A|\beta_1(z),$$

если выполняется утверждение (1) или (2) леммы 2.7 [1].

Из двух последних выделенных выражений вытекает, что

$$\beta(z) + |A|\beta_1(z) = -2|A| + 5.$$

Значит,

$$|A|(\beta_1(z) + 2) = 5 - \beta(z)$$

– равенство двух целых рациональных чисел, ибо $o(z) = 2$.

Допустим, что выполняется утверждение (2) леммы 2.7 [1], т. е. $\beta(1) = 1$. Тогда $\beta(z) = \pm 1$ и, значит, $|A|$ делит $5 + 1 = 6$ или $5 - 1 = 4$. Так как $|A|$ – подгруппа нечетного порядка большего 3, то мы получили противоречие.

Допустим теперь, что выполняется утверждение (1) леммы 2.7 [1], т. е. $\beta(1) = |A| + 1$. Тогда из утверждения (1) леммы 2.6 [1] вытекает, что

$$\widehat{\chi}_A = \rho_A + (|A| + 1)1_A.$$

Следовательно,

$$\alpha_A = (\widehat{\chi}_A, 1_A)_A = |A| + 2.$$

С другой стороны,

$$\widehat{\chi}_A = (\widehat{\chi}_{\Gamma^{(2)}})_A = (\eta_{AO_2((\ker\eta)_{\pi'})})_A + (\theta_{AO_2((\ker\eta)_{\pi'})})_A.$$

Поскольку характер $\theta_{AO_2((\ker\eta)_{\pi'})}$ точный и неприводимый степени $2|A| - 2$, то мы можем применить лемму 12 [9], по которой группа $AO_2((\ker\eta)_{\pi'})$ имеет такой неприводимый характер $\widehat{\theta}$, что $\widehat{\theta}_A = 2\rho_A - 2 \cdot 1_A$. По утверждению (2) леммы 2.6 [1]

$$(\theta_{AO_2((\ker\eta)_{\pi'})})_A = (\widehat{\theta}_A)\lambda^{-1} = (2\rho_A - 2 \cdot 1_A)\lambda^{-1} = 2\rho_A - 2\lambda^{-1}$$

для некоторого линейного характера λ подгруппы A . Отсюда мы видим, что

$$\alpha^{(1)} = ((\theta_{AO_2((\ker\eta)_{\pi'})})_A, 1_A)_A = ((2\rho_A - 2\lambda^{-1})_A, 1_A)_A \leq 2.$$

Поскольку $\eta(1) = \widehat{\chi}(1) - \theta(1) = (2|A| + 1) - (2|A| - 2) = 3$, то

$$\alpha^{(2)} = ((\eta_{AO_2((\ker\eta)_{\pi'})})_A, 1_A)_A \leq 3.$$

Тогда мы видим, что $\alpha_A = \alpha^{(1)} + \alpha^{(2)} = |A| + 2 \leq 5$. Отсюда следует, что $|A| \leq 3$, что противоречит условию (*).

Утверждения (1) и (2) леммы 2.7 [1] выполняться не могут.

Допустим, что выполняется ее утверждение (3), т. е.

$$\chi_C = \beta + 2|A|\beta_1,$$

$\beta(1) = \beta_1(1) = 1$. Так как $z \notin Z(\chi) = Z(\Gamma)$, то

$$\chi(z) = \beta(z) + 2|A|\beta_1(z) = \pm 2|A| \mp 1 \neq -2|A| + 5.$$

Аналогичным образом мы убеждаемся в том, что не могут выполняться утверждения (4), (5), (12) и (15) леммы 2.7 [1].

Мы доказали, что $\theta(1) \neq 2(|A| - 1)$.

Предположим, что $\theta(1) = 2|A| - 1 \neq 17$. По лемме 3.2 [1]

$$Y^{(q_1)} = O_{q_1}(Y^{(q_1)})C_{Y^{(q_1)}}(A), \quad 2|A| - 1 = (q_1)^{\alpha_1}, \alpha_1 \in \mathbf{N}.$$

Отсюда мы видим, что

$$AO_{q_1}(Y^{(q_1)}) \triangleleft \Gamma^{(q_1)}.$$

Так как мы рассматриваем случай, когда $Y^{(q_1)} \not\subseteq C$, то с применением теоремы Клиффорда и леммы 2.10 [1] мы убеждаемся в том, что характер $\theta_{AO_{q_1}(Y^{(q_1)})}$ неприводим. Тогда неприводим и характер $\theta_{O_{q_1}(Y^{(q_1)})}$ и поэтому существует элемент $1 \neq g \in Z(O_{q_1}(Y^{(q_1)}))$ простого порядка q_1 . Так как $Z(O_{q_1}(Y^{(q_1)})) \subseteq Z(\theta)$ и $g \in \ker \eta$, то с применением леммы 5 [9] мы убеждаемся в том, что $g \in Z(\Gamma^{(q_1)})$. Следовательно,

$$\widehat{\chi}(g) = \eta(g) + \theta(g) = 2 + (2|A| - 1)\xi, \quad \xi \neq 1,$$

и $\xi^{q_1} = 1$.

С другой стороны, поскольку $g \in C$, то мы можем применить лемму 2.7 [1]. Легко убедиться в том, что

$$\chi(g) \neq 2 + (2|A| - 1)\xi, \quad \xi \neq 1$$

для каждого такого ее утверждения, для которого $\chi(1) = 2|A| + 1$.

Предположим теперь, что $\theta(1) = 2|A| - 1 = 17$. Тогда $|A| = 9$, т. е. \bar{B} – простая $3'$ -группа. Ранее мы показали, что это невозможно.

Лемма доказана.

Нам осталось рассмотреть случаи, когда $\theta(1) = |A|$ или $\theta(1) = 2|A|$ для любой такой A -инвариантной подгруппы $Y^{(q_1)} \subseteq B$, что $B_{q_1} \subseteq Y^{(q_1)}$ и $B_{q_1} \not\subseteq C$.

Лемма 3.19. *Выполняется: $\theta(1) \neq 2|A|$.*

Доказательство. Пусть для некоторого простого числа $q_1 \in \pi(|B : C|)$ и для максимальной подгруппы $Y^{(q_1)}$ выполняется утверждение, что $\theta(1) = 2|A|$. Тогда $\eta(1) = 1$ и из леммы 3.2 [1] вытекает, что

$$A \ker \eta \triangleleft \Gamma^{(q_1)}, \quad Y^{(q_1)} = (\ker \eta)_{\pi} C_{Y^{(q_1)}}(A)$$

и характер $\theta_{A \ker \eta}$ точный. Также по лемме 3 [9] $[Y^{(q_1)}, A] \subseteq \ker \eta$.

Допустим, что характер θ не является неприводимым. По утверждению (1_{vi}) леммы 3.2 [1] подгруппа $Y^{(q_1)}$ абелева. Так как $N^{(q_1)} \subseteq Y^{(q_1)}$, то мы можем применить теорему Бернсайда. По ней подгруппа B содержит нормальное q_1 -дополнение R . Тогда $Z(B_q) \subseteq R$ и $\bar{R} = R/Z(B_q) \triangleleft \bar{B}$, что противоречит тому, что \bar{B} – простая группа.

Значит, характер θ неприводим.

По теореме Клиффорда

$$\theta_{Y^{(q_1)}} = \sum_{a \in A} (\beta)^a,$$

где β – неприводимый характер степени 2 подгруппы $Y^{(q_1)}$. Отсюда, из того, что $Y^{(q_1)} \subseteq Z(\eta)$, ибо $\eta(1) = 1$, а также из леммы 5 [9] и из леммы 5.11 [6] вытекает, что

$$Z(\theta_{Y^{(q_1)}}) = \cap_{a \in A} Z(\beta^a) = Z(Y^{(q_1)}); \quad \ker \theta = \cap_{a \in A} \ker \beta^a.$$

Так как $Z(\theta_{Y^{(q_1)}}) \subseteq Z(\theta)$ и $\ker \theta \subseteq Z(\theta) \subseteq Z(\eta)$, то по лемме 5 [9] $Z(\theta) = Z(\Gamma^{(q_1)})$. Поскольку $A \cap Z(\Gamma^{(q_1)}) = 1$, то мы можем записать, что

$$\ker \theta \subseteq Z(\theta) = Z(\Gamma^{(q_1)}) = Z(Y^{(q_1)}).$$

Так как по лемме 2.27(f) [6]

$$Z(\beta)/\ker \beta = Z(Y^{(q_1)}/\ker \beta),$$

то отсюда и из теоремы 14.23 [6] вытекает, что

$$|(Y^{(q_1)})^* = Y^{(q_1)}/Z(Y^{(q_1)}/\ker \beta)| = |Y^{(q_1)}/Z(\beta)| = 2^2 \cdot 3; \quad 2^3 \cdot 3; \quad 2^2 \cdot 3 \cdot 5.$$

Из выделенного выше выражения вытекает, что группа $Y^{(q_1)}$ изоморфна некоторой подгруппе из прямого произведения

$$\bar{Y}^{(q_1)} = Y^{(q_1)} / Z(Y^{(q_1)}) = \prod_{a \in A} Y^{(q_1)} / Z(\beta^a).$$

Отсюда следует, что, если простое число $q_1 \notin \pi(|(Y^{(q_1)})^*|)$, то подгруппа $B_{q_1} \subseteq Z(\beta)$. Поскольку подгруппа B_{q_1} – A -инвариантна, то $B_{q_1} = (B_{q_1})^a \subseteq (Z(\beta))^a$ для каждого элемента $a \in A$. Поэтому $B_{q_1} \subseteq Z(Y^{(q_1)})$ для $q_1 > 5$. Так как $(N^{(q_1)})_{\pi'} \subseteq Y^{(q_1)}$, то по теореме Бернсайда группа B содержит нормальное q_1 -дополнение R , что не так по лемме 3.17. Следовательно,

$$q_1 \in \pi(|(Y^{(q_1)})^*|)$$

и для всех простых делителей

$$r \in \pi' \setminus \pi(|(Y^{(q_1)})^*|) \cup q; \quad B_r \subseteq C.$$

Мы делаем вывод, что

$$q_1 \in \{2; 3; 5\}; \quad q_1 \neq q.$$

Понятно, что характер β может быть примитивным или импримитивным.

Предположим, что характер β примитивный.

Рассмотрим случай, когда факторгруппа $\Gamma^{(q_1)} / \ker \theta$ неразрешима. Тогда из утверждения (2iii) [10] и из того, что $Z(\theta) = Z(Y^{(q_1)}) \subseteq Z(\beta^a)$, $a \in A$, вытекает, что факторгруппа

$$\bar{Y}^{(q_1)} = Y^{(q_1)} / Z(Y^{(q_1)}) \cong (PSL(2; 5))^{a_1=1} \times \dots \times (PSL(2; 5))^{a|A}$$

либо

$$\bar{Y}^{(q_1)} \cong C_{Y^{(q_1)}}(A) / C_{Y^{(q_1)}}(A) \cap Z(Y^{(q_1)}) \cong (PSL(2; 5)), \quad (Y^{(q_1)})_{\{2; 3; 5\}} \subseteq C_{Y^{(q_1)}}(A).$$

Мы видим, что при выполнении последнего утверждения получим, что $B_{q_1} \subseteq C$, что противоречит выбору простого числа q_1 .

Стало быть, может выполняться только первое утверждение и $|(Y^{(q_1)})^*| = 2^2 \cdot 3 \cdot 5$, т. е.

$$(Y^{(q_1)})^* \cong PSL(2; 5)$$

и факторгруппа $\bar{Y}^{(q_1)}$ является прямым произведением $|A|$ копий группы $PSL(2; 5)$.

Предположим, что $C_{B_{q_1}}(A) = 1$ для каждого простого числа $q_1 \in \{2; 3; 5\}$. Тогда $C_{Y^{(q_1)}}(A) = C_{Y^{(q_1)}}(a) = 1$ для каждого элемента $a \in A$. Тогда подгруппа $Y^{(q_1)}$ является ядром группы Фробениуса $\Gamma^{(q_1)}$ и поэтому нильпотентна, что не так.

Значит, $C_{B_{q_1}}(A) \neq 1$ для некоторого простого числа $q_1 \in \{2; 3; 5\}$.

Допустим, что существует элемент $1 \neq c_1 \in C_{B_{q_1}}(A)$, $c_1 \notin Z(\beta)$ для некоторого простого числа $q_1 \in \{2; 3; 5\}$. Тогда $A \subseteq C_{(Y^{(q_1)})^*}(\langle c_1 \rangle Z(\beta) / Z(\beta))$. Следовательно,

$$\langle c_1 \rangle Z(\beta) / Z(\beta)^a = \langle c_1 \rangle Z(\beta) / Z(\beta)$$

для каждого элемента $a \in A$. Так как $\bigcap_{a \in A} \ker \beta^a = Z(Y^{(q_1)})$, то произведение

$$\bar{Y}^{(q_1)} = \prod_{a \in A} Y^{(q_1)} / Z(\beta^a) \cong \prod_{a \in A} (PSL(2; 5))^{a \in A}$$

не является прямым, ибо подгруппа

$$\langle c_1 \rangle Z(\beta) / Z(\beta) \subseteq (Y^{(q_1)} / Z(\beta^{a_i})) \cap (Y^{(q_1)} / Z(\beta^{a_j})), \quad i \neq j.$$

Это противоречит ранее доказанному утверждению о том, что группа $\bar{Y}^{(q_1)}$ является прямым произведением $|A|$ копий группы $PSL(2; 5)$.

Значит, $C_{B_{q_1}}(A) \subseteq Z(\beta)$ для каждого простого числа $q_1 \in \{2; 3; 5\}$.

Пусть $1 \neq c \in C_{B_{q_1}}(A)$. Тогда $aca^{-1} = c$ для каждого элемента $a \in A$ и, значит,

$$\theta(c) = \sum_{a \in A} \beta^a(c) = 2 \sum_{a \in A} \lambda^a(c) = 2 \sum_{a \in A} \lambda(aca^{-1}) = 2|A|\lambda(c), \quad \beta_{Z(\beta)} = 2\lambda, \quad \lambda(1) = 1.$$

Следовательно, $c \in Z(\theta)$. Поскольку c произвольный элемент $C_{B_{q_1}}(A)$ для произвольного простого числа q_1 , то

$$C_{Y^{(q_1)}}(A) \subseteq Z(\theta) = Z(\Gamma^{(q_1)}).$$

Из теоремы В [11] вытекает, что подгруппа $Y^{(q_1)}$ разрешима. Это противоречит рассматриваемому случаю, когда факторгруппа $\Gamma^{(q_1)}/\ker\theta$ и, значит, подгруппа $Y^{(q_1)}$ неразрешима.

Рассмотрим теперь случай, когда факторгруппа $\Gamma^{(q_1)}/\ker\theta$ разрешима. Тогда разрешима и подгруппа $Y^{(q_1)}$ и

$$(Y^{(q_1)})^* = Y^{(q_1)}/Z(\beta) \cong PSL(2;5).$$

По утверждению (2_{ii}) теоремы [10] факторгруппа

$$(Y^{(q_1)})^* \cong S_4; \quad Y^{(q_1)} = (Y^{(q_1)})_{\{2;3\}} \times (Y^{(q_1)})_{\{2;3\}'}, \quad (Y^{(q_1)})_{\{2;3\}'} \subseteq Z((Y^{(q_1)}))$$

или

$$(Y^{(q_1)})^* \cong A_4; \quad Y^{(q_1)} = (Y^{(q_1)})_{\{3\}} F(Y^{(q_1)}), \quad (Y^{(q_1)})_{\{2;3\}'} \subseteq Z((Y^{(q_1)}))$$

с абелевой подгруппой $F(Y^{(q_1)})_2$.

Допустим, что $q_1 > 3$. Как и чуть ранее убеждаемся в том, что группа B содержит нормальное q_1 -дополнение R , что не так.

Поэтому $q_1 \in \{2;3\}$ и $B_r \subseteq C$ для всех $r \in \{\pi(B) \setminus \{2;3;q\}\}$.

Следовательно, $(Y^{(q_1)})_{\{2;3\}'} \subseteq C$ и поэтому

$$\Gamma^{(q_1)} = A(Y^{(q_1)})_{\{2;3\}} \times (Y^{(q_1)})_{\{2;3\}'}, \quad (Y^{(q_1)})_{\{2;3\}'} \subseteq Z((\Gamma^{(q_1)})).$$

Так как группы S_4 и A_4 не являются нильпотентными, то $C_{B_{q_1}}(A) \neq 1$ для $q_1 = 2$ или для $q_1 = 3$. Ранее показано, что $C_{B_{q_1}}(A) \subseteq Z(\Gamma^{(q_1)})$.

Здесь следует заметить, что $C_{B_2}(A) \neq 1$. В самом деле, если это не так и подгруппа $C_2 = C_{B_2}(A) = 1$, то группа C имеет нечетный порядок. Тогда все ее неприводимые характеры имеют нечетную степень. Проанализировав утверждения леммы 2.7 [1], с учетом того, что $\chi(1) = 2|A| + 1$, мы убеждаемся в том, что все неприводимые компоненты характера χ_C линейные. Мы получаем, что подгруппа C абелева. По теореме В [11] группа B разрешима, что не так по лемме 3.17.

Так как $Y^{(q_1)} \subseteq N_B(C_{B_2}(A))$ и подгруппа $Y^{(q_1)}$ максимальная A -инвариантная в группе B , то $Y^{(q_1)} = N_B(C_{B_2}(A))$. Поскольку $C_{B_2}(A) \subseteq C \subseteq B$, то мы замечаем, что подгруппа $C_{B_2}(A) \subseteq Z(Y^{(q_1)})$. По теореме Бернсайда группа C содержит нормальное 2-дополнение $C_{2'}$ нечетного порядка. Как чуть ранее убеждаемся в том, что подгруппа $C_{2'}$ абелева. Понятно, что $(Y^{(q_1)})_{\{2;3\}'} \subseteq C_{2'}$.

Допустим, что $((Y^{(q_1)})_{\{2;3\}'})_{q'} \neq 1$. Тогда

$$L = \langle Y^{(q_1)}, C \rangle \subseteq C_B(((Y^{(q_1)})_{\{2;3\}'})_{q'}).$$

Поскольку подгруппа L является A -инвариантной и подгруппа $Y^{(q_1)}$ максимальна в группе B , то $L = Y^{(q_1)}$ или $L = B$. В первом случае мы получаем, что подгруппа $C \subseteq Y^{(q_1)}$ и, значит, абелева, что приводит к противоречию с простотой группы B , во втором случае мы получаем, что

$$1 \neq ((Y^{(q_1)})_{\{2;3\}'})_{q'} \subseteq Z(B),$$

что противоречит лемме 3.17.

Следовательно, $((Y^{(q_1)})_{\{2;3\}'})_{q'} = 1$ и, значит, подгруппа

$$Y^{(q_1)} = (Y^{(q_1)})_{\{2;3\}} \times Z(B_q)$$

максимальна в B и разрешима.

Нетрудно видеть, что

$$|\bar{B}| = |\bar{Y}^{(q_1)}| |\bar{C}| |\bar{B}_q|,$$

где $\bar{Y}^{(q_1)} = Y^{(q_1)}/Z(B_q)$, $\bar{C} = C/Z(B_q)$ и $\bar{B}_q = B_q/Z(B_q)$. Так как $\bar{Y}^{(q_1)}$ – максимальная $\{2;3\}$ -подгруппа в \bar{B} , подгруппа $\bar{C}_{2'}$ абелева и подгруппа \bar{B}_q примарная, а также по лемме 3.16 $\bar{Y}^{(q_1)} = Y^{(q_1)}/Z(B_q)$ является максимальной разрешимой в подгруппе \bar{B} , то мы можем утверждать, что каждая максимальная

в \bar{B} подгруппа разрешима. Это же мы можем сказать и о максимальных подгруппах группы B . По лемме 1 [12] группа B либо разрешима, либо группа \bar{B} изоморфна одной из следующих групп:

- (1) $L_2(p)$, где p – простое число, $p > 3$ и $p^2 - 1 \not\equiv 0 \pmod{5}$;
- (2) $L_2(2^p)$, где p – нечетное простое число;
- (3) $L_2(3^p)$, где p – нечетное простое число;
- (4) $L_3(3)$;
- (5) $Sz(2^p)$, где p – нечетное простое число.

Допустим, что $\bar{B} \cong L_2(p)$, p – простое число. Так как все ее подгруппы Силова нечетного порядка циклические, то и подгруппа $\bar{B}_q = B_q/Z(B_q)$ также циклическая. Получаем, что подгруппа B_q абелева, что не так, ибо она имеет нелинейный неприводимый характер $\hat{\chi}_{B_q}$.

То же самое можем повторить в случае, когда $\bar{B} \cong L_2(2^p)$, p – нечетное простое число.

Если $\bar{B} \cong L_2(3^p)$, где p – нечетное простое число, то все ее $3'$ -подгруппы Силова нечетного порядка циклические. Поскольку $q \neq 3$, то мы видим, что и этот случай невозможен.

Пусть $\bar{B} \cong L_3(3)$. В этой группе единственная $3'$ -подгруппа Силова порядка 13, т. е. тоже циклическая. Поэтому $q \neq 13$. Этот случай тоже невозможен.

Случай, когда $\bar{B} \cong Sz(2^p)$, где p – нечетное простое число рассматривался в начале доказательства леммы 3.18.

Случай, когда группа B разрешима противоречит лемме 3.17.

Осталось рассмотреть случай, когда характер β импримитивный. По утверждению (2_i) [10] подгруппа $S = (Y^{(q_1)})_2 \triangleleft Y^{(q_1)}$ и абелева, а также $Y^{(q_1)} = (Y^{(q_1)})_2 F(Y^{(q_1)})$. К тому же подгруппа $F(Y^{(q_1)})$ абелева тогда и только тогда, когда $(Y^{(q_1)})_2 \not\subseteq F(Y^{(q_1)})$. Нетрудно видеть, что подгруппа $(Y^{(q_1)})_2$ неабелева и что характер $\theta_{A(Y^{(q_1)})_2}$ неприводим. Ранее установлено, что $Z(\theta) = Z(\Gamma^{(q_1)}) = Z(Y^{(q_1)})$. Понятно, что $1 \neq Z((Y^{(q_1)})_2) = C_{B_2}(A)$ и что $Y^{(q_1)} = N_B(C_{B_2}(A))$ в силу максимальности подгруппы $Y^{(q_1)}$. Как чуть ранее убеждаемся в том, что $S = 1$ и, значит, подгруппа $Y^{(q_1)} = B_2$ и максимальна в группе B .

Оставшаяся часть доказательства аналогична предыдущему случаю за исключением, может быть, случая, когда $\bar{B} \cong L_2(3^p)$, где p – нечетное простое число. Простое число q может быть равным 3. Однако в этом случае подгруппа $B_2 \cong \bar{B}_2 = B_2 Z(B_q)/Z(B_q)$ абелева и не может иметь неприводимый характер β степени 2.

Лемма 3.19 доказана.

Осталось рассмотреть только случай, когда $\theta(1) = |A|$.

Лемма 3.20. Если $\theta(1) = |A|$, то $B = Y^{(r)} B_q$ для простого числа $r \in \pi(B)$ и $\bar{B} = (Y^{(r)}/Z(B_q)) \bar{B}_q$.

Доказательство проводится аналогично доказательству утверждения 3.15.5 из [4]. При этом неприводимый характер θ в этой работе играет роль характера θ' в той.

Лемма 3.21. Окончание доказательства теоремы (*).

Доказательство. Так как $\theta(1) = |A|$, то по лемме 3.2 [1]

$$Y^{(r)} = [Y^{(r)}, A] \times (C_{Y^{(r)}}(A) = C), \quad [Y^{(r)}, A] \subseteq Z(Y^{(r)}).$$

Тогда

$$\bar{B} = ([Y^{(r)}, A] \times (C_{Y^{(r)}}(A) = C))/Z(B_q) = [Y^{(r)}, A] \times C/Z(B_q).$$

Если $|[Y^{(r)}, A]| \neq 1$, то по теореме 3.9 [6] группа \bar{B} непростая.

Если же $|[Y^{(r)}, A]| = 1$, то $B = B_q C$ и по лемме 2.9 [1] $B = O_q(B)C$, что доказывает теорему (*).

Теорема (*) доказана.

Доказательство теоремы

Пусть π – множество нечетных простых чисел и G – конечная не π -замкнутая π - разрешимая группа с π -холловой TI -подгруппой H , имеющая точный неприводимый характер χ степени n . Так как H является TI -подгруппой в G , то

$$O_{\pi', \pi}(G) = HO_{\pi'}(G).$$

Отсюда по лемме Чунихина

$$G = N_G(H)O_{\pi'}(G).$$

Поскольку $O_{\pi',\pi}(G) \triangleleft G$, то отсюда, из теоремы Клиффорда и леммы 2.10 [1] следует, что характер $\chi_{O_{\pi',\pi}(G)}$ неприводим. Очевидно, что и характер $\chi_{O_{\pi'}(G)}$ также неприводим и H -инвариантный. По лемме 2.5 [1] группа $O_{\pi',\pi}(G)$ имеет характер $\widehat{\chi}$, который, легко видеть, точный. По лемме 2.1 [1] $C_{O_{\pi'}(G)}(h) = C_{O_{\pi'}(G)}(H) = C$ для каждого элемента $h \in H^\#$. Нетрудно заметить, что для $O_{\pi',\pi}(G)$, A , $O_{\pi'}(G)$, C , $\chi_{O_{\pi'}(G)}$ и n выполняется условие (*). Так как $n = q^\alpha = 2|H| + 1 > 7$, то $|H| > 3$. По теореме (*) $O_{\pi'}(G) = O_q(O_{\pi'}(G))C$. Тогда легко видеть, что $O_q(O_{\pi'}(G)) \subseteq O_q(G)$ и что

$$G = N_G(H)O_q(G).$$

Теорема доказана.

Литература

1. Ядченко А. А. О разрешимости и факторизации некоторых Π -разрешимых неприводимых линейных групп примарной степени. Часть I // Труды Института математики. 2022. Т. 30, № 1–2. С. 84–98.
2. Ядченко А. А. О разрешимости и факторизации некоторых Π -разрешимых неприводимых линейных групп примарной степени. Часть II // Труды Института математики. 2023. Т. 31, № 1. С. 77–89.
3. Ядченко А. А. О разрешимости и факторизации некоторых Π -разрешимых неприводимых линейных групп примарной степени. Часть III // Труды Института математики. 2023. Т. 31, № 2. С. 91–102.
4. Ядченко А. А. О разрешимости и факторизации некоторых Π -разрешимых неприводимых линейных групп примарной степени. Часть IV // Труды Института математики НАН Беларуси. 2024. Т. 32, № 2. С. 17–30.
5. Gorenstein D. Finite groups. New York: Harper and Row, 1968. 527 p.
6. Isaacs I. M. Character theory of finite groups. New York: Academic Press, 1976. 303 p.
7. Ядченко А. А. О конечных π -разрешимых линейных группах // Арифметическое и подгрупповое строение конечных групп. Минск: Наука и техника, 1986. С. 181–207.
8. Ядченко А. А. О спектрах p -элементов конечных комплексных p -разрешимых линейных групп // Математические заметки. 1991. Т. 50, № 3. С. 143–151.
9. Ядченко А. А. О Π -разрешимых неприводимых линейных группах с холловой TI -подгруппой нечетного порядка I // Труды Института математики. 2008. Т. 16, № 2. С. 118–130.
10. Ядченко А. А. О нормальных подгруппах и факторизации некоторых π -разрешимых неприводимых линейных групп // Труды Института математики. 2021. Т. 29, № 1–2. С. 149–164.
11. Beltran A. Actions with nilpotent fixed point subgroup // Archiv der Mathematik. 1997. Vol. 69, N 3. P. 177–184.
12. Ведерников В. А. Конечные группы, в которых каждая неразрешимая максимальная подгруппа холлова // Труды Института математики и механики УрО РАН. 2013. Т. 19, № 3. С. 71–82.

References

1. Yadchenko A. A. On the solvability and factorization of some π -solvable irreducible linear groups of primary degree. Part I. *Trudy Instituta matematiki*, 2022, vol. 30, no. 1–2, pp. 84–98 (in Russian).
2. Yadchenko A. A. On the solvability and factorization of some π -solvable irreducible linear groups of primary degree. Part II. *Trudy Instituta matematiki*, 2023, vol. 31, no. 1, pp. 77–89 (in Russian).
3. Yadchenko A. A. On the solvability and factorization of some π -solvable irreducible linear groups of primary degree. Part III. *Trudy Instituta matematiki*, 2023, vol. 31, no. 2, pp. 91–102 (in Russian).
4. Yadchenko A. A. On the solvability and factorization of some π -solvable irreducible linear groups of primary degree. Part IV. *Proceedings of the Institute of Mathematics of the NAS of Belarus*, 2024, vol. 32, no. 2, pp. 17–30 (in Russian).
5. Gorenstein D. *Finite groups*. New York, Harper and Row, 1968, 527 p.

6. Isaacs I. M. *Character theory of finite groups*. New York, Academic Press, 1976, 303 p.
7. Yadchenko A. A. *On finite π -solvable linear groups in The Arithmetic and Subgroup Structure of Finite Groups*. Minsk, Nauka i Tekhnika, 1986, pp. 181–207.
8. Yadchenko A. A. Spectra of p -elements of finite p -solvable complex linear groups. *Math notes*, 1991, vol. 50, no. 3. pp. 143–151 (in Russian).
9. Yadchenko A. A. On the π -solvable irreducible linear groups with Hall TI -subgroups of odd order I. *Trudy Instituta matematiki*, 2008, vol. 16, no. 2, pp. 118–130 (in Russian).
10. Yadchenko A. A. On the normal subgroups and factorization some π -solvable irreducible linear groups. *Trudy Instituta matematiki*, 2021, vol. 29, no. 1–2, pp. 149–164 (in Russian).
11. Beltran A. Actions with nilpotent fixed point subgroup. *Archiv der Mathematik*, 1997, vol. 69, no. 3, pp. 177–184.
12. Vedernikov V. A. Finite groups in which every non-solvable maximal subgroup is a Hall subgroup. *Trudy Instituta matematiki i mekhaniki UrO RAN*, 2013, vol. 19, no. 3, pp. 71–82 (in Russian).



ВЕЩЕСТВЕННЫЙ, КОМПЛЕКСНЫЙ
И ФУНКЦИОНАЛЬНЫЙ АНАЛИЗ
REAL, COMPLEX AND
FUNCTIONAL ANALYSIS



УДК 512.542

ТОПОЛОГИЧЕСКИЕ СТРУКТУРЫ НА ГРАДУИРОВАННЫХ МНОЖЕСТВАХ

А. Б. Антоневи́ч¹, М. Д. Ёжи́кова²

¹Институт математики НАН Беларуси, Минск, Беларусь

²Белорусский государственный университет, Минск, Беларусь
e-mail: antonevich@bsu.by, maya.yo1989banan@gmail.com

Поступила: 19.03.2025

Исправлена: 19.05.2025

Принята: 23.05.2025

Ключевые слова: градуированное множество, неархимедова метрика, sharp-топология, градуированная группа, градуированное векторное пространство, p -адический анализ, полином Тейлора, асимптотическая сходимость.

Аннотация. Группа L называется градуированной, если она представлена в виде объединения убывающей последовательности подгрупп L_m . Предложена общая схема введения так называемой sharp-метрики на таких группах, относительно которой алгебраические операции непрерывны и которая является неархимедовой. Показано, что такая группа всюду плотно вкладывается в полную группу, элементами которой являются ряды специального вида из элементов L . Аналогичные конструкции рассмотрены для градуированных колец и градуированных векторных пространств. В качестве примеров показано, что в конкретных частных случаях применение описанной конструкции приводит к построению p -адических чисел и к построению рядов Тейлора и Лорана.

TOPOLOGICAL STRUCTURES ON GRADED SETS

A. B. Antonevich¹, M. D. Yozhikova²

¹Institute of Mathematics of the National Academy of Sciences of Belarus, Minsk, Belarus

²Belarusian State University, Minsk, Belarus
e-mail: antonevich@bsu.by, maya.yo1989banan@gmail.com

Received: 19.03.2025

Revised: 19.05.2025

Accepted: 23.05.2025

Keywords: graded set, non-archimedean metric, sharp topology, graded group, graded vector space, p -adic analysis, Taylor polynomial, asymptotic convergence.

Abstract. A group L is called graded if it is represented as the union of a decreasing sequence of subgroups L_m . A general scheme for introducing the so-called sharp metric on such groups is proposed, with respect to which the algebraic operations are continuous and which is non-archimedean. It is shown that such a group is densely embedded in a complete group whose elements are series of a special type composed of elements of L . Similar constructions are considered for graded rings and graded vector spaces.

As examples, it is shown that in concrete special cases, the application of the described construction leads to the construction of p -adic numbers and to the construction of Taylor and Laurent series.

Введение

В анализе чаще всего используются топологические векторные пространства [1], в которых, по определению, топология согласована с алгебраическими операциями, т. е. сложение и умножение на число являются непрерывными в заданной топологии. Вместе с тем в ряде исследований встречаются *градуированные векторные пространства*. Это векторные пространства L , в которых задана двусторонняя последовательность векторных подпространств L_m , $m \in \mathbb{Z}$, такая, что

$$L_m \supset L_{m+1}, L = \bigcup_{m \in \mathbb{Z}} L_m. \quad (1)$$

На таких пространствах обычно нет топологии, согласованной со структурой векторного пространства, и к ним не может быть применена теория топологических векторных пространств.

Градуированные векторные пространства возникают при построении асимптотических разложений, исследовании уравнений с малым параметром [2; 3], в теории так называемых новых обобщенных функций (мнемофункций) [4] и других вопросах. При исследовании таких пространств обычно используется не только градуировка, но и более богатая структура *векторного пространства с градуированной топологией* – градуированного векторного пространства, в котором на каждом из выделенных подпространств L_m задана своя топология, согласованная с алгебраическими операциями.

Аналогично, в ряде вопросов возникают градуированные группы. Коммутативную группу L будем называть *градуированной*, если задана двусторонняя последовательность подгрупп L_m , $m \in \mathbb{Z}$, такая, что выполнено (1). В частности, градуированные векторные пространства являются градуированными группами.

В связи с появлением пространств с такой структурой в разных вопросах естественно провести систематическое исследование градуированных групп и градуированных пространств с целью приложений этой теории к решению конкретных задач.

В первой части данной работы предложен способ задания неархимедовой метрики на градуированных группах. Такой способ задания метрики на одной из алгебр мнемофункций был использован в [5] и, следуя терминологии из [5], будем называть введенные метрики *sharp-метриками*. Описано пополнение градуированной группы относительно введенной метрики. Показано, что в градуированных векторных пространствах операция умножения на число является разрывной в *sharp-метрике*.

Во второй части приведены примеры классических утверждений из анализа, в которых обычно не подчеркивается, что в них использована структура градуированного пространства, но они являются частными случаями общих утверждений о градуированных группах или пространствах. Также отмечены классические утверждения, которые используют не только градуировку, но и структуру пространства с градуированной топологией.

1. НЕАРХИМЕДОВЫ МЕТРИКИ НА ГРАДУИРОВАННЫХ ГРУППАХ

1.1. Sharp-метрики на пространствах последовательностей. Множество H будем называть *градуированным*, если в нем выделена последовательность подмножеств H_m , $m \in \mathbb{Z}$, такая, что

$$H_m \supset H_{m+1}, H = \bigcup_{m \in \mathbb{Z}} H_m.$$

Градуировка на множестве определяет некоторую иерархию элементов множества, похожую на отношение порядка – считается, что элемент из H_m в некотором смысле меньше элементов из $H \setminus H_m$.

