ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)

ВЕСЦІ
НАЦЫЯНАЛЬНАЙ АКАДЭМІІ НАВУК БЕЛАРУСІ

СЕРЫЯ ФІЗІКА-МАТЭМАТЫЧНЫХ НАВУК. 2019. Т. 55, № 4

ИЗВЕСТИЯ
НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК БЕЛАРУСИ

СЕРИЯ ФИЗИКО-МАТЕМАТИЧЕСКИХ НАУК. 2019. Т. 55, № 4

Журнал основан в 1965 г. как «Весці Акадэміі навук БССР. Серыя фізіка-матэматычных навук»,
с 1992 г. – «Весці Акадэміі навук Беларусі. Серыя фізіка-матэматычных навук»,
с 1998 г. – современное название

Выходит четыре раза в год

Учредитель – Национальная академия наук Беларуси

Журнал зарегистрирован в Министерстве информации Республики Беларусь,
свидетельство о регистрации № 392 от 18.05.2009

Издается при поддержке Белорусского физического общества

*Входит в Перечень научных изданий Республики Беларусь
для опубликования результатов диссертационных исследований,
включен в базу данных Российского индекса научного цитирования* (*РИНЦ*)

Адрес редакции:
ул. Академическая, 1, к. 118, 220072, г. Минск, Республика Беларусь.
Тел.: + 375 17 284-02-45; e-mail: fmvesti@mail.ru
vestifm.belnauka.by

ИЗВЕСТИЯ НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК БЕЛАРУСИ.
Серия физико-математических наук. 2019. Т. 55, № 4.
Выходит на русском, белорусском и английском языках

Редактор *Т. Е. Янчук* Компьютерная верстка *И. В. Счеснюк*

Подписано в печать 23.12.2019. Выход в свет 27.12.2019. Формат 60×84 $^{1}/_{8}$. Бумага офсетная.
Печать цифровая. Усл. печ. л. 14,88. Уч.-изд. л. 16,4. Тираж 62 экз. Заказ 321.
Цена: индивидуальная подписка – 11,81 руб., ведомственная подписка – 28,27 руб.

Издатель и полиграфическое исполнение:
Республиканское унитарное предприятие «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, 220141, г. Минск, Республика Беларусь

ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)

PROCEEDINGS
OF THE NATIONAL ACADEMY OF SCIENCES OF BELARUS

PHYSICS AND MATHEMATICS SERIES, 2019, vol. 55, no. 4

The Journal was founded in 1956 under the titles "Proceedings of the Academy of Sciences of BSSR. Physics and Mathematics Series", since 1992 – "Proceedings of the Academy of Sciences of Belarus. Physics and Mathematics Series", since 1998 – it comes under its actual title

Periodicity is 4 issues per annum

Founder is the National Academy of Sciences of Belarus

The journal is registered on May 18, 2009 by the Ministry of Information of the Republic of Belarus in the State Registry of Mass Media, reg. no. 392

It is published with support of the Belarusian Physical Society

The Journal is included in The List of Journals for Publication of the Results of Dissertation Researchin the Republic of Belarus and in the database of Russian Science Citation Index (*RSCI*)

Address of the Editorial Office:
1, Akademicheskaya Str., room 118, 220072, Minsk, Republic of Belarus.
Tel.: + 375 17 284-02-45; e-mail: fmvesti@mail.ru
vestifm.belnauka.by

PROCEEDING OF THE NATIONAL ACADEMY OF SCIENCES OF BELARUS.
Physics and Mathematics series, 2019, vol. 55, no. 4.
Printed in Russian, Belarusian and English languages

Editor *T. E. Yanchuk*
Computer imposition *I. V. Schasniuk*

It is sent of the press 23.12.2019. Appearance 27.12.2019. Format $60\times84^1/_8$. Offset paper.
The press digital. Printed pages 14,88. Publisher`s signatures 16,4. Circulation 62 copies. Order 321.
Price: individual subscription – 11,81 byn., departmental subscription – 28,27 byn.

Publisher and printing execution:

Republican unitary enterprise “Publishing House “Belaruskaya Navuka”.
Certificate on the state registration of the publisher, manufacturer,distributor of printing editions
No. 1/18 dated August 2, 2013. License for the press no. 02330/455 dated December 30, 2013.
Address: F. Scorina Str., 40, 220141, Minsk, Republic of Belarus.

ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)

СОДЕРЖАНИЕ

МАТЕМАТИКА

ФИЗИКА

УЧЕНЫЕ БЕЛАРУСИ

ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)

CONTENTS

MATHEMATICS

PHYSICS

SCIENTISTS OF BELARUS

ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)

МАТЕМАТИКА
MATHEMATICS

УДК 517.5
https://doi.org/10.29235/1561-2430-2019-55-4-391-405

Поступила в редакцию 15.07.2019
Received 15.07.2019

Е. А. Ровба, В. Ю. Медведева

Гродненский государственный университет им. Я. Купалы, Гродно, Беларусь

О РАЦИОНАЛЬНОЙ ИНТЕРПОЛЯЦИИ ФУНКЦИИ $|x|^{\alpha}$ ПО РАСШИРЕННОЙ СИСТЕМЕ УЗЛОВ ЧЕБЫШЕВА – МАРКОВА

Аннотация. В работе исследуются приближения функции $|x|^{\alpha}$, $\alpha > 0$ интерполяционными рациональными функциями Лагранжа на отрезке [–1,1]. В качестве узлов интерполирования выбираются нули четных рациональных функций Чебышева – Маркова и точка $x = 0$. Получено интегральное представление остатка интерполирования и оценка сверху рассматриваемых равномерных приближений. На их основании подробно изучаются:

а) полиномиальный случай; здесь авторы приходят к известному асимптотическому равенству М. Н. Ганзбурга;

б) в случае фиксированного числа геометрически различных полюсов получена оценка сверху соответствующих равномерных приближений, улучшающая известный результат К. Н. Лунгу;

в) при приближении общими интерполяционными рациональными функциями Лагранжа найдена оценка равномерных приближений и показано, что на концах отрезка [–1,1] ее можно улучшить.

Полученные результаты могут быть применены в теоретических исследованиях и численных методах.

Ключевые слова: рациональная дробь Чебышева – Маркова, рациональная интерполяция, функция со степенной особенностью

Для цитирования. Ровба, Е. А. О рациональной интерполяции функции $|x|^{\alpha}$ по расширенной системе узлов Чебышева – Маркова / Е. А. Ровба, В. Ю. Медведева // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2019. – Т. 55, № 4. – С. 391–405. https://doi.org/10.29235/1561-2430-2019-55-4-391-405

Y. A. Rovba, V. Yu. Medvedeva

Yanka Kupala State University of Grodno, Grodno, Belarus

RATIONAL INTERPOLATION OF THE FUNCTION $|x|^{\alpha}$ BY AN EXTENDED SYSTEM OF CHEBYSHEV – MARKOV NODES

Abstract. In this paper, we study the approximations of a function $|x|^{\alpha}$, $\alpha > 0$ by interpolation rational Lagrange functions on a segment [–1,1]. The zeros of the even Chebyshev – Markov rational functions and a point $x = 0$ are chosen as the interpolation nodes. An integral representation of an interpolation remainder and an upper bound for the considered uniform approximations are obtained. Based on them, a detailed study is made:

a) the polynomial case. Here, the authors come to the famous asymptotic equality of M. N. Hanzburg;

b) at a fixed number of geometrically different poles, the upper estimate is obtained for the corresponding uniform approximations, which improves the well-known result of K. N. Lungu;

c) when approximating by general Lagrange rational interpolation functions, the estimate of uniform approximations is found and it is shown that at the ends of the segment [–1,1] it can be improved.

The results can be applied in theoretical research and numerical methods.

Keywords: rational Chebyshev – Markov fraction, rational interpolation, function with power singularity

For citation. Rovba Y. A., Medvedeva V. Ju. Rational interpolation of the function $|x|^{\alpha}$ by an extended system of Chebyshev – Markov nodes. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2019, vol. 55, no. 4, pp. 391–405 (in Russian). https://doi.org/10.29235/1561-2430-2019-55-4-391-405

Приближение простейших функций, имеющих алгебраические особенности, имеет богатое историческое наследие. Первые исследования были посвящены приближению функции $|x|$ на отрезке $[-1,1]$ алгебраическими полиномами (см., напр., [1–4]). В 1913 г. С. Н. Бернштейн в [5] доказал, что существует число β, $\beta = 0{,}2801\ldots$ такое, что

$$\lim_{n\to\infty} nE_n = \beta,$$

где E_n – наилучшее равномерное приближение функции $|x|$ алгебраическими полиномами степени не выше n на отрезке $[-1,1]$. Позже в работе [6] он провел глубокий анализ наилучших равномерных полиномиальных приближений функции $|x|^\alpha$, $\alpha > 0$, на отрезке $[-1,1]$. Эти исследования были продолжены С. М. Никольским [7], Р. А. Райцином [8] и др.

Значительное число работ посвящено различным методам приближений простейших функций с алгебраической особенностью (см., напр., [9–11]). В этом ряду следует выделить исследования М. Н. Ганзбурга [12] и М. Реверса [13], посвященные интерполяции функции $|x|^\alpha$, $\alpha > 0$, на отрезке $[-1,1]$ по различным системам узлов Чебышева. Именно это направление привлекло внимание авторов настоящей статьи с точки зрения рациональной интерполяции. Как известно, для рациональной интерполяции представляют интерес узлы, являющиеся нулями соответствующей рациональной дроби Чебышева – Маркова.

Наилучшим равномерным рациональным приближениям функции $|x|^\alpha$ на отрезке $[-1,1]$ или функции x^α на отрезке $[-1,0]$ посвящено достаточно много работ (см., напр., [14–18]).

Задачу о приближении непрерывных функций с характерными особенностями рациональными функциями с фиксированным числом полюсов рассмотрел К. Н. Лунгу [19]. В [20] им были рассмотрены такие приближения функции x^α, $\alpha > 0$, на отрезке $[1,0]$. Позже это направление в рациональной аппроксимации получило развитие в [21–22] и др.[1]

Настоящая работа продолжает вышеназванные исследования. Рассматривается интерполирование функции $f(x) = |x|^\alpha$, $\alpha > 0$, по расширенной системе узлов Чебышева – Маркова на отрезке $[-1,1]$. Получено интегральное представление остатка интерполирования. Исходя из этого результата, найдены оценки сверху равномерных приближений рациональными интерполяционными функциями с такими узлами для различных случаев расположения полюсов аппроксимирующей рациональной функции.

Пусть $m_{2n}(x)$ – косинус-дробь Чебышева – Маркова

$$m_{2n}(x) = \cos\mu_{2n}(x), \tag{1}$$

где

$$\mu_{2n}(x) = \sum_{k=1}^{2n} \arccos\frac{x + a_k}{1 + a_k x},$$

$a_1, a_2, \ldots, a_{2n}$ – чисто мнимые числа либо нули, $k = 1,\ldots,2n$, причем

$$\text{а) } a_{n+k} = -a_k,\quad \operatorname{Im} a_k \ge 0,\quad k = 1,\ldots,n;$$

$$\text{б) } a_1 = a_2 = \ldots = a_r = 0,\quad r = \left[\frac{\alpha}{2}\right] + 1,\quad n > r. \tag{2}$$

Нетрудно проверить, что $m_{2n}(x)$ является четной рациональной функцией. В этом случае функция $m_{2n}(x)$ имеет $2n$ простых симметричных нулей на интервале $(-1,1)$:

$$-1 < x_{2n} < x_{2n-1} < \ldots < x_{n+1} < 0 < x_n < \ldots < x_2 < x_1 < 1; \tag{3}$$

$$x_{2n-k+1} = -x_k,\quad k = 1, 2, \ldots, n;$$

$$\mu_{2n}(x_k) = (2k-1)\frac{\pi}{2},\quad k = 1, 2, \ldots, 2n.$$

[1] Старовойтов, А. П. Аппроксимация рациональными функциями с заданным числом полюсов / А. П. Старовойтов // Белорусский гос. ун-т. – Минск, 1984. – 23 с. – Деп. в Бел. НИИНТИ 1984. – № 689-Бел Д84.