Выделим градуированные множества специального вида. Пусть задана двусторонняя последовательность множеств E_k , $k \in \mathbb{Z}$, каждое из которых содержит не менее двух точек, и в каждом выделена отмеченная точка, которую будем обозначать через 0. Декартово произведение

$$\hat{H} = \prod_{k \in \mathbb{Z}} E_k, \quad (2)$$

есть множество двусторонних последовательностей ξ вида

$$\xi = (\dots, \xi_{k-1}, \xi_k, \xi_{k+1}, \dots), \quad \xi_k \in E_k.$$

Декартово произведение счетного семейства множеств $H_m = \prod_{k \geq m} E_k$ будем рассматривать как подмножество в \hat{H} , отождествляя его с множеством последовательностей ξ , у которых $\xi_k = 0$ при $k < m$, т. е. вида

$$\xi = (0, \dots, 0, \xi_m, \xi_{m+1}, \dots), \quad \xi_k \in E_k, \quad (3)$$

Элементы ξ_k последовательности (3) будем называть координатами точки ξ . Последовательность, все члены которой есть отмеченная точка 0, будем обозначать через 0.

Интересующее нас градуированное множество, построенное с помощью заданной последовательности множеств E_k , есть

$$H = \bigcup_{m \in \mathbb{Z}} H_m \subset \widehat{H}. \quad (4)$$

Заметим, что здесь для градуировки выполнено условие

$$\bigcap_{m \in \mathbb{Z}} H_m = \{0\}.$$

Специальный вид рассматриваемого градуированного множества проявляется в том, что здесь не только $H_k \supset H_{k+1}$, но имеется и более сильная связь этих множеств:

$$H_k = E_k \times H_{k+1} \quad (5)$$

и вложение $H_{k+1} \rightarrow H_k$ действует по формуле

$$H_{k+1} \ni \xi \rightarrow (0, \xi) \in E_k \times H_{k+1}.$$

Порядком отличия элемента $\xi \in H$ с координатами ξ_k от элемента η с координатами η_k будем называть число

$$\nu(\xi, \eta) = \begin{cases} \min\{k : \xi_k \neq \eta_k\}, & \xi \neq \eta; \\ +\infty, & \xi = \eta. \end{cases}$$

Теорема 1.1. Пусть H есть градуированное множество вида (4).

1. Формула

$$\rho(\xi, \eta) = 2^{-\nu(\xi, \eta)} \quad (6)$$

задает *sharp*-метрику на H , для которой выполнено усиленное неравенство треугольника:

$$\rho(\xi, \eta) \leq \max\{\rho(\xi, \zeta), \rho(\zeta, \eta)\} \quad \forall \xi, \eta, \zeta. \quad (7)$$

2. Пространство H с метрикой (6) является полным. Для любой последовательности

$$\xi = (0, \dots, 0, \xi_m, \xi_{m+1}, \dots) \in H$$

последовательность из финитных последовательностей

$$\xi^n = (0, \dots, 0, \xi_m, \xi_{m+1}, \dots, \xi_n, 0, \dots) \quad (8)$$

сходится к ξ .

3. Замкнутый шар $B[0, 2^{-m}]$ совпадает с H_m и такие подмножества образуют базу окрестностей точки 0 в заданной топологии.

Доказательство. 1. Равенство $\rho(\xi, \eta) = 0$ выполнено только при $\xi = \eta$.

Пусть $\nu(\xi, \eta) = m$, $\nu(\eta, \zeta) = n$. Это означает, что $\xi_k = \eta_k$ при $k < m$ и $\eta_k = \zeta_k$ при $k < n$. Поэтому $\xi_k = \zeta_k$ при $k < \min\{m, n\}$, откуда получаем, что

$$\nu(\xi, \eta) \geq \min\{m, n\} = \min\{\nu(\xi, \eta), \nu(\eta, \zeta)\}.$$

Переходя к отрицательным степеням двойки, получаем (7).

2. Пусть $\xi^n = (0, \dots, \xi_k^n, \xi_{k+1}^n, \dots)$ есть последовательность Коши в пространстве H . Тогда для любого $\varepsilon = 2^{-m}$ существует такое N , что при $n \geq N$, $j \geq N$ выполнено $\rho(\xi^n, \xi^j) \leq 2^{-m}$. Согласно определению метрики получаем, что при $k < m$ соответствующие координаты совпадают: $\xi_k^n = \xi_k^j := x_k^0$. Тем самым определена последовательность $\xi_k^0 \in H$, которая и является пределом последовательности ξ^n .

3. Непосредственно следует из определений. \square

Замечание 1. Топология, порожденная заданной метрикой, на каждом H_m совпадает с топологией декартова произведения E_k , $k \geq m$, как топологических пространств с дискретной топологией. В частности, каждое E_k естественным образом вкладывается в H и метрика (6) порождает на E_k дискретную топологию.

При этом на всем пространстве H топология τ_{sh} , порожденная метрикой (6), сильнее топологии, индуцированной топологией декартова произведения (2) пространств с дискретной топологией. Действительно, в каждом E_k выберем точку $e_k \neq 0$ и зададим последовательность

$$\xi_k^n = \begin{cases} e_{-n}, & k = -n, \\ 0, & k \neq -n \end{cases}.$$

Эта последовательность сходится к 0 в топологии декартова произведения \hat{H} множеств с дискретной топологией, но не сходится в заданной метрике.

Замечание 2. При любом $p > 1$ формула $\rho_p(\xi, \eta) = p^{-v(\xi, \eta)}$ также задает неархимедову метрику на H , свойства этих метрик при разных p аналогичны.

Метрика, для которой выполнено неравенство (7), называется неархимедовой или ультраметрикой. Как известно ([6; 7]), геометрия пространств с неархимедовой метрикой обладает свойствами, нетипичными для евклидовых пространств. Для примера отметим следующие.

Любая точка шара в пространстве с неархимедовой метрикой является его центром.

Каждый шар в пространстве с неархимедовой метрикой распадается на непересекающиеся шары заданного меньшего радиуса.

1.2. Градуированные группы и sharp-метрики на них. Покажем, что градуировка на множествах с алгебраической структурой позволяет задать на них sharp-метрику.

Напомним некоторые известные факты и обозначения.

Пусть задана последовательность коммутативных групп G_k , $k \in \mathbb{Z}$. Через $\prod_{k \in \mathbb{Z}} G_k$ обозначается прямое декартово произведение групп G_k , т. е. множество последовательностей $\{g_k\}$, $g_k \in G_k$, с операцией покоординатного сложения.

Прямой суммой групп G_k называется множество $\bigoplus G_k$, состоящее из последовательностей $\{g_k\}$, $g_k \in G_k$, у которых только конечное множество элементов g_k отлично от нуля. Это множество является подгруппой в $\prod_{k \in \mathbb{Z}} G_k$.

Пусть G – коммутативная группа и G_0 – ее подгруппа. Классом смежности элемента $x \in G$ называется множество

$$[x] = \{y = x + x_0 : x_0 \in G_0\}.$$

На множестве классов смежности корректно определена операция сложения, это множество называется факторгруппой G по G_0 и обозначается G/G_0 .

Фундаментальной областью при действии G_0 на G называется подмножество $E \subset G$, содержащее по одному элементу из каждого класса эквивалентности.

По определению фундаментальной области, отображение

$$E \ni x \rightarrow [x] \in G/G_0$$

является биекцией, пусть $\varphi : G/G_0 \rightarrow E$ есть обратное отображение. Зададим биекцию

$$G \ni x \rightarrow (\varphi([x]), x - \varphi([x])) \in E \times G_0$$

и разложение $x = \varphi([x]) + \{x - \varphi([x])\}$.

Отметим, что эта биекция зависит от выбора фундаментальной области и в общем случае не является изоморфизмом групп.

В особом случае, когда в G существует подгруппа E , изоморфная G/G_0 , эта подгруппа является фундаментальной областью и при таком выборе фундаментальной области сложение в G переходит в покоординатное сложение в прямой сумме.

Перейдем к рассмотрению градуированных групп. Как было сказано выше, коммутативную группу L будем называть градуированной группой, если задана двусторонняя последовательность подгрупп L_m , $m \in \mathbb{Z}$, такая, что $L_m \supset L_{m+1}$ и $L = \bigcup_{m \in \mathbb{Z}} L_m$.

Выделим случай, когда выполнено условие

$$L_\infty := \bigcap_{m \in \mathbb{Z}} L_m = \{0\}. \quad (9)$$

Как было отмечено выше, в случае градуированной группы существуют биекции $L_m \sim \sim L_m/L_{m+1} \times L_{m+1}$, т. е. выполнено условие (5) и на ней определен порядок отличия элементов $\nu(x, y)$, который не зависит от выбора биекции и может быть введен непосредственно с использованием алгебраической структуры по формуле

$$\nu(x, y) = \min\{m : x - y \in L_m\} = \nu(x - y, 0).$$

При этом считаем, что $\nu(x, y) = +\infty$, если

$$x - y \in L_\infty = \bigcap_{m \in \mathbb{Z}} L_m.$$

В частности, $\nu(x, x) = +\infty$.

Поэтому при выполнении условия (9) формула

$$\rho_L(x, y) = p^{-\nu(x-y, 0)} \quad (10)$$

задает метрику на L и структура группы приводит к появлению дополнительных свойств sharp-метрики.

Функция $q : L \rightarrow \mathbb{R}^+$ на группе L , удовлетворяющая условиям

$$q(x + y) \leq q(x) + q(y); \quad q(0) = 0 \iff x = 0;$$

называется *нормированием*. Чтобы подчеркнуть аналогию с нормой, иногда используют обозначение $\|x\| := q(x)$. Однако отметим, что, в случае, когда L является также векторным пространством, нормирование не является нормой.

В случае градуированной группы функция

$$q(x) = \rho_L(x, 0) = p^{-\nu(x, 0)}$$

является нормированием, для которого выполняется *усиленное неравенство треугольника*:

$$q(x + y) \leq \max\{q(x), q(y)\}.$$

Из общих свойств нормирования следует

Лемма 1. *В градуированной группе L операции сложения и вычитания непрерывны в топологии, порожденной sharp-метрикой.*

Подгруппы L_m образуют базу окрестностей точки 0, а множества вида $x + L_m$ образуют базу окрестностей точки x .

1.3. Разложение элементов градуированной группы в ряды. Градуированная группа L с введенной метрикой в общем случае является неполным метрическим пространством. Пополнение этого пространства также является градуированной группой, получим описание такого пополнения.

Для каждого k выберем фундаментальную область $E_k \subset L_k$ при действии L_{k+1} на L_k .

Пусть H есть градуированное множество, построенное по последовательности фундаментальных областей E_k по описанному выше правилу, т. е. $H = \bigcup_m H_m$, где H_m есть множество всех последовательностей вида

$$\xi = (0, \dots, 0, \eta_m, \eta_{m+1}, \dots), \quad \eta_k \in E_k.$$

На этом множестве определена sharp-метрика ρ_H , относительно которой H является полным метрическим пространством. Все элементы последовательности ξ принадлежат группе L , что позволяет отождествить последовательность из H с (формальным) рядом из элементов группы L и считать, что

$$\xi = \sum_m \eta_k, \quad \eta_k \in E_k \subset L_k \subset L.$$

При этом финитным последовательностям вида (8) соответствуют частичные суммы

$$S_n = \sum_m^n \eta_k$$

ряда, являющиеся элементами из L . Согласно теореме 1.1 любой формальный ряд такого вида сходится в H с sharp -метрикой.

Пусть $\varphi_k : L_k/L_{k+1} \rightarrow E_k$ есть биекция, определяемая выбором фундаментальной области. Напомним, что с ее помощью строится биекция $\psi_k : L_k \rightarrow E_k \times L_{k+1}$, действующая на $x \in L_k$ по формуле

$$\psi_k(x) = (\varphi_k([x]_k), l_{k+1}(x)),$$

где $l_{k+1}(x) = x - \varphi_k([x]_k) \in L_{k+1}$.

Если $x \in L$, то $x \in L_m$ при некотором m , и тогда биекция ψ_m ставит в соответствие элементу $x \in L_m$ пару $(\varphi_m([x]_m), l_{m+1}(x))$, где $[x]_m \in G_m$, $\varphi_m([x]_m) \in E_m$, $l_{m+1}(x) = x - \varphi_m([x]_m) \in L_{m+1}$.

Далее, биекция ψ_{m+1} между L_{m+1} и декартовым произведением $E_{m+1} \times L_{m+2}$ ставит в соответствие элементу $l_{m+1}(x)$ пару $(\varphi_{m+1}([l_{m+1}(x)]_{m+1}), l_{m+2}(x))$, где $l_{m+2}(x) = l_{m+1}(x) - \varphi_{m+1}([l_{m+1}(x)]_{m+1}) \in L_{m+2}$. Продолжая описанный процесс, получаем последовательность $\pi(x) := (0, \dots, 0, x_m, x_{m+1}, \dots)$, $x_k \in E_k$, принадлежащую H_m .

Согласно сказанному выше, при $x \in L_m$ будем отождествлять последовательность $\pi(x)$ с рядом из элементов группы L и считать, что

$$\pi(x) = \sum_m^{\infty} x_k, \quad x_k \in E_k \subset L_k \subset L. \quad (11)$$

Теорема 1.2. Пусть L есть градуированная группа, в которой выполнено условие (9) и задана sharp -метрика ρ_L . Тогда множество H всех рядов вида (11) с sharp -метрикой ρ_H есть пополнение L как топологической группы.

Доказательство. То, что множество H всех рядов вида (11) с sharp -метрикой ρ_H является полным метрическим пространством и в нем любой ряд вида (11) сходится следует из теоремы 1.1.

Покажем, что построенное отображение $\pi : L \rightarrow H$ является изометрическим вложением и образ $\pi(L)$ есть всюду плотное подмножество в H , т. е. H есть пополнение L как метрического пространства. Если $\pi(x) = \pi(y)$, то $x - y \in L_k$ при всех k , т. е. $x - y \in L_{\infty}$. Поэтому в силу условия (9) отображение π инъективно, т. е. является вложением.

Согласно построению, если $x \in L_m \setminus L_{m+1}$, то при представлении (11) получаем, что $x_m \neq 0$. Отсюда следует, что

$$\rho_L(x, y) = \rho_H(\pi(x), \pi(y)),$$

т. е. вложение π является изометрическим.

Множество конечных сумм

$$x = \sum_m^N \eta_k, \quad \eta_k \in E_k \subset L_k \subset L,$$

принадлежит L и всюду плотно в H . Для таких x по построению получаем, что $\pi(x) = x$, откуда следует, что образ $\pi(L)$ всюду плотен в H .

Опишем операцию сложения $\hat{+}$ на множестве H , порожденную сложением в L , при которой построенное отображение $\pi : L \rightarrow H$ является непрерывным изометрическим гомоморфизмом групп. Пусть

$$\xi = \sum_{k=m}^{\infty} x_k \in H, \quad \eta = \sum_{k=m}^{\infty} y_k \in H. \quad (12)$$

На подмножестве $\pi(L)$ операция сложения рядов $\hat{+}$ определяется исходя из равенства

$$\pi(x) \hat{+} \pi(y) := \pi(x + y).$$

Покажем, что эта операция продолжается до операции сложения рядов из H .

Заметим, что частичные суммы этих рядов

$$X^n = \sum_{k=m}^n x_k, \quad Y^n = \sum_{k=m}^n y_k$$

являются элементами из L .

Если x и y из L и $\pi(x) = \xi, \pi(y) = \eta$, в силу непрерывности сложения получаем, что

$$\xi \hat{+} \eta = \lim_{n \rightarrow \infty} \pi(X^n + Y^n). \quad (13)$$

Суммы в правой части (13) определены в случае произвольных рядов из H , поскольку частичные суммы рядов являются элементами из L . Покажем, что предел в правой части (13) существует при всех ξ и η из H и эта формула задает операцию сложения $\hat{+}$ в H .

Результат почленного сложения

$$S_n := X^n + Y^n = \sum_{k=m}^n (x_k + y_k)$$

не является элементом из H , так как может оказаться, что $x_k + y_k \notin E_k$.

Поэтому, чтобы получить явное описание операции $\hat{+}$ на конечных суммах, нужно найти $\pi(S_n)$, т. е. построить разложение

$$S^n = \sum_k z_k^n, \quad z_k^n \in E_k.$$

Элемент $x_m + y_m$ принадлежит L_m и представляется в виде $x_m + y_m = \eta_m + l_{m+1}$, где $\eta_m \in E_m, l_{m+1} \in L_{m+1}$. Полагаем $z_m^n = \eta_m$, а слагаемое l_{m+1} переносим в следующий разряд.

Сумма $x_{m+1} + y_{m+1} + l_{m+1} \in L_{m+1}$ представляется в виде $x_{m+1} + y_{m+1} + l_{m+1} = \eta_{m+1} + l_{m+2}$, где $\eta_{m+1} \in E_{m+1}, l_{m+2} \in L_{m+2}$. Полагаем $z_{m+1}^n = \eta_{m+1}$, а слагаемое l_{m+2} переносим в следующий разряд.

Далее сумма $x_{m+2} + y_{m+2} + l_{m+2} \in L_{m+2}$ представляется в виде $x_{m+2} + y_{m+2} + l_{m+2} = \eta_{m+2} + l_{m+3}$, где $\eta_{m+2} \in E_{m+2}, l_{m+3} \in L_{m+3}$. Полагаем $z_{m+2}^n = \eta_{m+2}$, а слагаемое l_{m+3} переносим в следующий разряд. В результате получаем представление

$$S_n = \sum_{k=m}^{\infty} z_k^n, \quad z_k^n \in E_k.$$

Аналогично получаем, что

$$S_{n+j} = \sum_{k=m}^{\infty} z_k^{n+j}, \quad z_k^{n+j} \in E_k.$$

Поскольку

$$S_{n+j} - S_n = \sum_{k=n+1}^{n+j} (x_k + y_k) \in L_{n+1},$$

при переходе от S_n к S_{n+j} в полученном представлении изменяются только слагаемые с номерами $k > n$, т. е. $z_k^{n+j} = z_k^n := z_k$ при $n \geq k$. Тем самым определена последовательность $z_k \in E_k$ и из сказанного следует, что при $n \rightarrow \infty$ последовательность $\pi(S_n)$ сходится в смысле shap -метрики к ряду $z = \sum_{m}^{\infty} z_k$, т. е. $\xi \hat{+} \eta = z \in H_m$. Таким образом операция $\hat{+}$ задана корректно, множества H_m являются подгруппами в H при введенной операции и, следовательно, H является градуированной группой. \square

Замечание 3. Ввиду биекции между E_k и $G_k = L_k/L_{k+1}$ иногда вместо последовательности, элементы которой принадлежат E_k , удобнее использовать последовательность из соответствующих элементов факторгрупп G_k . При такой реализации множества H на нем определена операция покоординатного сложения, относительно которой H является группой, однако в общем случае при отображении π сложение в L не переходит в покоординатное сложение.

Особым является случай, когда при всех k группа L_k разлагается в прямую сумму своих подгрупп: $L_k = E_k \oplus L_{k+1}$. Тогда подгруппа E_k является фундаментальной областью и при таком выборе фундаментальных областей сложение $\hat{+}$ совпадает с покоординатным сложением в H .

Замечание 4. Если $L_{\infty} \neq \{0\}$, то фактор-группа L/L_{∞} градуирована последовательностью факторгрупп L_m/L_{∞} , для такой градуировки выполнено условие вида (9) и для факторгруппы выполнены утверждения, полученные с использованием этого условия.

Если $L_{\infty} \neq \{0\}$, то формула (10) задает полуметрику на L и при отображении π из равенств $\pi(x) = \pi(y)$ следует, что $x - y \in L_{\infty}$. Аналогично, в группе L остальные равенства также выполняются «с точностью до элемента из L_{∞} ». В частности, у последовательности частичных сумм заданного ряда существует много разных пределов и разность двух таких пределов принадлежит L_{∞} .

В ряде примеров используются группы, градуированные односторонней последовательностью подгрупп, т. е. имеющие вид

$$L = \bigcup_0^{+\infty} L_m,$$

где $L_m \supset L_{m+1}$ и $L = L_0$. Очевидно, что полученные утверждения с соответствующей корректировкой справедливы и в такой ситуации.

1.4. Sharp-метрики на градуированных кольцах и векторных пространствах. Пусть L – кольцо, аддитивная группа которого градуирована последовательностью подгрупп L_m , и пусть $q(x)$ есть нормирование, порожденное такой градуировкой.

Кольцо L называется *градуированным последовательностью подгрупп* L_m , если выполнено условие

$$q(xy) \leq q(x)q(y).$$

Из этого условия следует, что операция умножения в L непрерывна относительно sharp-метрики.

Кроме того, если $x, y \in L_m$, то произведение xy попадает в подгруппу L_{2m} . При $m \geq 0$ имеем $L_{2m} \subset L_m$ и подгруппа L_m является подкольцом в L . А при отрицательном m подгруппа L_{2m} шире чем L_m и L_m не является подкольцом.

Аналогично случаю операции сложения, операция умножения в L порождает операцию умножения в H , действующую на элементы (12) по формуле

$$\xi \hat{\times} \eta = \lim_{n \rightarrow \infty} \pi(X^n Y^n).$$

Как и в случае сложения, для получения более явного описания этой операции следует найти $\pi(X^n Y^n)$, т. е. построить разложение произведения

$$X^n Y^n = \sum_k z_k^n, \quad z_k^n \in E_k.$$

Поскольку

$$X^n Y^n = \sum x_k \sum y_j = \sum_{l=2m}^{2n} \left(\sum_{k+j=l} x_k y_j \right), \quad x_k y_j \in L_l,$$

аналогично случаю сложения получаем правило вычисления элементов z_k^n , причем при фиксированном k с ростом n также происходит стабилизация элементов z_k^n . В результате получаем утверждение.

Теорема 1.3. *Если L есть градуированное кольцо, то операция умножения в L порождает операцию умножения $\hat{\times}$ на градуированной группе H , это множество является топологическим кольцом относительно введенных операций $\hat{+}$ и $\hat{\times}$ и является пополнением исходного градуированного кольца L .*

Векторное пространство L над полем \mathbb{R} или \mathbb{C} называется градуированным, если задана двусторонняя последовательность подпространств L_m , $m \in \mathbb{Z}$, таких, что $L_m \supset L_{m+1}$ и $L = \bigcup_{m \in \mathbb{Z}} L_m$.

Такое векторное пространство является коммутативной группой по сложению и к нему применимы описанные конструкции и выполнены полученные выше утверждения. В частности, определен порядок $\nu(x, y)$ отличия элементов x и y и формула $\rho_L(x, y) = p^{-\nu(x, y)}$ задает sharp-метрику или полуметрику на L . Здесь появляется дополнительное свойство этой метрики: если число $\lambda \neq 0$, то $\rho_L(\lambda x, \lambda y) = \rho_L(x, y)$.

Как показано выше, задание для каждой фактор-группы $G_k = L_k/L_{k+1}$ биекций с подмножеством E_k в L_k позволяет построить вложение градуированной группы в соответствующее пространство H , для которого выполнены утверждения из теоремы 1.2. В случае векторных пространств факторгруппа $G_k = L_k/L_{k+1}$ является векторным пространством, которое может быть линейно вложено в L_k как векторное подпространство E_k , дополнительное к L_{k+1} . При таком вложении имеет место разложение $L_k \sim E_k \oplus L_{k+1}$ в прямую сумму векторных подпространств, что приводит к выполнению дополнительных свойств вложения π .

Теорема 1.4. *Пусть L есть градуированное векторное пространство, выполнено условие (9), для каждого L_k задано разложение векторных пространств $L_k = E_k \oplus L_{k+1}$ в прямую сумму*

векторных пространств и H есть векторное пространство, состоящее из всех (формальных) рядов вида (11), с операциями почленного сложения и умножения на число.

1. Пространство H является градуированным векторным пространством, полным относительно соответствующей *sharp*-метрики, в котором операция сложения непрерывна.

2. Произведение λx непрерывно зависит от x при фиксированном λ , но нет непрерывной зависимости этого произведения от λ и, следовательно, градуированное векторное пространство H с *sharp*-топологией не является топологическим векторным пространством.

3. Построенное выше отображение $\pi : L \rightarrow \hat{H}$ является линейным вложением, сохраняющим соответствующие *sharp*-метрики, что позволяет отождествлять L с векторным подпространством $\pi(L)$, всюду плотным в H , и следовательно, H есть пополнение L .

Доказательство. 1. Это общее свойство пространств H .

2. Пусть $x^n \rightarrow x^0$ в H . Если $\lambda \neq 0$, то $\rho_H(\lambda x^n, \lambda x^0) = \rho_H(x^n, x^0) \rightarrow 0$. Если $\lambda = 0$, то сходимость очевидна.

Пусть $\lambda_n \rightarrow \lambda_0, \lambda_n \neq \lambda_0$ и $x \neq 0$. Тогда

$$\rho_H(\lambda_n x, \lambda_0 x) = \rho_H((\lambda_n - \lambda_0)x, 0) = \rho_H(x, 0) \not\rightarrow 0.$$

3. Линейность отображения π возникает ввиду того, что при разложении векторного пространства в прямую сумму проекции на обе составляющие являются линейными отображениями. \square

2. ПРИМЕРЫ ИЗ АНАЛИЗА

2.1. Кольцо p -адических чисел. Наиболее известный пример градуированного множества и неархимедовой метрики, порожденной градуировкой, встречается в p -адическом анализе (см. [8;9]).

Покажем, что каждое положительное целое число $p \geq 2$ порождает градуировку кольца \mathbb{Q} и общая конструкция, описанная в первой части, приводит к построению p -адических чисел.

Для рационального числа $x \neq 0$ однозначно определено целое $\nu = \nu(x)$, такое, что имеется представление в виде $x = p^{\nu} \frac{t_1}{t_2}$, где t_1 и t_2 есть целые числа, такие, что t_1 не делится на p , а t_2 взаимно просто с p .

Очевидно представление $x = p^k \frac{s_1}{s_2}$, где s_1 и s_2 не делятся на p , которое в случае простого p и есть требуемое. В случае составного p требуемый вид получаем с помощью умножения числителя и знаменателя на такой множитель, что знаменатель делится на p . Например, при $p = 10$ получаем

$$\frac{15}{14} = \frac{15 \cdot 5}{14 \cdot 5} = \frac{1}{10} \cdot \frac{75}{7}.$$

Пусть

$$x = p^m \frac{t_1}{t_2}, \quad y = p^k \frac{s_1}{s_2},$$

где $k \geq m$, есть соответствующие представления двух рациональных чисел. Тогда

$$x + y = p^m \frac{t_1 s_2 + t_2 s_1 p^{k-m}}{t_2 s_2}$$

и в полученной дроби знаменатель взаимно прост с p . При этом может оказаться, что числитель делится на p и тогда в полученном представлении числа $x + y$ степень p не меньше, чем m , т. е.

$$\nu(x + y) \geq \min\{\nu(x), \nu(y)\}.$$

Для каждого $m \in \mathbb{Z}$ зададим множество рациональных чисел

$$L_m = \{x \in \mathbb{Q} : \nu(x) \geq m\}.$$

Из полученного неравенства следует, что L_m является подгруппой в \mathbb{Q} .

Если $x \neq 0$ и $m > \nu(x)$, то $x \notin L_m$, откуда следует, что $\bigcap_m L_m = \{0\}$.

Таким образом, каждому p соответствует своя градуировка группы \mathbb{Q} , для которой выполнено условие (9).

Согласно общему подходу, на \mathbb{Q} определено нормирование, заданное формулой $q(x) = p^{-v(x)}$, для него выполнено неравенство

$$q(x+y) \leq \max\{q(x), q(y)\}.$$

Для такого нормирования выполнено

$$q(xy) \leq q(x)q(y)$$

и, следовательно, \mathbb{Q} с указанной градуировкой является градуированным кольцом.

Заметим, что при простом p имеется равенство

$$q(xy) = q(x)q(y).$$

При составном p в представлении

$$xy = p^{m+k} \frac{t_1 s_1}{t_2 s_2}$$

может оказаться, что число $t_1 s_2$ делится на p , за счет чего получается строгое неравенство.

Теорема 1.3 содержит явное описание пополнения градуированного кольца в общем случае, применим эту теорему в рассматриваемом примере градуировки кольца \mathbb{Q} .

Пусть $\mathbb{Z}/p\mathbb{Z}$ есть кольцо вычетов по модулю p , будем отождествлять его с кольцом

$$\Gamma_p = \{0, 1, 2, \dots, p-1\}$$

с операциями сложения и умножения по модулю p . Для $n \in \mathbb{Z}$ через $\gamma(x) \in \Gamma_p$ обозначим вычет целого числа x по модулю p .

При составном p элемент $k \neq 0 \in \Gamma_p$ обратим, если k взаимно просто с p . В частности, если число p простое, то Γ_p является полем.

Лемма 2. При заданной градуировке кольца \mathbb{Q} все факторгруппы $G_k = L_k/L_{k+1}$ изоморфны группе Γ_p и являются кольцами.

Доказательство. Числа вида

$$p^k \gamma, \gamma \in \Gamma_p \tag{14}$$

принадлежат разным классам эквивалентности, т. е. определено вложение $\Gamma_p \ni \gamma \rightarrow [p^k \gamma] \in G_k$. В частности, L_{k+1} есть класс эквивалентности, соответствующий $\gamma = 0$.

Покажем, что если $x \in L_k \setminus L_{k+1}$, то класс $[x]$ содержит элемент вида (14).

Пусть $x = p^k \frac{t_1}{t_2}$ и $\gamma_1 = \gamma(t_1), \gamma_2 = \gamma(t_2)$. Поскольку t_2 взаимно просто с p , в кольце Γ_p существует обратный γ_3 к γ_2 .

Тогда

$$\frac{t_1}{t_2} = \frac{t_1 \gamma_3}{t_2 \gamma_3} = \frac{\gamma_1 \gamma_3 + p l_1}{1 + p l_2}, \quad l_1, l_2 \in \mathbb{Z}.$$

Поэтому

$$x - p^k \gamma_1 \gamma_3 = p^{k+1} \frac{l_1 + l_2}{1 + p l_2} \in L_{k+1}. \quad \square$$

Согласно общему подходу, чтобы построить соответствующее градуированное кольцо H и требуемое вложение $\pi: \mathbb{Q} \rightarrow H$, нужно для каждого G_k задать подмножество $E_k \subset L_k$, содержащее по одному представителю из каждого класса эквивалентности.

Для факторгруппы G_k естественно взять множество из наиболее простых представителей вида (14), т. е. положить

$$E_k = \{0, p^k, 2p^k, \dots, (p-1)p^k\} \subset L_k.$$

Тогда, согласно общей конструкции, искомое кольцо есть $H = \bigcup_{m \in \mathbb{Z}} H_m$, где H_m есть множество рядов вида

$$\xi = \sum_m a_k p^k, a_k \in \{0, 1, 2, \dots, p-1\}. \tag{15}$$

Построенное множество H с введенными операциями сложения и умножения обозначается \mathbb{Q}_p и называется *кольцом p -адических чисел*.

При вложении $\pi : \mathbb{Q} \rightarrow H$, построенном по общему правилу, получаем, что $\pi(x) = x$ для x , представимых в виде конечных сумм, в частности, это выполнено для положительных целых чисел. Остальным рациональным числам соответствуют бесконечные ряды, в частности, отрицательным целым числам соответствуют бесконечные ряды (с положительными коэффициентами). Например,

$$\pi(-1) = \sum_0^{\infty} (p-1)p^k.$$

Здесь общее правило задания сложения рядов из H становится наглядным. Например, при $p = 10$ в случае конечных сумм это обычное сложение десятичных разложений, которое осуществляется с помощью «переносов в старшие разряды» и применимо к бесконечным суммам.

При $m \geq 0$ подгруппы H_m являются подкольцами. Кольцо целых чисел \mathbb{Z} всюду плотно вкладывается в кольцо H_0 , которое обозначается \mathbb{Z}_p и называется *кольцом целых p -адических чисел*.

Таким образом, кольцо \mathbb{Q}_p получаем как частный случай градуированных колец H , и из теоремы 1.3 получаем известные свойства \mathbb{Q}_p .

Предложение 2.1. *Множество \mathbb{Q}_p , состоящее из формальных рядов вида (15), с введенными операциями сложения и умножения является кольцом. Если число p простое, то \mathbb{Q}_p является полем.*

Кольцо \mathbb{Q}_p с заданной неархимедовой sharp-метрикой является полным метрическим пространством, в котором алгебраические операции непрерывны и все формальные ряды (15) сходятся.

Отображение π есть изометрическое всюду плотное вложение кольца рациональных чисел \mathbb{Q} в \mathbb{Q}_p , откуда следует, что \mathbb{Q}_p есть пополнение \mathbb{Q} относительно заданной неархимедовой метрики. Кольцо \mathbb{Q}_p может быть введено также как пополнение кольца конечных сумм вида (15).

Заметим, что в большинстве источников \mathbb{Q}_p вводится именно как пополнение множества рациональных чисел, после чего строится представление p -адических чисел с помощью рядов.

2.2. Градуировки на \mathbb{Z} и системы остаточных классов. На кольце целых чисел \mathbb{Z} существует много разных градуировок следующего вида. Пусть задана возрастающая последовательность натуральных чисел M_k , $M_0 = 1$, и L_k есть подкольцо, состоящее из чисел, кратных M_k . Вложение $L_k \subset L_{k-1}$ имеет место, если M_k делится на M_{k-1} . При таком условии получаем градуировку кольца \mathbb{Z} односторонней последовательности подколец L_k , причем числа M_k представляются в виде произведения $M_k = p_1 p_2 \dots p_k$, где $p_j > 1$, и $\bar{p} = (p_1, p_2, \dots)$ есть произвольная последовательность натуральных чисел. Такие градуировки использовались, например, в [6].

При заданной градуировке рассматриваемого вида фактор-кольцо $G_k = L_{k-1}/L_k$ изоморфно кольцу Γ_{p_k} , наиболее естественной фундаментальной областью в L_{k-1} является подмножество

$$E_k = \{0, M_{k-1}, 2M_{k-1}, \dots, (p_k - 1)M_{k-1}\}. \quad (16)$$

После такого выбора фундаментальных областей, согласно общей конструкции, определено вложение $\pi : \mathbb{Z} \rightarrow H_{\bar{p}}$, где соответствующее множество $H_{\bar{p}}$ состоит из формальных рядов вида

$$\sum_{k=1}^{+\infty} a_k M_{k-1}, a_k \in \{0, 1, \dots, p_k - 1\} = \Gamma_{p_k}, \quad (17)$$

причем сложение и умножение в \mathbb{Z} порождают соответствующие операции сложения $\hat{+}$ и умножения $\hat{\times}$ в $H_{\bar{p}}$.

Теорема 2.1. *Каждая градуировка кольца \mathbb{Z} и выбор фундаментальных областей (16) порождает всюду плотное вложение кольца \mathbb{Z} в кольцо $H_{\bar{p}}$, состоящее из рядов вида (17), полное относительно соответствующей sharp-метрики.*

В частности, кольцо \mathbb{Z}_p целых p -адических чисел получаем, если $p_j \equiv p$ и $M_k = p^k$.

Как было сказано в замечании 3, множество $H_{\bar{p}}$ естественным образом отождествляется с множеством последовательностей со значениями в кольцах Γ_{p_k} , причем множество таких по-

следовательностей является полным в смысле sharp-метрики и является кольцом с операциями покомпонентного сложения и умножения.

Теорема 2.2. *Отображение $\gamma : \mathbb{Z} \rightarrow H_{\overline{p}}$, действующее по формуле*

$$\gamma : x \rightarrow (\gamma_1(x), \gamma_2(x), \dots), \quad \gamma_j(x) = x \pmod{p_j} \in \Gamma_{p_j},$$

т. е. ставящее в соответствие числу x последовательность из остатков от деления на числа p_j , задает вложение \mathbb{Z} в $H_{\overline{p}}$, при котором операции в \mathbb{Z} переходят в покомпонентные операции в $H_{\overline{p}}$.

Доказательство. То, что операции в \mathbb{Z} переходят в покомпонентные операции в $H_{\overline{p}}$ следует из свойств отображений γ_j . Отображение γ является вложением, так как при каждом заданном x последовательность $\gamma_j = x \pmod{p_j} \in \Gamma_j = \{0, 1, \dots, p_j - 1\}$ стабилизируется: если $p_j > x$, то $\gamma_j(x) = x$, откуда следует инъективность. \square

Особый интерес представляет случай, когда числа p_j при рассматриваемой градуировке попарно взаимно простые. Особые свойства вложения γ вытекают из следующего известного утверждения.

Китайская теорема об остатках. *Если натуральные числа p_1, p_2, \dots, p_k попарно взаимно просты, то для любых целых r_1, r_2, \dots, r_k , где $0 \leq r_j < p_j$, найдется число $N \in \mathbb{Z}$, которое при делении на p_j дает остаток r_j .*

Два числа $N_1 \in \mathbb{Z}$ и N_2 дают при делении на числа p_j те же остатки r_j тогда и только тогда, когда разность $N_1 - N_2$ делится на произведение $M_k = p_1 p_2 \dots p_k$.

Эта теорема эквивалентна утверждению, что если числа p_1, p_2, \dots, p_k попарно взаимно просты и $M_k = p_1 p_2 \dots p_k$, то кольцо Γ_{M_k} разлагается в прямую сумму

$$\Gamma_{M_k} = \bigoplus_1^k \Gamma_{p_j} \quad (18)$$

и такое разложение задается формулой

$$\Gamma_{M_k} \ni x \rightarrow (\gamma_1(x), \gamma_2(x), \dots, \gamma_k(x)),$$

где $\gamma_j(x) = x \pmod{p_j} \in \Gamma_{p_j}$. В частности,

$$\Gamma_{M_k} = \Gamma_{M_{k-1}} \bigoplus \Gamma_{p_k}$$

и переход от разложения $\Gamma_{M_{k-1}}$ к разложению Γ_{M_k} заключается в добавлении прямого слагаемого Γ_{p_k} .

Согласно китайской теореме об остатках, отображение, ставящее в соответствие числу x набор из k остатков

$$\{\gamma_1(x) = x \pmod{p_1}, \gamma_2(x) = x \pmod{p_2}, \dots, \gamma_k(x) = x \pmod{p_k}\}$$

определяет число с точностью до слагаемого, кратного M_k . Поэтому оно задает биекцию между такими последовательностями длины k и целыми числами x из заданного диапазона длины M_k , например, множество $\{x \in \mathbb{Z} : 0 \leq x < M_k\}$.

Такая биекция используется в так называемых *системах остаточных классов* [10].

Если числа x и y принадлежат заданному диапазону, то первые k остатков для чисел $x + y$ и xy вычисляются по формулам

$$\gamma_j(x + y) = \gamma_j(x) + \gamma_j(y), \quad \gamma_j(xy) = \gamma_j(x)\gamma_j(y).$$

Если при этом сумма $x + y$ и произведение xy также принадлежат заданному диапазону, то числа $x + y$ и xy однозначно определяются по полученным k остаткам. Поскольку каждый из этих остатков вычисляется независимо, без переносов в следующие разряды, числа $\gamma_j(x + y)$ и $\gamma_j(xy)$ можно вычислять одновременно на разных устройствах. Такое распараллеливание операций позволяет существенно ускорить процесс вычислений и на этом свойстве системы остаточных классов основана работа некоторых вычислительных машин [10].

Недостатком такого подхода является то, что он применим только в случае, когда результаты всех действий лежат в заданном диапазоне. Поэтому вычисления с помощью системы остаточных классов используются в специализированных вычислительных машинах, ориентированных на

определенные классы задач, когда известен диапазон рассматриваемых величин и быстродействие имеет принципиальное значение.

Если число p простое, то множество чисел $p\mathbb{Z}$, делящихся на p , является максимальным идеалом в \mathbb{Z} и факторкольцо $\mathbb{Z}/p\mathbb{Z}$ есть поле Γ_p . Верно и обратное – каждый максимальный идеал в \mathbb{Z} есть множество чисел, делящихся на некоторое простое число. Поэтому можно отметить особый случай, когда все простые числа занумерованы в последовательность p_j . Тогда преобразование γ ставит в соответствие числу x функцию на пространстве максимальных идеалов, значение которой в точке p_j принадлежит полю Γ_{p_j} , причем γ является гомоморфизмом колец.

Описанное представление чисел с помощью остатков используется и в ряде других вопросов. Например, в криптографии последовательность остатков $x(\bmod p_1), x(\bmod p_2), \dots, x(\bmod p_k)$ используется для шифрования чисел x из диапазона (p_k, M_k) , так как эти остатки однозначно определяют число x , но его невозможно найти по остаткам, не зная чисел p_j .

Отметим также, что разложение (18) помогает исследовать строение колец \mathbb{Q}_p и \mathbb{Z}_p при составных p . Например, разложение $\Gamma_6 = \Gamma_2 \oplus \Gamma_3$ приводит к тому, что кольцо \mathbb{Q}_6 является прямой суммой двух полей p -адических чисел:

$$\mathbb{Q}_6 = \mathbb{Q}_2 \oplus \mathbb{Q}_3.$$

Такие разложения были получены еще создателем теории p -адических чисел К. Гензелем [9].

2.3. Градуировка пространства гладких функций и формула Тейлора. Для пояснения постановки обсуждаемого ниже вопроса напомним, что, например, после введения понятия сходимости последовательности в курсе анализа следуют пояснения, что на рассматриваемом множестве имеется структура топологического пространства и понятие сходимости использует именно эту структуру. Но при этом обычно не обсуждается, какие структуры используются в формулировке утверждений о формуле Тейлора и рядах Тейлора.

Покажем, что в этих утверждениях также используются градуировка, sharp-метрика и конструкции, описанные в общем случае в первой части.

Напомним основные утверждения о формуле Тейлора, содержащиеся во всех учебниках по математическому анализу (например, [11]). Пусть F_0 есть пространство функций, определенных и бесконечно дифференцируемых на отрезке $[-1, 1]$. Для функции $f \in F_0$ определен ряд Тейлора

$$f \sim \sum_{k=0}^{\infty} a_k x^k, \quad \text{где } a_k = \frac{1}{k!} f^{(k)}(0). \quad (19)$$

Полином Тейлора $P_n(f; x) = \sum_{k=0}^n a_k x^k$ есть частичная сумма этого ряда, а функция $R_n(f; x) = f(x) - P_n(f; x)$ называется n -м остаточным членом.

Заметим, что ряд Тейлора может расходиться при всех $x \neq 0$, а в случае сходимости в некоторой окрестности нуля его сумма может не совпадать с $f(x)$. В частности, при заданном $x \neq 0$ остаточный член $R_n(f; x)$ может не стремиться к нулю при $n \rightarrow \infty$.

Основные утверждения о формуле Тейлора заключаются в следующем.

Утверждение 2.1. *Остаточный член $R_n(f; x)$ при $x \rightarrow 0$ есть бесконечно малая, порядок которой не меньше, чем $n + 1$:*

$$R_n(f; x) = O(x^{n+1}).$$

При выполнении этого утверждения говорят, что ряд (19) сходится к f асимптотически.

Утверждение 2.2. *Для остаточного члена формулы Тейлора выполнена оценка*

$$|R_n(f; x)| \leq \frac{M}{(n+1)!} |x|^{n+1},$$

где

$$M = \max_{-1 \leq t \leq 1} |f^{(n+1)}(t)|.$$

Естественная градуировка пространства F_0 задается последовательностью подпространств

$$F_m = \{f : f(x) = O(x^m) \text{ при } x \rightarrow 0\} = \{f(x) = x^m g(x) : g \in F_0\},$$

где m – неотрицательные целые числа. Здесь пересечение $F_\infty = \bigcap_{m \geq 0} F_m$ состоит из так называемых *плоских* функций, у которых все производные в точке 0 есть нули. Согласно теореме 1.4, градуировка позволяет поставить в соответствие функции f последовательность

$$\pi(f) = \{\hat{f} = (0, \dots, \xi_0, \xi_1, \xi_2, \dots)\}, \text{ где } \xi_k \in F_k/F_{k+1},$$

состоящую из элементов фактор-пространств.

В рассматриваемом примере каждое фактор-пространство $G_k = F_k/F_{k+1}$ изоморфно \mathbb{R} и его можно вложить в F_k как одномерное подпространство E_k , состоящее из наиболее простых представителей класса эквивалентности – функций вида $ax^k, a \in \mathbb{R}$. При таком выборе, согласно конструкции из теоремы 1.4, элемент $\psi(f)$, соответствующий функции $f \in F_0$, есть формальный ряд вида $\psi(f) = \sum_{k=0}^{\infty} a_k x^k$, где по построению получаем, что $R_m \in F_{m+1}$. Кроме того, здесь

$$a_k = \lim_{x \rightarrow 0} \frac{f(x) - P_{k-1}(x)}{x^k}.$$

Вычисляя предел по правилу Лопиталя, получаем, что $a_k = \frac{f^{(k)}(0)}{k!}$. Таким образом, в этом примере формальный ряд $\pi(f)$, построенный по общему правилу разложения, есть ряд Тейлора функции f .

В этом примере порядок функции f – число $\nu(f) = \min\{k : a_k \neq 0\}$, и это есть порядок стремления к нулю значений $f(x)$ при $x \rightarrow 0$, нормирование задается формулой $q(f) = 2^{-\nu(f)}$, и формула $\rho(f, g) = q(f - g)$ задает метрику на фактор-пространстве F_0/F_∞ и полуметрику на F_0 . По построению $\rho(f, P_m(f)) \rightarrow 0$ при $m \rightarrow \infty$, т. е. утверждение 2.1 говорит о том, что формальный ряд Тейлора сходится к f в этой метрике.

Таким образом, конструкция ряда Тейлора и утверждение 2.1 есть частный случай общих построений и утверждений из первой части, причем асимптотическая сходимость ряда Тейлора есть сходимость в смысле *sharp*-полуметрики на градуированном пространстве F_0 .

Отметим также известную теорему Бореля (см. [12]), которая утверждает, что для любой последовательности a_k (в том числе сколь угодно быстро возрастающей) существует такая бесконечно дифференцируемая функция f , что (19) есть ее ряд Тейлора. Отсюда получаем, что теорема Бореля эквивалентна утверждению, что фактор-пространство F_0/F_∞ с *sharp*-метрикой является полным метрическим пространством.

Как и в общем случае, в описанном способе разложения элементов и построения ряда есть произвол при выборе представителя из каждого класса эквивалентности – элемента фактор-пространства $G_k = F_k/F_{k+1}$. При другом выборе представителей получаем разложение функции в асимптотический ряд по другой системе функций. Например, если $h(x)$ – гладкая функция, такая, что $h(0) = 0, h'(x) \neq 0$, то в качестве представителей можно выбрать функции вида $b[h(x)]^k$ и f разлагается в асимптотически сходящийся ряд

$$f(x) \sim \sum_{k=0}^{\infty} b_k h(x)^k.$$

Если $t = h(x), x = g(t)$, где g есть функция, обратная к h , то такое разложение эквивалентно разложению в ряд по степеням t сложной функции $f(g(t))$. Заметим, что коэффициенты b_k вычисляются по известной формуле Фаа ди Брьюно.

Рассмотрим теперь, какая структура использована в утверждении 2.2.

На каждом из пространств F_m определена своя норма, заданная формулой

$$\|f\|_m = \max_{|x| \leq 1} \frac{|f(x)|}{|x|^m},$$

т. е. F_0 является градуированным нормированным пространством. С помощью этих норм неравенство из утверждения 2.2 можно записать в виде

$$\|R_m(f)\|_{m+1} \leq \frac{1}{(m+1)!} \|f^{(m+1)}\|_0.$$

Таким образом, если утверждение 2.1 говорит, что остаточный член $R_m(f)$ принадлежит F_{m+1} и использует только градуировку, то в более сильном утверждении 2.2 используется норма на подпространстве F_{m+1} и указан шар в смысле соответствующей нормы, в котором лежит остаточный член.

Приведенные рассуждения можно перенести на более широкое пространство. Пусть F есть пространство функций, представимых при $x \neq 0$ в виде

$$f(x) = \frac{h(x)}{g(x)},$$

где $h \in F_0, g \in F_0 \setminus F_\infty$, и $g(x) \neq 0$ при $x \neq 0$. Естественная градуировка пространства F задается аналогичной двусторонней последовательностью подпространств

$$F_m = \{f : f(x) = O(x^m) \text{ при } x \rightarrow 0\}, m \in \mathbb{Z}.$$

При отрицательных k , как и выше при положительных, фактор-пространство $G_k = F_k/F_{k+1}$ можно вложить в F_k как одномерное подпространство E_k , состоящее из наиболее простых представителей класса эквивалентности – функций вида $ax^k, a \in \mathbb{R}$. При таком выборе представителей, согласно конструкции из теоремы 1.4, элемент $\psi(f)$, соответствующий функции $f \in F_m$, есть ряд вида

$$\psi(f) = \sum_{k=m}^{\infty} a_k x^k,$$

где по построению получаем, что $R_n(x) := f(x) - \sum_{k=m}^n a_k x^k \in F_{n+1}$ и

$$a_k = \lim_{x \rightarrow 0} \frac{f(x) - P_{k-1}(x)}{x^k}.$$

Таким образом, при отрицательных m формальный ряд $\pi(f)$, построенный по общему правилу разложения, описанному в первой части, есть ряд Лорана функции f , а включение $R_n \in F_{n+1}$ есть аналог утверждения 2.1 в случае рядов Лорана и равносильно сходимости в неархимедовой метрике, порожденной градуировкой пространства F .

Аналогичная ситуация возникает при анализе гладких функций нескольких переменных. Для примера рассмотрим пространство $F_0(D)$ гладких функций двух переменных, определенных в замкнутом круге $D = \{x \in \mathbb{R}^2 : \|x\| \leq 1\}$ на плоскости. На этом пространстве имеется аналогичная градуировка из подпространств

$$F_m = \{f : f(x) = O(\|x\|^m)\},$$

которая позволяет ввести sharp-метрику.

В отличие от предыдущего примера, здесь фактор-пространство $G_1 = F_1/F_2$ двумерно и вкладывается в F_1 в виде подпространства E_1 , состоящего из линейных функций $a_{11}x_1 + a_{21}x_2$. Фактор-пространство $G_2 = F_2/F_3$ трехмерно и вкладывается в F_2 в виде подпространства E_2 , состоящего из квадратичных форм

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{33}x_2^2.$$

Фактор-пространство $G_3 = F_3/F_4$ имеет размерность 10 и вкладывается в F_3 в виде подпространства E_3 , состоящего из однородных полиномов третьей степени и т. д.

При такой конструкции формальный ряд $\psi(f)$, построенный согласно общей конструкции, есть ряд Тейлора гладкой функции двух переменных

$$\pi(f) = \sum_{k=0}^{\infty} d^k(x),$$

где слагаемые

$$d^k(x) = \sum_{|\alpha|=k} \frac{1}{\alpha_1! \alpha_2!} \frac{\partial^{|\alpha|} f(0)}{\partial^{\alpha_1} x_1 \partial^{\alpha_2} x_2} x_1^{\alpha_1} x_2^{\alpha_2}, \quad \alpha = (\alpha_1, \alpha_2), \quad |\alpha| = \alpha_1 + \alpha_2$$

есть дифференциалы функции f в точке 0.

Классическое утверждение об асимптотической сходимости частичных сумм такого ряда к f есть сходимость в смысле соответствующей sharp-метрики и является частным случаем общих утверждений.

Заключение. Приведенные примеры показывают, что структура градуированного пространства имеет весьма общий характер, так как фактически используется даже в классических конструкциях анализа. Это позволяет высказать предположение, что такая структура и утверждения, полученные в первой части работы, могут оказаться полезными при рассмотрении других задач. Более детальные исследования пространств с градуированной топологией и их использование при решении конкретных задач предполагается провести в последующих работах.

Отметим также, что разложения действительных чисел по заданному основанию p имеют вид

$$\sum_m^{\infty} \frac{a_k}{p^k}, \quad a_k \in E := \{0, 1, 2, \dots, p-1\},$$

и они внешне похожи на разложения (15) p -адических чисел. Однако такие разложения имеют принципиально другую природу, в частности, множество таких разложений устроено не так, как множество действительных чисел, поскольку имеются числа, которым соответствуют два разложения. Поэтому возникает вопрос о том, какую структуру имеет множество таких разложений.

Литература

1. Бурбаки Н. Топологические векторные пространства. М., 1959.
2. Олвер Ф. Введение в асимптотические методы и специальные функции. М.: Наука, 1978. 386 с.
3. Васильева А. В., Бутузов В. Ф. Асимптотические разложения решений сингулярно возмущенных уравнений. М., 1973.
4. Colombeau J. F. *New generalized functions and multiplication of distributions*. Amsterdam: North-Holland, 1984. 374 p.
5. Delcroix A., Scarpalezos D. Sharp Topologies on (C, E, P) -Algebras // *Nonlinear Theory of Generalized Function*, Chaptman & Hall, Research Notes of Mathematics. 1999. Vol. 401. P. 165–174.
6. Хренников А. Ю. Неархимедов анализ и его приложения. М.: Физматлит, 2003. 216 с.
7. Радына А. Я., Радына Я. М., Радына Я. В. Пачаткі неархімедавага аналізу: дапам. для студентаў мех.-мат. фак. Мінск: БДУ, 2010. 111 с.
8. Владимиров В. С., Волович И. В., Зеленов Е. И. P -адический анализ и математическая физика. М.: Наука, 1994. 352 с.
9. Каток С. В. P -адический анализ в сравнении с вещественным. М.: МЦНМО, 2004. 112 с.
10. Акушский И. Я., Юдицкий Д. И. Машинная арифметика в остаточных классах. М.: Советское радио, 1968. 439 с.
11. Кротов В. Г. Математический анализ: учеб. пособие. Минск: БГУ, 2017.
12. Нарасимхан Р. Анализ на действительных и комплексных многообразиях. М.: Мир, 1971. 232 с.

References

1. Bourbaki N. *Topological Vector Spaces*. Moscow, 1959 (in Russian).
2. Olver F. *Introduction to Asymptotic Methods and Special Functions*. Moscow, Nauka, 1978 (in Russian).
3. Vasil'eva A. V., Butuzov V. F. *Asymptotic Expansions of Solutions of Singularly Perturbed Equations*. Moscow, 1973 (in Russian).
4. Colombeau J. F. *New Generalized Functions and Multiplication of Distributions*. Amsterdam, North-Holland, 1984, 374 p.
5. Delcroix A., Scarpalezos D. Sharp Topologies on (C, E, P) -Algebras. *Nonlinear Theory of Generalized Functions, Chapman & Hall, Research Notes in Mathematics*, 1999, vol. 401, pp. 165–174.

6. Khrennikov A. Y. *Non-Archimedean Analysis and Its Applications*. Moscow, Fizmatlit, 2003 (in Belarussian).
7. Radyina A. Y., Radyina Y. M., Radyina Y. V. *Introduction to Non-Archimedean Analysis: Textbook for Students of Mechanics and Mathematics Faculty*. Minsk, BSU, 2010, 111 p. (in Belarussian).
8. Vladimirov V. S., Volovich I. V., Zelenov E. I. *P-adic Analysis and Mathematical Physics*. Moscow, Nauka, 1994, 352 p. (in Russian).
9. Katok S. V. *P-adic Analysis in Comparison with Real Analysis*. Moscow, MCCME, 2004, 112 p. (in Russian).
10. Akushsky I. Y., Yuditsky D. I. *Computer Arithmetic in Residue Classes*. Moscow, Soviet Radio, 1968, 439 p. (in Russian).
11. Krotov V. G. *Mathematical Analysis: Textbook for BSU*. Minsk, 2017 (in Russian).
12. Narasimhan R. *Analysis on Real and Complex Manifolds*. Moscow, Mir, 1971 (in Russian).

УДК 517.538.52+517.538.53

СУЩЕСТВОВАНИЕ И ЯВНЫЙ ВИД НЕЛИНЕЙНЫХ АППРОКСИМАЦИЙ ЭРМИТА–ЧЕБЫШЁВА

А. П. Старовойтов, И. В. Кругликов

Гомельский государственный университет им. Ф. Скорины, Гомель, Беларусь
e-mail: svoitov@gsu.by, igor.v.kruglikov@gmail.com

Поступила: 19.01.2025

Исправлена: 31.01.2025

Принята: 23.05.2025

Ключевые слова: ряды по многочленам Чебышёва, аппроксимации Эрмита–Паде, аппроксимации Паде–Чебышёва, тригонометрические аппроксимации Эрмита–Якоби, нелинейные аппроксимации Эрмита–Чебышёва.

Аннотация. В работе найдены достаточные условия существования тригонометрических аппроксимаций Эрмита–Якоби системы функций, являющихся суммами сходящихся рядов Фурье. Опираясь на эти результаты, установлены достаточные условия, при которых существуют нелинейные аппроксимации Эрмита–Чебышёва систем функций, представимых рядами Фурье по многочленам Чебышёва первого и второго рода. При выполнении найденных условий получены явные формулы для числителей и знаменателей тригонометрических аппроксимаций Эрмита–Якоби и нелинейных аппроксимаций Эрмита–Чебышёва первого и второго рода указанных систем функций.

EXISTENCE AND EXPLICIT FORM OF NONLINEAR HERMITE–CHEBYSHEV APPROXIMATIONS

A. P. Starovoitov, I. V. Kruglikov

F. Skorina Gomel State University, Gomel, Belarus
e-mail: svoitov@gsu.by, igor.v.kruglikov@gmail.com

Received: 19.01.2025

Revised: 31.01.2025

Accepted: 23.05.2025

Keywords: series in Chebyshev polynomials, Hermite–Padé approximations, Padé–Chebyshev approximations, trigonometric Hermite–Jacobi approximations, nonlinear Hermite–Chebyshev approximations.

Abstract. In this paper, sufficient conditions for the existence of trigonometric Hermite–Jacobi approximations of a system of functions that are sums of convergent Fourier series are found. Based on these results, sufficient conditions are established under which nonlinear Hermite–Chebyshev approximations of systems of functions representable by Fourier series in Chebyshev polynomials of the first and second kind exist. When the found conditions are met, explicit formulas are obtained for the numerators and denominators of trigonometric Hermite–Jacobi approximations and nonlinear Hermite–Chebyshev approximations of the first and second kind of the specified systems of functions.

1. Введение

Пусть система $\mathbf{f}^{ch1} = (f_1^{ch1}, \dots, f_k^{ch1})$ состоит из функций, представимых рядами Фурье по многочленам Чебышёва $T_n(x) = \cos(n \arccos x)$ первого рода

$$f_j^{ch1}(x) = \frac{a_0^j}{2} + \sum_{l=1}^{\infty} a_l^j T_l(x), \quad j = 1, \dots, k, \quad (1)$$

с действительными коэффициентами, которые сходятся при всех $x \in [-1, 1]$. Через \mathbb{Z}_+^k обозначим множество всех k -мерных мультииндексов $\vec{m} = (m_1, \dots, m_k)$, являющихся упорядоченным набором k целых неотрицательных чисел. Число $m := m_1 + \dots + m_k$ назовем порядком мультииндекса $\vec{m} = (m_1, \dots, m_k)$.

Зафиксируем индекс $n \in \mathbb{Z}_+^1$ и мультииндекс $\vec{m} = (m_1, \dots, m_k) \in \mathbb{Z}_+^k$ и рассмотрим аналог задачи Эрмита–Паде для \mathbf{f}^{ch1} (см. [1, гл. 4, §1, задача А]):

Задача 1.1 (A^{ch1}). Для системы \mathbf{f}^{ch1} найти такой тождественно не равный нулю многочлен $Q_m^{ch1}(x; \mathbf{f}^{ch1}) = Q_{n, \vec{m}}^{ch1}(x; \mathbf{f}^{ch1}) = \sum_{p=0}^m u_p T_p(x)$ и такие многочлены $P_j^{ch1}(x; \mathbf{f}^{ch1}) = P_{n_j, n, \vec{m}}^{ch1}(x; \mathbf{f}^{ch1}) = \sum_{p=0}^{n_j} v_p^j T_p(x)$, $n_j = n + m - m_j$, чтобы для $j = 1, \dots, k$

$$Q_m^{ch1}(x; \mathbf{f}^{ch1}) f_j^{ch1}(x) - P_j^{ch1}(x; \mathbf{f}^{ch1}) = \sum_{l=n+m+1}^{\infty} \tilde{a}_l^j T_l(x).$$

Определение 1.2. Если пара (Q_m^{ch1}, P^{ch1}) , где $P^{ch1} = (P_1^{ch1}, \dots, P_k^{ch1})$, является решением задачи A^{ch1} , то рациональные дроби

$$\pi_j^{ch1}(x; \mathbf{f}^{ch1}) = \pi_{n_j, n, \vec{m}}^{ch1}(x; \mathbf{f}^{ch1}) = \frac{P_j^{ch1}(x; \mathbf{f}^{ch1})}{Q_m^{ch1}(x; \mathbf{f}^{ch1})}, \quad j = 1, \dots, k,$$

будем называть *линейными аппроксимациями Эрмита–Чебышёва 1-го рода* для мультииндекса (n, \vec{m}) и системы \mathbf{f}^{ch1} .

Определение 1.3. *Нелинейными аппроксимациями Эрмита–Чебышёва 1-го рода* для мультииндекса (n, \vec{m}) и системы \mathbf{f}^{ch1} назовем рациональные дроби

$$\hat{\pi}_j^{ch1}(x; \mathbf{f}^{ch1}) = \hat{\pi}_{n_j, n, \vec{m}}^{ch1}(x; \mathbf{f}^{ch1}) = \frac{\hat{P}_j^{ch1}(x; \mathbf{f}^{ch1})}{\hat{Q}_m^{ch1}(x; \mathbf{f}^{ch1})},$$

где многочлены $\hat{Q}_m^{ch1}(x; \mathbf{f}^{ch1})$, $\hat{P}_j^{ch1}(x; \mathbf{f}^{ch1})$ ($n_j = n + m - m_j$), степени которых не превышают соответственно m и n_j , подобраны так, чтобы

$$f_j^{ch1}(x) - \frac{\hat{P}_j^{ch1}(x; \mathbf{f}^{ch1})}{\hat{Q}_m^{ch1}(x; \mathbf{f}^{ch1})} = \sum_{l=n+m+1}^{\infty} \hat{a}_l^j T_l(x), \quad j = 1, \dots, k.$$

При $k = 1$ основные свойства линейных и нелинейных аппроксимаций Эрмита–Чебышёва (в этом случае их называют линейными и нелинейными аппроксимациями Паде–Чебышёва; дополнительно о терминологии см. [2]) описаны достаточно подробно (прежде всего см. [2–4] и приведенную там литературу, а также [5–13]). Например, известно, что линейная аппроксимация Паде–Чебышёва всегда существует, но, вообще говоря, не единственна. Нелинейная аппроксимация Паде–Чебышёва не всегда существует, но в случае существования всегда единственна. Аналогичные свойства справедливы и для аппроксимаций Эрмита–Чебышёва [9; 11]. Имеются примеры систем функций \mathbf{f}^{ch1} , для которых нелинейные аппроксимации Эрмита–Чебышёва существуют, но не являются линейными аппроксимациями Эрмита–Чебышёва (см. [2; 9; 11; 14; 15]).

Рассмотрим теперь другой тип аппроксимаций Эрмита–Чебышёва. Предположим, что система $\mathbf{f}^{ch2} = (f_1^{ch2}, \dots, f_k^{ch2})$ состоит из функций, представимых рядами Фурье по многочленам Чебышёва $U_n(x) = \frac{1}{\sqrt{1-x^2}} \sin(n \arccos x)$ второго рода

$$f_j^{ch2}(x) = \sum_{l=1}^{\infty} b_l^j U_l(x), \quad j = 1, \dots, k, \quad (2)$$

с действительными коэффициентами, которые сходятся при всех $x \in [-1, 1]$. Если вместо рядов (1) взять ряды (2), то конструкции, аналогичные предыдущим, приводят к линейным и нелинейным аппроксимациям Эрмита–Чебышёва 2-го рода. Постановка задачи Эрмита–Паде для рядов (2) следующая:

Задача 1.4 (A^{ch2}). Найти многочлен $Q_m^{ch2}(x; \mathbf{f}^{ch2}) = Q_{n, \vec{m}}^{ch2}(x; \mathbf{f}^{ch2})$, $\deg Q_m^{ch2} \leq m$, тождественно не равный нулю, и многочлены $P_j^{ch2}(x; \mathbf{f}^{ch2}) = P_{n_j, n, \vec{m}}^{ch2}(x; \mathbf{f}^{ch2})$, $\deg P_j^{ch2} \leq n_j$, $n_j = n + m - m_j$, чтобы для $j = 1, \dots, k$

$$Q_m^{ch2}(x; \mathbf{f}^{ch2}) f_j^{ch2}(x) - P_j^{ch2}(x; \mathbf{f}^{ch2}) = \sum_{l=n+m+1}^{\infty} \tilde{b}_l^j U_l(x).$$

Определение 1.5. Если пара (Q_m^{ch2}, P^{ch2}) , где $P^{ch2} = (P_1^{ch2}, \dots, P_k^{ch2})$, является решением задачи A^{ch2} , то рациональные дроби

$$\pi_j^{ch2}(x; \mathbf{f}^{ch2}) = \pi_{n_j, n, \vec{m}}^{ch2}(x; \mathbf{f}^{ch2}) = \frac{P_j^{ch2}(x; \mathbf{f}^{ch2})}{Q_m^{ch2}(x; \mathbf{f}^{ch2})}, \quad j = 1, \dots, k,$$

будем называть *линейными аппроксимациями Эрмита–Чебышёва 2-го рода* для мультииндекса (n, \vec{m}) и системы \mathbf{f}^{ch2} .

Определение 1.6. *Нелинейными аппроксимациями Эрмита–Чебышёва 2-го рода для мультииндекса (n, \vec{m}) и системы \mathbf{f}^{ch2} назовем алгебраические рациональные функции*

$$\widehat{\pi}_j^{ch2}(x; \mathbf{f}^{ch2}) = \widehat{\pi}_{n_j, n, \vec{m}}^{ch2}(x; \mathbf{f}^{ch2}) = \frac{\widehat{P}_j^{ch2}(x; \mathbf{f}^{ch2})}{\widehat{Q}_m^{ch2}(x; \mathbf{f}^{ch2})},$$

где многочлены $\widehat{Q}_m^{ch2}(x; \mathbf{f}^{ch2})$, $\widehat{P}_j^{ch2}(x; \mathbf{f}^{ch2})$ ($n_j = n + m - m_j$), степени которых не превышают соответственно m и n_j , подобраны так, чтобы

$$f_j^{ch2}(x) - \frac{\widehat{P}_j^{ch2}(x; \mathbf{f}^{ch2})}{\widehat{Q}_m^{ch2}(x; \mathbf{f}^{ch2})} = \sum_{l=n+m+1}^{\infty} \widehat{b}_l^j U_l(x), \quad j = 1, \dots, k.$$

Рассмотрим теперь две системы $\mathbf{f}^1 = (f_1^1, \dots, f_k^1)$, $\mathbf{f}^{t1} = (f_1^{t1}, \dots, f_k^{t1})$ степенных и тригонометрических рядов

$$f_j^1(z) = \frac{a_0^j}{2} + \sum_{l=1}^{\infty} a_l^j z^l, \quad f_j^{t1}(x) = \frac{a_0^j}{2} + \sum_{l=1}^{\infty} a_l^j \cos lx,$$

ассоциированных с системой \mathbf{f}^{ch1} , и две системы $\mathbf{f}^2 = (f_1^2, \dots, f_k^2)$, $\mathbf{f}^{t2} = (f_1^{t2}, \dots, f_k^{t2})$ степенных и тригонометрических рядов

$$f_j^2(z) = \sum_{l=1}^{\infty} b_l^j z^l, \quad f_j^{t2}(x) = \sum_{l=1}^{\infty} b_l^j \sin lx,$$

ассоциированных с системой \mathbf{f}^{ch2} .

При $k = 1$ в [2], опираясь на свойства операторов Фабера, установлено, что вопрос о существовании нелинейных аппроксимаций Паде–Чебышёва 1-го рода решается с помощью известных результатов о классических аппроксимациях Паде соответствующего ассоциированного степенного ряда. Результаты, полученные в данной работе, позволяют сделать вывод о том, что вопрос о существовании нелинейных аппроксимаций Эрмита–Чебышёва 1-го рода аналогичным образом решается с помощью известных результатов об аппроксимациях Эрмита–Паде соответствующей ассоциированной системы степенных рядов. Основные теоремы работы получены нами без привлечения теории операторов Фабера, а предложенный метод решения поставленной задачи оказался применимым и для аппроксимаций Эрмита–Чебышёва 2-го рода. Если доказательство основного результата в [2] проиллюстрировать схемой: *аппроксимации Паде степенного ряда \iff преобразования Фабера \iff аппроксимации Паде–Чебышёва*, то схема наших рассуждений выглядит так: *аппроксимации Эрмита–Паде системы степенных рядов \iff тригонометрические аппроксимации Эрмита–Якоби \iff аппроксимации Эрмита–Чебышёва*. Роль преобразований Фабера в нашей схеме рассуждений занимают тригонометрические аппроксимации Эрмита–Якоби, которым посвящен отдельный параграф.

В дальнейшем будем рассматривать только нелинейные аппроксимации Эрмита–Чебышёва, а основной задачей исследований данной работы является нахождение условий на коэффициенты рядов (1) и (2), при которых нелинейные аппроксимации Эрмита–Чебышёва 1-го и 2-го рода существуют. В случае их существования будем искать явный вид таких аппроксимаций. Доказательство основных теорем работы существенно опирается на установленную в [9–11] связь между нелинейными аппроксимациями Эрмита–Чебышёва и соответствующими тригонометрическими аппроксимациями Эрмита–Якоби.

Описанию условий существования и единственности линейных аппроксимации Эрмита–Чебышёва 1-го и 2-го рода посвящена другая наша работа [16].

2. Аппроксимации Эрмита–Паде и Эрмита–Якоби

Приведем некоторые известные факты теории аппроксимаций Эрмита–Паде и Эрмита–Якоби, которые понадобятся в дальнейшем.