Для функции $f(x) = |x|^\alpha$, $\alpha > 0$, построим интерполяционную рациональную функцию Лагранжа с узлами в точках $x_1, x_2, \ldots, x_{2n}$ и в точке $x_0 = 0$:

$$L_{2n}(x,f) = \sum_{k=0}^{2n} f(x_k) l_k(x), \tag{4}$$

где

$$l_k(x) = \frac{x m_{2n}(x)}{(x - x_k)\left(x m_{2n}(x)\right)'_{x=x_k}}, \quad k = 0, 1, \ldots, 2n.$$

Для оценки остатка интерполирования введем следующие величины:

$$\varepsilon_{2n}(x,a) = |x|^\alpha - L_{2n}(x,f), \ \ x \in [-1,1];$$
$$\varepsilon_{2n}(a) = \max_{-1 \le x \le 1} \left|\varepsilon_{2n}(x,a)\right|. \tag{5}$$

Т е о р е м а 1. *Для приближений функции* $f(x) = |x|^\alpha$, $\alpha > 0$, *на отрезке* $[-1,1]$ *интерполяционными рациональными функциями Лагранжа* (4) *при условиях* (2) *справедливы соотношения:*

$$1) \ \varepsilon_{2n}(x,a) = \frac{4}{\pi} x^2 \sin\frac{\pi\alpha}{2} \cdot m_{2n}(x) \int_0^1 \frac{u^{\alpha-1}}{(1-u^2)^{\frac{\alpha}{2}}} \frac{1}{x^2(1-u^2)+u^2} \frac{du}{\psi_n(u) + \psi_n^{-1}(u)}; \tag{6}$$

$$2) \ \varepsilon_{2n}(a) \le \frac{4}{\pi}\left|\sin\frac{\pi\alpha}{2}\right| \left|\int_0^1 \frac{u^{\alpha-1}}{(1-u^2)^{\frac{\alpha}{2}}} \frac{du}{\left|\psi_n(u)\right| + \left|\psi_n^{-1}(u)\right|}\right|, \tag{7}$$

где

$$x \in [-1,1], \quad \psi_n(u) = \prod_{k=1}^{n} \frac{u - \beta_k}{u + \beta_k}, \quad \beta_k = \frac{1}{\sqrt{1 + |a_k|^2}}, \quad k = 1, 2, \ldots, n, \quad n > \frac{\alpha}{2}.$$

Причем, если полюсы функции $L_{2n}(x,f)$ *имеют четную кратность, то оценка* (7) *является точной, т. е. имеет место знак равенства.*

Д о к а з а т е л ь с т в о. Из соотношения (4) имеем

$$L_{2n}(x,f) = \sum_{k=1}^{n} x_k^{\alpha-1} \frac{x m_{2n}(x)}{(x-x_k) m'_{2n}(x_k)} - \sum_{k=n+1}^{2n} (-x_k)^{\alpha-1} \frac{x m_{2n}(x)}{(x-x_k) m'_{2n}(x_k)}. \tag{8}$$

Легко видеть, что справедливо тождество

$$1 \equiv \sum_{k=1}^{2n} \frac{m_{2n}(x)}{(x-x_k) m'_{2n}(x_k)}. \tag{9}$$

Полагаем, что $x \in (0,1]$. Умножим обе части тождества (9) на $f(x)$ и из полученного равенства вычтем почленно равенство (8). Получим

$$\varepsilon_{2n}(x,a) = f(x) - L_{2n}(x,f) = x m_{2n}(x) \cdot \left[\sum_{k=1}^{n} \frac{x^{\alpha-1} - x_k^{\alpha-1}}{(x-x_k) m'_{2n}(x_k)} + \sum_{k=n+1}^{2n} \frac{x^{\alpha-1} + (-x_k)^{\alpha-1}}{(x-x_k) m'_{2n}(x_k)}\right] =$$
$$= x m_{2n}(x) \cdot \left[S_1(x) + S_2(x)\right], \tag{10}$$

где

$$S_1(x) = \sum_{k=1}^{n} \frac{x^{\alpha-1} - x_k^{\alpha-1}}{(x-x_k) m'_{2n}(x_k)}, \quad S_2(x) = \sum_{k=n+1}^{2n} \frac{x^{\alpha-1} + (-x_k)^{\alpha-1}}{(x-x_k) m'_{2n}(x_k)}.$$

В плоскости с разрезом по промежутку $\{z : z = iy, -\infty < y < 0\}$ определим ветвь аналитической функции $f(z) = z^{\alpha}$ так, что $f(1) = 1$. Тогда нетрудно проверить, что

$$S_1(x) = \frac{1}{2\pi i}\int\limits_{\Gamma_\delta}\frac{x^{\alpha-1} - z^{\alpha-1}}{(x-z)m_{2n}(z)}dz, \quad S_2(x) = \frac{1}{2\pi i}\int\limits_{\Gamma'_\delta}\frac{x^{\alpha-1} + (-z)^{\alpha-1}}{(x-z)m_{2n}(z)}dz, \tag{11}$$

где контур $\Gamma_\delta = \overline{\Gamma}_\delta^{(1)} \cup \overline{\Gamma}_\delta^{(2)} \cup \overline{\Gamma}_\delta^{(3)}$,

$$\Gamma_\delta^{(1)} = \{z : z = iy, \delta < y < +\infty\},$$

$$\Gamma_\delta^{(2)} = \left\{z : |z| = \delta, -\frac{\pi}{2} \le \arg z \le \frac{\pi}{2}\right\},$$

$$\Gamma_\delta^{(3)} = \{z : z = iy, -\infty < y < -\delta\},$$

$\delta, 0 < \delta < x_n$ – достаточно малое положительное число;
и контур $\Gamma'_\delta = \Gamma_\delta^{(3)} \cup \overline{\Gamma}_\delta^{(4)} \cup \Gamma_\delta^{(1)}$,

$$\Gamma_\delta^{(4)} = \left\{z : |z| = \delta, \frac{\pi}{2} \le \arg z \le 3\frac{\pi}{2}\right\}.$$

Заметим, что так как $x \in (0, 1]$, $\alpha > 0$ и $|m_{2n}(0)| = 1$, то

$$\lim_{\delta\to 0}\int\limits_{\Gamma_\delta^{(2)}}\frac{x^{\alpha-1} - z^{\alpha-1}}{(x-z)m_{2n}(z)}dz = 0, \quad \lim_{\delta\to 0}\int\limits_{\Gamma_\delta^{(4)}}\frac{x^{\alpha-1} - (-z)^{\alpha-1}}{(x-z)m_{2n}(z)}dz = 0.$$

Тогда из равенств (11) следует

$$S_1(x) = -\frac{1}{2\pi i}\int\limits_{-i\infty}^{+i\infty}\frac{x^{\alpha-1} - z^{\alpha-1}}{(x-z)m_{2n}(z)}dz, \quad S_2(x) = \frac{1}{2\pi i}\int\limits_{-i\infty}^{+i\infty}\frac{x^{\alpha-1} + (-z)^{\alpha-1}}{(x-z)m_{2n}(z)}dz. \tag{12}$$

Преобразуем интеграл $S_1(x)$. С этой целью сделаем замену $z = it$ и разобьем полученный интеграл на два:

$$S_1(x) = -\frac{1}{2\pi}\int\limits_{-\infty}^{+\infty}\frac{x^{\alpha-1} - (it)^{\alpha-1}}{(x-it)m_{2n}(it)}dt = -\frac{1}{2\pi}\left[\int\limits_{-\infty}^{0}\frac{x^{\alpha-1} - (it)^{\alpha-1}}{(x-it)m_{2n}(it)}dt + \int\limits_{0}^{+\infty}\frac{x^{\alpha-1} - (it)^{\alpha-1}}{(x-it)m_{2n}(it)}dt\right].$$

Теперь в первом интеграле сделаем замену $t = -v$:

$$S_1(x) = -\frac{1}{2\pi}\left[\int\limits_{0}^{+\infty}\frac{x^{\alpha-1} - (-iv)^{\alpha-1}}{(x+iv)m_{2n}(-iv)}dv + \int\limits_{0}^{+\infty}\frac{x^{\alpha-1} - (it)^{\alpha-1}}{(x-it)m_{2n}(it)}dt\right].$$

Пользуясь четностью дроби Чебышева – Маркова $m_{2n}(x)$, найдем

$$S_1(x) = -\frac{1}{2\pi}\int\limits_{0}^{+\infty}\left[\frac{x^{\alpha-1} - (-iz)^{\alpha-1}}{x+iz} + \frac{x^{\alpha-1} - (iz)^{\alpha-1}}{x-iz}\right]\frac{dz}{m_{2n}(iz)}. \tag{13}$$

Аналогичные преобразования проведем с интегралом $S_2(x)$ (см. (12)) и получим, что

$$S_2(x) = \frac{1}{2\pi}\int\limits_{0}^{+\infty}\left[\frac{x^{\alpha-1} + (iz)^{\alpha-1}}{x+iz} + \frac{x^{\alpha-1} + (-iz)^{\alpha-1}}{x-iz}\right]\frac{dz}{m_{2n}(iz)}. \tag{14}$$

На основании полученных равенств (13) и (14) имеем

$$H_n(x) = S_1(x) + S_2(x) = \frac{1}{2\pi}\int\limits_{0}^{+\infty}\left[\frac{(-iz)^{\alpha-1} + (iz)^{\alpha-1}}{x+iz} + \frac{(iz)^{\alpha-1} + (-iz)^{\alpha-1}}{x-iz}\right]\frac{dz}{m_{2n}(iz)}. \tag{15}$$

После несложных преобразований придем к выражению

$$H_n(x)=\frac{x}{\pi}\int_0^{+\infty}\frac{z^{\alpha-1}\cdot\left((-i)^{\alpha-1}+i^{\alpha-1}\right)}{x^2+z^2}\frac{dz}{m_{2n}(iz)}=\frac{2x}{\pi}\sin\frac{\pi\alpha}{2}\cdot\int_0^{+\infty}\frac{z^{\alpha-1}}{x^2+z^2}\frac{dz}{m_{2n}(iz)}. \tag{16}$$

Тогда (см. равенства (10), (15) и (16)) имеем

$$\varepsilon_{2n}(x,a)=\frac{2x^2}{\pi}\sin\frac{\pi\alpha}{2}\cdot m_{2n}(x)\int_0^{+\infty}\frac{z^{\alpha-1}}{x^2+z^2}\frac{dz}{m_{2n}(iz)}. \tag{17}$$

Теперь займемся преобразованием интеграла, стоящего в правой части этого равенства:

$$J_n(x)=\int_0^{+\infty}\frac{z^{\alpha-1}}{x^2+z^2}\frac{dz}{m_{2n}(iz)}. \tag{18}$$

Сделаем замену

$$t=i\sqrt{\frac{1-iz}{1+iz}},\quad z=-i\frac{1+t^2}{1-t^2},\quad dz=\frac{-4ti}{(1-t^2)^2}dt. \tag{19}$$

Образом промежутка $(0, +\infty)$ является дуга единичной окружности

$$\Gamma=\left\{t: t=e^{i\phi},\ 0<\phi<\frac{\pi}{2}\right\}$$

с обходом по часовой стрелке.