Пусть $\mathbf{f} = (f_1, \dots, f_k)$ – набор из k степенных рядов

$$f_j(z) = \sum_{l=0}^{\infty} f_l^j z^l, \quad j = 1, \dots, k, \tag{3}$$

с комплексными коэффициентами. Зафиксируем $n \in \mathbb{Z}_+^1$ и $\vec{m} = (m_1, \dots, m_k) \in \mathbb{Z}_+^k$ и рассмотрим хорошо известную задачу Эрмита–Паде [1, гл. 4, §1, задача А]:

Задача 2.1А. Найти тождественно не равный нулю многочлен $Q_m(z; \mathbf{f}) = Q_{n, \vec{m}}(z; \mathbf{f})$, $\deg Q_m \leq m$, и многочлены $P_j(z; \mathbf{f}) = P_{n_j, n, \vec{m}}(z; \mathbf{f})$, $\deg P_j \leq n_j$, $n_j = n + m - m_j$, чтобы для $j = 1, \dots, k$

$$Q_m(z; \mathbf{f})f_j(z) - P_j(z; \mathbf{f}) = O(z^{n+m+1}). \quad (4)$$

Здесь и далее под $O(z^p)$ понимаем степенной ряд вида $c_1 z^p + c_2 z^{p+1} + \dots$

Определение 2.2. Если многочлены $Q_m(z; \mathbf{f}), P_1(z; \mathbf{f}), \dots, P_k(z; \mathbf{f})$ являются решением задачи А (решение задачи А всегда существует [1]), то рациональные дроби

$$\pi_j(z; \mathbf{f}) = \pi_{n_j, n, \vec{m}}(z; \mathbf{f}) = \frac{P_j(z; \mathbf{f})}{Q_m(z; \mathbf{f})}, \quad j = 1, \dots, k,$$

называют *аппроксимациями Эрмита–Паде* для мультииндекса (n, \vec{m}) и системы \mathbf{f} .

Для системы экспонент дроби $\{\pi_j(z; \mathbf{f})\}_{j=1}^k$ впервые введены в рассмотрение Ш. Эрмитом в работе [17], посвященной доказательству трансцендентности числа e . В случае произвольной системы \mathbf{f} условиями (4) они определяются, вообще говоря, не однозначно [1; 18]. Особый интерес представляют системы функций \mathbf{f} , для которых $\{\pi_j(z; \mathbf{f})\}_{j=1}^k$ определяются однозначно для любого мультииндекса (n, \vec{m}) . Такими системами являются, например, совершенные системы [1, гл. 4, §1]. Существуют системы \mathbf{f} , отличные от совершенных, для которых $\{\pi_j(z; \mathbf{f})\}_{j=1}^k$ также определяются однозначно [18; 19].

Введем в рассмотрение кратные аналоги дробей К. Якоби [20] (подробнее см. [21]).

Определение 2.3. Рациональные функции вида

$$\hat{\pi}_j(z; \mathbf{f}) = \hat{\pi}_{n_j, n, \vec{m}}(z; \mathbf{f}) = \frac{\hat{P}_j(z; \mathbf{f})}{\hat{Q}_m(z; \mathbf{f})}, \quad j = 1, \dots, k,$$

где алгебраические многочлены $\hat{Q}_m(z; \mathbf{f}) = \hat{Q}_{n, \vec{m}}(z; \mathbf{f})$, $\hat{P}_j(z; \mathbf{f}) = \hat{P}_{n_j, n, \vec{m}}^j(z; \mathbf{f})$ имеют степени соответственно не выше m и n_j , $n_j = n + m - m_j$, будем называть *аппроксимациями Эрмита–Якоби* для мультииндекса (n, \vec{m}) и системы \mathbf{f} , если

$$f_j(z) - \frac{\hat{P}_j(z; \mathbf{f})}{\hat{Q}_m(z; \mathbf{f})} = O(z^{n+m+1}).$$

В отличие от аппроксимаций Эрмита–Паде аппроксимации Эрмита–Якоби могут не существовать [11].

Введем новые обозначения. Для индекса $n \in \mathbb{Z}_+^1$ и мультииндекса $\vec{m} = (m_1, \dots, m_k)$, $m \neq 0$, рассмотрим определитель

$$H_{n, \vec{m}}(\mathbf{f}) = \det \begin{pmatrix} H^1 \\ H^2 \\ \vdots \\ H^k \end{pmatrix} = \begin{vmatrix} f_{n-m_1+1}^1 & f_{n-m_1+2}^1 & \cdots & f_{n_1}^1 \\ \cdots & \cdots & \cdots & \cdots \\ f_n^1 & f_{n+1}^1 & \cdots & f_{n+m-1}^1 \\ \cdots & \cdots & \cdots & \cdots \\ f_{n-m_k+1}^k & f_{n-m_k+2}^k & \cdots & f_{n_k}^k \\ \cdots & \cdots & \cdots & \cdots \\ f_n^k & f_{n+1}^k & \cdots & f_{n+m-1}^k \end{vmatrix},$$

при $m_j \neq 0$ состоящий из блоков

$$H^j = \begin{pmatrix} f_{n-m_j+1}^j & f_{n-m_j+2}^j & \cdots & f_{n_j}^j \\ f_{n-m_j+2}^j & f_{n-m_j+3}^j & \cdots & f_{n_j+1}^j \\ \cdots & \cdots & \cdots & \cdots \\ f_n^j & f_{n+1}^j & \cdots & f_{n+m-1}^j \end{pmatrix},$$

расположенных друг над другом. При $l < 0$ считаем, что $f_l^j = 0$. По определению считаем, что при $m_j = 0$ определитель $H_{n, \vec{m}}(\mathbf{f})$ не содержит блок H^j . При $k = 1$ определитель $H_{n, \vec{m}}(\mathbf{f})$ является хорошо известным определителем Адамара (см. [4]).

В [22] (см. также [11]) доказан кратный аналог теоремы Якоби [4; 20]: если для мультииндекса (n, \vec{m}) и системы функций \mathbf{f} определитель $H_{n, \vec{m}}(\mathbf{f}) \neq 0$, то аппроксимации Эрмита–Якоби $\{\widehat{\pi}_j(z; \mathbf{f})\}_{j=1}^k$ существуют, определяются единственным образом и каждая из них тождественно совпадает с соответствующей аппроксимацией Эрмита–Паде, т. е.

$$\widehat{\pi}_j(z; \mathbf{f}) = \pi_j(z; \mathbf{f}), \quad j = 1, \dots, k.$$

В таком случае, опираясь на результаты работы [18], можно описать явный вид числителей и знаменателя дробей $\{\widehat{\pi}_j(z; \mathbf{f})\}_{j=1}^k$. Для этого введем необходимые обозначения.

При $m_j \neq 0$ к матрице H^j добавим в качестве последнего столбца столбец

$$\left(f_{n_j+1}^j \quad f_{n_j+2}^j \quad \dots \quad f_{n+m}^j \right)^T.$$

В результате получим матрицу F^j порядка $m_j \times (m+1)$. Располагая F^j друг над другом согласно своего номера, построим новую матрицу порядка $m \times (m+1)$

$$F_{n, \vec{m}}(\mathbf{f}) = \left[F^1 \quad F^2 \quad \dots \quad F^k \right]^T := \begin{pmatrix} f_{n-m_1+1}^1 & f_{n-m_1+2}^1 & \dots & f_{n_1+1}^1 \\ f_{n-m_1+2}^1 & f_{n-m_1+3}^1 & \dots & f_{n_1+2}^1 \\ \dots & \dots & \dots & \dots \\ f_n^1 & f_{n+1}^1 & \dots & f_{n+m}^1 \\ \dots & \dots & \dots & \dots \\ f_{n-m_k+1}^k & f_{n-m_k+2}^k & \dots & f_{n_k+1}^k \\ f_{n-m_k+2}^k & f_{n-m_k+3}^k & \dots & f_{n_k+2}^k \\ \dots & \dots & \dots & \dots \\ f_n^k & f_{n+1}^k & \dots & f_{n+m}^k \end{pmatrix}.$$

Определим также функциональные матрицы-строки порядка $1 \times (m+1)$

$$E(z) = \left(z^m \quad z^{m-1} \quad \dots \quad z \quad 1 \right),$$

$$E_{m_j}(z) = \left(\sum_{l=0}^{n-m_j} f_l^j z^{m+l} \quad \sum_{l=0}^{n-m_j+1} f_l^j z^{m+l-1} \quad \dots \quad \sum_{l=0}^{n_j} f_l^j z^l \right).$$

Если к матрице $F_{n, \vec{m}}(\mathbf{f})$ в качестве последней строки добавить соответственно строки $E(z)$, $E_{m_j}(z)$ и $\left(f_{n+l}^j \quad f_{n+l+1}^j \quad \dots \quad f_{n+m+l}^j \right)$, то получим квадратные матрицы, определители которых обозначим соответственно через $D(n, \vec{m}; z)$, $G_j(n, \vec{m}; z)$ и $d_{n, \vec{m}, l}^j$.

Если $H_{n, \vec{m}}(\mathbf{f}) \neq 0$, то $F_{n, \vec{m}}(\mathbf{f})$ является матрицей полного ранга и тогда при подходящем выборе нормирующего множителя знаменателя и числителя дробей $\pi_j(z; \mathbf{f})$ и $\widehat{\pi}_j(z; \mathbf{f})$ при $j = 1, \dots, k$ могут быть представлены в виде [18]

$$Q_m(z; \mathbf{f}) = \widehat{Q}_m(z; \mathbf{f}) = D(n, \vec{m}; z) = \det \left[F^1 \quad F^2 \quad \dots \quad F^k \quad E(z) \right]^T, \quad (5)$$

$$P_j(z; \mathbf{f}) = \widehat{P}_j(z; \mathbf{f}) = G_j(n, \vec{m}; z) = \det \left[F^1 \quad F^2 \quad \dots \quad F^k \quad E_{m_j}(z) \right]^T, \quad (6)$$

кроме того,

$$Q_m(z; \mathbf{f}) f_j(z) - P_j(z; \mathbf{f}) = \sum_{l=1}^{\infty} d_{n, \vec{m}, l}^j z^{n+m+l}. \quad (7)$$

Если коэффициенты рядов (3) – действительные числа, то определители $d_{n, \vec{m}, l}^j$ принимают действительные значения, а $Q_m(z; \mathbf{f})$ и $P_j(z; \mathbf{f})$ являются алгебраическими многочленами с действительными коэффициентами [18]. Из (5)–(7) следует, что при $j = 1, \dots, k$

$$f_j(z) - \widehat{\pi}_j(z; \mathbf{f}) = \sum_{l=1}^{\infty} \widehat{d}_{n, \vec{m}, l}^j z^{n+m+l},$$

в которых коэффициенты $\widehat{d}_{n, \vec{m}, l}^j$ – действительные числа.

3. Тригонометрические аппроксимации Эрмита–Паде и Эрмита–Якоби

Пусть $\mathbf{f}^t = (f_1^t, \dots, f_k^t)$ – система тригонометрических рядов

$$f_j^t(x) = \frac{a_0^j}{2} + \sum_{l=1}^{\infty} (a_l^j \cos lx + b_l^j \sin lx), \quad j = 1, \dots, k, \quad (8)$$

с действительными коэффициентами. Считаем, что ряды (8) сходятся при всех $x \in \mathbb{R}$ и каждый ряд определяет функцию, заданную на всей действительной прямой.

Для системы \mathbf{f}^t существует [9] тождественно не равный нулю тригонометрический многочлен $Q_m^t(x; \mathbf{f}^t) = Q_{n, \vec{m}}^t(x; \mathbf{f}^t)$, $\deg Q_m^t \leq m$, и такие тригонометрические многочлены $P_j^t(x; \mathbf{f}^t) = P_{n_j, n, \vec{m}}^t(x; \mathbf{f}^t)$, $\deg P_j^t \leq n_j$, $n_j = n + m - m_j$, что для $j = 1, \dots, k$

$$Q_m^t(x; \mathbf{f}^t) f_j^t(x) - P_j^t(x; \mathbf{f}^t) = \sum_{l=n+m+1}^{\infty} (\tilde{a}_l^j \cos lx + \tilde{b}_l^j \sin lx). \quad (9)$$

Определение 3.1. Если многочлены $Q_m^t(x; \mathbf{f}^t)$, $P_j^t(x; \mathbf{f}^t)$, $j = 1, \dots, k$, удовлетворяют условиям (9), то тригонометрические рациональные дроби

$$\pi_j^t(x; \mathbf{f}^t) = \pi_{n_j, n, \vec{m}}^t(x; \mathbf{f}^t) = \frac{P_j^t(x; \mathbf{f}^t)}{Q_m^t(x; \mathbf{f}^t)}, \quad j = 1, \dots, k,$$

будем называть *тригонометрическими аппроксимациями Эрмита–Паде (совместными аппроксимациями Эрмита–Фурье)* для мультииндекса (n, \vec{m}) и системы \mathbf{f}^t .

Отметим, что условиями (9) тригонометрические аппроксимации Эрмита–Паде определяются, вообще говоря, не однозначно. Задача нахождения условий, при которых они определяются однозначно, подробно исследуется в работах [5; 9; 10].

Определим теперь тригонометрические аналоги аппроксимаций Эрмита–Якоби.

Определение 3.2. Рациональные функции

$$\hat{\pi}_j^t(x; \mathbf{f}^t) = \hat{\pi}_{n_j, n, \vec{m}}^t(x; \mathbf{f}^t) = \frac{\hat{P}_j^t(x; \mathbf{f}^t)}{\hat{Q}_m^t(x; \mathbf{f}^t)}, \quad j = 1, \dots, k,$$

где $\hat{Q}_m^t(x; \mathbf{f}^t) = \hat{Q}_{n, \vec{m}}^t(x; \mathbf{f}^t)$, $\hat{P}_j^t(x; \mathbf{f}^t) = \hat{P}_{n_j, n, \vec{m}}^t(x; \mathbf{f}^t)$ – тригонометрические многочлены, степени которых соответственно не выше m и n_j , $n_j = n + m - m_j$, будем называть *тригонометрическими аппроксимациями Эрмита–Якоби* для мультииндекса (n, \vec{m}) и системы \mathbf{f}^t , если

$$f_j^t(x) - \frac{\hat{P}_j^t(x; \mathbf{f}^t)}{\hat{Q}_m^t(x; \mathbf{f}^t)} = \sum_{l=n+m+1}^{\infty} (\hat{a}_l^j \cos lx + \hat{b}_l^j \sin lx), \quad j = 1, \dots, k.$$

В отличие от тригонометрических аппроксимаций Эрмита–Паде тригонометрические аппроксимации Эрмита–Якоби могут не существовать [2; 5; 14]. Известны примеры систем функций \mathbf{f}^t , для которых тригонометрические аппроксимации Эрмита–Якоби существуют, но не являются тригонометрическими аппроксимациями Эрмита–Паде [9; 11; 23].

Теорема 3.3. Предположим, что мультииндекс (n, \vec{m}) , $m \neq 0$, удовлетворяет условию $n \geq \max\{m_j : 1 \leq j \leq k\}$, а $\mathbf{f}^1 = (f_1^1, \dots, f_k^1)$, $\mathbf{f}^{t1} = (f_1^{t1}, \dots, f_k^{t1})$ – две ассоциированные с \mathbf{f}^{ch1} системы степенных и тригонометрических рядов. Тогда для существования тригонометрических аппроксимаций Эрмита–Якоби $\{\hat{\pi}_j^t(x; \mathbf{f}^{t1})\}_{j=1}^k$ достаточно, чтобы для системы степенных рядов \mathbf{f}^1 выполнялись следующие условия:

- 1) существуют алгебраические аппроксимации Эрмита–Якоби $\{\hat{\pi}_j(z; \mathbf{f}^1)\}_{j=1}^k$;
- 2) каждый степенной ряд f_j^1 системы \mathbf{f}^1 имеет радиус сходимости $R_j^1 > 1$;
- 3) рациональные дроби $\hat{\pi}_j(z; \mathbf{f}^1)$, $j = 1, \dots, k$, не имеют полюсов в $\bar{D} = \{z : |z| \leq 1\}$.

Если для системы \mathbf{f}^1 выполнены условия 1)–3), то при соответствующей нормировке для всех $j = 1, \dots, k$ справедливы равенства:

$$\hat{Q}_m^t(x; \mathbf{f}^{t1}) = \hat{Q}_m(e^{ix}; \mathbf{f}^1) \overline{\hat{Q}_m(e^{ix}; \mathbf{f}^1)} = \sum_{l=0}^m q_l \cos lx,$$

$$\widehat{P}_j^t(x; \mathbf{f}^1) = \operatorname{Re} \left\{ \widehat{P}_j(e^{ix}; \mathbf{f}^1) \overline{\widehat{Q}_m(e^{ix}; \mathbf{f}^1)} \right\} = \sum_{l=0}^{n_j} p_l^j \cos lx,$$

$$f_j^1(x) - \widehat{\pi}_j^t(x; \mathbf{f}^1) = \sum_{l=n+m+1}^{\infty} d_l^j \cos lx,$$

где q_l, p_l^j, d_l^j – действительные числа.

Теорема 3.4. *Предположим, что мультииндекс (n, \vec{m}) , $m \neq 0$, удовлетворяет условию $n \geq \max\{m_j; 1 \leq j \leq k\}$, а $\mathbf{f}^2 = (f_1^2, \dots, f_k^2)$, $\mathbf{f}^1 = (f_1^1, \dots, f_k^1)$ – две ассоциированные с \mathbf{f}^{ch2} системы степенных и тригонометрических рядов. Тогда для существования тригонометрических аппроксимаций Эрмита–Якоби $\{\widehat{\pi}_j^t(x; \mathbf{f}^2)\}_{j=1}^k$ достаточно, чтобы для системы степенных рядов \mathbf{f}^2 выполнялись условия:*

- 1) существуют алгебраические аппроксимации Эрмита–Якоби $\{\widehat{\pi}_j(z; \mathbf{f}^2)\}_{j=1}^k$;
- 2) каждый степенной ряд f_j^2 системы \mathbf{f}^2 имеет радиус сходимости $R_j^2 > 1$;
- 3) рациональные дроби $\widehat{\pi}_j(z; \mathbf{f}^2)$, $j = 1, \dots, k$, не имеют полюсов в $\overline{D} = \{z : |z| \leq 1\}$.

Если для системы \mathbf{f}^2 выполнены условия 1)–3), то при соответствующей нормировке для всех $j = 1, \dots, k$ справедливы равенства:

$$\widehat{Q}_m^t(x; \mathbf{f}^2) = \widehat{Q}_m(e^{ix}; \mathbf{f}^2) \overline{\widehat{Q}_m(e^{ix}; \mathbf{f}^2)} = \sum_{l=0}^m \widehat{q}_l \cos lx, \quad (10)$$

$$\widehat{P}_j^t(x; \mathbf{f}^2) = \operatorname{Im} \left\{ \widehat{P}_j(e^{ix}; \mathbf{f}^2) \overline{\widehat{Q}_m(e^{ix}; \mathbf{f}^2)} \right\} = \sum_{l=0}^{n_j} \widehat{p}_l^j \sin lx, \quad (11)$$

$$f_j^2(x) - \widehat{\pi}_j^t(x; \mathbf{f}^2) = \sum_{l=n+m+1}^{\infty} \widehat{d}_l^j \sin lx, \quad (12)$$

где $\widehat{q}_l, \widehat{p}_l^j, \widehat{d}_l^j$ – действительные числа.

Замечание 3.5. *Условия 2) и 3) в теоремах 3.3 и 3.4 не являются существенными в следующем смысле. Если они не выполнены, например, для \mathbf{f}^1 , то можно перейти к другой системе функций $\mathbf{f}_r^1 = (f_1^1(rz), \dots, f_k^1(rz))$, для которой при достаточно малом r , $0 < r < 1$, они будут выполнены.*

Замечание 3.6. *В формулировках теорем 3.3 и 3.4 предполагается, что мультииндекс удовлетворяет условию $n \geq \max\{m_j; 1 \leq j \leq k\}$. Это означает, что мы рассматриваем только «верхнюю», часть общей тригонометрической таблицы Паде (см. [1, гл. 4, §1]), включая ее главную диагональ. Это условие на мультииндекс позволяет существенно упростить нахождение явных формул для аппроксимаций Эрмита–Паде и поэтому встречается в ряде работ, посвященных данной тематике (см., например, [2; 8; 24–26]).*

Остановимся только на доказательстве теоремы 3.4. Теорема 3.3 доказывается аналогично. Так как выполнено условие 1) и коэффициенты рядов (2) действительные числа, то дроби $\{\widehat{\pi}_j(z; \mathbf{f}^2)\}_{j=1}^k$ определены, их числители и знаменатель являются многочленами с действительными коэффициентами и в некоторой окрестности нуля

$$f_j^2(z) - \widehat{\pi}_j(z; \mathbf{f}^2) = \sum_{l=n+m+1}^{\infty} \widehat{d}_l^j z^l, \quad j = 1, \dots, k. \quad (13)$$

Выполнение условий 2) и 3) позволяет в качестве такой окрестности взять открытый круг с центром в нуле, радиус которого больше 1. Очевидно, что $f_j^2(x) = \operatorname{Im} f_j^2(e^{ix})$. Тогда, полагая в (13) $z = e^{ix}$, а затем приравнявая мнимые части нового равенства, получим

$$f_j^2(x) - \operatorname{Im} \left\{ \widehat{\pi}_j(e^{ix}; \mathbf{f}^2) \right\} = \sum_{l=n+m+1}^{\infty} \widehat{d}_l^j \sin lx. \quad (14)$$

Покажем, что при $n \geq \max\{m_j; 1 \leq j \leq k\}$

$$\widehat{\pi}_j^t(x; \mathbf{f}^2) = \operatorname{Im} \left\{ \widehat{\pi}_j(e^{ix}; \mathbf{f}^2) \right\}. \quad (15)$$

Знаменатель $\widehat{Q}_m(z; \mathbf{f}^2)$ и числитель $\widehat{P}_j(z; \mathbf{f}^2)$ дроби $\widehat{\pi}_j(z; \mathbf{f}^2)$ – алгебраические многочлены с действительными коэффициентами. Предположим, что они представляются в виде

$$\widehat{Q}_m(z; \mathbf{f}^2) = \sum_{l=0}^m \widehat{q}_l z^l, \quad \widehat{P}_j(z; \mathbf{f}^2) = \sum_{l=0}^{n_j} \widehat{p}_l^j z^l.$$

Тогда при $z = e^{ix}$

$$\begin{aligned} \operatorname{Im} \left\{ \widehat{\pi}_j(e^{ix}; \mathbf{f}^2) \right\} &= \frac{1}{2i} \left(\frac{\widehat{P}_j(e^{ix}; \mathbf{f}^2)}{\widehat{Q}_m(e^{ix}; \mathbf{f}^2)} - \overline{\frac{\widehat{P}_j(e^{ix}; \mathbf{f}^2)}{\widehat{Q}_m(e^{ix}; \mathbf{f}^2)}} \right) = \\ &= \frac{1}{2i} \frac{\sum_{l=0}^{n_j} \widehat{p}_l^j e^{ilx} \sum_{s=0}^m \widehat{q}_s e^{-isx} - \sum_{l=0}^{n_j} \widehat{p}_l^j e^{-ilx} \sum_{s=0}^m \widehat{q}_s e^{isx}}{\sum_{s=0}^m \widehat{q}_s e^{isx} \sum_{l=0}^m \widehat{q}_l e^{-ilx}} = \\ &= \frac{\sum_{l=0}^{n_j} \sum_{s=0}^m \widehat{p}_l^j \widehat{q}_s \sin(l-s)x}{\sum_{s=0}^m \sum_{l=0}^m \widehat{q}_s \widehat{q}_l \cos(l-s)x}. \end{aligned} \quad (16)$$

Если $n \geq \max\{m_j : 1 \leq j \leq k\}$, то $n_j \geq m$. Поэтому тригонометрический многочлен, стоящий в числителе дроби (16), имеет степень не выше n_j , а многочлен, стоящий в знаменателе, имеет степень не выше m . Отсюда и из (14) делаем вывод о том, что справедливо равенство (15). Тогда из (14)–(16) вытекает справедливость равенств (10)–(12). Теорема 3.4 доказана.

4. Существование нелинейных аппроксимаций Эрмита–Чебышёва

Теоремы 3.3 и 3.4 позволяют найти достаточные условия, при которых существуют нелинейные аппроксимации Эрмита–Чебышёва и описать их явный вид.

Теорема 4.1. Пусть $n \geq \max\{m_j : 1 \leq j \leq k\}$, $m \neq 0$, а для системы \mathbf{f}^1 степенных рядов, ассоциированных с системой \mathbf{f}^{ch1} , выполняются условия 1)–3) теоремы 3.3. Тогда существуют нелинейные аппроксимации Эрмита–Чебышёва первого рода $\{\widehat{\pi}_j^{ch1}(x; \mathbf{f}^{ch1})\}_{j=1}^k$ и для их знаменателя и числителей справедливы представления

$$\begin{aligned} \widehat{Q}_m^{ch1}(x; \mathbf{f}^{ch1}) &= \widehat{Q}_m(e^{i \arccos x}; \mathbf{f}^1) \overline{\widehat{Q}_m(e^{i \arccos x}; \mathbf{f}^1)}, \\ \widehat{P}_j^{ch1}(x; \mathbf{f}^{ch1}) &= \operatorname{Re} \left\{ \widehat{P}_j(e^{i \arccos x}; \mathbf{f}^1) \overline{\widehat{Q}_m(e^{i \arccos x}; \mathbf{f}^1)} \right\}, \end{aligned}$$

где многочлены Эрмита–Якоби $\widehat{Q}_m(z; \mathbf{f}^1)$, $\widehat{P}_j(z; \mathbf{f}^1)$ совпадают с многочленами Эрмита–Паде $Q_m(z; \mathbf{f}^1)$, $P_j(z; \mathbf{f}^1)$ и в том случае, когда определитель $H_{n, \vec{m}}(\mathbf{f}^1) \neq 0$, находятся по формулам (5), (6), в которых $\mathbf{f} = \mathbf{f}^1$.

Теорема 4.2. Пусть $n \geq \max\{m_j : 1 \leq j \leq k\}$, $m \neq 0$, а для системы \mathbf{f}^2 степенных рядов, ассоциированных с системой \mathbf{f}^{ch2} , выполняются условия 1)–3) теоремы 3.4. Тогда существуют нелинейные аппроксимации Эрмита–Чебышёва второго рода $\{\widehat{\pi}_j^{ch2}(x; \mathbf{f}^{ch2})\}_{j=1}^k$ и для их знаменателя и числителей справедливы представления

$$\widehat{Q}_m^{ch2}(x; \mathbf{f}^{ch2}) = \widehat{Q}_m(e^{i \arccos x}; \mathbf{f}^2) \overline{\widehat{Q}_m(e^{i \arccos x}; \mathbf{f}^2)}, \quad (17)$$

$$\widehat{P}_j^{ch2}(x; \mathbf{f}^{ch2}) = \frac{1}{\sqrt{1-x^2}} \operatorname{Im} \left\{ \widehat{P}_j(e^{i \arccos x}; \mathbf{f}^2) \overline{\widehat{Q}_m(e^{i \arccos x}; \mathbf{f}^2)} \right\}, \quad (18)$$

где многочлены Эрмита–Якоби $\widehat{Q}_m(z; \mathbf{f}^2)$, $\widehat{P}_j(z; \mathbf{f}^2)$ совпадают с многочленами Эрмита–Паде $Q_m(z; \mathbf{f}^2)$, $P_j(z; \mathbf{f}^2)$ и в том случае, когда определитель $H_{n, \vec{m}}(\mathbf{f}^2) \neq 0$ находятся по формулам (5), (6), в которых $\mathbf{f} = \mathbf{f}^2$.

Докажем теорему 4.2. Рассмотрим ассоциированную с системой \mathbf{f}^{ch2} систему \mathbf{f}^2 . Поскольку выполняются условия теоремы 3.4, то существуют тригонометрические аппроксимации Эрмита–Якоби $\{\widehat{\pi}_j^{f^2}(x; \mathbf{f}^2)\}_{j=1}^k$ и справедливы формулы (10)–(12). Заменяем в этих равенствах x на $\arccos x$, а затем разделим почленно равенства (11) и (12) на $\sqrt{1-x^2}$. Поскольку на отрезке $[-1, 1]$ справедливы тождества

$$f_j^{f^2}(\arccos x) = \sqrt{1-x^2} f_j^{ch2}(x), \quad j = 1, \dots, k,$$

то в результате получим

$$\widehat{Q}_m^j(\arccos x; \mathbf{f}^2) = \widehat{Q}_m(e^{i\arccos x}; \mathbf{f}^2) \overline{\widehat{Q}_m(e^{i\arccos x}; \mathbf{f}^2)} = \sum_{l=0}^m \widehat{q}_l T_l(x),$$

$$\frac{1}{\sqrt{1-x^2}} \widehat{P}_j^l(\arccos x; \mathbf{f}^2) = \frac{1}{\sqrt{1-x^2}} \operatorname{Im} \left\{ \widehat{P}_j(e^{i\arccos x}; \mathbf{f}^2) \overline{\widehat{Q}_m(e^{i\arccos x}; \mathbf{f}^2)} \right\} = \sum_{l=0}^{n_j} \widehat{p}_l^j U_l(x),$$

$$f_j^{ch2}(x) - \frac{1}{\sqrt{1-x^2}} \widehat{\pi}_j^l(\arccos x; \mathbf{f}^2) = \sum_{l=n+m+1}^{\infty} \widehat{d}_l^j U_l(x).$$

Отсюда следует существование аппроксимаций Эрмита–Чебышёва $\{\widehat{\pi}_j^{ch2}(x; \mathbf{f}^{ch2})\}_{j=1}^k$ и справедливость равенств (17), (18). Теорема 4.2 доказана. Теорема 4.1 доказывается аналогично.

Замечание 4.3. При $k = 1$ теорема 4.1 доказана в [2] (без описания явного вида аппроксимаций Паде–Чебышёва). Утверждения теоремы 4.2 являются новыми и представляют самостоятельный интерес в том числе в случае $k = 1$: нелинейные аппроксимации Паде–Чебышёва второго рода в такой же мере, как и нелинейные аппроксимации Паде–Чебышёва первого рода (см. [27–30]) могут быть востребованы в различных приложениях, а также при проведении научных и технических расчетов (подробнее см. [2; 31; 32]). Отметим также, что для нелинейных аппроксимаций Эрмита–Чебышёва второго рода задача нахождения условий их существования до настоящего времени не исследовалась.

Работа выполнена при финансовой поддержке Министерства образования Республики Беларусь в рамках Государственной программы научных исследований на 2021–2025 годы.

Литература

1. Никишин Е. М., Сорокин В. Н. Рациональные аппроксимации и ортогональность. М.: Наука, 1988.
2. Суетин С. П. О существовании нелинейных аппроксимаций Паде–Чебышёва для аналитических функций // Математические заметки. 2009. Т. 86, № 2. С. 290–303.
3. Гончар А. А., Рахманов Е. А., Суетин С. П. Аппроксимации Паде–Чебышёва для многозначных аналитических функций, вариация равновесной энергии и S -свойство стационарных компактов // Успехи мат. наук. 2011. Т. 66, № 6. С. 3–36.
4. Бейкер мл. Дж., Грейвс-Моррис. П. Аппроксимации Паде. 1. Основы теории. 2. Обобщения и приложения. М.: Мир, 1986.
5. Лабыч Ю. А., Старовойтов А. П. Тригонометрические аппроксимации Паде функций с регулярно убывающими коэффициентами Фурье // Математический сб. 2009. Т. 200, № 7. С. 107–130.
6. Geddes K. O. Block structure in the Chebyshev–Padé table // SIAM J. Numer. Anal. 1981. Vol. 18, N 5. P. 844–861.
7. Адуков В. М., Ибряева О. Л. Асимптотическое поведение знаменателей аппроксимаций Паде–Чебышёва для последней промежуточной строки. Рациональный случай // Вестник ЮУрГУ. Сер. Математика. Физика. Химия. 2005. Т. 6, № 6. С. 11–18.
8. Ибряева О. Л. Достаточное условие единственности линейной аппроксимации Паде–Чебышёва // Известия Челябинского научного центра. 2002. № 4. С. 1–5.
9. Старовойтов А. П., Кечко Е. П., Оснач Т. М. О существовании тригонометрических аппроксимаций Эрмита–Якоби и нелинейных аппроксимаций Эрмита–Чебышёва // Журнал Белорусского государственного университета. Математика. Информатика. 2023. № 2. С. 6–17.
10. Старовойтов А. П., Кечко Е. П., Оснач Т. М. Существование и единственность совместных аппроксимаций Эрмита–Фурье // Проблемы физики, математики и техники. 2023. № 2 (55). С. 68–73.
11. Старовойтов А. П., Кругликов И. В., Оснач Т. М. Рациональные аппроксимации степенных, тригонометрических рядов и рядов по многочленам Чебышёва // Журнал Белорусского государственного университета. Математика. Информатика. 2024. № 3. С. 6–21.
12. Mason J. C., Crampton A. Laurent–Padé approximants to four kinds of Chebyshev polynomial expansions. I. Maehly type approximants // Numer. Algorithms. 2005. Vol. 38. P. 3–18.

13. Mason J. C., Crampton A. Laurent–Padé approximants to four kinds of Chebyshev polynomial expansions. II. Clenshaw–Lord type approximants // *Numer. Algorithms*. 2005. Vol. 38. P. 19–29.
14. Суетин С. П. Аппроксимации Паде и эффективное аналитическое продолжение степенного ряда // *Успехи мат. наук*. 2002. Т. 57, № 1. С. 45–142.
15. Суетин С. П. Вопросы сходимости аппроксимаций Паде–Фабера: дисс. ... канд. физ.-мат. наук. М.: МГУ, 1981.
16. Старовойтов А. П., Кругликов И. В. Достаточные условия единственности и явный вид линейных аппроксимаций Эрмита–Чебышёва // *Проблемы физики, математики и техники*. 2025. № 1(62). С. 67–72.
17. Hermite C. Sur la fonction exponentielle. Paris: C.R. Akad. Sci., 1873. Vol. 77. P. 18–293.
18. Старовойтов А. П., Рябченко Н. В. О детерминантных представлениях многочленов Эрмита–Паде // *Труды Московского математического общества*. 2022. Т. 83, № 1. С. 17–35.
19. Старовойтов А. П., Рябченко Н. В. О единственности решений задач Эрмита–Паде // *Вестці Національної академії наук. Сер. фізика-математичних наук*. 2019. Т. 55, № 4. С. 445–456.
20. Jacobi C. Uber die Darstellung einer Reihe gegebner Werthe durch eine gebrochne rationale Function // *J. Reine Angew. Math*. 1846. Vol. 30. P. 127–156.
21. Антекарев А. И., Буслаев В. И., Мартинес-Финкельштейн А., Суетин С. П. Аппроксимации Паде, непрерывные дроби и ортогональные многочлены // *Успехи мат. наук*. 2011. Т. 66, № 6 (402). С. 37–122.
22. Оснач Т. М., Рябченко Н. В., Старовойтов А. П. Аналог теоремы Якоби для одновременной эрмитовской интерполяции нескольких функций // *Проблемы физики, математики и техники*. 2023. № 1 (54). С. 89–92.
23. Németh G., Páris G. The Gibbs phenomenon in generalized Padé approximation // *J. Math. Phys*. 1985. Vol. 26, N 6. P. 1175–1178.
24. De Bruin M. G. Convergence of the Padé table for ${}_1F_1(1; c; x)$ // *K. Nederl. Akad. Wetensch., Ser. A*. 1976. Vol. 79. P. 408–418.
25. Антекарев А. И. Об аппроксимациях Паде к набору $\{{}_1F_1(1, c; \lambda_i z)\}_{i=1}^k$ // *Вестник МГУ. Сер. 1, Математика. Механика*. 1981. № 2. С. 58–62.
26. Старовойтов А. П. Аппроксимации Эрмита–Паде функций Миттаг–Леффлера // *Труды Математического института имени В. А. Стеклова РАН*. 2018. Т. 301. С. 241–258.
27. Andrianov I. V. Application of Padé approximants in perturbation methods // *Adv. in Mech*. 1991. Vol. 14, N 2. P. 3–25.
28. Andrianov I. V., Awrejcewicz J. New trends in asymptotic approaches: summation and interpolation methods // *Appl. Mech. Rev*. 2001. Vol. 54, N 1. P. 69–92.
29. Книжнерман Л. А. Выделение полюсов потенциальных полей с помощью разложения в ряды Фурье–Чебышёва // *Изв. АН СССР. Сер. Физика Земли*. 1984. № 11. С. 119–123.
30. Ермохин К. М. Продолжение геофизических полей в область источников аномалий методом аппроксимации цепными дробями // *Геофизика*. 2007. № 1. С. 51–55.
31. Прохоров Г. В., Колбеев В. В., Желнов К. И., Леденев М. А. Математический пакет Maple V Release 4: Руководство пользователя. Калуга: Облиздат, 1998.
32. Tee T. W., Trefethen L. N. A rational spectral collocation method with adaptively transformed Chebyshev grid points // *SIAM J. Sci. Comput*. 2006. Vol. 28, N 5. P. 1798–1811.

References

1. Nikishin E. M., Sorokin V. N. *Rational approximations and orthogonality*. Moscow, Nauka, 1988 (in Russian).
2. Suetin S. P. On the existence of nonlinear Padé–Chebyshev approximations for analytical functions. *Mathematical Notes*, 2009, vol. 86, iss. 2, pp. 290–303 (in Russian).
3. Gonchar A. A., Rakhmanov E. A., Suetin S. P. Padé–Chebyshev approximations for multivalued analytical functions, equilibrium energy variation and the S-property of stationary compacts. *Russian Mathematical Surveys*, 2011, vol. 66, iss. 6, pp. 3–36 (in Russian).

4. Baker Jr. G., Graves-Morris P. *Padé Approximations. 1. Fundamentals of theory. 2. Generalizations and applications*. Moscow, Mir, 1986 (in Russian).
5. Labych Yu. A., Starovoitov A. P. Trigonometric Padé approximations of functions with regularly decreasing Fourier coefficients. *Sbornik: Mathematics*, 2009, vol. 200, iss. 7, pp. 107–130 (in Russian).
6. Geddes K. O. Block structure in the Chebyshev–Padé table. *SIAM J. Numer. Anal.*, 1981, vol. 18, iss. 5, pp. 844–861.
7. Adukov V. M., Ibryaeva O. L. The asymptotic behavior of the denominators of the Padé–Chebyshev approximations for the last intermediate row. The rational case. *Bulletin of the South Ural State University, series «Mathematics. Mechanics. Chemistry»*, 2005, vol. 6, iss. 6, pp. 11–18 (in Russian).
8. Ibryaeva O. L. Sufficient condition for the uniqueness of the linear Padé–Chebyshev approximation. *News of the Chelyabinsk Scientific Center*, 2002, iss. 4, pp. 1–5 (in Russian).
9. Starovoitov A. P., Kechko E. P., Osnach T. M. On the existence of trigonometric Hermite–Jacobi approximations and nonlinear Hermite–Chebyshev approximations. *Journal of the Belarusian State University. Mathematics and Informatics*, 2023, iss. 2, pp. 6–17 (in Russian).
10. Starovoitov A. P., Kechko E. P., Osnach T. M. Existence and uniqueness of consistent Hermite–Fourier approximations. *Problems of Physics, Mathematics and Technics*, 2023, iss. 2 (55), pp. 68–73 (in Russian).
11. Starovoitov A. P., Kruglikov I. V., Osnach T. M. Rational approximations of power series, trigonometric series and series of Chebyshev polynomials. *Journal of the Belarusian State University. Mathematics and Informatics*, 2025, vol. 61, iss. 3, pp. 6–21 (in Russian).
12. Mason J. C., Crampton A. Laurent–Padé approximants to four kinds of Chebyshev polynomial expansions. I. Maehly type approximants. *Numer. Algorithms*, 2005, vol. 38, pp. 3–18.
13. Mason J. C., Crampton A. Laurent–Padé approximants to four kinds of Chebyshev polynomial expansions. II. Clenshaw–Lord type approximants. *Numer. Algorithms*, 2005, vol. 38, pp. 19–29.
14. Suetin S. P. Padé approximations and effective analytical continuation of the power series. *Russian Mathematical Surveys*, 2002, vol. 57, iss. 1, pp. 45–142 (in Russian).
15. Suetin S. P. *Questions of convergence of the Padé–Faber approximations*. Ph. D. thesis. Moscow: MGU, 1981 (in Russian).
16. Starovoitov A. P., Kruglikov I. V. Uniqueness and explicit form of linear Hermite–Chebyshev approximations. *Problems of Physics, Mathematics and Technics*, 2025, iss. 1(62), pp. 67–72 (in Russian).
17. Hermite C. *Sur la fonction exponentielle*. Paris, C. R. Akad. Sci., 1873, vol. 77, pp. 18–293.
18. Starovoitov A. P., Ryabchenko N. V. On determinant representations of Hermite–Padé polynomials. *Transactions of the Moscow Mathematical Society*, 2022, vol. 83, iss. 1, pp. 17–35 (in Russian).
19. Starovoitov A. P., Ryabchenko N. V. Uniqueness of the solutions of the Hermite–Padé problems. *Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics Series*, 2019, vol. 55, iss. 4, pp. 445–456 (in Russian).
20. Jacobi C. Über die Darstellung einer Reihe gegebner Werthe durch eine gebrochne rationale Function. *J. Reine Angew. Math.*, 1846, vol. 30, pp. 127–156.
21. Aptekarev A. I., Buslaev V. I., Martinez–Finkelshtein A., Suetin S. P. Padé approximants, continued fractions, and orthogonal polynomials. *Russian Mathematical Surveys*, 2011, vol. 66, iss. 6 (402), pp. 37–122 (in Russian).
22. Osnach T. M., Ryabchenko N. V., Starovoitov A. P. An analogue of Jacobi’s theorem for simultaneous Hermitian interpolation of several functions. *Problems of Physics, Mathematics and Technics*, 2023, iss. 1 (54), pp. 89–92 (in Russian).
23. Németh G., Páris G. The Gibbs phenomenon in generalized Padé approximation. *J. Math. Phys.*, 1985, vol. 26, iss. 6, pp. 1175–1178.
24. De Bruin M. G. Convergence of the Padé table for ${}_1F_1(1; c; x)$. *K. Nederl. Akad. Wetensch., Ser. A.*, 1976, vol. 79, pp. 408–418.
25. Aptekarev A. I. On approximations of the Padé to the set $\{{}_1F_1(1, c; \lambda_i z)\}_{i=1}^k$. *Moscow University Mathematics Bulletin, Moscow University Mechanics Bulletin*, 1981, iss. 2, pp. 58–62 (in Russian).
26. Starovoitov A. P. Hermite–Padé approximations of Mittag–Leffler functions. *Proceedings of the Steklov Institute of Mathematics*, 2018, vol. 301, pp. 241–258 (in Russian).

27. Andrianov I. V. Application of Padé approximants in perturbation methods. *Adv. in Mech.*, 1991, vol. 14, iss. 2, pp. 3–25.
28. Andrianov I. V., Awrejcewicz J. New trends in asymptotic approaches: summation and interpolation methods. *Appl. Mech. Rev.*, 2001, vol. 54, iss. 1, pp. 69–92.
29. Knizhnerman L. A. Identification of potential field poles using Fourier–Chebyshev series expansion. *Proceedings of the USSR Academy of Sciences. Earth Physics Series*, 1984, iss. 11, pp. 119–123 (in Russian).
30. Ermokhin K. M. Continuation of geophysical fields into the area of anomaly sources by the method of approximation by chain fractions. *Geophysics*, 2007, iss. 1, pp. 51–55 (in Russian).
31. Prokhorov G. V., Kolbeev V. V., Zhelnov K. I., Ledenev M. A. *Maple V Release 4 Math Package: User's Guide*. Kaluga: Oblizdat, 1998 (in Russian).
32. Tee T. W., Trefethen L. N. A rational spectral collocation method with adaptively transformed Chebyshev grid points. *SIAM J. Sci. Comput.*, 2006, vol. 28, iss. 5, pp. 1798–1811.



УДК 519.21+519.6

**ПРИБЛИЖЕННАЯ ФОРМУЛА ДЛЯ МАТЕМАТИЧЕСКИХ ОЖИДАНИЙ
ОТ РЕШЕНИЯ СТОХАСТИЧЕСКОГО ДИФФЕРЕНЦИАЛЬНОГО УРАВНЕНИЯ
С ДРЕЙФОМ**

А. В. Жерело

*Белорусский государственный университет, Минск, Беларусь
e-mail: zherelo@bsu.by*

Поступила: 04.04.2025

Исправлена: 13.05.2025

Принята: 23.05.2025

Ключевые слова: случайный процесс, стохастическое дифференциальное уравнение, интеграл Ито, математическое ожидание, приближенные вычисления, слабые аппроксимации.

Аннотация. В работе рассмотрен случай стохастического дифференциального уравнения в смысле Ито с дрейфом. Для рассматриваемого уравнения построена формула приближенного вычисления математических ожиданий от его решения. Для построенной формулы получена оценка погрешности. Проведен численный эксперимент.

**APPROXIMATE FORMULA FOR MATHEMATICAL EXPECTATIONS OF A SOLUTION
OF A STOCHASTIC DIFFERENTIAL EQUATION WITH DRIFT**

A. V. Zherelo

*Belarusian State University, Minsk, Belarus
e-mail: zherelo@bsu.by*

Received: 04.04.2025

Revised: 13.05.2025

Accepted: 23.05.2025

Keywords: random process, stochastic differential equation, Ito integral, mathematical expectation, approximate calculations, weak approximations.

Abstract. The paper considers the case of a stochastic differential equation in the sense of Ito with drift. For the equation under consideration, a formula for the approximate calculation of mathematical expectations from its solution is constructed. An estimate of the error of the constructed formula is obtained. A numerical experiment is performed.

1. Введение

Исследователи различных направлений науки, например, таких как экономика, физика, генетика, все чаще обращают свое внимание на необходимость учета влияния случайности на исследуемые явления. Желание оценить влияние случайных факторов на получаемые результаты служит мотивом к переходу от детерминированных моделей к моделям, учитывающим случайную природу явлений. В ряде случаев модель явления можно описать с помощью стохастического дифференциального уравнения (СДУ). Здесь необходимо отметить, что точно решается только относительно небольшой класс таких уравнений (например, см. [1–3]). Как следствие, особое значение приобретают методы численного решения СДУ и приближенного вычисления математических ожиданий от функционалов от решений таких уравнений.

В численных методах можно выделить два основных подхода: сильный и слабый. Сильный связан с нахождением решения потраекторно. Здесь, как правило, используются широкий класс методов Монте-Карло (например, см. [4]). Данные методы позволяют построить решение с достаточно небольшой погрешностью, но требуют значительных вычислительных затрат. Еще

одной особенностью использования методов Монте-Карло является то, что при моделировании генерируются траектории, приближающие решение уравнения, тогда как во многих задачах нас интересует прежде всего некоторый интегральный показатель, что для обеспечения необходимой точности приводит к еще большему росту вычислительных затрат. Слабый подход связан с поиском приближений, близких к решению СДУ в некотором смысле, например, близких моментов до заданного порядка или характеристического функционала. При таком подходе, как правило, необходимо провести предварительные аналитические вычисления, что в последствии компенсируется высокой скоростью вычисления [5–7].

Особый интерес представляют СДУ с дрейфом, поскольку в случае его нелинейности решение такого уравнения требует симуляции еще большего количества траекторий для сильных методов и значительно усложняется применение слабых методов.

В данной работе предложена формула для приближенного вычисления математических ожиданий от функционалов от решений СДУ с дрейфом, развивающая слабый подход.

Далее будем рассматривать стохастическое дифференциальное уравнение, записанное в соответствующей интегральной форме:

$$X_t = X_0 + \int_0^t \alpha(X_{s-}, s-) ds + \int_0^t \beta(X_{s-}, s-) dW_s, \quad (1)$$

где W_t – процесс Винера, $t \in [0, 1]$, $X_0 \in \mathbb{R}$, α и β – функции, гладко зависящие от своих переменных.

Здесь будем полагать, что интеграл по W_t в правой части уравнения является стохастическим интегралом в смысле Ито. Также будем предполагать, что на функции α и β наложены условия существования сильного решения (например, см. [2; 3]):

$$\begin{aligned} |\alpha(y_1, t) - \alpha(y_2, t)|^2 + |\beta(y_1, t) - \beta(y_2, t)|^2 &\leq K_1 |y_1 - y_2|^2, \\ |\alpha(y, t)|^2 + |\beta(y, t)|^2 &\leq K_2 (1 + |y|^2), \end{aligned}$$

для $\forall y_1, y_2 \in \mathbb{R}$, где $K_1, K_2 \in \mathbb{R}$.

Задачей исследования является построение приближенной формулы для вычисления математических значений функционалов вида $\mathbb{E}[G] \equiv \mathbb{E}[G(X_{(\cdot)})]$, где $G(x)$ это некоторый функционал, гладко зависящий от своего аргумента. Здесь символ (\cdot) указывает на то, что функционал G может зависеть не только от значения процесса в некоторой точке X_s , $s \in [0, t]$, но и от его траектории (например, функционал может содержать интеграл от X_s).

Здесь и далее будем предполагать, что для указанного функционала существует сильная производная по его аргументу, которую будем обозначать $G'(x)$. Также будем предполагать, что эта производная ограничена некоторым $C \in \mathbb{R}$.

1.1. Построение приближенной формулы

Сначала построим приближенную формулу, точную для моментов третьей степени вида $\mathbb{E}[\xi_1 \xi_2 \xi_3]$, где $\xi_j = \{t_1, t_2, t_3, W_{t_4}, W_{t_5}, W_{t_6}\}$. Для ее построения был адаптирован подход к построению формул произвольной степени точности для вычисления функционалов от гауссовских мер, предложенный в работах Л. А. Яновича, А. Д. Егорова, П. И. Соболевского и др. (например, см. [5]). Отличие предлагаемого далее подхода заключается в том, что поскольку рассматриваемое уравнение является уравнением с дрейфом, возникает необходимость, во-первых, учитывать значения моментов, содержащих время как один из множителей, а во-вторых, структура используемых функций должна быть удобна для осуществления вычислений во вспомогательном уравнении, которое будет использовано в дальнейшем для построения приближенной формулы для функционала от решения стохастического уравнения с дрейфом.

Лемма 1.1. *Формула (2) точна для математических ожиданий мономов третьей степени вида $\mathbb{E}[G] = \mathbb{E}[\xi_1 \xi_2 \xi_3]$, где $\xi_j = \{1, t_1, t_2, t_3, W_{t_4}, W_{t_5}, W_{t_6}\}$ за исключением случая $\xi_1 \xi_2 \xi_3 = t_1 t_2 t_3$.*

$$\mathbb{E}[G] \approx J(G) \equiv \frac{1}{2} \sum_{j=1}^2 A_j \int_0^1 \int_0^1 \int_{-1}^1 G(\rho_{j1}(\cdot, u_1) + \rho_{j2}(\cdot, u_2), \rho(\cdot, v)) du_1 du_2 dv, \quad (2)$$

$$\begin{aligned} \rho_{jk}(s, u_k) &= a_{jk} 1_{[u_k, 1]}(s), k = 1, 2, \quad \rho(s, v) = \text{sign}(v) 1_{[|v|, 1]}(s), \\ A_1 + A_2 &= 1, \\ a_{11} &= \frac{1}{2} \left(1 - \sqrt{-\frac{A_2}{A_1}} \right), \quad a_{12} = \frac{1}{2} \left(1 + \sqrt{-\frac{A_2}{A_1}} \right), \\ a_{21} &= \frac{1}{2} \left(1 - \sqrt{-\frac{A_1}{A_2}} \right), \quad a_{22} = \frac{1}{2} \left(1 + \sqrt{-\frac{A_1}{A_2}} \right). \end{aligned}$$

Лемму 1.1 не сложно доказать непосредственно подставив $\rho_{k,j}(t, u_k) + \rho_{k,j}(t, u_k)$ и $\rho(t, v)$ вместо t и W_t в приведенные в формулировке леммы мономы.

Символ точки в формуле (2) указывает на то, что функционал G может зависеть не только от заданного момента времени t , но и от траектории случайных процессов за промежуток $[0, t]$.

В данной работе в качестве параметров построенной приближенной формулы были выбраны следующие значения:

$$A_1 = \frac{4}{3}, A_2 = -\frac{1}{3}, a_{11} = \frac{1}{4}, a_{12} = \frac{3}{4}, a_{21} = -\frac{1}{2}, a_{22} = \frac{3}{2}.$$

Для построения приближенной формулы для решения уравнения (1) введем и решим вспомогательные уравнения вида (3), полученные с помощью подстановки функций $\rho(t, v)$ и $\rho_{j1}(t, u_1) + \rho_{j2}(t, u_2)$ в (1):

$$\begin{aligned} Y_j(t, u_1, u_2, v) &= X_0 + \int_0^t \alpha(Y_j(s-, u_1, u_2, v), s-) d(\rho_{j1}(s, u_1) + \rho_{j2}(s, u_2)) + \\ &+ \int_0^t \beta(Y_j(s-, u_1, u_2, v), s) d\rho(s, v). \end{aligned} \quad (3)$$

Решим уравнение (3) с помощью метода последовательных приближений. Обозначим $Y_{jt}^{(k)} = Y_j^{(k)}(t, u_1, u_2, v)$ значение $Y_j(t, u_1, u_2, v)$ на итерации с номером k .

Пусть $Y_j^{(0)}(t, u_1, u_2, v) = X_0$. На первом шаге получаем

$$\begin{aligned} Y_{jt}^{(1)}(t, u_1, u_2, v) &= X_0 + \int_0^t \alpha(X_0, s-) d(\rho_{j1}(s, u_1) + \rho_{j2}(s, u_2)) + \int_0^1 \beta(X_0, s) d\rho(s, v) = \\ &= X_0 + \alpha(X_0, u_1) \rho_{j1}(t, u_1) + \alpha(X_0, u_2) \rho_{j2}(t, u_2) + \beta(X_0, |v|) \rho(t, v). \end{aligned}$$

Далее подставим $Y_{jt}^{(1)}$ в уравнение (3) и воспользуемся следующими свойствами функций ρ и ρ_{jk} :

$$\rho_{j1}(u_1-, u_1) = 0, \quad \rho_{j2}(u_2-, u_2) = 0, \quad \rho(|v|-, v) = 0,$$

и получим

$$\begin{aligned} Y_{jt}^{(2)} &= X_0 + \alpha(X_0 + \alpha(X_0, u_2) \rho_{j2}(u_1-, u_2) + \beta(X_0, |v|) \rho(u_1-, v), u_1) \rho_{j1}(t, u_1) + \\ &+ \alpha(X_0 + \alpha(X_0, u_1) \rho_{j1}(u_2-, u_1) + \beta(X_0, |v|) \rho(u_2-, v), u_2) \rho_{j2}(t, u_2) + \\ &+ \beta(X_0 + \alpha(X_0, u_1) \rho_{j1}(|v|-, u_1) + \alpha(X_0, u_2) \rho_{j2}(|v|-, u_2), |v|) \rho(t, v). \end{aligned}$$

Далее можно вычислить $Y_{jt}^{(3)}$, если принять во внимание значения функций ρ , ρ_{jk} в зависимости от соотношения между переменными u_1, u_2 и v .

Если $u_1 < u_2 < |v|$, тогда

$$\begin{aligned} \rho_{j2}(u_1-, u_2) &= 0, & \rho(u_1-, v) &= 0, \\ \rho_{j1}(u_2-, u_1) &= a_{j1}, & \rho(u_2-, v) &= 0, \\ \rho_{j1}(|v|-, u_1) &= a_{j1}, & \rho(|v|-, u_2) &= a_{j2}. \end{aligned}$$

Если $u_2 < u_1 < |v|$, тогда

$$\begin{aligned} \rho_{j2}(u_1-, u_2) &= a_{j2}, & \rho(u_1-, v) &= 0, \\ \rho_{j1}(u_2-, u_1) &= 0, & \rho(u_2-, v) &= 0, \\ \rho_{j1}(|v|-, u_1) &= a_{j1}, & \rho(|v|-, u_2) &= a_{j2}. \end{aligned}$$

Если $u_2 < |v| < u_1$, тогда

$$\begin{aligned} \rho_{j2}(u_1-, u_2) &= a_{j2}, & \rho(u_1-, v) &= \text{sign}(v), \\ \rho_{j1}(u_2-, u_1) &= 0, & \rho(u_2-, v) &= 0, \\ \rho_{j1}(|v|-, u_1) &= 0, & \rho(|v|-, u_2) &= a_{j2}. \end{aligned}$$

Если $u_1 < |v| < u_2$, тогда

$$\begin{aligned} \rho_{j2}(u_1-, u_2) &= 0, & \rho(u_1-, v) &= 0, \\ \rho_{j1}(u_2-, u_1) &= a_{j1}, & \rho(u_2-, v) &= \text{sign}(v), \\ \rho_{j1}(|v|-, u_1) &= a_{j1}, & \rho(|v|-, u_2) &= 0. \end{aligned}$$

Если $|v| < u_1 < u_2$, тогда

$$\begin{aligned} \rho_{j2}(u_1-, u_2) &= 0, & \rho(u_1-, v) &= \text{sign}(v), \\ \rho_{j1}(u_2-, u_1) &= a_{j1}, & \rho(u_2-, v) &= \text{sign}(v), \\ \rho_{j1}(|v|-, u_1) &= 0, & \rho(|v|-, u_2) &= 0. \end{aligned}$$

Если $|v| < u_2 < u_1$, тогда

$$\begin{aligned} \rho_{j2}(u_1-, u_2) &= a_{j2}, & \rho(u_1-, v) &= \text{sign}(v), \\ \rho_{j1}(u_2-, u_1) &= 0, & \rho(u_2-, v) &= \text{sign}(v), \\ \rho_{j1}(|v|-, u_1) &= 0, & \rho(|v|-, u_2) &= 0. \end{aligned}$$

Если $u_2 = u_1$, тогда

$$\rho_{j2}(u_1-, u_2) = 0, \quad \rho_{j2}(u_2-, u_1) = 0.$$

Если $u_1 = |v|$, тогда

$$\rho_{j1}(|v|-, u_1) = 0, \quad \rho(u_1-, v) = 0.$$

Если $u_2 = |v|$, тогда

$$\rho_{j2}(|v|-, u_2) = 0, \quad \rho(u_2-, v) = 0.$$

Перейдем к третьему шагу и выпишем участвующие в уравнении для $Y_{jt}^{(3)}$ значения функций $\alpha(Y_{ju_1-}^{(2)}, u_1)$, $\alpha(Y_{ju_2-}^{(2)}, u_2)$, $\beta(Y_{j|v|-, |v|)$, возникающие после подстановки $Y_{jt}^{(2)}$ в уравнение (3) с учетом последующего интегрирования.

Для $\alpha(Y_{ju_1-}^{(2)}, u_1)$ получим:

$$\begin{aligned} u_1 < u_2 < |v| & : \alpha(X_0, u_1), \\ u_2 < u_1 < |v| & : \alpha(X_0 + \alpha(X_0, u_2)a_{j2}, u_1), \\ u_2 < |v| < u_1 & : \alpha(X_0 + \alpha(X_0, u_2)a_{j2} + \\ & \quad + \beta(X_0 + \alpha(X_0, u_2)a_{j2}, |v|)\text{sign}(v), u_1), \\ u_1 < u_2 < |v| & : \alpha(X_0, u_1), \\ |v| < u_1 < u_2 & : \alpha(X_0 + \beta(X_0, |v|)\text{sign}(v), u_1), \\ |v| < u_2 < u_1 & : \alpha(X_0 + \alpha(X_0 + \\ & \quad + \beta(X_0, |v|)\text{sign}(v), u_2)a_{j2} + \beta(X_0, |v|)\text{sign}(v), u_1); \end{aligned}$$

для $\alpha(Y_{ju_2-}^{(2)}, u_2)$:

$$\begin{aligned} u_1 < u_2 < |v| & : \alpha(X_0 + \alpha(X_0, u_1)a_{j1}, u_2), \\ u_2 < u_1 < |v| & : \alpha(X_0, u_2), \\ u_2 < |v| < u_1 & : \alpha(X_0, u_2), \\ u_1 < u_2 < |v| & : \alpha(X_0 + \alpha(X_0, u_1)a_{j1} + \\ & \quad + \beta(X_0 + \alpha(X_0, u_1)a_{j1}, |v|)\text{sign}(v), u_2), \\ |v| < u_1 < u_2 & : \alpha(X_0 + \alpha(X_0 + \beta(X_0, |v|)\text{sign}(v), u_1)a_{j1} + \\ & \quad + \beta(X_0, |v|)\text{sign}(v), u_2), \\ |v| < u_2 < u_1 & : \alpha(X_0 + \beta(X_0, |v|)\text{sign}(v), u_2); \end{aligned}$$

для $\beta(Y_{j|v|}^{(2)}, |v|)$:

$$\begin{aligned} u_1 < u_2 < |v| & : \beta(X_0 + \alpha(X_0, u_1)a_{j1} + \alpha(X_0 + \alpha(X_0, u_1)a_{j1}, u_2)a_{j2}, |v|), \\ u_2 < u_1 < |v| & : \beta(X_0 + \alpha(X_0 + \alpha(X_0, u_2)a_{j2}, u_1)a_{j1} + \alpha(X_0, u_2)a_{j2}, |v|), \\ u_2 < |v| < u_1 & : \beta(X_0 + \alpha(X_0, u_2)a_{j2}, |v|), \\ u_1 < u_2 < |v| & : \beta(X_0 + \alpha(X_0, u_1)a_{j1}, |v|), \\ |v| < u_1 < u_2 & : \beta(X_0, |v|), \\ |v| < u_2 < u_1 & : \beta(X_0, |v|). \end{aligned}$$

Таким образом, $Y_t^{(3)}$ имеет вид

$$\begin{aligned} Y_{jt}^{(3)} = & X_0 + \alpha \left(X_0 + \alpha(X_0 + \beta(X_0, |v|)\text{sign}(v)1_{(|v|,1]}(u_2), u_2) a_{j2}1_{(u_2,1]}(u_1) + \right. \\ & \left. + \beta(X_0 + \alpha(X_0, u_2)a_{j2}1_{(u_2,1]}(|v|), |v|) \text{sign}(v)1_{(|v|,1]}(u_1) + u_1 \right) \rho_{j,1}(t, u_1) + \\ & + \alpha \left(X_0 + \alpha(X_0 + \beta(X_0, |v|)\text{sign}(v)1_{(|v|,1]}(u_1), u_1) a_{j1}1_{(u_1,1]}(u_2) + \right. \\ & \left. + \beta(X_0 + \alpha(X_0, u_1)a_{j1}1_{(u_1,1]}(|v|), |v|) \text{sign}(v)1_{(|v|,1]}(u_2), u_2 \right) \rho_{j,2}(t, u_2) + \\ & + \beta \left(X_0 + \alpha(X_0 + \alpha(X_0, u_2)a_{j2}1_{(u_2,1]}(u_1), u_1) a_{j1}1_{(u_1,1]}(|v|) + \right. \\ & \left. + \alpha(X_0 + \alpha(X_0, u_1)a_{j1}1_{(u_1,1]}(u_2), u_2) a_{j2}1_{(u_2,1]}(|v|), |v| \right) \rho(t, v). \end{aligned}$$

Осуществив подстановку полученного выражения для $Y_{jt}^{(3)}$ в формулу (3) и проделав аналогичные приведенным выше вычисления, не трудно убедиться, что $Y_{jt}^{(4)} = Y_{jt}^{(3)}$, т. е. $Y_{jt} = Y_j(t, u_1, u_2, v) = Y_{jt}^{(3)}$ является решением уравнения (3).

Приближенную формулу для математического ожидания функционала от решения уравнения (1) получим подставив $Y_j(t, u_1, u_2, v)$ в формулу (2):

$$J(G, Y) = \frac{1}{2} \sum_{j=1}^2 A_j \int_0^1 \int_0^1 \int_{-1}^1 G(Y_j(\cdot, u_1, u_2, v), \cdot) du_1 du_2 dv. \quad (4)$$

Справедлива следующая теорема.

Теорема 1.2. Точность приближенной формулы (4) имеет порядок $O(t^{3/2})$, т. е.

$$|\mathbb{E}[G(X_{(\cdot)})] - J[G, Y]| = (|A_1| + |A_2|)C\sqrt{K_2(1 + X_0^2)}t^{3/2} + o(t^{3/2}) = O(t^{3/2}).$$

Доказательство. Введем следующие обозначения: $\tilde{X}_t = X_t - X_0$ и $\tilde{Y}_{jt} = Y_{jt} - X_0$, тогда, пользуясь представлением, приведенным в [5], можем записать следующее равенство:

$$\begin{aligned} |\mathbb{E}[G(X_{(\cdot)})] - J(G, Y)| &= \mathbb{E} \left[G(X_0) + \int_0^1 \int_0^t G'_s(X_0 + \tau \tilde{X}_{(\cdot)}) \tilde{X}_s d\tau ds \right] - \\ &- \frac{1}{2} \sum_{j=1}^2 A_j \int_0^1 \int_0^1 \int_{-1}^1 \left\{ G(X_0) + \int_0^1 \int_0^t G'_s(X_0 + \tau \tilde{Y}_{j(\cdot)}) \tilde{Y}_{js} d\tau ds \right\} du_1 du_2 dv = (*) \end{aligned}$$

и введем процесс $X_t^{(1)}$ вида

$$X_t^{(1)} = X_0 + \int_0^t \alpha(X_0, s-) ds + \int_0^t \beta(X_0, s-) dW_s,$$

тогда

$$\begin{aligned} (*) &= \left| \mathbb{E} \left[\int_0^1 \int_0^t d\tau ds \times \left\{ G'_s(X_0 + \tau \tilde{X}_{(\cdot)}) \tilde{X}_s - G'_s(X_0 + \tau \tilde{X}_{(\cdot)}) \tilde{X}_s^{(1)} + \right. \right. \right. \\ &\left. \left. \left. + G'_s(X_0 + \tau \tilde{X}_{(\cdot)}) \tilde{X}_s^{(1)} - \frac{1}{2} \sum_{j=1}^2 A_j \int_0^1 \int_0^1 \int_{-1}^1 G'_s(X_0 + \tau \tilde{Y}_{j(\cdot)}) \tilde{Y}_{js} du_1 du_2 dv \right\} \right] \right| = (**), \end{aligned}$$

следовательно, выражение выше можно переписать в виде

$$\begin{aligned} (**) &= \left| \int_0^1 \int_0^t d\tau ds \times \mathbb{E} \left[G'_s(X_0 + \tau \tilde{X}_{(\cdot)}) (\tilde{X}_s - \tilde{X}_s^{(1)}) + \right. \right. \\ &\left. \left. + \frac{1}{2} \sum_{j=1}^2 A_j \int_0^1 \int_0^1 \int_{-1}^1 G'_s(X_0 + \tau \tilde{X}_{(\cdot)}) \tilde{X}_s^{(1)} - G'_s(X_0 + \tau \tilde{Y}_{j(\cdot)}) \tilde{X}_s^{(1)} + \right. \right. \\ &\left. \left. + G'_s(X_0 + \tau \tilde{Y}_{j(\cdot)}) \tilde{X}_s^{(1)} - G'_s(X_0 + \tau \tilde{Y}_{j(\cdot)}) \tilde{Y}_{js} du_1 du_2 dv \right] \right| \leq \\ &\leq \int_0^1 \int_0^t d\tau ds \times \left\{ C \left(\mathbb{E} \left[\left(X_s - X_s^{(1)} \right)^2 \right] \right)^{1/2} + \sum_{j=1}^2 |A_j| C \mathbb{E} [|\tilde{X}_s^{(1)}|] + \right. \\ &\left. + \left| \frac{1}{2} \sum_{j=1}^2 A_j \int_0^1 \int_0^1 \int_{-1}^1 G'_s(X_0 + \tau \tilde{Y}_{j(\cdot)}) \left(\mathbb{E}[\tilde{X}_s^{(1)}] - \tilde{Y}_{js} \right) du_1 du_2 dv \right| \right\} = (*). \end{aligned}$$

Используя условия, накладываемые на функции α и β , не трудно показать, что

$$\left(\mathbb{E} \left[\left(X_s^{(1)} - X_0 \right)^2 \right] \right)^{1/2} = \sqrt{K_2(1 + X_0^2)} s^{1/2} + o(s^{1/2})$$

и

$$\left(\mathbb{E} \left[\left(X_s - X_s^{(1)} \right)^2 \right] \right)^{1/2} = \sqrt{K_1 K_2(1 + X_0^2)} s + o(s) = o(s^{1/2}),$$

(см. [2; 3]), поэтому

$$\begin{aligned} (*) &= (|A_1| + |A_2|) C \sqrt{K_2(1 + X_0^2)} t^{3/2} + o(t^{3/2}) + \\ &+ \int_0^1 \int_0^t \left| \frac{1}{2} \sum_{j=1}^2 A_j \int_0^1 \int_0^1 \int_{-1}^1 G'_s(X_0 + \tau \tilde{Y}_{j(\cdot)}) \times \left(\int_0^s \alpha(X_0, z) dz - \tilde{Y}_{js} \right) du_1 du_2 dv \right| d\tau ds \leq \\ &\leq (|A_1| + |A_2|) C \sqrt{K_2(1 + X_0^2)} t^{3/2} + o(t^{3/2}) + \int_0^1 \int_0^t \frac{1}{2} (|A_1| + |A_2|) C O(s) d\tau ds = \\ &= (|A_1| + |A_2|) C \sqrt{K_2(1 + X_0^2)} t^{3/2} + o(t^{3/2}) = O(t^{3/2}). \quad \square \end{aligned}$$

2. Численный эксперимент

Для демонстрации возможностей предложенной формулы рассмотрим уравнение вида

$$X_t = 1 + \int_0^t \cos(X_{s-}) ds + \int_0^t X_{s-} dW_s$$

и вычислим математическое ожидание его решения $\mathbb{E}[X_t]$.

В таблице приведены результаты вычисления для $t = 0,01$ и $t = 0,1$ с использованием формулы (4) и метода Монте-Карло. При моделировании траекторий использовалась схема Мильштейна (см. [8]), а параметр N указывает на количество выполненных независимых розыгрышей траекторий процесса. Необходимо отметить, что каждая строка в столбцах, соответствующих схеме Мильштейна, является результатом осреднения по независимо смоделированному набору траекторий. Характерной особенностью метода Монте-Карло является то, что при каждом независимом вычислении искомого значения функционала получается новое значение.