Как известно [23, с. 48], косинус-дробь Чебышева – Маркова $m_{2n}(x)$, заданная на отрезке $[-1,1]$, связана с косинус-дробью Бернштейна $M_{2n}(y)$, заданной на $\mathbb{R}$, следующим соотношением:

$$m_{2n}(x)=M_{2n}\left(\sqrt{\frac{1-x}{1+x}}\right), \tag{20}$$

где

$$M_{2n}(y)=\frac{1}{2}\left(\chi_{2n}(y)+\chi_{2n}^{-1}(y)\right),\quad \chi_{2n}(y)=\prod_{k=1}^{2n}\frac{y-z_k}{y-\overline{z_k}}, \tag{21}$$

z_k – корни уравнения $z^2+\dfrac{1+a_k}{1-a_k}=0$, у которых $\operatorname{Im} z_k>0,\quad k=1,2,\ldots,2n.$

Следовательно, в нашем случае

$$z_k=i\frac{1+a_k}{\sqrt{1+|a_k|^2}},\quad k=1,2,\ldots,2n. \tag{22}$$

Подставляя соотношения (19) в интеграл (18) и учитывая (20), получим

$$J_n(x)=4i\int_\Gamma\left(-i\frac{1+t^2}{1-t^2}\right)^{\alpha-1}\frac{tdt}{\left((1+t^2)^2-x^2(1-t^2)^2\right)\mathrm{M}_{2n}(it)}.$$

Теперь сделаем еще одну замену:

$$t=u+i\sqrt{1-u^2},\quad u=\frac{1}{2}\left(t+\frac{1}{t}\right),\quad dt=-i\frac{u+i\sqrt{1-u^2}}{\sqrt{1-u^2}}du.$$

Рассматриваемое преобразование является обратным к функции Жуковского. Поэтому дуга Γ отображается в отрезок $[0,1]$. Нетрудно посчитать, что

$$\frac{1+t^2}{1-t^2}=\frac{iu}{\sqrt{1-u^2}}$$

и интеграл $J_n(x)$ имеет вид

$$J_n(x)=\int_0^1 \frac{u^{\alpha-1}}{(1-u^2)^{\frac{\alpha}{2}}}\frac{du}{\left(x^2(1-u^2)+u^2\right)\cdot M_{2n}\left(iu-\sqrt{1-u^2}\right)}. \tag{23}$$

Займемся теперь преобразованием функции $M_{2n}\left(iu-\sqrt{1-u^2}\right)$. Обращаясь к представлению (21) и учитывая (22), а также симметричный выбор параметров a_k, $k=1,2,\ldots,2n$, заметим, что в выражении для $M_{2n}(y)$ будут присутствовать множители следующего вида:

$$\chi^{(k)}(y)=\frac{y-i\frac{1+a_k}{\sqrt{1+|a_k|^2}}}{y+i\frac{1-a_k}{\sqrt{1+|a_k|^2}}}\cdot\frac{y-i\frac{1-a_k}{\sqrt{1+|a_k|^2}}}{y+i\frac{1+a_k}{\sqrt{1+|a_k|^2}}}.$$

Преобразовывая их, придем к выражению

$$\chi^{(k)}(y)=\left(y^2-1-i\frac{2y}{\sqrt{1+|a_k|^2}}\right)\Bigg/\left(y^2-1+i\frac{2y}{\sqrt{1+|a_k|^2}}\right).$$

Сделав замену $y=i\left(u+i\sqrt{1-u^2}\right)$, получим

$$\chi^{(k)}(y)=\frac{u-\beta_k}{u+\beta_k},\quad \beta_k=\frac{1}{\sqrt{1+|a_k|^2}},\quad k=1,2,\ldots,n.$$

Воспользовавшись этим представлением, найдем

$$M_{2n}\left(iu-\sqrt{1-u^2}\right)=\frac{1}{2}\left(\psi_n(u)+\psi_n^{-1}(u)\right),$$

где

$$\psi_n(u)=\prod_{k=1}^{n}\frac{u-\beta_k}{u+\beta_k}.$$

Подставив данное выражение в (23), будем иметь

$$J_n(x)=2\int_0^1 \frac{u^{\alpha-1}}{(1-u^2)^{\frac{\alpha}{2}}}\frac{1}{x^2(1-u^2)+u^2}\frac{du}{\psi_n(u)+\psi_n^{-1}(u)}.$$

Тогда из равенства (17) вытекает

$$\varepsilon_{2n}(x,a)=\frac{4}{\pi}x^2\sin\frac{\pi\alpha}{2}\cdot m_{2n}(x)\int_0^1 \frac{u^{\alpha-1}}{(1-u^2)^{\frac{\alpha}{2}}}\frac{1}{x^2(1-u^2)+u^2}\frac{du}{\psi_n(u)+\psi_n^{-1}(u)}.$$

Очевидно, что

$$\frac{x^2}{x^2(1-u^2)+u^2}\le 1,\quad x\in[-1,1],\quad u\in[0,1],$$

причем при любом $u \in [0, 1]$ максимум этой функции достигается в точке $x = 1$. Следовательно,

$$\varepsilon_{2n}(a) \le \frac{4}{\pi}\left|\sin\frac{\pi\alpha}{2}\right|\int_0^1 \frac{u^{\alpha-1}}{(1-u^2)^{\frac{\alpha}{2}}}\frac{du}{\left|\psi_n(u)\right| + \left|\psi_n^{-1}(u)\right|}, \quad n > \frac{\alpha}{2}+1.$$

Легко видеть, что в этом неравенстве имеет место знак равенства, если предположить, что функция $\psi n(u)$ неотрицательна на [0,1], т. е. числа $\beta_1, \beta_2, \ldots, \beta_n$ имею четную кратность. Таким образом, теорема 1 доказана.

Рассмотрим некоторые приложения оценки (7).

1. Полиномиальный случай. В этом случае все числа $a_k = 0$, $k = 1,2,\ldots,n$, и

$$\psi_n(u) = \left(\frac{u-1}{u+1}\right)^n.$$

Следовательно,

$$\varepsilon_{2n}(0) = \varepsilon_{2n,0} = \frac{4}{\pi}\left|\sin\frac{\pi\alpha}{2}\right|\int_0^1 \frac{u^{\alpha-1}}{(1-u^2)^{\frac{\alpha}{2}}}\frac{du}{\left(\frac{1-u}{1+u}\right)^n + \left(\frac{1-u}{1+u}\right)^{-n}}.$$

В интеграле справа сделаем замену $u = \dfrac{t}{2n}$:

$$\varepsilon_{2n,0} = \frac{4}{\pi}\left|\sin\frac{\pi\alpha}{2}\right|\frac{1}{(2n)^\alpha}\int_0^{2n} \frac{t^{\alpha-1}}{\left(1-(t/2n)^2\right)^{\frac{\alpha}{2}}}\frac{dt}{\left(\frac{1-t/2n}{1+t/2n}\right)^n + \left(\frac{1-t/2n}{1+t/2n}\right)^{-n}}.$$

Так как подынтегральная функция равномерно относительно $t \in [0, 2n]$ сходится при $n \to \infty$ к функции

$$\frac{t^{\alpha-1}}{e^t + e^{-t}},$$

то нетрудно показать, что по теореме 1 из [24, с. 695]

$$\lim_{n\to\infty}(2n)^\alpha \varepsilon_{2n,0} = \frac{4}{\pi}\left|\sin\frac{\pi\alpha}{2}\right|\int_0^{+\infty}\frac{t^{\alpha-1}}{e^t + e^{-t}}dt.$$

Этот результат получен в работе [12].

2. Случай заданного числа полюсов. Обозначим через A_{2n} множество точек $(a_1, a_2, \ldots, a_{2n}) \in R^{2n}$, удовлетворяющих условиям (2).

Пусть $n > r$, $r = \left[\dfrac{\alpha}{2}\right] + 1$, $n_1 = n - r$ и q – произвольное целое число, $0 \le q \le n_1$; $A_{2n,2q}$ есть множество точек $a = (a_1, a_2, \ldots, a_{2n}) \in A_{2n}$ и таких, что среди этих чисел $a_1, a_2, \ldots, a_n$ имеется не больше q различных отличных от нуля и кратность каждой точки не больше $\left[\dfrac{n_1}{q+1}\right]$. Полагаем

$$\varepsilon_{2n} = \inf_{a \in A_n} \varepsilon_{2n}(a),$$

$$\varepsilon_{2n,2q} = \inf_{a \in A_{2n,2q}} \varepsilon_{2n}(a). \tag{24}$$

Т е о р е м а 2. *При любых целых n и q,* $0 \le q < n$, $n > r$, *справедливо неравенство*

$$\varepsilon_{2n,\,2q} \le C(\alpha)\left|\sin\frac{\pi\alpha}{2}\right|\cdot \inf_{1<t<+\infty}\left(e^{-\frac{n}{2t}}+\left(\frac{q+1}{n}\right)^{\alpha}\left(1+\frac{t}{q+1}\right)^{-2q\alpha}\right), \tag{25}$$

и где $C(\alpha)$ *– некоторая положительная постоянная, зависящая только от* α[1].

Д о к а з а т е л ь с т в о. Пусть

$$m=\left[\frac{n_1}{q+1}\right],\quad n_1=n-r.$$

Определим числа β_k следующим образом:

$$\beta_k=\xi^{2k},\quad \xi\in(0,1),\quad k=0,1,\ldots,q.$$

Рассмотрим следующий интеграл из оценки (7):

$$I_n=\int_0^1\frac{u^{\alpha-1}}{(1-u^2)^{\frac{\alpha}{2}}}\frac{|\psi_n(u)|}{1+\psi_n^2(u)}du. \tag{26}$$

Легко показать, что в условиях теоремы 2 будем иметь

$$I_n\le\int_0^1 u^{\alpha-1}\prod_{k=0}^{q}\left|\frac{u-\beta_k}{u+\beta_k}\right|^m du.$$

Далее, разобьем этот интеграл на два:

$$I_n\le\int_0^{\beta_q} u^{\alpha-1}\prod_{k=0}^{q}\left|\frac{u-\beta_k}{u+\beta_k}\right|^m du+\int_{\beta_q}^{1} u^{\alpha-1}\prod_{k=0}^{q}\left|\frac{u-\beta_k}{u+\beta_k}\right|^m du=:I_n^{(1)}+I_n^{(2)}. \tag{27}$$

Оценим каждый из этих интегралов:

$$I_n^{(1)}\le\int_0^{\beta_q}u^{\alpha-1}\left(\frac{\beta_q-u}{\beta_q+u}\right)^m du\,.$$

В интеграле справа сделаем замену $u=\dfrac{\beta_q}{2m}t$:

$$I_n^{(1)}\le\left(\frac{\beta_q}{2m}\right)^{\alpha}\int_0^{2m}t^{\alpha-1}\left(\frac{1-t/2m}{1+t/2m}\right)^m dt.$$

Нетрудно проверить, что

$$\frac{1-u}{1+u}\le e^{-2u},\quad u\in(0,1].$$

Следовательно,

$$I_n^{(1)}<\Gamma(\alpha)\left(\frac{\beta_q}{2m}\right)^{\alpha}, \tag{28}$$

где $\Gamma(\alpha)$ – эйлеров интеграл второго рода.

[1] Здесь и далее через $C(\alpha)$, $C_1(\alpha)$, $C_2(\alpha)$,... будем обозначать положительные постоянные, зависящие только от параметра α, $\alpha > 0$.