Результаты численных экспериментов

$t = 0,01$			
Формула (4)	Схема Мильштейна		
	$N = 100000$	$N = 1000000$	$N = 5000000$
1,005366940	1,005337034	1,005520395	1,005409730
	1,005512123	1,005516914	1,005365776
	1,005171719	1,005347186	1,005468750
$t = 0,1$			
Формула (4)	Схема Мильштейна		
	$N = 10000$	$N = 100000$	$N = 1000000$
1,050521622	1,048873096	1,052577854	1,054480922
	1,052423041	1,054184545	1,054275370
	1,047200251	1,054789066	1,053922429

3. Заключение

В работе предложена формула приближенного вычисления математических ожиданий от решения стохастического дифференциального уравнения, в котором присутствует дрейф. Произведена оценка погрешности полученной формулы. Формула (4) относится к классу слабых аппроксимаций и основана на аппроксимации моментов от случайных процессов. Для своего применения формула требует проведения предварительных символьных вычислений, однако затраты на эти вычисления компенсируются в последствии за счет высокой скорости вычислений. Еще одним преимуществом является то, что вычисление приближенного значения с использованием формулы (4), требует значительно меньшего количества памяти и вычислительных ресурсов, поскольку при использовании схемы метода Монте-Карло необходимо произвести осреднение значения на основе всего множества генерируемых траекторий.

Литература

1. Гухман И. И., Скороход А. В. Теория случайных процессов. М.: Наука, 1975. Т. 3.
2. Øksendal B. Stochastic Differential Equations: An Introduction with Applications. Springer, 2003.
3. Applebaum D. Levy processes and stochastic calculus. Cambridge University Press, 2009.
4. Kloeden P. E., Platen E. Numerical solution of stochastic differential equations. Springer, 1999.
5. Egorov A. D., Sobolevsky P. I., Yanovich L. A. Functiona Integrals; Approximate Evaluation and Applications. Dordrecht: Kluwer Acad. Publ., 1993.
6. Egorov A., Zherelo A. Approximation of functional integrals with respect to measure generated by solutions of stochastic differential equations // Monte Carlo Methods Appl. 2004. Vol. 10. P. 257–264.

7. Егоров А. Д., Жидков Е. П., Лобанов Ю. Ю. Введение в теорию и приложения функциональных интегралов. М.: Физматлит, 2006.

8. Мильштейн Г. Н. Приближенное интегрирование стохастических дифференциальных уравнений // Теория вероятностей и ее применения. 1974. Т. 19, № 3. С. 583–588.

References

1. Gihman I. I., Skorokhod A. V. *Theory of random processes. Vol. 3.* Moscow, Nauka, 1975 (in Russian).

2. Øksendal B. *Stochastic Differential Equations: An Introduction with Applications.* Springer, 2003.

3. Applebaum D. *Levy processes and stochastic calculus.* Cambridge University Press, 2009.

4. Kloeden P. E., Platen E. *Numerical solution of stochastic differential equations.* Springer, 1999.

5. Egorov A. D., Sobolevsky P. I., Yanovich L. A. *Functiona Integrals; Approximate Evaluation and Applications.* Dordrecht, Kluwer Acad. Publ., 1993.

6. Egorov A., Zherelo A. Approximation of functional integrals with respect to measure generated by solutions of stochastic differential equations. *Monte Carlo Methods Appl.*, 2004, vol. 10, pp. 257–264.

7. Egorov A. D., Zhidkov E. P., Lobanov Yu. Yu. *Introduction in a theory and applications of functional integrals.* Moscow, Phismatlit, 2006 (in Russian).

8. Milshtein G. N. Approximate integration of stochastic differential equations. *Theory of Probability & Its Applications*, 1974, vol. 19, no 3. pp. 583–588.

УДК 512.542

ОБ ОДНОЙ СТОХАСТИЧЕСКОЙ МОДЕЛИ ДОЙНОГО СТАДА

А. Жерело¹, Г. Крылов¹, В. Оболонкин², М. Йоргенсен³, Р. Спелман⁴

¹Белорусский государственный университет, Минск, Беларусь

²Отдел общественного здравоохранения, Совет по здравоохранению округа Нортленд,
Новая Зеландия

³Университет Вайкато, Гамильтон, Новая Зеландия

⁴Livestock Improvement Corporation, Гамильтон, Новая Зеландия
e-mail: zherelo@bsu.by, krylov@bsu.by

Поступила: 12.04.2025

Исправлена: 19.05.2025

Принята: 23.05.2025

Ключевые слова: случайный процесс, стохастическое дифференциальное уравнение, интеграл Ито, математическое ожидание, процесс Пуассона, метод Монте-Карло, генная динамика.

Аннотация. В работе предложена стохастическая модель, описывающая динамику биаллельного полиморфизма в дойном стаде. Предложенная модель предполагает, что стадо формируется в условиях контролируемого спаривания. Система уравнений, описывающих модель основана на использовании случайных процессов, содержащих скачки. Построен дискретный аналог стохастической системы, для которого проведена Монте-Карло симуляция.

ABOUT ONE STOCHASTIC MODEL OF A DAIRY HERD

A. Zherelo¹, G. Krylov¹, V. Obolonkin², M. Jorgensen³, R. Spelman⁴

¹Belarusian State University, Minsk, Belarus

²Public Health Unit, Northland District Health Board, New Zealand

³University of Waikato, Hamilton, New Zealand

⁴Livestock Improvement Corporation, Hamilton, New Zealand
e-mail: zherelo@bsu.by, krylov@bsu.by

Received: 12.04.2025

Revised: 19.05.2025

Accepted: 23.05.2025

Keywords: random process, stochastic differential equation, Ito integral, mathematical expectation, approximate calculations, Poisson process, Monte-Carlo method, gene dynamics.

Abstract. The paper proposes a stochastic model, which describes the dynamics of biallelic polymorphism in a dairy herd. The proposed model assumes that the herd is formed under controlled mating conditions. The system of equations describing the model is based on the use of random processes containing jumps. A discrete analog of the stochastic system is constructed, for which a Monte Carlo simulation is performed.

1. Введение

Основополагающие работы генетики Г. Харди [1] и В. Вайнберга [2] по устойчивому распределению аллелей были основаны на законе Менделя.

Важнейшим результатом упомянутых исследований является равновесие Харди–Вайнберга, которое предполагает, что если аллель A встречается в популяции в некотором интересующем локусе с вероятностью p , а аллель B с вероятностью $q = 1 - p$, то наблюдаемые частоты генотипов в этом локусе равны p^2 для AA , $2pq$ для AB и q^2 для BB , причем выполняется данное соотношение в случае выполнения ряда условий (см. таблицу). Однако в популяциях домашних животных, используемых в современном сельском хозяйстве, указанные условия, как правило, не выполняются. В качестве примера можно привести дойное стадо Новой Зеландии, которое состоит из более чем $4 \cdot 10^6$ животных, сгруппированных приблизительно в $1,1 \cdot 10^4$ стад. Очевидно, что спаривание в этом национальном стаде далеко не случайно. В 2013–2014 гг. для осеменения было использовано чуть более 3700 быков, причем в 80 % спариваний всего национального стада было использовано менее 100 лучших быков. За исключением случайных мутаций популяции в целом далеки от пропорций

равновесия Харди–Вайнберга в любом локусе. Частоты генотипов могут быть различными даже для одного локуса в случае различных исследуемых стад или регионов. В таблице приведены условия, которые приводят к равновесию Харди–Вайнберга в некоторой идеальной модельной популяции и условия, существующие в молочной промышленности Новой Зеландии (как и в других странах, где используются промышленные подходы в сельском хозяйстве).

Условия Харди–Вайнберга и дойное стадо Новой Зеландии

Идеальная популяция	Индустриальное производство
Большая однородная популяция	Неоднородные стада и регионы
Случайное спаривание	Контролируемое спаривание
Отсутствие мутаций	Мутации случаются
Отсутствие отбора	Отбор существует
Отсутствие миграции	Экспорт/импорт генов

В действительности такие устойчивые результаты не представляют особого интереса для принятия решений по разведению животных, поскольку владелец стада стремится изменить производительные характеристики животных с целью повышения получаемой прибыли. Для начала упомянем наиболее распространенные модели в генетике, используемые для описания генной динамики в идеальной популяции.

Одной из моделей, оказавших существенное влияние на развитие моделирования в рассматриваемом в данной статье направлении, является генетическая модель эволюции (см. работы С. Райта [3] и Р. Фишера [4]). В данной модели предполагается, что размер исследуемой популяции является фиксированным, рассматривается динамика бинарных аллелей (A, B) и упомянутая модель является дискретной, участвующие особи – равно приспособленные, а моделируемые процессы – марковские. Модель является идеальной, поскольку фактически в ней полагается, что каждое из животных следующего поколения случайным образом выбирает родителей из предыдущего поколения, чего не происходит в реальных условиях. Значимым результатом исследования нейтральной модели Райта–Фишера без мутаций является то, что за конечное время произойдет либо закрепление, либо элиминация некоторого заданного аллеля.

Модель Райта–Фишера является слишком идеализированной и ее дальнейшее развитие осуществлялось за счет изменения ее основных предположений, приближающих модель к реальным условиям, например, за счет добавления большего количества типов аллелей (см. [5]) или использования плавающей общей численности популяции [6; 7]. Более усовершенствованные модели используют более сложные законы распределения (например, распределение Пуассона) для числа потомков особи (см. [5]), пытаются учесть наличие мутаций и/или отбора (см. [8]), добавить в модель процесс смертности (см. [9]) или являются моделями диффузионного типа (см. [10]). Ознакомиться с рядом из обозначенных моделей можно в [5; 11; 12].

Необходимо отметить, что одной из целей приведенных выше исследований было построение таких моделей, для которых возможно было бы получить аналитические представления для интересующих исследователей параметров или построить соответствующие асимптотические разложения для них (например, см. [13]), что неизбежно приводит к упрощению модели, и накладываемых на нее условий, и, как следствие, модель в меньшей степени соответствует процессу, для которого она была построена. В частности, при исследовании генной динамики зачастую предполагается, что популяция может иметь неограниченный размер, что, естественно, не соответствует случаю дойного стада на ферме, которое, как правило, не превышает нескольких сотен голов, и находится под управлением владельца стада как с точки зрения численности, так и с точки зрения показателей продуктивности особей. Далее мы будем рассматривать случай стада ограниченного размера в условиях стохастического управления.

В этой статье мы концентрируем внимание на коровах на конкретной ферме и, более того, в конкретном стаде, но упомянутый выше факт, что несколько десятков лучших быков оплодотворяют миллионы новозеландских коров, говорит о тесной связи всех фермерских стад. Такая ориентация на небольшое количество самцов еще больше контрастирует с ситуацией, которую предполагают классические популяционные модели.

Еще одной сложностью при рассмотрении моделей, ориентированных на получение аналитических представлений, является то, что они ориентированы прежде всего на получение аналитических представлений для описания эволюции средних значений интересующих переменных, в то время как в реальных приложениях существенное значение имеет также информация о, например, дисперсии или форме распределений исследуемых переменных. Это приводит к необходимости создания модели, близкой к реальным условиям, т. е., как следствие, к стохастической модели.

Благодаря успехам в области генетики и селекции разведение животных перестало быть процессом трудно управляемым и мало предсказуемым, а основным инструментом, используемым, в частности, в молочной промышленности для улучшения генетических качеств стада, стало искусственное оплодотворение, что позволяет осознанно выбрать родителей для следующего поколения. Именно с целью понимания последующего влияния вносимых в популяцию контролируемых человеком изменений с помощью искусственного разведения, строятся современные модели.

Для определения управляющего воздействия при искусственном разведении при выборе производителей чрезвычайно важными стали геномные методы [14]. Благодаря технологии статистического геномного прогнозирования, используемой для оценки животных, становятся все более доступными обширные наборы данных генотипирования.

Модель, которую мы разрабатываем в данной статье, является довольно общей в том смысле, что она допускает целый ряд корректировок. Например, можно рассмотреть другие виды, а также более сложные правила суммы генотипов, возможно, учитывающие зависящий от генотипа (геномный) отбор или другие отклонения от предположения о равной приспособленности. Представленная модель используется для описания динамики аллелей в одном локусе у каждого животного стада в условиях контролируемого спаривания.

При разработке моделей, представленных ниже, были приняты в расчет условия, являющиеся специфичными для сезонного управления новозеландским молочным стадом. Такие условия были учтены в модели потому, что одной из целей данного исследования была оценка времени, которое необходимо потратить на внедрение желательного геномного признака в рассматриваемое дойное стадо или на снижение доли нежелательного признака. В качестве примера желаемого признака можно выделить геномный вариант, благоприятно влияющий на состав молока, а в качестве нежелательного – ген «маленького теленка», обнаруженный усилиями Livestock Improvement Corporation (например, см. [15]), снижающий продуктивность, время жизни животного, а также приводящий к проблемам с фертильностью.

На данный момент существует достаточно много исследований в данном направлении (например, см. [16–18]). Однако очевидно, что число обнаруженных и описанных нежелательных генов будет возрастать в процессе новых знаний и понимания причин и механизмов развития распространенных заболеваний. Каждый бык, выбранный для искусственного осеменения, несет в себе некоторые нежелательные гены. Разработанная модель может помочь в анализе рисков путем моделирования различных сценариев для оптимизации стратегии искусственного осеменения в смысле закрепления генов, улучшающих производительность животных, и подавления вредоносных генов в популяции.

2. Описание модели

В данном разделе мы представляем стохастическую модель эволюции генотипа при условии ограничения размера стада некоторым максимальным числом особей и в предположении, что стадо изолировано. Для целей численного моделирования на основе указанной модели будет представлена ее модификация на случай дискретного времени, а также предложен вариант, допускающий приток животных из внешнего источника в предположении неограниченности внешнего источника по количеству доступных особей. В заключение, мы намечаем направления для возможного дальнейшего развития.

Исследуемое стадо фактически состоит из двух подстад: основного (производственного) стада и подменного стада.

Основное стадо состоит из взрослых коров, обеспечивающих производство молока. Подменное стадо состоит из молодых животных до момента их включения в основное стадо или удаления

по причине заболевания, смерти и по прочим причинам, а также включает некоторое количество взрослых животных, временно исключенных из основного стада.

Важно отметить, что использованные в уравнениях (2) и ниже индексы i и j не описывают какое-то конкретное животное, а используются для обозначения занимаемых животным в стадах мест (подобно номерам игроков в команде). Свойства животного в i -й в момент времени t будем считать функцией времени, которую будем обозначать $f_i(t)$. Новое животное добавляется в стадо либо в случае, когда стадо имеет размер меньше установленного ограничения, либо в качестве замены удаленному животному, если текущий размер стада является максимальным. В последнем случае вводимое в стадо животное займет первое незанятое положение. Если i -е животное заменяется в момент времени t_* , это обычно вызывает разрыв или «скачок» в функции $f_i(t)$ в $t = t_*$. Обращаем внимание читателей, что при таких предположениях стохастические процессы в нашей модели могут иметь скачки в случае замены животных.

Здесь мы будем полагать, что стохастические процессы со скачками могут быть представлены в виде интегральной формы стохастического дифференциального уравнения, опирающегося на использование обобщения интеграла Ито на случай стохастически разрывных процессов ([19; 20]).

Отметим, что скачки в случайных процессах часто возникают и в других областях, таких как, финансы, например, при наблюдении и моделировании поведения акций на рынке либо при моделировании поведения портфеля акций в случае принятия решения о замене одного или нескольких активов на другие.

Приведем далее ряд предположений, необходимых для построения модели стада.

2.1. Предположения для модели с непрерывным временем

Будем полагать, что:

1. Количество коров в основном стаде в начальный момент времени $t_0 = 0$ равно N_0 и никогда не превышает этого значения. $N(t)$ обозначает количество животных в основном стаде в момент времени t , $t \in [0, T]$, где $T \in \mathbb{R}$ – некоторое заданное значение.

2. Количество коров в подменном стаде обозначается $M(t)$, в момент времени $t_0 = 0$ равно M_0 и $M(t) \leq M_0$, $t \in [0, T]$.

3. В начальный момент времени $t_0 = 0$ возраст коров в основном стаде генерируется настраиваемым генератором случайных чисел.

4. Генотип потомства следует из генотипа его родителей, определенного через правило суммирования. В этой статье мы рассмотрим только случай суммирования по Менделю, задаваемый правилом (1).

5. С целью упрощения модели с точки зрения моделирования причин удаления животного из основного стада будем считать, что данный процесс подчиняется пуассоновскому процессу интенсивностью λ_D .

6. Аналогично удалению животного из основного стада, для подменного стада также будем полагать, что удаление подчиняется процессу Пуассона интенсивностью λ_d . Отметим, что удаление из подменного стада единообразно описывает и удаление животного по причине смерти либо заболевания, и по причине перевода его в основное стадо.

7. Перемещение животных из подменного стада в основное и удаление из стад происходят один раз в год и моделируются следующим образом: исходя из предположений 1–6, моделируется набор животных текущего года, которые будут удалены из стада; затем вакансии в основном стаде заполняются случайно выбранными животными подменного стада (переменная ξ), достигшими возраста t_{\min} , до заданного максимального размера основного стада; если это окажется невозможным из-за отсутствия животных подходящего возраста в подменном стаде, то заменяется максимально возможное количество животных. Очевидно, что может возникнуть ситуация, когда особей из подменного стада будет недостаточно для заполнения основного стада до максимального размера.

8. Замена животных в подменном стаде также производится один раз в год за счет добавления новорожденных телят, полученных от коров основного стада до определенного ранее размера. Очевидно, что если это окажется невозможным из-за недостатка телят, размер подменного стада уменьшится.

9. Для осеменения всего стада используется детерминированная последовательность известного генотипа, материалом, полученным от одного или нескольких быков-производителей.

10. Приток животных извне отсутствует.

Предположения 5 и 6 введены с целью унификации модели и упрощения процессов численного моделирования. Такой подход позволяет существенно снизить ресурсы, необходимые для проведения симуляции, однако, в случае необходимости, их можно заменить на процессы, которые более точно описывают управление стадом.

Параметры λ_D, λ_d выбираются путем оценки средней продолжительности жизни животных в основном и подменном стадах либо на основе общей практики, либо на основе детального анализа кривой выживаемости. В Новой Зеландии фактическая смерть животного как причина удаления его из любого стада является крайне редкой, но плохое состояние здоровья животного может привести к решению об удалении животного из любого стада.

2.2. Правило суммы одного генотипа

Для построения формальной математической модели представим генотип животного в интересующем локусе как число из $\{-1, 0, 1\}$ и введем термин «индекс генозиготы» (GZI). Здесь -1 и 1 обозначают две гомозиготы (AA и BB) и 0 для гетерозиготных (AB) генотипов соответственно.

Предполагается, что тип наследования в одном локусе менделевский, описываемый следующим правилом суммирования (P задает вероятность каждого результата):

$$\begin{aligned}
 (-1) \dot{+} (-1) &= -1, P = 1; \\
 0 \dot{+} (-1) &= \begin{cases} -1 & P = 0,5 \\ 0 & P = 0,5 \end{cases}; \\
 1 \dot{+} (-1) &= 0, P = 1; \\
 0 \dot{+} 0 &= \begin{cases} -1, & P = 0,25 \\ 0, & P = 0,5 \\ 1, & P = 0,25 \end{cases}; \\
 1 \dot{+} 0 &= \begin{cases} 0, & P = 0,5 \\ 1, & P = 0,5 \end{cases}; \\
 1 \dot{+} 1 &= 1, P = 1.
 \end{aligned} \tag{1}$$

Здесь $\dot{+}$ – коммутативная инфиксная операция, задающая случайное значение генотипа потомка как случайной функции двух переменных соответствующих родительских генотипов.

2.3. Модель с непрерывным временем: интегральная форма

Целью данной статьи было предложить максимально общую модель, которая позволит гибко представлять различные типы управляемых популяций животных. Чтобы выразить такую модель на языке стохастических дифференциальных уравнений, будем полагать, что время непрерывно.

Модели с непрерывным временем могут демонстрировать близость к наблюдаемой динамике стада, но требуют некоторой модификации предположений 6 и 7 в приведенном выше списке предположений относительно случайного времени скачка для процессов удаления из стада.

Будем использовать индекс j для вышеупомянутой формальной позиции в основном стаде и индекс i для позиции в подменном стаде. В нашей стохастической модели взаимодействия двух стад определим по два элементарных эволюционных процесса для каждого из индексов.

В модели процессы $D_j(t)$ или $d_i(t)$ описывают процесс изменения значения генотипа (genezygosis index – GZI) в позициях j или i в основном и подменном стаде. Процессы $A_j(t)$ и $a_i(t)$ отражают процесс изменения возраста животного в позициях. Отметим, что процессы участвуют и в принятии решений относительно характеристик животных в стаде.

Для моделирования процесса удаления животных из стада используются дополнительные независимые процессы Пуассона P_{D_j} и P_{d_i} . Изменения $D_j(t)$, $A_j(t)$ и $d_i(t)$, $a_i(t)$ происходят точно в моменты скачков соответствующего пуассоновского процесса.

Тогда, учитывая предыдущее обсуждение и условия, мы можем записать систему стохастических уравнений, описывающую эволюцию во времени ансамбля коров в основном и подменном стадах в следующем виде:

$$\begin{aligned} D_j(t) &= D_j(0) + \int_0^t (-D_j(s-) + d_{\xi_j}(s-)) dP_{D_j}(s + A_j(0)), \\ d_i(t) &= d_i(0) + \int_0^t (-d_i(s-) + f(D_{\eta_i}(s-), S_{\zeta_i})) dP_{d_i}(s + a_i(0)), \\ A_j(t) &= A_j(0) + t - \int_0^t (A_j(s-) - a_{\xi_j}(s-)) dP_{D_j}(s + A_j(0)), \\ a_i(t) &= a_i(0) + t - \int_0^t a_i(s-) dP_{d_i}(s + a_i(0)), \end{aligned} \quad (2)$$

где $t \in [0, T]$, а другие величины в уравнении (2) определяются следующим образом:

D_j – значение *GZI* (аллеля) для j -й коровы в основном стаде;

P_{D_j} – процесс Пуассона с параметром λ_D , который определяет показатель удаления коровы из основного стада;

A_j – возраст j -й коровы основного стада;

d_i – значение аллеля для i -й коровы подменного стада;

P_{d_i} – процесс Пуассона с параметром λ_d , который определяет показатель удаления коровы из подменного стада;

a_i – возраст i -й коровы подменного стада;

$f(\cdot, \cdot)$ – генотип теленка женского пола как случайная функция родительских генотипов.

Здесь $f(\cdot, \cdot)$ обычно задается правилом суммирования (1), но возможны и другие правила;

η_i – случайная величина, соответствующая случайному выбору коровы из основного стада, которая впоследствии родит самку, которая, в свою очередь, будет помещена в подменное стадо;

ξ – правило для выбора животного из подменного стада, в случае, если животное в основном стаде необходимо заменить;

$\{S_k\}$ – набор значений аллелей для быков-производителей;

величина ζ задает правило для выбора производителя для осеменения из набора $\{S_k\}$.

Опишем более детально систему (2) для j -й позиции основного стада. Пусть в момент времени s произойдет скачок процесса Пуассона для первого уравнения, что означает, что было принято решение об удалении из j -й позиции находящегося на ней животного. Тогда новое значение локуса формируется за счет суммы $-D_j(s-) + d_{\xi}(s-)$, которая позволяет установить новое значение локуса в соответствии с значением правилом ξ . В уравнение для возраста животного в основном стаде в (2) скачок того же процесса Пуассона приводит к изменению в j -й позиции переменной $A_j(s)$, с учетом естественного увеличения возраста (параметр t в этом уравнении). Аналогично в модель введены уравнения для подменного стада. Исключение составляет новое значение генотипа в локусе, которое определяется правилом суммирования (1). Для определения нового значения генотипа животного в подменном стаде используется функция суммирования генотипов f через аргумент $D_{\eta_i}(s-)$.

Для упрощения моделирования здесь и далее мы предполагаем, что все переменные и процессы взаимно независимы. Также полагаем, что интегралы в правой части системы это стохастические интегралы в смысле Ито (см., напр., [21]).

Важно отметить, что стохастический интеграл в смысле Ито по процессу Пуассона, вообще говоря, не определен, однако, принимая во внимание ограниченность выражений под знаком интеграла и тот факт, что $P(t) = \tilde{P}(t) + \lambda t$, где $P(t)$ – процесс Пуассона с параметром λ , а $\tilde{P}(t)$ – это компенсированный процесс Пуассона, соответствующий $P(t)$, являющийся мартингалом, для

которого можно определить интеграл Ито, можно говорить о том, что предложенная модель представляет собой строго определенную систему стохастических дифференциальных уравнений, представленных в соответствующем интегральном виде.

Существуют различные подходы к моделированию, например, аналогичные используемым в эволюционной теории игр [22], которые, как и текущая модель, позволяют использовать при моделировании методы Монте-Карло, однако использованная здесь форма в виде интегральных уравнений позволяет использовать дополнительно приближенные методы слабого типа (например, см. [23; 24]), которые в ряде случаев получают приближенные значения не только для средних, но и для функционалов, заданных на решении представленной системы. Возможность применения таких методов может стать предметом последующих исследований.

2.4. Модель в дискретной форме

При проведении компьютерного моделирования зачастую удобно работать с моделью, использующей дискретное время.

Здесь необходимо обратить внимание на то, что в рассматриваемой задаче обычно отсутствует информация о времени события в стаде, однако доступна информация, предоставляемая ежемесячно или ежегодно. Мы дискретизируем предложенную выше модель по последовательности фиксированных моментов времени $0 = \tau_0 < \tau_1 < \dots < \tau_L = T$.

Будем полагать, что в каждом из L интервалов $(\tau_{l-1}, \tau_l]$ вероятность более чем одного скачка процесса Пуассона (управляющее воздействие) в пределах интервала пренебрежимо мала, или, в качестве альтернативы, несколько скачков можно заменить одним скачком, величина которого равна сумме отдельных скачков. Тогда мы приходим к следующей дискретной модели:

$$\begin{aligned} D_j(t) &= D_j(0) + \sum_{l=1}^L (-D_j(\tau_{l-1}) + d_\xi(\tau_{l-1})) \text{sign}[\mathcal{P}(\lambda_D A_j(\tau_{l-1}))] I_{[0,t)}(\tau_l), \\ d_i(t) &= d_i(0) + \sum_{l=1}^L (-d_i(\tau_{l-1}) + f(D_\eta(\tau_{l-1}), S_\zeta)) \text{sign}[\mathcal{P}(\lambda_d a_i(\tau_{l-1}))] I_{[0,t)}(\tau_l), \\ A_j(t) &= A_j(0) + \sum_{l=1}^L (1 + (a_\xi(\tau_{l-1}) - A_j(\tau_{l-1})) \text{sign}[\mathcal{P}(\lambda_D A_j(\tau_{l-1}))]) I_{[0,t)}(\tau_l), \\ i(t) &= i(0) + \sum_{l=1}^L (1 - i(\tau_{l-1}) \text{sign}[\mathcal{P}(\lambda_d a_i(\tau_{l-1}))]) I_{[0,t)}(\tau_l). \end{aligned} \quad (3)$$

Здесь $I_{[0,t)}(u)$ равна 1, если $u \in [0, t)$ и 0 в противном случае, а $\mathcal{P}(k)$ – независимые случайными величинами Пуассона.

2.5. Модель с притоком животных

При проведении моделирования в описанной в предыдущем разделе модели наблюдается «эффекта вымирания», который упоминался в литературе. На практике такой эффект сокращения численности популяции не наблюдается, в связи с тем, что фермер компенсирует недостаток животных за счет их покупки на рынке.

Ниже мы предлагаем способ учета такого способа управления стадом в предложенной выше модели. Будем полагать, что новое животное всегда помещается в подменное стадо. Такое условие позволяет включить приток животных извне просто за счет изменения значения переменной η в уравнении (3) не внося существенных изменений в используемые уравнения. Для переменной η предположим, что ее нулевое значение будет соответствовать рождению самца животного, а для рожденных особей женского пола значение переменной по-прежнему равно индексу матери в основном стаде. Тогда уравнения для подменного стада примут вид:

$$\begin{aligned}
d_i(t) &= d_i(0) + \sum_{l=1}^L (-d_i(\tau_{l-1}) + \text{sign}(\eta) f(D_\eta(\tau_{l-1}), S_\zeta) + \\
&\quad + (1 - \text{sign}(\eta)) D_M) \text{sign}[P_d(\lambda_d a_i(\tau_{l-1}))] I_{[0,t)}(\tau_l), \\
a_i(t) &= a_i(0) + \sum_{l=1}^L \left(1 + ((1 - \text{sign}(\eta)) a_M - \right. \\
&\quad \left. - \text{sign}(\eta) a_i(\tau_{l-1})) \text{sign}[P_d(\lambda_d a_i(\tau_{l-1}))] \right) I_{[0,t)}(\tau_l). \tag{4}
\end{aligned}$$

В этих уравнениях мы вводим случайную величину D_M , которая задает распределение смоделированного аллеля у животных с рынка. Распределение для D_M необходимо определить экспериментально или каким-либо еще образом. a_M – это аналогичный предыдущему тип случайной величины, которая задает распределение возраста коров на рынке.

Из уравнений (4) следует, что при $\eta = 0$, т. е. в случае рождения самца, множитель $1 - \text{sign}$ – ненулевой вклад в члены с индексом M , что используется как признак для включения животных извне. В случае же, когда $\eta > 0$, данный множитель равен нулю и, соответственно, исключает влияние внешнего притока животных и приводит систему к исходному виду (3). Еще одним техническим преимуществом последней предложенной модели является то, что теперь размер подменного стада постоянен, что несколько упрощает численное моделирование.

В данной статье мы не намерены приводить результаты моделирования последней предложенной модели и ограничиваемся только ее формулировкой.

3. Результаты моделирования динамики генотипа

Приведем далее результаты моделирования в предположении, что суммирование генотипа осуществляется согласно (1) и при условии контролируемого спаривания. Как и ранее, будем использовать $-1, 0, 1$ при кодировании генотипа. При моделировании для каждого животного мы начинаем с коровы с генотипом -1 (гомозиготный с аллелем, который необходимо устранить). Самка, полученное от нее потомство и так далее осеменяются последовательностью производителей с известным генотипом. Зададим дискретизацию по времени с постоянным интервалом в один год, то есть полагаем $\tau_l - \tau_{l-1} = 1$ (год), $l = 0, 1, \dots, L$. В начальный момент времени τ_0 также предполагается, что распределение аллелей в подменном стаде такое же, как и в основном. Это приводит, как мы увидим позже, к двухлетней задержке при переключении к новому значению.

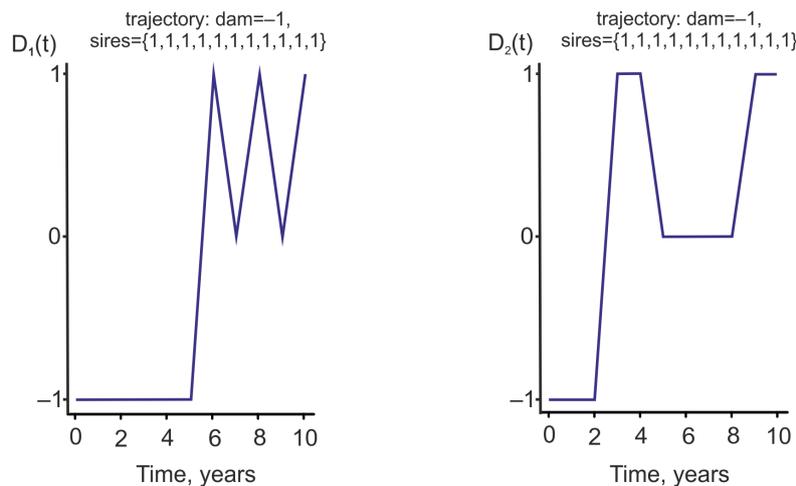


Рис. 1. Две случайно выбранные траектории стохастического процесса D_1, D_2 для заданной позиции в основном стаде

Сначала рассмотрим динамику безусловного переключения генотипа в состояние 1 (гомозиготный с аллелем, который будет введен). Для достижения цели будем использовать последовательность производителей, которые имеют генотип 1. На рис. 1 приведены две реализации единичной траектории случайного процесса для такого случая.

Стоит отметить, что каждый график показывает зависимость GZI в конкретной, скажем в j -й, позиции в основном стаде. На рис. 1 только прыжки можно интерпретировать как замену животного, тогда как горизонтальные линии могут указывать как на замену коровы на животное с тем же генотипом, так и на отсутствие замены.

Подчеркнем, что отдельные траектории (рис. 1) и пучки траекторий (рис. 2, 3) перехода из одного состояния в другое заметно отличаются от среднего (рис. 4).

Далее, на рис. 2 мы строим набор из 100 траекторий с небольшим возмущением, чтобы на графике можно было выделить вероятность каждой отдельной траектории. Плотность линий позволяет визуальнo оценить численность животных в отдельных состояниях $\{-1, 0, 1\}$.

Особое внимание стоит обратить на рис. 3. На нем, как и на рис. 2, рассмотрено 50 траекторий, и всего у одного из производителей значение GZI равно -1 , однако это оказывает существенное влияние на переход к желаемому значению в основном стаде. Напомним также, что почти исчезающее значение GZI , равное -1 , в конце периода моделирования на рис. 2 и 3 не означает, что значение -1 было полностью исключено из стада, т. е. существуют траектории, у которых значение по прошествии 10 лет будет равно -1 .

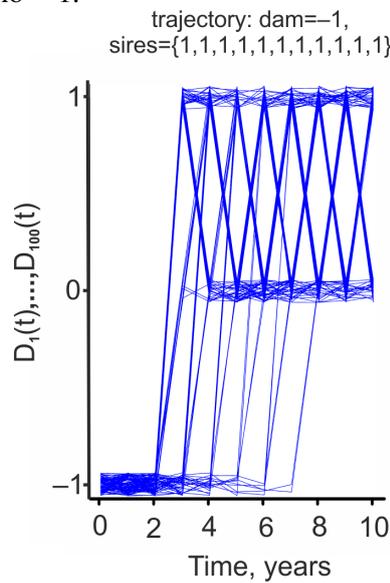


Рис. 2. Набор из 100 траекторий стохастического процесса D_1, \dots, D_{100} (небольшое случайное возмущение было добавлено на каждом временном шаге для лучшей визуализации индивидуальной траектории), все быки в последовательности с $GZI = 1$

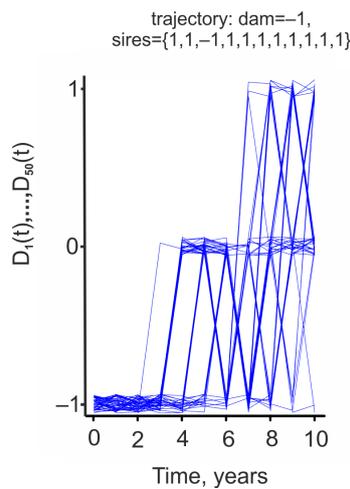


Рис. 3. Набор из 50 траекторий стохастического процесса D_1, \dots, D_{50} , один бык в последовательности (третий) с $GZI = -1$

На рис. 4 мы показываем временную эволюцию среднего значения и дисперсии для случая 1000 стад. Каждая симуляция начинается со стада из 200 гомозиготных самок в основном стаде (все начальные GZI равны -1) и затем стада развиваются независимо. Среднее значение D определяется

стандартным способом как $\text{mean } D(t) \equiv \bar{D}(t) = 1/N \sum_j^N D_j(t)$, где N – общее количество животных во всех стадах. Дисперсия σ^2 рассчитывается как дисперсия среднего для стад.

Чтобы записать формулу для дисперсии, добавим индекс k к $D_j(t)$ для подсчета заданного количества стад, тогда

$$\sigma^2 = \frac{1}{N_{\text{herds}} - 1} \times \sum_{k=1}^{N_{\text{herds}}} \left(\bar{D}(t) - \frac{1}{N_{\text{size}}} \sum_{j=1}^{N_{\text{size}}} D_j^k(t) \right)^2,$$

где N_{herds} – это количество моделируемых стад; N_{size} – это количество коров в стаде (при моделировании использовались стада размером 100, 200 или 400 животных).

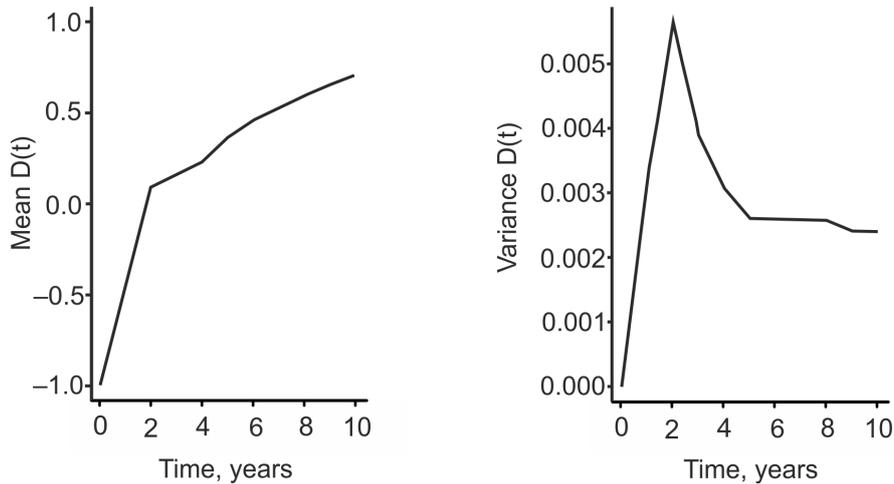


Рис. 4. Среднее (слева) и дисперсия (справа) значения GZI для 1000 стад, размер каждого основного стада составляет 200, размер подменного стада составляет 100 в начальный момент времени, $\lambda_D = 0,114$, $\lambda_d = 0,25$

Важной особенностью, обнаруженной в ходе этого моделирования, является то, что размер стада со временем достаточно быстро падает и стабилизируется на уровне, меньшем заданного в качестве верхней границы. На рис. 5 показан график для среднего значения размера стада, что похоже на «частичное вымирание».

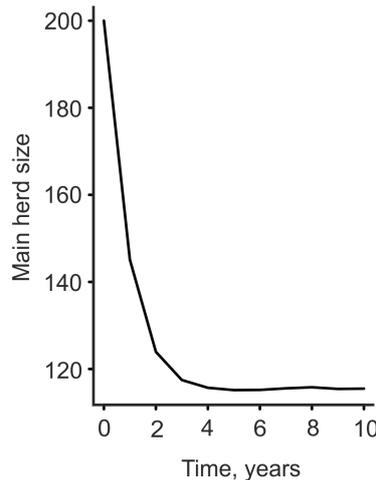


Рис. 5. Зависимость среднего размера основного стада от времени

Это можно объяснить тем фактом, что положительные колебания числа рожденных особей женского пола ограничиваются заданным предельным размером стада, в то время как редкие сильные отрицательные колебания численности рождающихся самок создают дефицит коров в подменном стаде и серьезно влияют на последующую динамику размера стада. Таким образом, использование подменного стада сильно влияет на основное стадо. Представленные результаты по основному стаду следует рассматривать только как демонстрацию тенденций, потому что в данном исследовании мы не стремились изучать влияние управления подменным стадом или пытаться как-то оптимизировать его.

В реальности, на первый взгляд неприятный «эффект вымирания» компенсируется притоком животных с рынка и может быть использован для более быстрого переключения всего стада на заданный аллель за счет введения в стадо животных, у которых интересующее значение уже присутствует, что, в свою очередь, отразится на последующем воспроизводстве потомства.

Интенсивность процесса Пуассона λ_D для основного стада напрямую влияет на уровень «частичного вымирания», поскольку, как это можно увидеть из графиков на рис. 6–10, на которых мы демонстрируем переключение от определенного знака (был выбран равным -1) аллеля в значение 1 , скорость переключения зависела от значений λ_D . На рисунках графики для λ_D , соответствующих вероятности, равной $0,8$, для коровы дожить в основном стаде до $4, 6, 8, 10, 12$ лет, обозначены красным, голубым, зеленым, синим, черным цветами и метками $1, 2, 3, 4, 5$ соответственно. Моделирование было проведено по 10000 стад.

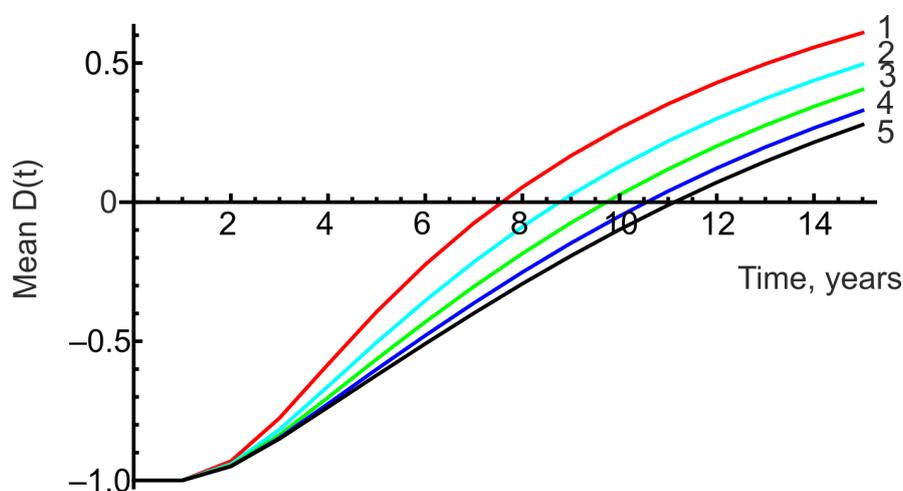


Рис. 6. Среднее значение GZI во времени для 10000 стад, размер каждого основного стада составляет 400 голов, размер подменного стада в начальный момент времени составляет 200 голов

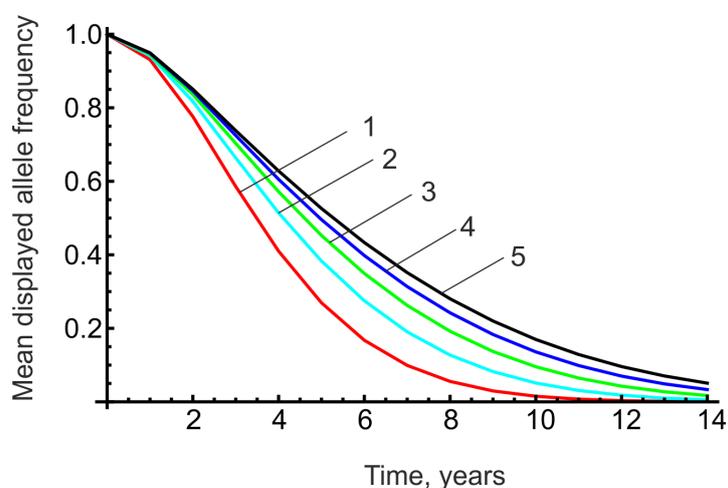


Рис. 7. Средняя частота в популяции замещающего аллеля по времени для эволюции 10000 стад, размер каждого основного стада составляет 400 , размер подменного стада в начальный момент времени составляет 200 . Последовательность цветов (и числовых меток) соответствует разным значениям средней продолжительности жизни коровы в основном стаде (см. подробности в тексте)

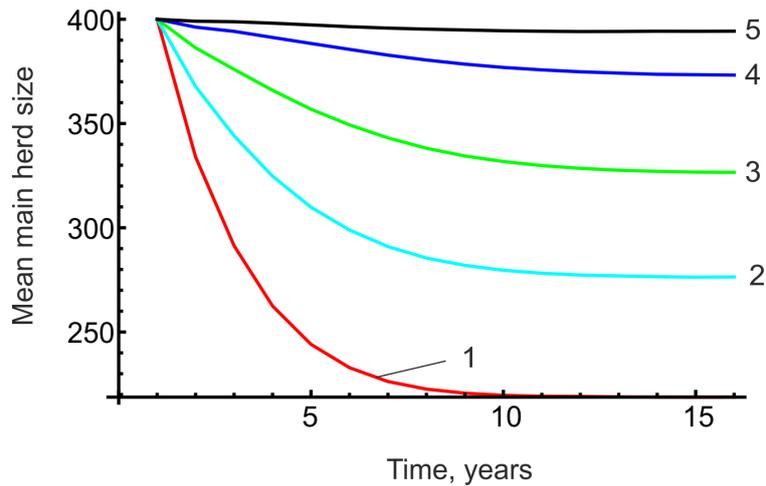


Рис. 8. Изменение среднего размера основного стада во времени для 10000 стад, размер каждого основного стада составляет 400 голов, размер подменного стада в начальный момент времени составляет 200 голов. Последовательность цветов (и числовых меток) соответствует разным значениям средней продолжительности жизни коровы в основном стаде (см. подробности в тексте)

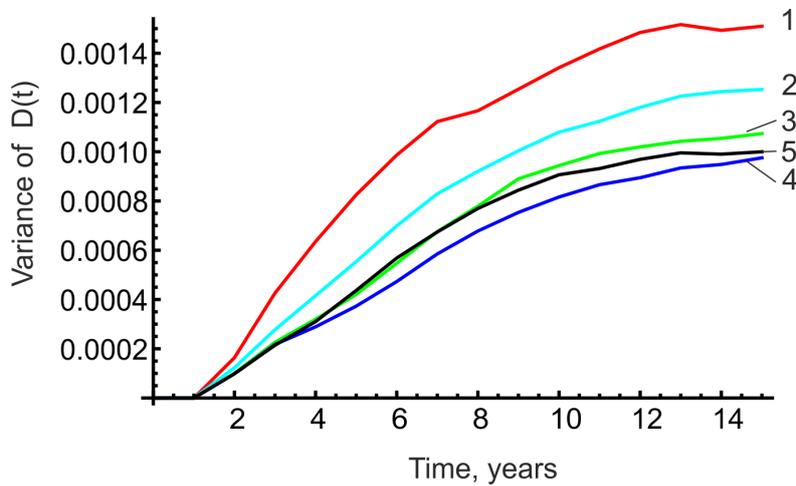


Рис. 9. Дисперсия среднего изменения значения GZI для 10000 стад, размер каждого основного стада составляет 400 голов, размер подменного стада в начальный момент времени составляет 200 голов. Последовательность цветов (и числовых меток) соответствует разным значениям средней продолжительности жизни коровы в основном стаде (см. подробности в тексте)

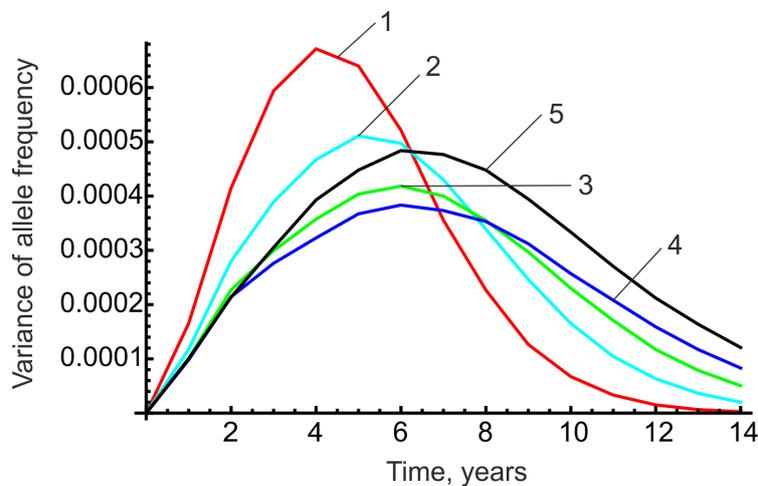


Рис. 10. Дисперсия частоты замещаемого аллеля для 10000 стад, размер каждого основного стада составляет 400 голов, размер подменного стада в начальный момент времени составляет 200 голов. Последовательность цветов (и числовых меток) соответствует разным значениям средней продолжительности жизни коровы в основном стаде (см. подробности в тексте)

Одной из интересных особенностей предложенной модели дойного стада является чувствительность к деталям перехода животного из подменного стада в основное. Рассмотрим две модели управления процессом исключения и замещения животных в стаде.

В модели 1 мы заполняем основное стадо до предельного размера сначала из подменного стада, затем случайным образом осуществляем удаление животных из подменного стада из числа оставшихся.

В модели 2 сначала делаем удаление коров из стада (и основного, и подменного), и лишь потом животных, удаленных из подменного стада, вводим в основное стадо.

Оказывается, что эти две немного отличающиеся процедуры дают заметно различное поведение, что показано на рис. 11 для средних значений, для дисперсии – рис. 12 и для среднего размера основного стада – рис. 13. Динамика среднего значения GZI очень близка, отклонения немного отличаются, но средний размер стада указывает на существенное влияние схемы управления.

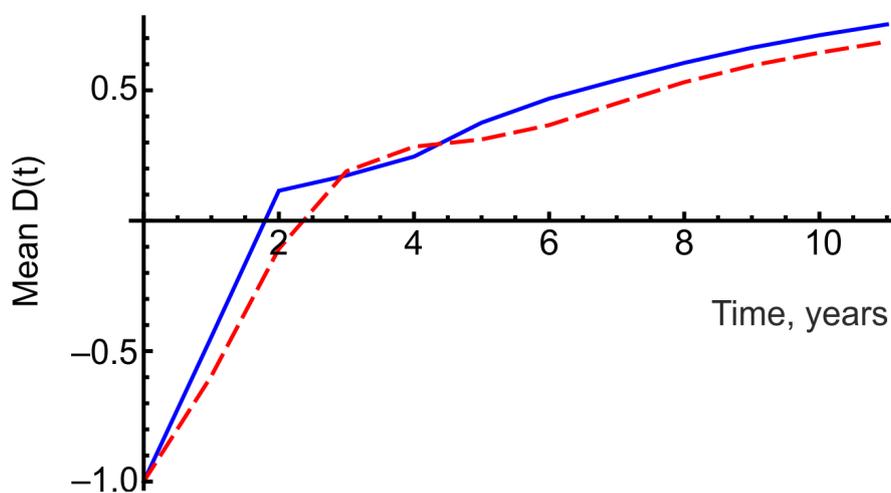


Рис. 11. Сравнение среднего значения аллеля основного стада GZI по времени для двух моделей. Пунктирная красная линия – первая модель, сплошная синяя линия – вторая модель. Все параметры такие же, как на рис. 4

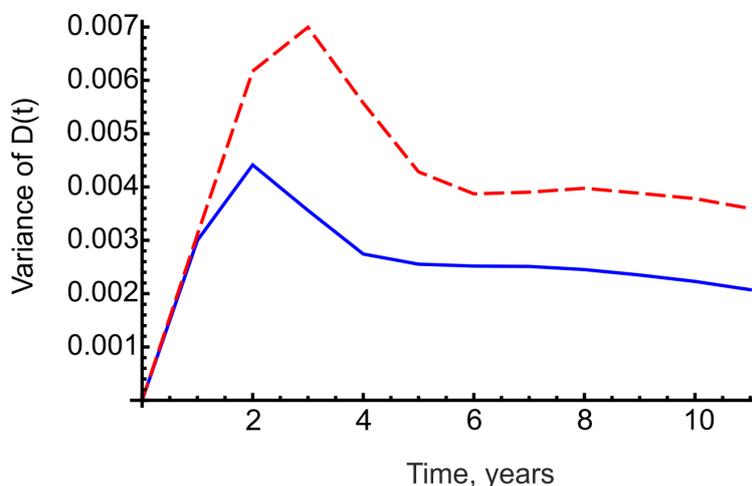


Рис. 12. Сравнение усредненной дисперсии аллеля основного стада GZI по времени для двух моделей. Пунктирная красная линия – первая модель, сплошная синяя линия – вторая модель

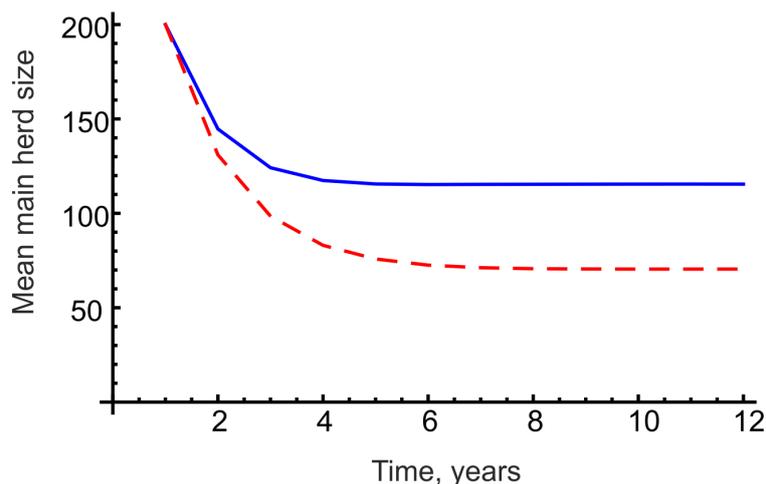


Рис. 13. Сравнение среднего размера основного стада для двух моделей. Пунктирная красная линия – первая модель, сплошная синяя линия – вторая модель

Рис. 12, 13 обращают внимание на то, что необходимо очень внимательно относиться к правилам, используемым в управлении стадом, поскольку на первый взгляд близкие правила приводят к существенным различиям.

4. Заключение

В данной работе для целей моделирования временной эволюции биаллельного полиморфизма дойного стада была предложена модель, основанная на использовании стохастических дифференциальных уравнений. Модель включает ряд дополнительных условий: разделение стада на основное и подменное стадо, типичное время жизни животного в стаде, ограничение размера основного и подменного стада, приток замены животных в основном стаде.

Моделирование показало, что при установлении максимального размера стада локальные колебания новорожденных животных будут сильно влиять на динамику системы и приводить к заметному уменьшению размера стада («частичному вымиранию»). Для устранения этого явления, которое не наблюдается в реальных фермерских ситуациях, была предложена модификация модели с учетом притока животных извне. Моделирование показало, что политика управления животными в стаде может иметь большое значение для достижения целей оптимизации.

Благодарности

Мы выражаем благодарность Джеку Хуперу из Livestock Improvement Corporation за прочтение черновиков этой статьи и за ценные замечания по существующей в Новой Зеландии практике управления стадом.

Литература

1. Hardy G. H. Mendelian proportions in a mixed population // *Science*. 1908. N 28. P. 49–50.
2. Weinberg W. Über den nachweis der vererbung beim menschen // *W. Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg*. 1908. N 64. P. 368–382.
3. Wright S. Evolution in Mendelian populations // *Genetics*. 1931. N 16. P. 97–159.
4. Fisher R. A. *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press, 1931.
5. Feng S. *Poisson-Dirichlet Distribution and Related Topics: Models and asymptotic behaviors*. Berlin, Heidelberg: Springer-Verlag, 2010.
6. Donnelly P., Kurtz T. G. Genealogical processes for Fleming–Viot models with selection and recombination // *Annals of Applied Probability*. 1999. Vol. 9, N 4. P. 1091–1148.
7. Kaj I., Krone S. M. The coalescent process in a population with stochastically varying size // *Journal of Applied Probability*. 2003. Vol. 40, N 1. P. 33–48.

8. Steinsaltz D., Evans S. N., Wachter K. W. A generalized model of mutation-selection balance with applications to aging // *Advances in Applied Mathematics*. 2005. Vol. 35, N 1. P. 16–33.
9. Moran P. A. P. Random processes in genetics // *Mathematical Proceedings of the Cambridge Philosophical Society*. 1958. Vol. 54. P. 60–71.
10. Feller W. Diffusion processes in genetics // *Proc. Second Berkeley Symp. Math. Statist. Prob.*, 1951. P. 227–246.
11. Dawson D. A. "Stochastic population dynamics. Lectures at Summer School in Probability, at PIMS-UBC" [Electronic resource]. – Mode of access: <https://personal.math.ubc.ca/~db5d/SummerSchool09/LectureNotes.html>. – Date of access: 12.04.2025.
12. Pfaffelhuber P., Pennings P., Hermisson J. "Tutorial: Population genetics. Mathematics and BioSciences Group, University of Vienna" [Electronic resource]. – Mode of access: https://www.mabs.at/fileadmin/user_upload/p_mabs/popgen2009.pdf. – Date of access: 12.04.2025.
13. Draghi J. A., Parsons T. L., Plotkin J. B. Epistasis increases the rate of conditionally neutral substitution in an adapting population // *Genetics*. 2011. Vol. 187. P. 1139–1152.
14. Harris B., Johnson D. Genomic predictions for New Zealand dairy bulls and integration with national genetic evaluation // *Journal of Dairy Science*. 2010. Vol. 93. P. 1243–1252.
15. Inbreeding and recessive genes [Electronic resource]. – Mode of access: <https://www.lic.co.nz/products-and-services/artificial-breeding/inbreeding-and-recessive-genes/>. – Date of access: 12.04.2025.
16. Grisart B., Farnir F., Karim L., Cambisano N., Kim J. J. et al. Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition // *Proceedings of the National Academy of Sciences of the United States of America*. 2004. Vol. 101. P. 2398–2403.
17. Allais-Bonnet A., Grohs C., Medugorac I., Krebs S., Djari A. et al. Novel Insights into the Bovine Polled Phenotype and Horn Ontogenesis in Bovidae // *PLoS ONE*. 2013. Vol. 8. Art. e63512. <https://doi.org/10.1371/journal.pone.0063512>
18. Morris A. P., Zeggini E. An evaluation of statistical approaches to rare variant analysis in genetic association studies // *Genetic epidemiology*. 2010. Vol. 34. P. 188–193.
19. Øksendal B. K., Sulem A. *Applied stochastic control of jump diffusions*. Springer, 2005.
20. Applebaum D. *Lévy processes and stochastic calculus*. Cambridge University Press, 2009.
21. Gardiner C. W. *Handbook of stochastic methods: for Physics, Chemistry and the Natural Sciences*. Springer, 2004.
22. Smith M. *Evolution and the Theory of Games*. Cambridge: Cambridge University Press, 1982.
23. Egorov A. D., Sobolevsky P. I., Yanovich L. F. *Functional Integrals: Approximate Evaluation and Applications*. Springer, 1993.
24. Zherelo A. Approximate Calculation of Mathematical Expectations on Processes with a Drift // *Nonlinear Phenomena in Complex Systems*. 2015. Vol. 18, N 2. P. 207–214.

References

1. Hardy G. H. Mendelian proportions in a mixed population. *Science*, 1908, no. 28, pp. 49–50.
2. Weinberg W. Über den nachweis der vererbung beim menschen. *W. Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg*, 1908, no. 64, pp. 368–382.
3. Wright S. Evolution in Mendelian populations. *Genetics*, 1931, no. 16, pp. 97–159.
4. Fisher R. A. *The Genetical Theory of Natural Selection*. Oxford, Clarendon Press, 1931.
5. Feng S. *Poisson-Dirichlet Distribution and Related Topics: Models and asymptotic behaviors*. Berlin, Heidelberg, Springer-Verlag, 2010.
6. Donnelly P., Kurtz T. G. Genealogical processes for Fleming–Viot models with selection and recombination. *Annals of Applied Probability*, 1999, vol. 9, no. 4, pp. 1091–1148.
7. Kaj I., Krone S. M. The coalescent process in a population with stochastically varying size. *Journal of Applied Probability*, 2003, vol. 40, no. 1, pp. 33–48.
8. Steinsaltz D., Evans S. N., Wachter K. W. A generalized model of mutation-selection balance with applications to aging. *Advances in Applied Mathematics*, 2005, vol. 35, no. 1, pp. 16–33.

9. Moran P. A. P. Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 1958, vol. 54, pp. 60–71.
10. Feller W. Diffusion processes in genetics. *Proc. Second Berkeley Symp. Math. Statist. Prob.*, 1951, pp. 227–246.
11. Dawson D. A. "Stochastic population dynamics. Lectures at Summer School in Probability, at PIMS-UBC". – URL: <https://personal.math.ubc.ca/~db5d/SummerSchool09/LectureNotes.html> (accessed 12.04.2025)
12. Pfaffelhuber P., Pennings P., Hermisson J. "Tutorial: Population genetics. Mathematics and BioSciences Group, University of Vienna". – URL: https://www.mabs.at/fileadmin/user_upload/p_mabs/popgen2009.pdf (accessed 12.04.2025).
13. Draghi J. A., Parsons T. L., Plotkin J. B. Epistasis increases the rate of conditionally neutral substitution in an adapting population. *Genetics*, 2011, vol. 187, pp. 1139–1152.
14. Harris B., Johnson D. Genomic predictions for New Zealand dairy bulls and integration with national genetic evaluation. *Journal of Dairy Science*, 2010, vol. 93, pp. 1243–1252.
15. *Inbreeding and recessive genes*. – URL: <https://www.lic.co.nz/products-and-services/artificial-breeding/inbreeding-and-recessive-genes/> (accessed 12.04.2025).
16. Grisart B., Farnir F., Karim L., Cambisano N., Kim J. J. et al. Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. *Proceedings of the National Academy of Sciences of the United States of America*, 2004, vol. 101, pp. 2398–2403.
17. Allais-Bonnet A., Grohs C., Medugorac I., Krebs S., Djari A. et al. Novel Insights into the Bovine Polled Phenotype and Horn Ontogenesis in Bovidae. *PLoS ONE*, 2013, vol. 8, art. e63512. <https://doi.org/10.1371/journal.pone.0063512>
18. Morris A. P., Zeggini E. An evaluation of statistical approaches to rare variant analysis in genetic association studies. *Genetic epidemiology*, 2010, vol. 34, pp. 188–193.
19. Øksendal B. K., Sulem A. *Applied stochastic control of jump diffusions*. Springer, 2005.
20. Applebaum D. *Lévy processes and stochastic calculus*. Cambridge University Press, 2009.
21. Gardiner C. W. *Handbook of stochastic methods: for Physics, Chemistry and the Natural Sciences*. Springer, 2004.
22. Smith M. *Evolution and the Theory of Games*. Cambridge, Cambridge University Press, 1982.
23. Egorov A. D., Sobolevsky P. I., Yanovich L. F. *Functional Integrals: Approximate Evaluation and Applications*. Springer, 1993.
24. Zherelo A. Approximate Calculation of Mathematical Expectations on Processes with a Drift. *Nonlinear Phenomena in Complex Systems*, 2015, vol. 18, no. 2, pp. 207–214.



ТЕОРИЯ ВЕРОЯТНОСТЕЙ И
МАТЕМАТИЧЕСКАЯ СТАТИСТИКА
PROBABILITY THEORY AND
MATHEMATICAL STATISTICS



UDC 511.42

STEADY-STATE ANALYSIS OF THE MULTI-SERVER RETRIAL QUEUEING
SYSTEM WITH HETEROGENEOUS SERVERS AND PHASE TYPE DISTRIBUTION
OF SERVICE TIMES

M. Liu, A. N. Dudin

Belarusian State University, Minsk, Belarus
e-mail: liumei19910101@126.com, dudin@bsu.by

Received: 21.04.2025

Revised: 23.05.2025

Accepted: 23.05.2025

Keywords: Markovian arrival process, retrials, heterogeneous servers, phase-type distribution, asymptotically quasi-Toeplitz Markov chains.

Abstract. A multi-server retrial queueing system with heterogeneous servers is analysed. The service times have a phase-type distribution with different irreducible representations. Customer arrival to the system is defined by a Markovian arrival process. When all servers are busy at an arrival moment, the customer moves to the virtual place called orbit to retry to reach the servers in exponentially distributed periods of time. The total retrial rate from the orbit infinitely increases when the number of customers residing in orbit grows. Upon arrival or retrial from the orbit, a customer occupies the server having the minimal number among all idle servers, if any. The dynamics of the system states is described by a multidimensional Markov chain having the special block structure of the infinitesimal generator. The explicit expression for this is presented. Ergodicity condition is derived. The expressions for computation of the key performance characteristics of the system are given. Numerical results, which highlight dependencies of these measures on the mean arrival rate for the system and its particular cases, when the arrivals are described by the stationary Poisson process or (and) service times follow the exponential distribution, are presented.

СТАЦИОНАРНЫЙ АНАЛИЗ МНОГОЛИНЕЙНОЙ СИСТЕМЫ МАССОВОГО
ОБСЛУЖИВАНИЯ С НЕОДНОРОДНЫМИ ПРИБОРАМИ И РАСПРЕДЕЛЕНИЕМ
ВРЕМЕНИ ОБСЛУЖИВАНИЯ ФАЗОВОГО ТИПА

Лю Мэй, А. Н. Дудин

Белорусский государственный университет, Минск, Беларусь
e-mail: liumei19910101@126.com, dudin@bsu.by

Поступила: 21.04.2025

Исправлена: 23.05.2025

Принята: 23.05.2025

Ключевые слова: Марковский процесс поступления, повторные попытки, неоднородные серверы, распределение фазового типа, асимптотически квазитеплицевы цепи Маркова.

Аннотация. Анализируется многолинейная система обслуживания с повторными вызовами и неоднородными серверами. Длительности обслуживания имеют фазовое распределение с различными неприводимыми представлениями. Поступление запросов в систему определяется марковским процессом поступления. Когда все серверы заняты в момент поступления, запрос помещается в виртуальное место, называемое орбитой, чтобы повторить попытку достичь серверов через экспоненциально распределенные периоды времени. Общая скорость повторных вызовов с орбиты бесконечно увеличивается с ростом числа запросов, находящихся на орбите. При поступлении или повторных вызовах с орбиты запрос занимает сервер с минимальным номером среди всех свободных серверов, если таковые имеются. Динамика состояний системы описывается многомерной цепью Маркова, имеющей специальную блочную структуру инфинитезимального генератора. Представлено явное выражение для генератора. Выведено условие эргодичности. Приведены выражения для вычисления ключевых характеристик производительности системы. Представлены численные результаты, иллюстрирующие зависимости характеристик производительности системы от средней скорости поступления заявок для системы и ее частных случаев, когда поступления описываются стационарным пуассоновским процессом или (и) времена обслуживания подчиняются экспоненциальному распределению.

1. Introduction

Retrial queues fit well for the mathematical description of various real-world systems, telecommunication networks, including the mobile cellular networks, and contact centers in particular. Analysis of such queues is much more involved than the study of the queues with customer loss or buffers for waiting in case of the lack of available servers having the same types of arrival and service processes. This is explained by to the state inhomogeneous behavior of the stochastic process describing the dynamics of the system. This is the reason why the retrial queues are investigated in a far less extent.

The fundamental results obtained for the multi-server retrial queues of the $M/M/N$ type (this means that the arrivals are described by the stationary Poisson process and service times follow the exponential distribution) are presented in the well-known [1]. However, due to the significant change of the character of the flows in communication networks and customers service time during the last decades, essentially more adequate model of arrivals in the modern real-world systems is the MAP (Markovian Arrival Process), see, e. g., [2–5]. This process well describes the modern correlated bursty flows. Essentially more general distribution of service time than the exponential one is Phase-Type (PH) distribution, see, e. g., [5; 6] which allows to fit not only the mean service time, but also higher moments, including the variance. Taking these circumstances into account, the $BMAP/PH/N$ type retrial queue was studied in [7]. But only the aspects relating to the ergodicity condition of the system are considered there. More comprehensive analysis of the $BMAP/PH/N$ type retrial queue was given [8] where, besides to the proof of the sufficient condition for the ergodicity in cases of the classical retrial policy and the constant retrial rate, the effective algorithms for the computation of the stationary distribution of the system states and the main performance measures were presented.

Essential assumption imposed in [1] and [8] is that the servers are identical, while they can be heterogeneous in some real-world system. In this paper, we significantly weaken this restrictive assumption. In the papers [9] and [10], this assumption was already weakened and the servers are not identical. In that papers, it was supposed that service times have the exponential distributions with different parameters. In the present paper, we suppose that service times have the Phase-Type distributions with different parameters.

Generalization from the case of the exponential distribution to the case of the Phase-Type distribution has the practical importance because the former one allows to fit only the average value of the real service time while the latter one allows to fit simultaneously many initial moments of the distribution of the real service time, and the variance of this time in particular. From the theoretical point of view, this generalization leads to the necessity of construction and analysis of the multi-dimensional continuous-time Markov chain with more involved structure of the blocks on the infinitesimal generator. Here this analysis is successfully implemented.

The outline of the presentation is as follows. The mathematical model under study is described in Section 2. The random process describing the dynamics of the considered system is introduced in Section 3 as the multi-dimensional continuous-time Markov chain and the explicit expression for the generator of the chain is presented there. The sufficient condition for the ergodicity of this Markov chain is derived in Section 4. Formulas for computation of the values of the key performance measures of the system are given in Section 5. The illustrative numerical results are presented in Section 6. Section 7 contains some concluding remarks.

2. The mathematical model

We consider an N -server queueing system. The primary customers arrive to the system according to a MAP (Markovian Arrival Process). We denote the directing process of the MAP by ν_t , $t \geq 0$. The state space of the irreducible continuous-time Markov chain ν_t is $\{0, 1, \dots, W\}$. The intensities of transitions of the process ν_t are defined as the entries of the square matrices D_0 and D_1 of size $\bar{W} = W + 1$. The matrix D_0 contains the intensities of transitions at which customers do not arrive. The matrix D_1 contains the intensities of transitions at which customer arrives to the system. The matrix $D(1) = D_0 + D_1$ is an infinitesimal generator of the process ν_t . The vector θ that is the unique solution to the system of equations $\theta D(1) = 0$, $\theta e = 1$ defines the stationary distribution of the process ν_t . Here

and thereafter e is a column vector of an appropriate size consisting of 1's and $\mathbf{0}$ is a row vector of an appropriate size consisting of zeroes.

The average (fundamental) arrival rate λ of the *MAP* is defined as $\lambda = \boldsymbol{\theta} D_1 e$. The coefficient c_{var} of variation of intervals between customer arrivals is defined by $c_{\text{var}} = 2\lambda\boldsymbol{\theta}(-D_0)^{-1}e - 1$. The coefficient of correlation c_{cor} of successive intervals between arrivals is computed as $c_{\text{cor}} = (\lambda\boldsymbol{\theta}(-D_0)^{-1}D_1(-D_0)^{-1}e - 1)/c_{\text{var}}^2$.