Теперь перейдем к оценке интеграла $I_n^{(2)}$:

$$I_n^{(2)} < \frac{1}{\alpha} \cdot \max_{u \in [\beta_q, 1]} \prod_{k=0}^{q} \left| \frac{u - \beta_k}{u + \beta_k} \right|^m . \tag{29}$$

Предположим, что

$$u \in \left(\xi^{2i}, \xi^{2i-1} \right], \ 1 \le i \le q, \ \ (q \ge 1).$$

Будем иметь

$$\prod_{k=0}^{q} \left| \frac{u - \beta_k}{u + \beta_k} \right| \le \prod_{k=0}^{i-1} \frac{\xi^{2k} - \xi^{2i}}{\xi^{2k} + \xi^{2i}} \cdot \prod_{k=i}^{q} \frac{\xi^{2i-1} - \xi^{2k}}{\xi^{2i-1} + \xi^{2k}} \le \prod_{k=0}^{q} \frac{1 - \xi^{2k+1}}{1 + \xi^{2k+1}} \le \exp\left(-2 \sum_{k=0}^{q} \xi^{2k+1} \right) = \exp\left(-2 \frac{1 - \xi^{2q+2}}{\xi^{-1} - 1} \right). \tag{30}$$

Эта же оценка справедлива и для

$$u \in \left(\xi^{2i+1}, \xi^{2i} \right), \ 0 \le i \le q - 1.$$

Таким образом, из (28) и (30) на основании (27) получим, что

$$I_n < \Gamma(\alpha) \left(\frac{\xi^{2q}}{2^m} \right)^{\alpha} + \frac{1}{\alpha} \exp\left(-2 \frac{1 - \xi^{2q+2}}{\xi^{-1} - 1} \right).$$

Будем полагать, что $\xi^{2q+2} \le 2^{-1}$. Тогда из последнего неравенства имеем, что

$$I_n < \Gamma(\alpha) \left(\frac{(q+1)\xi^{2q}}{n_1} \right)^{\alpha} + \frac{1}{\alpha} \exp\left(-\frac{m}{2(q+1)(\xi^{-1} - 1)} \right).$$

Обозначаем

$$(q+1)(\xi^{-1} - 1) = t, \quad \xi^{-1} = 1 + \frac{t}{q+1}.$$

Следовательно,

$$I_n \le \Gamma(\alpha) \left(\frac{q+1}{n_1} \left(1 + \frac{t}{q+1} \right)^{-2q} \right)^{\alpha} + \frac{1}{\alpha} \exp\left(-\frac{n_1}{2t} \right).$$

Осталось заметить, что $n_1 = n - r$, и перейти к оценке (7) с учетом (24):

$$\varepsilon_{2n, 2q} \le C(\alpha) \left| \sin \frac{\pi\alpha}{2} \right| \cdot \left(\exp\left(-\frac{n}{2t} \right) + \left(\frac{q+1}{n} \right)^{\alpha} \left(1 + \frac{t}{q+1} \right)^{-2q\alpha} \right). \tag{31}$$

Поскольку $\xi \in (0, 1)$ и $\xi^{2q+2} \le 2^{-1}$, то легко видеть, что t может быть произвольным из промежутка $(1, +\infty)$. Следовательно, можем перейти к точной нижней грани в (31), т. е.

$$\varepsilon_{2n, 2q} \le C(\alpha) \left| \sin \frac{\pi\alpha}{2} \right| \cdot \inf_{i < t < +\infty} \left(\exp\left(-\frac{n}{2t} \right) + \left(\frac{q+1}{n} \right)^{\alpha} \left(1 + \frac{t}{q+1} \right)^{-2q\alpha} \right).$$

Теорема 2 доказана.

Фиксируем теперь целое число q, $q \ge 0$, и положим в (31)

$$t=\frac{n}{2(2q+1)\alpha\ln n},\quad n>n_0,$$

где n_0 – некоторое натуральное число, зависящее от q и α. Тогда получим

$$\varepsilon_{2n,2q}\le C(\alpha,q)\left|\sin\frac{\pi\alpha}{2}\right|\frac{\ln^{2q\alpha}n}{n^{(2q+1)\alpha}},\quad n>1, \tag{32}$$

где C(α, q) – некоторая положительная постоянная, зависящая только от q и α.

С л е д с т в и е. *Если q – фиксированное неотрицательное целое число, то справедлива оценка* (32).

З а м е ч а н и е 1. Полученная оценка (32) несколько точнее, чем соответствующая оценка в [20, см. следствие 1].

3. Общий рациональный случай. Т е о р е м а 3. *Для приближений функции* $|x|^{\alpha}$, $\alpha>0$, *интерполяционными рациональными функциями* (4) *справедливо неравенство*

$$\varepsilon_{2n}\le C_1(\alpha)\left|\sin\frac{\pi\alpha}{2}\right|\sqrt{n}e^{-\pi\sqrt{n\alpha}},\quad n\ge 1. \tag{33}$$

Причем для приближений на концах отрезка [–1,1] *при некоторых* a_k, $k=1,2,\ldots,2n$, *имеет место следующая оценка:*

$$\varepsilon_{2n}(\pm 1,a)\le C_2(\alpha)\left|\sin\frac{\pi\alpha}{2}\right|\cdot\sqrt{n}e^{-\pi\sqrt{2n\alpha}},\quad n\ge r. \tag{34}$$

Д о к а з а т е л ь с т в о. Обратимся к оценке (7). Очевидно, что

$$\varepsilon_{2n}(a)<\frac{4}{\pi}\left|\sin\frac{\pi\alpha}{2}\right|\int_0^1\frac{u^{\alpha-1}}{(1-u^2)^{\frac{\alpha}{2}}}|\psi_n(u)|du. \tag{35}$$

Рассмотрим интеграл справа:

$$J_n=\int_0^1\frac{u^{\alpha-1}}{(1-u^2)^{\frac{\alpha}{2}}}|\psi_n(u)|du.$$

Проведем следующие несложные преобразования. Замечая, что $n>r$ (см. (2)), полагаем $n=r+n_1$, $n_1\ge 1$. Тогда

$$J_n=\int_0^1\frac{u^{\alpha-1}}{(1-u^2)^{\frac{\alpha}{2}}}\left(\frac{1-u}{1+u}\right)^r\left|\prod_{k=r+1}^{n}\frac{u-\beta_k}{u+\beta_k}\right|du<\int_0^1 u^{\alpha-1}\left|\prod_{k=1}^{n_1}\frac{u-\beta_{r+k}}{u+\beta_{r+k}}\right|du.$$

Теперь поставим задачу об оценке точной нижней грани рассматриваемого интеграла по $\beta_{r+1},\beta_{r+2},\ldots,\beta_n$. Воспользуемся одним результатом работы [25] (см. также [17]) о том, что для любых $\gamma>0$ и $n\in N$ существуют числа β_k, $k=r+1,r+2,\ldots,r+n$, такие, что

$$u^{\gamma}\left|\prod_{k=1}^{n}\frac{u-\beta_{r+k}}{u+\beta_{r+k}}\right|\le C(\gamma)e^{-\pi\sqrt{n\gamma}},\quad u\in[0,1],$$

где $C(\gamma)$ – некоторая положительная постоянная, зависящая от γ. Тогда будем иметь

$$J_n\le\int_0^1 u^{-1+\frac{1}{\sqrt{n_1}}}\left|u^{\alpha-\frac{1}{\sqrt{n_1}}}\prod_{k=1}^{n_1}\frac{u-\beta_{r+k}}{u+\beta_{r+k}}\right|du\le C\left(\alpha-\frac{1}{\sqrt{n_1}}\right)\exp\left(-\pi\sqrt{\left(\alpha-\frac{1}{\sqrt{n_1}}\right)n_1}\right)\cdot\int_0^1 u^{-1+\frac{1}{\sqrt{n_1}}}du.$$

Заметим, из [25] следует, что постоянная $C\left(\alpha-\frac{1}{\sqrt{n_1}}\right)$ непрерывно зависит от параметра $\alpha-\frac{1}{\sqrt{n_1}}$ и, следовательно,

$$C\left(\alpha-\frac{1}{\sqrt{n_1}}\right)\le C_3(\alpha).$$

Далее,

$$\sqrt{n_1\alpha}-\sqrt{n_1\alpha-\sqrt{n_1}}=\frac{\sqrt{n_1}}{\sqrt{n_1\alpha}+\sqrt{n_1\alpha-\sqrt{n}_1}}<\frac{1}{\sqrt{\alpha}},\quad n_1>\frac{1}{\alpha^2}.$$

В итоге будем иметь, что

$$J_n\le C_4(\alpha)\sqrt{n_1}e^{-\pi\sqrt{n_1\alpha}}\le C_5(\alpha)\sqrt{n}e^{-\pi\sqrt{n\alpha}}.$$

Подставляя эту оценку в (35), получим

$$\varepsilon_{2n}=\inf_{\{a_1,a_2,\dots,a_n\}}\varepsilon_{2n}(a)\le C_1(\alpha)\left|\sin\frac{\pi\alpha}{2}\right|\sqrt{n}e^{-\pi\sqrt{n\alpha}},\quad n>r.$$

Неравенство (33) доказано.

Для доказательства оценки (34) воспользуемся равенством (6). Очевидно,

$$\varepsilon_{2n}(\pm1,a)=\frac{4}{\pi}\sin\frac{\pi\alpha}{2}\cdot\int_0^1\frac{u^{\alpha-1}}{(1-u^2)^{\alpha/2}}\frac{\psi_n(u)}{1+\psi_n^2(u)}du. \tag{36}$$

Далее, рассмотрим интеграл

$$I_n=\int_0^1\frac{u^{\alpha-1}}{(1-u^2)^{\alpha/2}}\frac{\psi_n(u)}{1+\psi_n^2(u)}du.$$

Представим его в виде

$$I_n=I_n^{(1)}-I_n^{(2)}, \tag{37}$$

где

$$I_n^{(1)}=\int_0^1\frac{u^{\alpha-1}}{(1-u^2)^{\alpha/2}}\psi_n(u)du,$$

$$I_n^{(2)}=\int_0^1\frac{u^{\alpha-1}}{(1-u^2)^{\alpha/2}}\frac{\psi_n^3(u)}{1+\psi_n^2(u)}du.$$

Займемся первым интегралом $I_n^{(1)}$. Его можно представить в виде

$$I_n^{(1)}=\int_0^1\phi(u)u^{\alpha-1}\prod_{k=1}^{n_1}\frac{u-\beta_{r+k}}{u+\beta_{r+k}}du, \tag{38}$$

где

$$\phi(u)=\frac{1}{(1-u^2)^{\alpha/2}}\left(\frac{u-1}{u+1}\right)^r,\quad r=\left[\frac{\alpha}{2}\right]+1,\quad n=n_1+r,\quad n>r.$$

Введем новые обозначения:

$$\prod_{k=1}^{n_1}\frac{u-\beta_{r+k}}{u+\beta_{r+k}}=\prod_{k=r+1}^{n}\frac{u-\beta_k}{u+\beta_k}=\psi_n(u).$$

Для оценки интеграла $I_n^{(1)}$ введем функцию

$$\Phi\left(\beta_{r+1},\beta_{r+2},\ldots,\beta_n\right)=\int_0^1\phi(t)t^{\alpha-1}\psi_n^2(t)dt.$$

Нетрудно показать, что функция $\Phi\left(\beta_{r+1},\beta_{r+2},\ldots,\beta_n\right)$ достигает наименьшего значения в некоторой точке $\left(\beta_{r+1}^*,\beta_{r+2}^*,\ldots,\beta_n^*\right)\in(0,1)^{n-r}$, причем числа $\beta_{r+1}^*,\beta_{r+2}^*,\ldots,\beta_n^*$ попарно различны. Запишем необходимое условие экстремума. Прежде всего, найдем частные производные:

$$\frac{\partial\Phi}{\partial\beta_m}=\int_0^1 2\phi(t)\cdot t^{\alpha-1}\psi_n(t)\prod_{\substack{k=r+1\\k\neq m}}^{n}\frac{t-\beta_k}{t+\beta_k}\frac{-2t}{(t+\beta_m)^2}dt=$$

$$=2\int_0^1\phi(t)t^{\alpha}\psi_n^2(t)\frac{-2dt}{(t+\beta_m)(t-\beta_m)}=2\int_0^1\phi(t)t^{\alpha}\psi_n^2(t)\frac{1}{\beta_m}\left(\frac{1}{t+\beta_m}-\frac{1}{t-\beta_m}\right)dt,\quad m=r+1,r+2,...,n.$$

Следовательно,

$$\int_0^1\phi(t)t^{\alpha}\psi_n^{*2}(t)\frac{dt}{t+\beta_m^*}=\int_0^1\phi(t)t^{\alpha}\psi_n^{*2}(t)\frac{dt}{t-\beta_m^*},\tag{39}$$

$$\psi_n^*(t)=\prod_{k=r+1}^{n}\frac{t-\beta_k^*}{t+\beta_k^*},\quad m=r+1,r+2,...,n.$$

Далее, легко видеть, что имеют место представления

$$\frac{\psi_n^*(t)-\psi_n^*(0)}{t}=\sum_{m=r+1}^{n}\frac{A_m}{t+\beta_m^*};\tag{40}$$

$$\frac{\psi_n^*(-t)-\psi_n^*(0)}{-t}=\sum_{m=r+1}^{n}\frac{A_m}{-t+\beta_m^*};\quad\psi_n^*(-t)=\left(\psi_n^*(t)\right)^{-1};$$

$$\frac{1-\psi_n^*(0)\psi_n^*(t)}{t\psi_n^*(t)}=\sum_{m=r+1}^{n}\frac{A_m}{t-\beta_m^*},\tag{41}$$

где $A_{r+1},A_{r+2},...,A_n$ – некоторые числа.