The servers are independent of each other. The service time of a customer by n -th server, $n = \overline{1, N}$, is governed by the continuous-time Markov chain (directing process) $\eta_t^{(n)}$. This process has an absorbing state 0 and the set $\{1, \dots, M^{(n)}\}$ of transient states. The initial state of the process $\eta_t^{(n)}$ at the epoch of starting the service is chosen among the transient states with the probabilities defined by the entries of the row-vector $\boldsymbol{\beta}^{(n)} = (\beta_1^{(n)}, \dots, \beta_{M^{(n)}}^{(n)})$. The transitions of the process $\eta_t^{(n)}$ inside the set of transient states do not lead to service completion and are defined by the entries of the irreducible matrix $S^{(n)}$ of size $M^{(n)}$. The diagonal entries of this matrix are negative. Their modules define the rates of the exit of the process $\eta_t^{(n)}$ from its transient states. The non-diagonal entries define the intensities of transitions inside the set of the transient states. The rates of transition to the absorbing state, which lead to service completion, are defined by the entries of the column vector $S_0^{(n)} = -S^{(n)}e$.

The m th initial moment $b_m^{(n)}$ of the distribution of the service time in the n th server is computed as

$$b_m^{(n)} = m! \boldsymbol{\beta}^{(n)} ((-S)^{(n)})^{-m} e, \quad m \geq 1.$$

The value μ_n defined by the formula $\mu_n^{-1} = \boldsymbol{\beta}^{(n)} (-S^{(n)})^{-1} e$, $n = \overline{1, N}$, is the mean service rate in the n th server. The value $\frac{b_2^{(n)} - (b_1^{(n)})^2}{(b_1^{(n)})^2}$ is the squared coefficient of variation of the service time in the n th server.

If the arriving customer meets all servers being idle, the customer enters the first server to receive the service. If the first server is busy, then the customer enters the idle server with the minimum number. If all servers are busy, then the customer goes to the virtual place called orbit. Capacity of the orbit is unlimited. These customers are said to be repeated customers. These customers try their luck later until they will be served. We assume that the total flow of retrials from the orbit is such that the probability of generating the retrial attempt in the small interval $(t, t + \Delta t)$ is equal to $\alpha_i \Delta t + o(\Delta t)$ when the orbit size (the number of customers on the orbit) is equal to i , $i > 0$, $\alpha_i = 0$. We do not fix the explicit dependence of the intensities α_i on i . We assume the infinitely increasing retrial rate: $\lim_{i \rightarrow \infty} \alpha_i = \infty$. This holds true, in particular, for classic retrial strategy where $\alpha_i = i\alpha$ and the linear strategy $\alpha_i = i\alpha + \gamma$.

Our goal is to obtain the sufficient condition for existence of stationary state distribution of the system, this distribution itself, and the expressions for the key performance measures of the system via this stationary distribution.

3. The random process defining the dynamics of the system

Let, at the moment t , $t > 0$,

i_t be the number of customers on the orbit, $i_t \geq 0$;

$\eta_t^{(n)}$ be the state of the underlying process of the service in the n th server, $n = \overline{1, N}$. This state belongs to the set $\{1, \dots, M^{(n)}\}$ if this server is busy and is assumed to be 0 if the server is idle;

ν_t be the state of the directing process of the *MAP*, $\nu_t = \overline{0, W}$.

Let \mathcal{R} be the state space of the process $\{\eta_t^{(1)}, \dots, \eta_t^{(N)}\}$ defining the phases of service in all servers of the system:

$$\mathcal{R} = \{(r^{(1)}, \dots, r^{(N)}) : r^{(n)} = \overline{0, M^{(n)}}, n = \overline{1, N}\}.$$

Consider the continuous time multi-dimensional process

$$\xi_t = \{i_t, \eta_t^{(1)}, \dots, \eta_t^{(N)}, \nu_t\}, \quad t \geq 0.$$

It is easy to see that this process is an irreducible continuous-time Markov chain.

Let us define the stationary probabilities of this Markov chain as the limits

$$\pi(i, r^{(1)}, \dots, r^{(N)}, \nu) = \lim_{t \rightarrow \infty} P\{i_t = i, (\eta_t^{(1)}, \dots, \eta_t^{(N)}) = (r^{(1)}, \dots, r^{(N)}) \in \mathcal{R}, \nu_t = \nu\},$$

$$i \geq 0, \nu = \overline{0, W}.$$

Sufficient condition for existence of these limits will be presented in Theorem 4.1 below.

Let us enumerate the states of the chain ζ_t , $t \geq 0$, in the lexicographic order and form the row-vector

$$\pi(i, r^{(1)}, \dots, r^{(N)}) = (\pi(i, r^{(1)}, \dots, r^{(N)}, 0), \dots, \pi(i, r^{(1)}, \dots, r^{(N)}, W))$$

of the stationary probabilities $\pi(i, r^{(1)}, \dots, r^{(N)}, \nu)$, and the row-vectors π_i , consisting of the vectors $\pi(i, r^{(1)}, \dots, r^{(N)})$, $i \geq 0$.

Note that the size of the vectors π_i is equal to $K = (W + 1)\hat{M}$ where $\hat{M} = \prod_{n=1}^N (M^{(n)} + 1)$.

Define also the infinite-dimensional probability vector $\pi = (\pi_0, \pi_1, \pi_2, \dots)$.

For the use in the sequel, introduce the following notation:

I is an identity matrix of appropriate dimension (when needed, the dimension is identified with a subscript);

$O_{n \times n'}$ denotes zero matrices with n rows and n' columns; \otimes and \oplus are the symbols of the Kronecker product and sum of matrices, see, e. g., [11];

$$J_n = \begin{pmatrix} O_{1 \times 1} & O_{1 \times M^{(n)}} \\ O_{M^{(n)} \times 1} & I_{M^{(n)}} \end{pmatrix}, n = \overline{1, N};$$

$$\bigotimes_{l=r}^m J_l = J_r \otimes J_{r+1} \otimes \dots \otimes J_m, r \leq m, m = \overline{1, N}; J = \bigotimes_{l=1}^N J_l;$$

$\mathbf{f}^{(n)}$ is column vector of size $(M^{(n)} + 1)$ having the first entry equal to 1 and other entries equal to 0;

$\delta_{i,j}$ is Kronecker delta. It is equal to 1 if $i = j$ and 0, otherwise;

Λ is the diagonal matrix with diagonal entries defined by the diagonal entries of the matrix D_0 ;

$$G_n = \begin{pmatrix} O_{1 \times 1} & O_{1 \times M^{(n)}} \\ S_0^{(n)} & S^{(n)} \end{pmatrix}, n = \overline{1, N};$$

$$G = \sum_{n=1}^N \left(\begin{matrix} I_{n-1} \otimes G_n \otimes I_{\prod_{l=n+1}^N (M^{(l)}+1)} \\ \prod_{l=1}^{n-1} (M^{(l)}+1) \end{matrix} \right);$$

$$\Gamma_n = \bigotimes_{l=1}^{n-1} J_l \otimes G_n \otimes \bigotimes_{l=n+1}^N J_l, n = \overline{1, N};$$

H_n is the diagonal matrix with the diagonal entries coinciding with the corresponding diagonal entries of the matrix G_n ;

$$H = \sum_{n=1}^N \left(\bigotimes_{l=1}^{n-1} J_l \otimes H_n \otimes \bigotimes_{l=n+1}^N J_l \right);$$

$C = -(H \oplus \Lambda)^{-1}$. The matrix $H \oplus \Lambda$ is nonsingular as the irreducible sub-generator;

$$B_n = \begin{pmatrix} O_{1 \times 1} & \beta^{(n)} \\ O_{M^{(n)} \times 1} & O_{M^{(n)} \times M^{(n)}} \end{pmatrix}, n = \overline{1, N};$$

$$\tilde{I}_\beta = \sum_{k=1}^N \left(\bigotimes_{l=1}^{k-1} J_l \otimes B_k \otimes I_{\prod_{l=k+1}^N (M^{(l)}+1)} \right);$$

$$\tilde{I} = (I_{\hat{M}} - J);$$

the product of numbers $\prod_{l=a}^b c_l$ or matrices $\prod_{l=a}^b C_l$ is supposed to be equal to 1 or I , correspondingly, if $b < a$. The same relates to the Kronecker products.

The following statements hold true:

Lemma 3.1. *If the vector π of stationary probabilities exists, then it satisfies the system of equilibrium equations*

$$\pi \mathbf{Q} = \mathbf{0}, \pi \mathbf{e} = 1$$

where $\mathbf{0}$ is the infinite-size row-vector consisting of zeroes and the matrix \mathbf{Q} , which is the infinitesimal generator of the chain $\zeta_t, t \geq 0$, has the following structure:

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{00} & \mathbf{Q}_{01} & O & O & \cdots \\ \mathbf{Q}_{10} & \mathbf{Q}_{11} & \mathbf{Q}_{12} & O & \cdots \\ O & \mathbf{Q}_{21} & \mathbf{Q}_{22} & \mathbf{Q}_{23} & \cdots \\ O & O & \mathbf{Q}_{32} & \mathbf{Q}_{33} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

where the blocks $\mathbf{Q}_{i,j}, i, j \geq 0, j = \{\max\{0, i-1\}, i, i+1\}$, of the matrix \mathbf{Q} have size K and are defined as follows:

$$\begin{aligned} \mathbf{Q}_{i,i+1} &= J \otimes D_1, \quad \mathbf{Q}_{i,i-1} = \alpha_i \tilde{I}_\beta \otimes I_{\bar{W}}, \\ \mathbf{Q}_{i,i} &= I_{\bar{M}} \otimes D_0 - \alpha_i \tilde{I} \otimes I_{\bar{W}} + G \otimes I_{\bar{W}} + \tilde{I}_\beta \otimes D_1. \end{aligned}$$

Proof. The presented form of the blocks $\mathbf{Q}_{i,j}$ is easy explained if the intuitive meaning of some denotations is taken into account. In particular:

the matrix J_n is used to distinguish the residing of the process $\eta_t^{(n)}$ in the absorbing and transient states during which service is not provided and provided, correspondingly;

the matrix J highlights the states of the vector process $\boldsymbol{\eta}_t = \{\eta_t^{(1)}, \eta_t^{(2)}, \dots, \eta_t^{(N)}\}$ at which all servers are busy;

the matrix \tilde{I} highlights the states of the vector process $\boldsymbol{\eta}_t$ at which not all servers are busy;

the matrix G_n describes transition rates of the process $\eta_t^{(n)}$ in its state space;

the matrix G contains the rates of all possible transitions of the process $\boldsymbol{\eta}_t$;

the matrix B_n is used to instal the initial state of the underlying process $\eta_t^{(n)}$ at the service beginning moment in the n th server;

the matrix \tilde{I}_β describes transition rates of the process $\boldsymbol{\eta}_t$ at the moment of service beginning at the available server having the minimal number. Here, the matrix B_k defines the installment of service

namely in the k th server, $k = \overline{1, N}$. This matrix is preceded by the Kronecker product $\bigotimes_{l=1}^{k-1} J_l$ that guarantees

that the previous $k-1$ servers are busy and service cannot start in these servers. Additionally this matrix is multiplied from the right in Kronecker manner by the Kronecker product $I \prod_{l=k+1}^N (M^{(l)}+1)$ that

shows that no transitions occur in underlying processes of customers service in the servers having the numbers $k+1, k+2, \dots, N$.

The increase in the number of customers in orbit occurs at the moment of a customer arrival to the system (with the rates defined by the entries of the matrix D_1) when all servers are busy. Therefore, we evidently obtain that $\mathbf{Q}_{i,i+1} = J \otimes D_1$.

The decrease in the number of customers in orbit occurs at the moment when one of the customers staying in orbit retries to enter the service (with the rate α_i if i customers stay in the orbit) when there are available servers and the server with the minimal number is occupied. The transition rates of the process $\boldsymbol{\eta}_t$ in this scenario are defined by the matrix \tilde{I}_β . Any transition in the underlying process of arrivals is not possible. Therefore, we evidently obtain that $\mathbf{Q}_{i,i-1} = \alpha_i \tilde{I}_\beta \otimes I_{\bar{W}}$.

The diagonal entries of the diagonal blocks $\mathbf{Q}_{i,i}$ are negative. Their modules define the exit rate of the Markov chain $\zeta_t, t \geq 0$, from the corresponding states. The non-diagonal entries of the diagonal blocks $\mathbf{Q}_{i,i}$ are non-negative. They define the rates of the Markov chain ζ_t transitions that maintain the value of the number i of customers in the orbit. There exist four scenarios of such exits or transitions.

One scenario corresponds to the exit or transition of the underlying process of arrivals the rates of which are defined by the matrix D_0 . No transition of the process $\boldsymbol{\eta}_t$ are allowed in this scenario. This scenario explains the first summand $I_{\bar{M}} \otimes D_0$ in the expression for $\mathbf{Q}_{i,i}$.

The second scenario of the exit from the current state due to successful retrial of a customer from the orbit. This scenario explains the second summand $-\alpha_i \tilde{I} \otimes I_{\bar{W}}$ in the expression for $\mathbf{Q}_{i,i}$.

The third scenario corresponds to the exit or transition of the underlying process η_t of service, the rates of which are defined by the matrix G . No transition of the process ν_t are allowed in this scenario. This scenario explains the third summand in the form $G \otimes I_{\bar{W}}$.

The last scenario of transition of the Markov chain ζ_t corresponds to the new customer arrival, the rates of which are defined by the matrix D_1 and an immediate admission of this customer for service. Transitions probabilities of the process η_t at this arrival moment are defined by the matrix \tilde{I}_β . The proof of the formula for the block $\mathbf{Q}_{i,i}$ and of Lemma 3.1 is finished. \square

Lemma 3.2. *Markov chain ξ_t belongs to the class of asymptotically quasi-Toeplitz Markov chains, see [12].*

Proof. According to the definition of the asymptotically quasi-Toeplitz Markov chains given in [12], we have to prove the existence of the limits

$$Y_0 = \lim_{i \rightarrow \infty} R_i^{-1} \mathbf{Q}_{i,i-1}, Y_2 = \lim_{i \rightarrow \infty} R_i^{-1} \mathbf{Q}_{i,i+1}, Y_1 = \lim_{i \rightarrow \infty} R_i^{-1} \mathbf{Q}_{i,i} + I$$

where R_i is a diagonal matrix with diagonal entries defined as the moduli of the corresponding diagonal entries of the matrix $\mathbf{Q}_{i,i}$, $i \geq 0$. It can be easily verified that $R_i = \alpha_i \bar{I} - H \oplus \Lambda$.

Then, by direct calculations with account of the imposed above assumption that $\lim_{i \rightarrow \infty} \alpha_i = \infty$, it can be verified that

$$Y_0 = \tilde{I}_\beta \otimes I_{\bar{W}}, Y_2 = C(J \otimes D_1), Y_1 = C\left(\sum_{k=1}^N \Gamma_k \oplus D_0\right) + I.$$

Lemma 3.2 is proven. \square

4. Ergodicity condition

Theorem 4.1. (i) *The Markov chain ζ_t is ergodic if the inequality*

$$\lambda < \sum_{k=1}^N \mu_k \quad (1)$$

is fulfilled.

(ii) *The Markov chain ζ_t is non-ergodic if inequality (1) has an opposite sign.*

Proof. (i) As follows from [12], the sufficient condition for ergodicity of the AQTMC ξ_n , $n \geq 1$, is the fulfillment of the inequality

$$\mathbf{x}Y_2\mathbf{e} < \mathbf{x}Y_0\mathbf{e}, \quad (2)$$

where \mathbf{x} is the unique solution of the system

$$\mathbf{x}(Y_0 + Y_1 + Y_2) = \mathbf{x}, \quad \mathbf{x}\mathbf{e} = 1. \quad (3)$$

Calculating the vector \mathbf{x} from system (3) and substituting the expression obtained into inequality (2) after some algebra we get inequality (1).

Statement (ii) of the theorem follows from (1) and the results of [12]. \square

Remark 4.2. *Condition for ergodicity is easy tractable: the average arrival rate λ is less than the sum of the mean service rates in all servers of the system.*

The numerically stable algorithm for computation of vectors π_i , $i \geq 0$, can be found in [12].

5. Performance measures

As soon as the vectors π_i , $i \geq 0$, have been calculated, we are able to find various performance measures of the system.

Let us introduce the following denotations:

\mathcal{R}_k , $k = \overline{1, N}$, is the set of the states $\{r^{(1)}, \dots, r^{(N)}\} \in \mathcal{R}$ of the process $\{\eta_t^{(1)}, \dots, \eta_t^{(N)}\}$ such that $r^{(l)} > 0$ for $l = \overline{1, k-1}$, $r^{(k)} = 0$.

For the fixed set $(r^{(1)}, \dots, r^{(N)})$, $(r^{(1)}, \dots, r^{(N)}) \in \mathcal{R}$, the value $l(r^{(1)}, \dots, r^{(N)})$ defines the number of states having nonzero value of the components $r^{(n)}$, $n = \overline{1, N}$:

$$l(r^{(1)}, \dots, r^{(N)}) = \sum_{n=1}^N (1 - \delta_{r^{(n)}, 0}).$$

The average number L_{orbit} of customers in the orbit is computed by

$$L_{\text{orbit}} = \sum_{i=1}^{\infty} i \pi_i e.$$

The probability $P_{\text{empty-orbit}}$ that the orbit is empty at an arbitrary moment is computed by

$$P_{\text{empty-orbit}} = \pi_0 e.$$

The average number N_{busy} of busy servers at an arbitrary moment is computed by

$$N_{\text{busy}} = \sum_{i=0}^{\infty} \sum_{(r^{(1)}, \dots, r^{(N)}) \in \mathcal{R}} l(r^{(1)}, \dots, r^{(N)}) \pi(i, r^{(1)}, \dots, r^{(N)}) e.$$

The probability $P_0^{(n)}$ that the n th server, $n = \overline{1, N}$, is idle at an arbitrary moment is computed by

$$P_0^{(n)} = \sum_{i=0}^{\infty} \pi_i \left(\bigotimes_{r=1}^{n-1} e_{M^{(r)+1}} \otimes \mathbf{f}^{(n)} \otimes \bigotimes_{r=n+1}^N e_{M^{(r)+1}} \otimes e_{\bar{W}} \right).$$

The row vector defining the stationary probability distribution \mathbf{p}_n of the status of the n th server at an arbitrary moment is given by formula

$$\mathbf{p}_n = \sum_{i=0}^{\infty} \pi_i \left(\bigotimes_{r=1}^{n-1} e_{M^{(r)+1}} \otimes I_{M^{(n)+1}} \otimes \bigotimes_{r=n+1}^N e_{M^{(r)+1}} \otimes e_{\bar{W}} \right), \quad n = \overline{1, N}.$$

The output rate φ_n from the n th server is defined by

$$\varphi_n = \mathbf{p}_n \begin{pmatrix} 0 \\ S_0^{(n)} \end{pmatrix}, \quad n = \overline{1, N}.$$

Relation $\lambda = \sum_{n=1}^N \varphi_n$ can be used for control of accuracy of computation of the stationary distribution of the system states.

The probability $P_0^{(\text{serv})}$ that all servers are idle at an arbitrary moment is computed by

$$P_0^{(\text{serv})} = \sum_{i=0}^{\infty} \pi_i \left(\bigotimes_{n=1}^N \mathbf{f}^{(n)} \otimes e_{\bar{W}} \right).$$

The probability P_{imm} that an arbitrary customer will succeed to enter the service immediately upon arrival is computed by

$$P_{\text{imm}} = \frac{1}{\lambda} \sum_{i=0}^{\infty} \pi_i (\tilde{I}_{\beta} \otimes D_1) e.$$

The share of customers, which start service immediately upon arrival by the k th server, is computed by

$$Z_k = \frac{1}{\lambda} \sum_{i=0}^{\infty} \sum_{(r^{(1)}, \dots, r^{(N)}) \in \mathcal{R}_k} (\pi(i, r^{(1)}, \dots, r^{(N)}) \otimes D_1) e, \quad k = \overline{1, N}.$$

6. Numerical results

To illustrate the feasibility and outcome of the presented algorithms as well to show the effect of correlation in arrival process, we briefly consider the following example.

Let initially the *MAP*-input be characterized by the matrices

$$D_0 = \begin{pmatrix} -1.35164 & 0 \\ 0 & -0.04387 \end{pmatrix},$$

$$D_1 = \begin{pmatrix} 1.34265 & 0.00899 \\ 0.02443 & 0.01944 \end{pmatrix}.$$

This arrival process has the coefficient of correlation of two successive intervals between arrivals $c_{\text{cor}} = 0.2$, and the squared coefficient of variation of the intervals between customer arrivals $c_{\text{var}} = 13.4$. In the presented experiment, we will vary the average rate of the *MAP* λ what is done by multiplying the matrices D_0 and D_1 by the same scalar.

In parallel, we present the results of computation for the model where the arrival flow is defined as the stationary Poisson process with the same intensity. Let us assume that the total number N of servers be equal to 3 and $M^{(1)} = 2, M^{(2)} = 2, M^{(3)} = 3$. The retrial rates are defined by $\alpha_0 = 0, \alpha_i = i\alpha, \alpha = 1, i > 0$.

We will denote the *PH*-distributions of service times on three devices as $PH_1^{(\text{serv})}, PH_2^{(\text{serv})}, PH_3^{(\text{serv})}$.

$PH_1^{(\text{serv})}$ – the 2nd order hyperexponential distribution with $c_{\text{var}}^2 = 4.55$ – is characterized by the following vector and matrix:

$$\beta^{(1)} = (0.1, 0.9), S^{(1)} = \begin{pmatrix} -2 & 0 \\ 0 & -18 \end{pmatrix}.$$

$PH_2^{(\text{serv})}$ – the 2nd order hyperexponential distribution with $c_{\text{var}}^2 = 4.54$ – is characterized by the following vector and matrix:

$$\beta^{(2)} = (0.1, 0.9), S^{(2)} = \begin{pmatrix} -1.5 & 0 \\ 0 & -13.5 \end{pmatrix}.$$

$PH_3^{(\text{serv})}$ – the 3rd order hyperexponential distribution with $c_{\text{var}}^2 = 4.28$ – is characterized by the following vector and matrix:

$$\beta^{(3)} = \left(\frac{1}{15}, \frac{2}{15}, \frac{12}{15}\right), S^{(3)} = \begin{pmatrix} -0.2 & 0 & 0 \\ 0 & -0.4 & 0 \\ 0 & 0 & -2.4 \end{pmatrix}.$$

We can calculate the service rates at the corresponding servers as follows:

$$\mu_1 = 10, \mu_2 = 7.5, \mu_3 = 1$$

Let us assume that the total number N of servers be equal to 3. When we consider the service time to be exponential, assuming service rates at the corresponding servers be $\mu_1 = 10, \mu_2 = 7.5$ and $\mu_3 = 1$, correspondingly. Fig. 1 shows the behavior of the value L_{orbit} depending on the input rate λ . Fig. 2 shows the behavior of the value N_{busy} depending on the input rate λ . Fig. 3 shows the behavior of the value P_{imm} depending on the input rate λ . Fig. 4 shows the behavior of the value $P_0^{(n)}$ depending on the input rate λ under different numbers of servers. Fig. 5–7 show the behavior of the value $P_0^{(n)}$ depending on the input rate λ when $n = 1, 2, 3$, respectively, with different models.

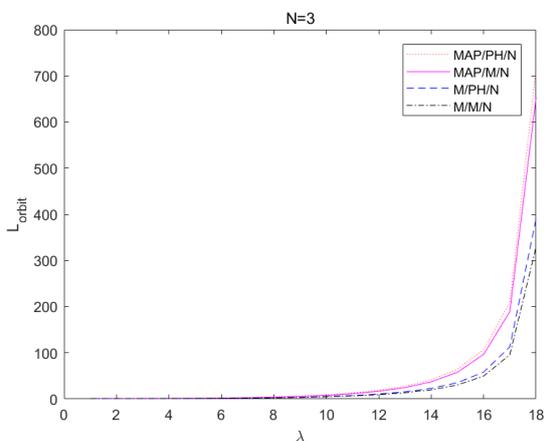


Fig. 1. Dependence of the number L_{orbit} on the input rate λ when $N = 3$ with different models

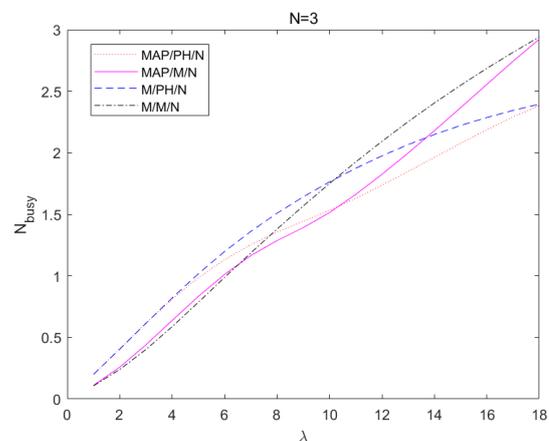


Fig. 2. Dependence of the number N_{busy} on the input rate λ when $N = 3$ with different models

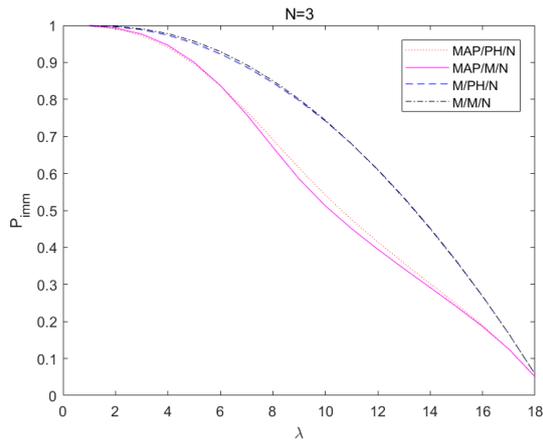


Fig. 3. Dependence of the number P_{imm} on the input rate λ when $N = 3$ with different models

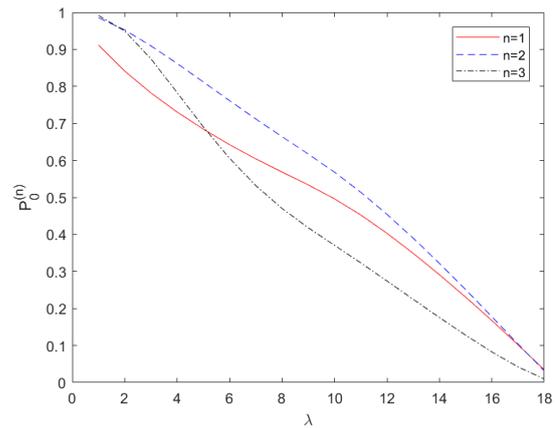


Fig. 4. Dependence of the number $P_0^{(n)}$ on the input rate λ when $N = 3$

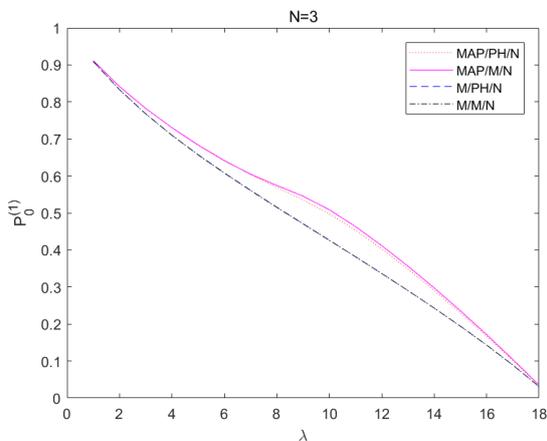


Fig. 5. Dependence of the number $P_0^{(1)}$ on the input rate λ when $N = 3$ with different models

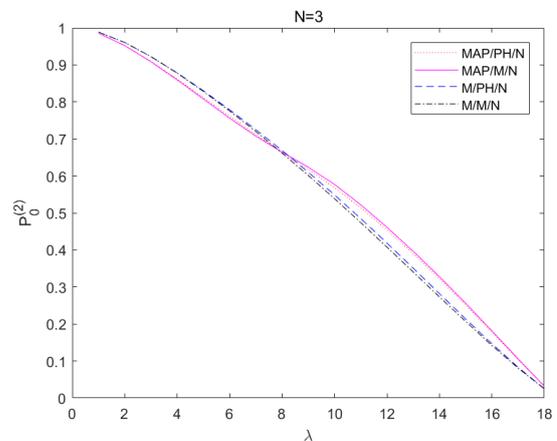


Fig. 6. Dependence of the number $P_0^{(2)}$ on the input rate λ when $N = 3$ with different models

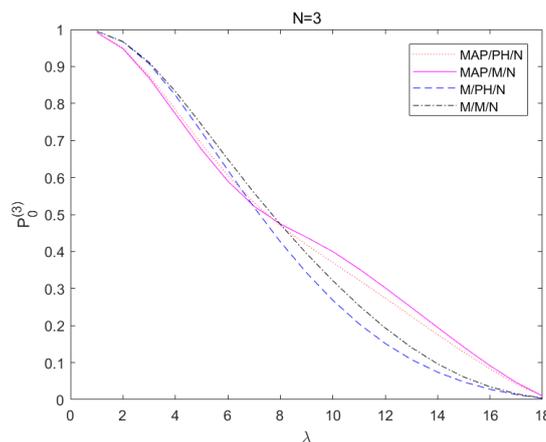


Fig. 7. Dependence of the number $P_0^{(3)}$ on the input rate λ when $N = 3$ with different models

7. Conclusion

In this paper, the algorithmic analysis of the $MAP/PH/N$ retrial queue with heterogeneous servers is presented. The obtained results are numerically illustrated in brief.

This research has received support by the Belarusian Republican Foundation for Fundamental Research (grant F25UZB-016) and the Ministry of Higher Education, Science and Innovations of the Republic of Uzbekistan (grant FL-8824063218).

References

1. Falin G. I., Templeton J. G. *Retrial queues*. Routledge, New York, NY, USA, 2023.
2. Lucantoni D. New results on the single server queue with a batch Markovian arrival process. *Communication in Statistics-Stochastic Models*, 1991, vol. 7, pp. 1–46.
3. Chakravarthy S. R. The batch Markovian arrival process: A review and future work. *Adv. Probab. Theory Stoch. Process*, 2001, vol. 1, pp. 21–49.
4. Chakravarthy S. R. *Introduction to Matrix-Analytic Methods in Queues 1: Analytical and Simulation Approach – Basics*. ISTE Ltd, London and John Wiley and Sons, New York, 2022.
5. Dudin A. N., Klimenok V. I., Vishnevsky V. M. *The theory of queueing systems with correlated flows*. Springer Nature, 2020.
6. Neuts M. *Matrix-geometric solutions in stochastic models*. North Chelmsford, Courier Corporation, 1994.
7. He Q. M., Li H., Zhao Y. Q. Ergodicity of the $BMAP/PH/s/s + K$ retrial queue with PH -retrial times. *Queueing Systems*, 2000, vol. 35, pp. 323–347.
8. Breuer L., Dudin A., Klimenok V. A retrial $BMAP/PH/N$ system. *Queueing Systems*, 2002, vol. 40, no. 4, pp. 433–457.
9. Liu M. Analysis of a Queue System with Repeated Calls, Heterogeneous Devices, and a Markov Arrival Process. *Informatics*, 2020, vol. 17, no. 1, pp. 48–57.
10. Liu M., Dudin A. Analysis of Retrial Queue with Heterogeneous Servers and Markovian Arrival Process. *Applied Probability and Stochastic Processes*, Springer, 2020, pp. 29–49.
11. Graham A. Kronecker products and matrix calculus with applications. *Courier Dover Publications*, Mineola, 2018.
12. Klimenok V. I., Dudin A. N. Multi-dimensional asymptotically quasi-Toeplitz Markov chains and their application in queueing theory. *Queueing Systems*, 2006, vol. 54, pp. 245–259.

ПРАВИЛА ДЛЯ АВТОРОВ

1. Для публикации в журнале принимаются ранее не опубликованные в других изданиях научные статьи. Обзорные статьи публикуются по решению редколлегии.

2. Объем статьи не более 20 журнальных страниц (с учетом таблиц и рисунков, а также списка литературы на 2 языках), объем краткого сообщения – до 5 страниц.

3. Статьи в Журнале публикуются на русском или английском языке.

4. Статья должна быть подготовлена в системе \LaTeX по образцу, находящемуся на сайте Журнала. Не допускается использование в \TeX -файлах «нестандартных» \TeX -команд (т. е. команд, не входящих в стандартную поставку \LaTeX), а также переопределение стандартных команд. При подаче статьи автору необходимо указывать рубрику Журнала, к которому относится статья.

5. Текст статьи начинается с индекса УДК, затем следуют название статьи, инициалы и фамилии авторов, а также краткая аннотация (не более 15 строк) и ключевые слова (5–10 слов). Аннотация не должна содержать ссылок на формулы и литературу статьи. Сведения об авторе (авторах), название статьи, аннотация и ключевые слова пишутся на русском и английском языках, остальные элементы оформляются на русском (английском) языке.

6. Ссылки на литературу в тексте нумеруются в порядке их упоминания и даются в квадратных скобках. Использование ссылок на неопубликованные работы не допускается. Необходимо приводить два списка ссылок на использованную в статье литературу – «Литература» и «References».

7. Если присутствует информация о финансировании (поддержке грантами проектов и т. п.), ее следует размещать в конце статьи.

GUIDELINES FOR AUTHORS

1. Scientific articles not previously published in other publications are accepted for publication in the journal. Review articles are published by decision of the editorial board.

2. The volume of the article is no more than 20 journal pages, taking into account tables and figures, the volume of a short message is up to 5 pages.

3. Articles in the Journal are published in Russian or English.

4. The article must be prepared in the \LaTeX system according to the model located on the Journal's website. It is not allowed to use "non-standard" in \TeX files \TeX commands (i. e. commands not included in the standard \LaTeX distribution), and also override standard commands. When submitting an article, the author must indicate the category of the Journal to which the article belongs.

5. The text of the paper must begin with the UDC, the title, and the name(s) of the author(s) preceded by initials followed by a short abstract (no more than 15 lines) and keywords (5–10 words). The abstract must not contain numbered references to the formulas and bibliography items. Information about the author (authors), title of the article, abstract and keywords are written in Russian and English, other elements are written in Russian (English).

6. The references in the text are numbered in order of their appearance and in square brackets. The use of references to unpublished works is not permitted. It is necessary to provide two lists of references to the literature used in the article – "References" and "Literature". An example of bibliographic descriptions is given in the model located on the Journal's website.

7. If there is information about financial support for the research (grant support for projects, etc.), it should be placed at the end of the article.

ТРУДЫ ИНСТИТУТА МАТЕМАТИКИ НАН БЕЛАРУСИ
2025. ТОМ 33, № 1

Редактор Т. П. Петрович
Компьютерная верстка И. В. Блинец

Подписано в печать 23.06.2025. Формат 60×84 ¹/₈. Усл. печ. л. 14,18. Уч.-изд. л. 8,27.
Государственное научное учреждение «Институт математики Национальной академии наук
Беларуси».
220072, г. Минск, ул. Сурганова, д. 11.
Свидетельство о государственной регистрации издателя, изготовителя, распространителя
печатных изданий № 1 / 257 от 2 апреля 2014 г.

Отпечатано в Республиканском унитарном предприятии «Издательский дом «Беларуская
навука». Свидетельство о государственной регистрации издателя, изготовителя,
распространителя печатных изданий № 1 / 18 от 02.08.2013. ЛП № 02330 / 455 от 30.12.2013.
Ул. Ф. Скорины, 40. 220084, г. Минск.
Тираж 30 экз. Заказ