На основании равенств (39)–(41) заключаем, что

$$\int_0^1\phi(t)\cdot t^{\alpha}\psi_n^{*2}(t)\frac{\psi_n^*(t)-\psi_n^*(0)}{t}dt=\int_0^1\phi(t)\cdot t^{\alpha}\psi_n^{*2}(t)\frac{1-\psi_n^*(0)\psi_n^*(t)}{t\psi_n^*(t)}dt.$$

Отсюда найдем, что

$$\int_0^1\phi(t)t^{\alpha-1}\psi_n^*(t)dt=\int_0^1\phi(t)t^{\alpha-1}\psi_n^{*3}(t)dt.\tag{42}$$

Оценим интеграл справа в (42)

$$J_n = \int_0^1 \phi(t) t^{\alpha-1} \psi_n^{*3}(t) dt.$$

Нетрудно видеть, что

$$\left|J_n\right| \le \int_0^1 \phi(t) t^{\alpha-1} \psi_n^{*2}(t) dt. \tag{43}$$

Опираясь на работу [25] и действуя по той же схеме, что и при доказательстве неравенства (33), можно показать, что существуют числа $\beta_{r+1}, \beta_{r+2}, ..., \beta_n$ такие, что при $n > r$

$$\int_0^1 \phi(t) t^{\alpha-1} \left(\prod_{k=r+1}^{n} \frac{t-\beta_k}{t+\beta_k} \right)^2 dt \le C_6(\alpha) \sqrt{n} e^{-\pi\sqrt{2n\alpha}}. \tag{44}$$

Тогда, учитывая, что правая часть в неравенстве (43) является наименьшим значением функции $\Phi(\beta_{r+1}, \beta_{r+2}, \ldots, \beta_n)$, для нее тем более справедлива оценка (44), т. е.

$$\left|J_n\right| \le C_6(\alpha) \sqrt{n} e^{-\pi\sqrt{2n\alpha}}, \quad n > r. \tag{45}$$

Из равенства (42) и выражения для интеграла $I_n^{(1)}$ (см. (37), (38)) следует, что в случае, когда $\beta_k = \beta_k^*, \ k = r+1, r+2, ..., n,$

$$\left|I_n^{(1)}\right| = \left|J_n\right| \le C_6(\alpha) \sqrt{n} e^{-\pi\sqrt{2n\alpha}}, \quad n > r. \tag{46}$$

Естественно, что теперь интеграл $I_n^{(2)}$ (см. (37)) будет иметь вид

$$I_n^{(2)} = \int_0^1 \frac{u^{\alpha-1}}{(1-u^2)^{\alpha/2}} \frac{\psi_n^{*3}(u)}{1+\psi_n^{*2}(u)} du.$$

После несложных выкладок придем к тому, что и для этого интеграла также справедлива оценка (46), т. е.

$$\left|I_n^{(2)}\right| \le C_6(\alpha) \sqrt{n} e^{-\pi\sqrt{2n\alpha}}, \quad n > r.$$

В результате из равенства (37) следует, что при $\beta_k = \beta_k^*, \quad k = r+1, r+2, ..., n,$

$$\left|I_n\right| \le 2C_6(\alpha) \sqrt{n} e^{-\pi\sqrt{2n\alpha}}, \quad n > r.$$

Таким образом, неравенство (34), а с ним и теорема 3, доказаны.

З а м е ч а н и е 2. Выскажем предположение, что, скорее всего, оценка (34) имеет место на всем отрезке [–1,1].

Благодарности. Авторы выражают искреннюю признательность профессору А. А. Пекарскому за полезное обсуждение результатов данной работы.

Acknowledgments. The authors are sincerely grateful to Professor A. A. Pekarsky for useful discussion of the results of the present work.

Список использованных источников

1. Lebesgue, H. Sur l'approximation des fonctions / H. Lebesgue // Bull. Sci. Math. – 1898. – № 22. – P. 278–287.

2. Landay, E. Über die Approximation einer stetigen Funktion durch eine ganze rationale Funktion / E. Landay // Rend. Circ. Mat. Palermo. – 1908. – Vol. 25, №. 1. – P. 337–345. https://doi.org/10.1007/bf03029135

3. De la Vallee Poussin, Ch. J. Sur l'approximation des fonctions d'une variable réelle et de leurs dérivées par des polynomes et des suites limitées de Fourier / Ch. J. De la Vallee Poussin // Bull. Ac. de Belgique. – 1908. – № 3. – P. 3–64.

4. Бернштейн, С. Н. Доказательство теоремы Вейерштрасса, основанное на теории вероятностей / С. Н. Бернштейн // Сочинения. – 1912. – Т. 1. – С. 105–106.

5. Бернштейн, С. Н. О наилучшем приближении $|x|$ посредством многочленов данной степени / С. Н. Бернштейн // Собр. соч.: в 2 т. – М.: Изд-во Акад. наук СССР, 1952. – Т. 1. – С. 157–206.

6. Бернштейн, С. Н. О наилучшем приближении $|x|^p$ при помощи многочленов весьма высокой степени / С. Н. Бернштейн // Изв. Акад. наук СССР. Сер.мат. – 1938. – Т. 2, вып. 2. – С. 169–190.

7. Никольский, С. М. О наилучшем приближении многочленами в среднем функции $|a–x|^s$ / С. М. Никольский // Изв. АН СССР. Сер. мат. – 1947. – Т. 11, № 2. – С. 139–180.

8. Райцин, Р. А. О наилучшем среднеквадратическом приближении многочленами и целыми функциями конечной степени функций, имеющих алгебраическую особую точку / Р. А. Райцин // Изв. вузов. Математика. – 1969. – № 4. – С. 59–61.

9. Brutman, L. On the Divergence of Lagrange Interpolation to $|x|$ / L. Brutman, E. Passow // J. Approx. Theory. – 1995. – Vol. 81, № 1. – P. 127–135. https://doi.org/10.1006/jath.1995.1037

10. Byrne, G. J. On Lagrange interpolation with equidistant nodes / G. J. Byrne, T. M. Mills, S. J. Smith // Bull. Aust. Math. Soc. – 1990. – Vol. 42, № 1. – P. 81–89. https://doi.org/10.1017/s0004972700028161

11. Li, X. Local convergence of Lagrange interpolation associated with equidistant nodes / X. Li, E. B. Saff // J. Approx. Theory. – 1994. – Vol. 78, № 2. – P. 213–225. https://doi.org/10.1006/jath.1994.1073

12. Ganzburg, M. I. The Bernstein Constant and Polynomial Interpolation at the Chebyshev Nodes / M. I. Ganzburg // J. Approx. Theory. – 2002. – Vol. 119, № 2. – P. 193–213. https://doi.org/10.1006/jath.2002.3729

13. Revers, M. On the asymptotics of polynomial interpolation to $|x|^\alpha$ at the Chebyshev nodes / M. Revers // J. Approx. Theory. – 2013. – Vol. 165. – P. 70–82.

14. Ganelius, T. H. Rational approximation to x^α on [0,1] / T. H. Ganelius // Anal. Math. – 1979. – № 5. – P. 19–33. https://doi.org/10.1007/bf02079347

15. Andersson, J.-E. Rational approximation to functions like x^α in integral norms / J.-E. Andersson // Anal. Math. – 1988. – № 14, № 1. – P. 11–25.

16. Вячеславов, Н. С. О приближении функции $|x|$ рациональными функциями / Н. С. Вячеславов // Мат. заметки. – 1974. – Т. 16, № 1. – С. 163–171.

17. Вячеславов, Н. С. О наименьших уклонениях функции sign x и ее первообразных от рациональных функциях в метриках L_p, $0 <p \leq \infty$ / Н. С. Вячеславов // Мат. сб. – 1977. – Т. 103 (145), № 1 (5). – С. 23–36.

18. Шталь, Г. Наилучшие равномерные рациональные аппроксимации $|x|$ на [–1,1] / Г. Шталь // Мат. сб. – 1992. – Т. 183, № 8. – С. 85–118.

19. Лунгу, К. Н. О наилучших приближениях рациональными функциями с фиксированным числом полюсов / К. Н. Лунгу // Мат. сб. – 1971. – Т. 86 (128), № 2 (10). – С 314–324.

20. Лунгу, К. Н. О наилучших приближениях рациональными функциями с фиксированным числом полюсов / К. Н. Лунгу // Сиб. мат. журн. – 1984. – Т. 25, № 2. – С 151–160.

21. Ровба, Е. А. О приближении рациональными функциями с заданным числом полюсов / Е. А. Ровба // Современные проблемы теории функций: материалы Всесоюз. шк. по теории функций, Баку, 21 мая – 1 июня 1977 г. / Бакин. гос. ун-т. – Баку, 1980. – С. 234–239.

22. Ровба, Е. А. О приближении периодических аналитических функций с характерными особенностями рациональными функциями / Е. А. Ровба // Вес. АН БССР. Сер. фiз.-мат. навук. – 1974. – № 6. – С. 43–49.

23. Русак, В. Н. Рациональные функции как аппарат приближения / В. Н. Русак. – Минск: БГУ, 1979. – 176 с.

24. Фихтенгольц, Г. М. Курс дифференциального и интегрального исчисления / Г. М. Фихтенгольц. – СПб.: Лань, 1997. – Т. 2. – 800 с.

25. Пекарский, А. А. Построение экстремальных произведений Бляшке / А. А. Пекарский, Е. В. Ковалевская // Вес. Гродн. гос. ун-та им. Я. Купалы. Сер. 2. – 2017. – Т. 7, № 1. – С. 6–13.

References

1. Lebesgue H. Sur I'approximation des fonctions. *Bulletin des Sciences Mathématiques*, no. 22, 1898, pp. 278–287.

2. Landay E. Über die Approximation einer stetigen Funktion durch eine ganze rationale Funktion. *Rendiconti del Circolo Matematico di Palermo*, 1908, vol. 25, no. 1, pp. 337–345. https://doi.org/10.1007/bf03029135

3. De la Vallee Poussin Ch. J. Sur l'approximation des fonctions d'une variable réelle et de leurs dérivées par des polynomes et des suites limitées de Fourier. *Bull. Ac. de Belgique*, 1908, no. 3, pp. 3–64.

4. Bernshtein S. N. *Proof of the WeierstrassTheorem, based on probability theory. Vol. 1.* Writings, 1912, pp. 105–106.

5. Bernshtein S. N. *On the best approximation $|x|$ by polynomials of a given degree. Vol. 1.* Moscow, Publishing House Acad. sciences USSR, 1952, pp. 157–206 (in Russian).

6. Bernshtein S. N. Sur la meilleure approximation de $|x|^p$ par des polynomes de degres tre's eleves. *Izvestiya AN SSSR. Seriya matematicheskaya* [Mathematics of the USSR-Izvestiya], 1938, vol. 2, no. 2, pp. 169–190 (in Russian).

7. Nikolskii S. M. On the best mean approximation by polynomials of the functions $|a–x|^s$. *Izvestiya AN SSSR. Seriya matematicheskaya* [Mathematics of the USSR-Izvestiya], 1947, vol. 11, pp. 139–180 (in Russian).

8. Raitsin R. A. On the best approximation in the mean by polynomials and entire functions of finite degree of functions having an algebraic singularity. *Izvestiya vysshikh uchebnykh zavedenii. Matematika* [Russian Mathematics], 1969, no. 13, pp. 59–61 (in Russian).

9. Brutman L., Passow E. On the divergence of Lagrange interpolation to $|x|$. *Journal of Approximation Theory*, 1995, vol. 81, no. 1, pp. 127–135. https://doi.org/10.1006/jath.1995.1037

10. Byrne G. J., Mills T. M., Smith S. J. On Lagrange interpolation with equidistant nodes. *Bulletin of the Australian Mathematical Society,* 1990, vol. 42, no. 1, pp. 81–89. https://doi.org/10.1017/s0004972700028161

11. Li X., Saff E. B. Local convergence of Lagrange interpolation associated with equidistant nodes. *Journal of Approximation Theory,* 1994, vol. 78, no. 2, pp. 213–225. https://doi.org/10.1006/jath.1994.1073

12. Ganzburg M. The Bernstein Constant and Polynomial Interpolation at the Chebyshev Nodes. *Journal of Approximation Theory,* 2002, vol. 119, no. 2, pp. 193–213. https://doi.org/10.1006/jath.2002.3729

13. Revers M. On the asymptotics of polynomial interpolation to $|x|^\alpha$ at the Chebyshev nodes. *Journal of Approximation Theory,* 2013, vol. 165, pp. 70–82.

14. Ganelius T. H. Rational approximation to x^α on [0,1]. *Analysis Mathematica*, 1979, no. 5, pp. 19–33. https://doi.org/10.1007/bf02079347

15. Andersson J.-E. Rational approximation to functions like x^α in integral norms. *Analysis Mathematica*, 1988, vol. 14, no. 1, pp. 11–25. https://doi.org/10.1007/bf02350637

16. Vyacheslavov N. S. Approximation of the function $|x|$ by rational functions. *Mathematical Notes*, 1974, vol. 16, no. 1, pp. 680–685.

17. Vyacheslavov N. S. On the least deviations of the function sign x and its primitives from the rational functions in the Lp metrics, $0 < p \leq \infty$. *Mathematics of the USSR-Sbornik*, 1977, vol. 32, no. 1, pp. 19–31. https://doi.org/10.1070/sm1977v032n01abeh002313

18. Shtal' G. Best uniform rational approximation of $|x|$ on $[-1,1]$. *Russian Academy of Sciences. Sbornik Mathematics*, 1993, vol. 76, no. 2, pp. 461–487. https://doi.org/10.1070/sm1993v076n02abeh003422

19. Lungu K. N. On best approximations by rational functions with a fixed number of poles. *Mathematics of the USSR-Sbornik*, 1971, vol. 15, no. 2, pp. 313–324. https://doi.org/10.1070/sm1971v015n02abeh001547

20. Lungu K. N. Best approximations by rational functions with a fixed number of poles. *Siberian Mathematical Journal*, 1984, vol. 25, no. 2, pp. 289– 296.https://doi.org/10.1007/bf00971467

21. Rovba E. A. On approximation by rational functions with a given number of poles. *Sovremennye problem teorii funktsii: materialy Vsesoyuznoi shkoly po teorii funktsii, Baku, 21 maya – 1 iyunya 1977 g.* [Modern Problems of the Theory of Functions: Materials All-Union. schools on the theory of functions, Baku, May 21 – June 1, 1977]. Baku, 1980, pp. 234–239 (in Russian).

22. Rovba E. A. On approximation of periodic analytic functions with characteristic features of rational functions. *Vestsi Akademii navuk BSSR. Seryya fizika-matematychnykh navuk = Proceedings of the Academy of Sciences of BSSR. Physics and Mathematics series*, 1974, no. 6, pp. 43–49 (in Russian).

23. Rusak V. N. *Rational Functions as an Approximation Apparatus*. Minsk, BSU, 1979. 176 p. (in Russian).

24. Fikhtengol'ts G. M. *Course of Differential and Integral Calculus. Vol. 2.* St. Petersburg, Lan` Publ., 1997. 800 p. (in Russian).

25. Pekarskii A. A, Kovalevskaya E. V. Building extreme works Blaschke. *Vesnik Grodnenskogo gosudarstvennogo universiteta imeni Yanki Kupaly. Seriya 2 = Vesnik of the Yanka Kupala Grodno State University. Series 2*, 2017, vol. 7, no. 1, pp. 6–13 (in Russian).

Информация об авторах

Ровба Евгений Алексеевич – доктор физико-математических наук, профессор, заведующий кафедрой фундаментальной и прикладной математики, Гродненский государственный университет им. Я. Купалы (ул. Ожешко, 22, 230023, г. Гродно, Республика Беларусь). E-mail: rovba.ea@gmail.com

Медведева Виктория Юрьевна – магистрант, Гродненский государственный университет им. Я. Купалы (ул. Ожешко, 22, 230023, г. Гродно, Республика Беларусь). E-mail: Medvedeva_VJ_97@mail.ru

Information about the authors

Evgeniy A. Rovba – Dr. Sc. (Physics and Mathematics), Professor, Head of the Department of Fundamental and Applied Mathematics, Faculty of Mathematics and Informatics, Yanka Kupala State University of Grodno (22, Ozheshko Str., 230023, Grodna, Republic of Belarus). E-mail: rovba.ea@gmail.com

Victoria Yu. Medvedeva – Undergraduate, Yanka Kupala State University of Grodno (22, Ozheshko Str., 230023, Grodno, Republic of Belarus). E-mail: Medvedeva_VJ_97@.mail.ru

ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)
УДК 517.956.3
https://doi.org/10.29235/1561-2430-2019-55-4-406-412

Поступила в редакцию 27.11.2019
Received 27.11.2019

В. И. Корзюк[1,2], С. Н. Наумовец[3], В. А. Севастюк[1]

[1]Институт математики Национальной академии наук Беларуси, Минск, Беларусь
[2]Белорусский государственный университет, Минск, Беларусь
[3]Брестский государственный технический университет, Брест, Беларусь

КЛАССИЧЕСКОЕ РЕШЕНИЕ СМЕШАННОЙ ЗАДАЧИ ДЛЯ ОДНОМЕРНОГО ВОЛНОВОГО УРАВНЕНИЯ С ПРОИЗВОДНЫМИ ВТОРОГО ПОРЯДКА В ГРАНИЧНЫХ УСЛОВИЯХ

Аннотация. Рассмотрена смешанная задача для одномерного волнового уравнения с производными второго порядка в граничных условиях. Методом характеристик найдено классическое решение указанной задачи в аналитическом виде. Доказана его единственность при выполнении соответствующих условий согласования.

Ключевые слова: дифференциальные уравнения, гиперболические уравнения, частные производные, граничные условия, условия Коши, условия согласования, классическое решение

Для цитирования. Корзюк, В. И. Классическое решение смешанной задачи для одномерного волнового уравнения с производными второго порядка в граничных условиях / В. И. Корзюк, С. Н. Наумовец, В. А. Севастюк // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2019. – Т. 55, № 4. – С. 406–412. https://doi.org/10.29235/1561-2430-2019-55-4-406-412

V. I. Korzyuk[1,2], S. N. Naumavets[3], V. A. Sevastyuk[1]

[1]Institute of Mathematics of the National Academy of Sciences of Belarus, Minsk, Belarus
[2]Belarusian State University, Minsk, Belarus
[3]Brest State Technical University, Brest, Belarus

CLASSICAL SOLUTION OF THE MIXED PROBLEM FOR A ONE-DIMENSIONAL WAVE EQUATION WITH SECOND-ORDER DERIVATIVES AT BOUNDARY CONDITIONS

Abstract. This paper considers the mixed problem for a one-dimensional wave equation with second-order derivatives at boundary conditions. Using the method of characteristics, a classical solution to this problem is found in analytical form. Its uniqueness is proved under the relevant compatibility conditions.

Keywords: differential equations, hyperbolic equations, partial derivatives, boundary conditions, Cauchy conditions, agreement conditions, classical solution

For citation. Korzyuk V. I., Naumavets S. N., Sevastyuk V. A. Classical solution of the mixed problem for a one-dimensional wave equation with second-order derivatives at boundary conditions. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2019, vol. 55, no. 4, pp. 406–412 (in Russian). https://doi.org/10.29235/1561-2430-2019-55-4-406-412

Введение. В [1] в аналитическом виде представлено классическое решение в полуполосе первой смешанной задачи, где на части границы заданы условия с производными второго порядка. В [2] рассмотрена задача, где на границе задано условие в виде производной, порядок которой превышает порядок основного дифференциального уравнения. Задачи рассмотрены для одномерного волнового уравнения. В настоящей работе находится классическое решение смешанной задачи, где на всей границе присутствуют дифференциальные операторы второго порядка.

Метод характеристик, кроме построения искомого решения, также позволяет проанализировать качественную картину самого решения. Близкими к данной статье являются исследования [3, 4] и другие, где рассматриваются классические решения смешанных задач. Постановка задачи и полученные для нее результаты являются новыми. В решении смешанных задач могут быть обобщения как в сторону рассмотрения других уравнений, так и задания новых граничных условий.

Постановка задачи. На замыкании $\overline{Q}=[0,\infty)\times[0,l]$ области $Q=(0,\infty)\times(0,l)$ двух независимых переменных $\boldsymbol{x}=(x_0,x_1)\in\overline{Q}\subset\Re^2$ рассмотрим волновое уравнение

$$\left(\partial_{x_0}^2-a^2\partial_{x_1}^2\right)u(\boldsymbol{x})=f(\boldsymbol{x}),\quad \boldsymbol{x}=(x_0,x_1)\in\overline{Q}\subset\Re^2, \tag{1}$$

где a^2, l – положительные действительные числа, $\partial_{x_0}^2=\partial^2/\partial x_0^2$, $\partial_{x_1}^2=\partial^2/\partial x_1^2$ – частные производные по x_0 и x_1 второго порядка. К уравнению (1) на границе ∂Q области Q присоединяются условия типа Коши

$$u(0,x_1)=\varphi(x_1),\quad \left(\partial_{x_0}^2+b^{(1)}(x_1)\partial_{x_0}+b^{(0)}(x_1)\right)u(0,x_1)=\psi(x_1),\quad b^{(1)}(x_1)\neq 0,\quad x_1\in[0,l], \tag{2}$$

граничные условия

$$\partial_{x_1}^2u(x_0,0)=\mu^{(1)}(x_0),\quad \partial_{x_1}^2u(x_0,l)=\mu^{(2)}(x_0),\quad x_0\in[0,\infty). \tag{3}$$

Здесь $f:\boldsymbol{x}\to f(\boldsymbol{x})$, $\boldsymbol{x}\in\overline{Q}$, – заданная на $\overline{Q}$ функция, $\varphi:x_1\to\varphi(x_1)\in\Re$, $x_1\in[0,l]$, ψ: $x_1\to\psi(x_1)\in\Re$, $x_1\in[0,l]$, $b:x_1\to b(x_1)\in\Re$, $x_1\in[0,l]$, – функции на $[0,l]$, $\mu^{(j)}:x_0\to\mu^{(j)}(x_0)\in\Re$, $x_0\in[0,\infty)$, – заданные функции на $[0,\infty]$, $j=1,2$. Гладкость всех указанных функций будет уточнена ниже.

Кроме того, функции f, φ, ψ и $\mu^{(j)}$, $j=1,2$, удовлетворяют следующим однородным условиям согласования:

$$\mu^{(1)}(0)-d^2\varphi(0)+\frac{1}{a^2}f(0,0)=0, \tag{4}$$

$$\mu^{(2)}(0)-d^2\varphi(0)+\frac{1}{a^2}f(0,1)=0. \tag{5}$$

Классическое из $C^2\left(\overline{Q}\right)$ решение задачи (1)–(3). Построение решения задачи (1)–(3) будем осуществлять из общего решения u класса $C^2(\overline{Q})$ уравнения (1). Как известно, это решение есть сумма общего решения $u^{(0)}\in C^2(\overline{Q})$ однородного уравнения

$$\left(\partial_{x_0}^2-a^2\partial_{x_1}^2\right)u^{(0)}(\boldsymbol{x})=0,\quad \boldsymbol{x}\in\overline{Q}, \tag{6}$$

и частного решения $v_p\in C^2(\overline{Q})$ неоднородного уравнения (1).

Построение решения задачи (1)–(3) будет осуществляться пошагово, начиная с условий Коши (2). В качестве частного решения возьмем частное решение, построенное для уравнения (1) в работе [3], а именно:

$$v_p(\boldsymbol{x})=v_p^{(m)}(\boldsymbol{x}),\quad \boldsymbol{x}\in Q^{(m)}=\left(\frac{(m-1)l}{a},\frac{ml}{a}\right),\ m=1,2,3,..., \tag{7}$$

где

$$v_p(\boldsymbol{x})=f^{(1,m)}\left(x_1-ax_0\right)+f^{(2,m)}(x_1+ax_0)-\frac{1}{4a^2}\int\limits_{l-ml}^{x_1-ax_0}dy\int\limits_{ml}^{x_1+ax_0}f\left(\frac{z-y}{2a},\frac{z+y}{a}\right)dz,\quad \boldsymbol{x}\in\overline{Q^{(m)}}, \tag{8}$$

$\overline{Q^{(m)}}$ – замыкание подобласти $Q^{(m)}$, функции $f^{(j,m)}$ из класса C^2.

Т е о р е м а 1. *Пусть правая часть уравнения* (1) $f\in C^1(\overline{Q})$. *Тогда функция* v_p, *определенная формулами* (7) *и* (8) *при соответствующем выборе* $f^{(j,m)}$, $j=1,2$, $m\in\aleph$, *принадлежит классу* $C^2(\overline{Q})$, *является решением уравнения* (1) *и удовлетворяет условиям*

$$\begin{gathered}v_p(0,x_1)=\partial_{x_0}v_p(0,x_1)=0,\\ \partial_{x_0}^2v_p(0,x_1)=f(0,x_1),\quad x_1\in[0,l].\end{gathered} \tag{9}$$

Доказательство см. в [3].

Как известно, общее решение уравнения (1) из класса $C^2(\overline{Q})$ представимо в виде

$$u(\boldsymbol{x}) = g^{(1)}(x_1 - ax_0) + g^{(2)}(x_1 - ax_0) + v_p(\boldsymbol{x}), \quad \boldsymbol{x} \in \overline{Q}, \tag{10}$$

где $g^{(j)}$ ($j = 1,2$) – произвольные функции из класса C^2. Точнее, $g^{(1)} \in C^2\left(\left(-\infty, l\right]\right)$, $g^{(2)} \in C^2\left(\left[0, \infty\right)\right)$ для $\boldsymbol{x} \in \overline{Q}$, если условиться, что $a > 0$.

Функции $g^{(j)}$ выбираем такими, чтобы решение (10) удовлетворяло условиям (2) и (3). Подставляя в (10) в условия типа Коши (2), получим систему двух уравнений

$$\begin{gathered} g^{(1)}(x_1) + g^{(2)}(x_1) = \varphi(x_1), \quad x_1 \in [0, l], \\ a^2 d^2 g^{(1)}(x_1) + a^2 d^2 g^{(2)}(x_1) - ab^{(1)}(x_1) dg^{(1)}(x_1) + ab^{(1)}(x_1) dg^{(2)}(x_1) + \\ + b^{(0)}(x_1) g^{(1)}(x_1) + b^{(0)}(x_1) g^{(2)}(x_1) = \psi(x_1), \quad x_1 \in [0, l]. \end{gathered} \tag{11}$$

Из системы (11) находим значения $g^{(j)}(x_1)$ функций $g^{(j)}$, $j = 1,2$. Так как по условию $b^{(1)}(x_1) \neq 0$ для $x_1 \in [0, l]$, то

$$\begin{aligned} g^{(1)}(z) = g^{(1,0)}(z) = \frac{1}{2}\varphi(z) - \frac{1}{2a}\int_0^z \frac{1}{b^{(1)}(\xi)}\left[\psi - a^2 d^2 \varphi - b^{(0)}\varphi\right](\xi) d\xi - C, \\ g^{(2)}(z) = g^{(2,0)}(z) = \frac{1}{2}\varphi(z) + \frac{1}{2a}\int_0^z \frac{1}{b^{(1)}(\xi)}\left[\psi - a^2 d^2 \varphi - b^{(0)}\varphi\right](\xi) d\xi + C, \end{aligned} \tag{12}$$

где $z \in [0, l]$, C – произвольная постоянная из множества $\Re$, d – оператор обыкновенной производной, $d^2 = d \cdot d$, $d^3 = d \cdot d \cdot d$ и так далее.

Для других z значения $g^{(j)}(z)$, $j = 1,2$, определяем из граничных условий (3). Для этого введем обозначения

$$\begin{gathered} g^{(1)}(z) = g^{(1,k)}(z), \quad \text{если } z \in [-kl, -(k-1)l], \\ g^{(2)}(z) = g^{(2,k)}(z), \quad \text{если } z \in [kl, (k+1)l], \end{gathered} \tag{13}$$

где $k = 1,2,3,\ldots$. Подставляя представление решения (10) в первое условие (3), получим обыкновенное дифференциальное уравнение

$$d^2 g^{(1)}(z) = \mu^{(1)}\left(-\frac{z}{a}\right) - d^2 g^{(2)}(-z) - \partial_{x_1}^2 v_p\left(-\frac{z}{a}, x_1 = 0\right). \tag{14}$$

Используя второе условие из (3), получим уравнения

$$d^2 g^{(2)}(z) = \mu^{(2)}\left(\frac{z-l}{a}\right) - d^2 g^{(1)}(2l - z) - \partial_{x_1}^2 v_p\left(\frac{z-l}{a}, l\right). \tag{15}$$

Уравнения (14) и (15) представляют собой уравнения с аргументами со сдвигом. Учитывая сказанное и то, что значения, согласно формулам (12), определены на отрезке $[0,l]$, целесообразно формулы (14) и (15) записать в виде

$$\begin{aligned} d^2 g^{(1,k)}(z) = \mu^{(1)}\left(-\frac{z}{a}\right) - d^2 g^{(2,k-1)}(-z) - \partial_{x_1}^2 v_p\left(-\frac{z}{a}, 0\right), \quad z \in [-kl, -(k-1)l], \\ d^2 g^{(2,k)}(z) = \mu^{(2)}\left(\frac{z-l}{a}\right) - d^2 g^{(1,k-1)}(2l - z) - \partial_{x_1}^2 v_p\left(\frac{z-l}{a}, l\right), \quad z \in [kl, (k+1)l]. \end{aligned} \tag{16}$$

В пределах указанных отрезков интегрируем уравнения (16). В результате получим соотношения

$$
\begin{aligned}
&g^{(1,k)}(z)=\int_0^z (z-\xi)\mu^{(1)}\left(-\frac{\xi}{a}\right)d\xi-\int_0^z (z-\xi)\partial_{x_1}^2 v_p\left(-\frac{\xi}{a},0\right)d\xi-\\
&\quad -g^{(2,k-1)}(-z)+C^{(1,k)}z+C^{(2,k)},\quad z\in[-kl,-(k-1)l],\\
&g^{(2,k)}(z)=\int_0^z (z-\xi)\mu^{(2)}\left(\frac{\xi-l}{a}\right)d\xi-\int_0^z (z-\xi)\partial_{x_1}^2 v_p\left(\frac{\xi-l}{a},l\right)d\xi-\\
&\quad -g^{(1,k-1)}(2l-z)+\tilde{C}^{(1,k)}z+\tilde{C}^{(2,k)},\quad z\in[kl,(k+1)l],
\end{aligned}
\tag{17}
$$

где $C^{(j,k)}$, $\tilde{C}^{(j,k)}$ – произвольные постоянные из $\mathfrak{R}$, которые появились в результате интегрирования уравнения (16), $k=1,2,3,\ldots$.

Из формул (7), (8), (12) и (17) видно, что значения $g^{(j,k)}(z)$ функций $g^{(j,k)}$, $j=1,2$, $k=1,2,3,\ldots$, определяются через заданные функции f, φ, ψ и $\mu^{(1)}$ и $\mu^{(2)}$. Если предположить, что $f\in C^1(\bar{Q})$, $\varphi\in C^3([0,l])$, $\psi\in C^1([0,l])$, $\mu^{(j)}\in C([0,\infty))$, то функции $g^{(1,k)}$ будут из класса $C^2([-kl,-(k-1)l])$, а $g^{(2)}\in C^2([kl,(k+1)l])$, $k=1,2,3,\ldots$. В целом, чтобы функции $g^{(1)}\in C^2((-\infty,l])$ и $g^{(2)}\in C^2([0,\infty))$, определенные через эти функции согласно формулам (13), необходимо и достаточно, чтобы для частично определенных функций $g^{(j,k)}$ в общих точках соприкосновения выполнялись равенства

$$
d^p g^{(1,k+1)}(-kl)=d^p g^{(1,k)}(-kl),\quad p=0,1,2;\ \ k=0,1,2,\ldots,
\tag{18}
$$

$$
d^p g^{(2,k)}(kl)=d^p g^{(2,k-1)}(kl),\quad p=0,1,2;\ k=1,2,\ldots .
\tag{19}
$$

Л е м м а 1. *Равенства* (18) *и* (19) *для каждого* $p=0,1,2$ *в отдельности выполняются тогда и только тогда, когда они выполняются для какого-нибудь* k.

Д о к а з а т е л ь с т в о непосредственно следует из формул (12) и (17), так как они представляют собой рекуррентную зависимость между $g^{(j,k+1)}$ и $g^{(j,k)}$.

Чтобы дальше исследовать равенства (18) и (19), вычислим производные первого порядка функций (12) и (17):

$$
\begin{aligned}
dg^{(1,0)}(z)&=\frac{1}{2}d\varphi(z)-\frac{1}{2ab^{(1)}(z)}\left[\psi(z)-a^2d^2\varphi(z)-b^{(0)}(z)\varphi(z)\right],\\
dg^{(2,0)}(z)&=\frac{1}{2}d\varphi(z)+\frac{1}{2ab^{(1)}(z)}\left[\psi(z)-a^2d^2\varphi(z)-b^{(0)}(z)\varphi(z)\right],
\end{aligned}
\tag{20}
$$

$$
\begin{aligned}
dg^{(1,k)}(z)&=\int_0^z \mu^{(1)}\left(-\frac{\xi}{a}\right)d\xi+dg^{(2,k-1)}(-z)-\int_0^z \partial_{x_1}^2 v_p\left(-\frac{\xi}{a},0\right)d\xi+C^{(1,k)},\\
dg^{(2,k)}(z)&=\int_0^z \mu^{(2)}\left(\frac{\xi-l}{a}\right)d\xi+dg^{(1,k-1)}(2l-z)-\int_0^z \partial_{x_1}^2 v_p\left(\frac{\xi-l}{a}\right)d\xi+\tilde{C}^{(1,k)}.
\end{aligned}
\tag{21}
$$

Рассмотрим сначала равенства (18) для $k=0$ и равенства (19) для $k=1$. Запишем их через заданные функции задачи (1)–(3).

Пусть $p=1$. Так как производные $dg^{(j,0)}(z)$, $j=1,2$, представленные формулами (20), не содержат произвольных постоянных, а представляются единственным образом через заданные функции задачи (1)–(3), то, подставляя их в формулы (21), определим единственным образом и константы $C^{(1,1)}$ и $\tilde{C}^{(1,1)}$ через заданные функции. Продолжая данный процесс дальше, мы единственным образом определим все константы $C^{(1,k)}$ и $\tilde{C}^{(1,k)}$ через заданные функции задачи (1)–(3) и для $k=2,3,4,\ldots$.

А теперь пусть $p=0$. Из соотношений (12) и (17), используя условие согласования (18) для $k=0$, имеем равенство

$$g^{(1,1)}(l)=-\frac{1}{2}\varphi\left(0\right)-C+C^{(2,1)}=g^{(1,0)}(0)=\frac{1}{2}\varphi(0)-C.$$

Отсюда $C^{(2,1)}=\varphi(0)$. Аналогично, из условия (19) в случае $k=1$ из соотношений (12) и (17) следует равенство

$$g^{(2,1)}(l)=\int_0^l(z-\xi)\mu^{(2)}\left(\frac{\xi-l}{a}\right)d\xi-\frac{1}{2}\varphi(l)+\frac{1}{2a}\int_0^l\frac{\left[\psi-a^2d^2\varphi-b^{(0)}\varphi\right](\xi)}{b^{(1)}(\xi)}d\xi+$$

$$+C-\int_0^l(z-\xi)\partial_{x1}^2v_p\left(\frac{\xi-l}{a},l\right)d\xi+\tilde{C}^{(1,1)}l+\tilde{C}^{(2,1)}=$$

$$=g^{(2,0)}(l)=\frac{1}{2}\varphi(l)+\frac{1}{2a}\int_0^l\frac{\left[\psi-a^2d^2\varphi-b^{(0)}\varphi\right](\xi)}{b^{(1)}(\xi)}d\xi+C.$$

Из последнего равенства следует, что константа $\tilde{C}^{(2,1)}$ также определяется единственным образом через значения заданных функций задачи (1)–(3), а именно:

$$\tilde{C}^{(2,1)}=\varphi(l)-\int_0^l(z-\xi)\mu^{(2)}\left(\frac{\xi-l}{a}\right)d\xi+\int_0^l(z-\xi)\partial_{x1}^2v_p\left(\frac{\xi-l}{a},l\right)d\xi-\tilde{C}^{(1,1)}l.$$

Продолжая данный процесс дальше, из условий согласования (18) и (19) согласно полученным выражениям $C^{(2,k)}$ и $\tilde{C}^{(2,k)}$ для других $k=2,3,\ldots$, которые определяются единственным образом.

Полученный результат сформулируем в виде леммы.

Л е м м а 2. *Если функции*

$$f\in C^1(\overline{Q}),\ \ \varphi\in C^3\left(\left[0,l\right]\right),\ \ b^{(0)},b^{(1)},\ \psi\in C^1\left(\left[0,l\right]\right),\ \ \mu^{(j)}\in C\left(\left[0,\infty\right)\right),\ \ j=1,2,$$

то функции $g^{(j,k)}$, $j=1,2$, $k=0,1,2,\ldots$, *представленные формулами* (12) *и* (17)*, записываются в виде*

$$g^{(j,k)}(z)=\tilde{g}^{(j,k)}(z)+(-1)^jC,\quad j=1,2,\quad k=0,1,2,\ldots, \tag{22}$$

тогда и только тогда, когда для них выполняются условия согласования (18) *и* (19) *для* $p=0,1$, *где функции* $\tilde{g}^{(j,k)}$ *из класса* $C^2\left(\left[(-1)^jkl,(-1)^j(k+1)l\right]\right)$ *определяются единственным образом через заданные функции задачи* (1)–(3)*,* C *– произвольная константа из* $\Re$.

Кроме этого, при выполнении условий леммы 2 функции $g^{(j)}$, $j=1,2$, определенные формулами (13), принадлежат классу C^1 и выражаются формулами (22). Справедлива

Л е м м а 3. *При выполнении условий леммы 2 функция* $g^{(1)}\in C^1\left((-\infty,l]\right)$, *а* $g^{(2)}\in C^1\left([0,\infty)\right)$ *тогда и только тогда, когда для составляющих их функций* $g^{(j,k)}$, $j=1,2$, $k=0,1,2,\ldots$, *выполняются условия согласования* (18) *и* (19) *для* $p=0,1$. *Кроме этого, при соответствующем выборе единственным способом констант* $C^{(j,k)}$ *и* $\tilde{C}^{(j,k)}$, *входящих в определение функций* $g^{(j,k)}$, *справедливы представления*

$$g^{(j)}(z)=\tilde{g}^{(j)}(z)+(-1)^jC,\quad j=1,2, \tag{23}$$

где функции определены единственным образом через заданные функции задачи (1)–(3)*,* C *– произвольная константа из* $\Re$.

Т е о р е м а 2. *Пусть заданные функции задачи* (1)–(3) *удовлетворяют следующим условиям гладкости:*

$$f\in C^1(\overline{Q}),\quad \varphi\in C^3\left(\left[0,l\right]\right),\quad \psi\in C^1\left(\left[0,l\right]\right),\quad \mu^{(j)}\in C\left(\left[0,\infty\right)\right),\quad j=1,2,\ \ b^{(0)},b^{(1)}\in C^1\left(\left[0,l\right]\right).$$

При выполнении этих условий функция вида (10) *является единственным классическим решением из класса* $C^2(\overline{Q})$ *тогда и только тогда, когда имеют место однородные условия согласования* (18) *для* $p=0,1$ *и* $k=0$, (19) *– для* $p=0,1$ *и* $k=1$, *однородные условия согласования* (4) *и* (5), *где частное решение* v_p *определяется формулами* (7), (8), *функции* $g^{(j)}$ (j = 1,2) – *формулами* (12), (13), (17).

Д о к а з а т е л ь с т в о. Согласно теореме 1, если $f \in C^1(\bar{Q})$, то частное решение $v_p \in C^2(\bar{Q})$.

Из леммы 3 следует, что функция $g^{(1)}$ – из класса $C^1\left((-\infty, l]\right)$, $g^{(2)}$ – из класса $C^1\left([0,\infty)\right)$. Как было сказано ранее, эти функции дважды непрерывно дифференцируемы на соответствующих областях определения тогда и только тогда, когда для p = 2 дополнительно будут выполняться условие (18) для k = 0 и условие (19) для k = 1. Распишем эти условия в явном виде через заданные функции, используя формулы (12) и (16). Для этого вычислим вторые производные соотношений (12). В результате получим

$$\begin{aligned} d^2 g^{(1,0)}(z) &= \frac{1}{2} d^2\varphi(z) + \frac{1}{2a\left(b^{(1)}(z)\right)^2} \cdot db^{(1)}(z)\left[\psi(z) - a^2 d^2\varphi(z) - b^0 v(z)\varphi(z)\right] - \\ &- \frac{1}{2ab^{(1)}(z)}\left[d\psi(z) - a^2 d^3\varphi(z) - db^{(0)}(z)\varphi(z) - b^{(0)}(z) d\varphi(z)\right], \\ d^2 g^{(2,0)}(z) &= \frac{1}{2} d^2\varphi(z) - \frac{1}{2a\left(b^{(1)}(z)\right)^2} \cdot db^{(1)}(z)\left[\psi(z) - a^2 d^2\varphi(z) - b^0(z)\varphi(z)\right] + \\ &+ \frac{1}{2ab^{(1)}(z)}\left[d\psi(z) - a^2 d^3\varphi(z) - db^{(0)}(z)\varphi(z) - b^{(0)}(z) d\varphi(z)\right]. \end{aligned} \tag{24}$$

Пусть в равенстве (18) p = 2, k = 0. В этом случае, используя первое соотношение $d^2 g^{(1,1)}(0)$ из (16), а также $d^2 g^{(1,0)}$ из (24), данное равенство запишется в виде

$$\mu^{(1)}(0) - \frac{1}{2} d^2\varphi(0) - \partial_{x1}^2 v_p(0,0) = \frac{1}{2} d^2\varphi(0).$$

Так как согласно уравнению (1) и последнему условию из (9)

$$\partial_{x1}^2 v_p(0,0) = \frac{1}{a^2}\partial_{x0}^2 v_p(0,0) - \frac{1}{2a^2} f(0,0) = -\frac{1}{a^2} f(0,0),$$

мы получили однородное условие согласования (4)

$$\mu^{(1)}(0) - d^2\varphi(0) + \frac{1}{a^2} f(0,0) = 0. \tag{25}$$

Аналогично, из равенства (19) для p = 2, k = 1 получаем однородное условие согласования (5)

$$\mu^{(2)}(0) - d^2\varphi(l) + \frac{1}{a^2} f(0,l) = 0. \tag{26}$$

Таким образом, согласно лемме 3 и равенствам (18) и (19), функции $g^{(1)} \in C^2\left((-\infty, l]\right)$, $g^{(2)} \in C^2\left([0,\infty)\right)$. Следовательно, функция u, представленная формулой (10), принадлежит классу $C^2(\bar{Q})$ и является решением задачи (1)–(3). Это решение является единственным, что следует из леммы 3, а также из представления (23) для функций $g^{(j)}$, j = 1,2.

Все остальные утверждения теоремы 2 следуют из предыдущих рассуждений.

З а м е ч а н и е. В теореме 2 сформулированы утверждения через выполнение необходимых и достаточных условий согласования (18) для p = 0,1 и k = 0 и (19) для p = 0,1 и k = 1. Из предыдущих рассуждений следует, что они эквивалентны специально выбранным единственным образом коэффициентам $C^{(j,k)}$ *и* $\tilde{C}^{(j,k)}$, j = 1,2; k = 0,1,2,…, в формулах (17). Эти коэффициенты, начиная с k = 1, можно выписать в явном виде. Тогда в формулировке теоремы вместо условий согласования (18) и (19) для p = 0,1 будут необходимыми и достаточными единственным образом выбранные эти выражения, которые фигурируют в качестве констант $C^{(j,k)}$ *и* $\tilde{C}^{(j,k)}$, входящих в формулы (17) функций $g^{(j,k)}$.

Список использованных источников

1. Корзюк, В. И. Классическое решение первой смешанной задачи одномерного волнового уравнения с условиями типа Коши / В. И. Корзюк, И. С. Козловская, С. Н. Наумовец // Вес. Нац. акад. Навук Беларусі. Сер. фіз.-мат. навук. – 2015. – № 1. – С. 7–20.

2. Корзюк, В. И. Классическое решение смешанной задачи для одномерного волнового уравнения с производными высокого порядка в граничных условиях / В. И. Корзюк, С. Н. Наумовец // Докл. Нац. акад. наук Беларуси. – 2016. – Т. 60, № 3. – С. 11–17.

3. Корзюк, В. И. О классическом решении второй смешанной задачи для одномерного волнового уравнения / В. И. Корзюк, С. Н. Наумовец, В. А. Севастюк // Тр. Ин-та математики. – 2018. – Т. 26, № 1. – С. 35–42.

4. Корзюк, В. И. Метод характеристического параллелограмма решения второй смешанной задачи для одномерного волнового уравнения / В. И. Корзюк, С. Н. Наумовец, В. П. Сериков // Тр. Ин-та математики. – 2018. – Т. 26, № 1. – С. 43–53.

References

1. Korzyuk V. I., Kozlovskaya I. S., Naumavets S. N. Classical solution to the first mixed problem for one-dimensional wave equation with conditions of cauchy type. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2015, no. 1, pp. 7–20 (in Russian).

2. Korzyuk V. I., Naumavets S. N. Classical solution of mixed problem for one-dimensional wave equation with derivatives of high order in the boundary conditions. *Doklady Natsional'noi akademii nauk Belarusi = Doklady of the National Academy of Sciences of Belarus*, 2016, vol. 60, no. 3, pp. 11–17 (in Russian).

3. Korzyuk V. I., Naumavets S. N., Sevastyuk V. A. On the classical solution of the second mixed problem for a one-dimensional wave equation. *Trudy Instituta matematiki = Proceedings of the Institute of Mathematics*, 2018, vol. 26, no. 1, pp. 35–42 (in Russian).

4. Korzyuk V. I., Naumavets S. N., Serikov V. P. The method of the characteristic parallelogram of the solution of the second mixed problem for the one-dimensional wave equation. *Trudy Instituta matematiki = Proceedings of the Institute of Mathematics*, 2018, vol. 26, no. 1, pp. 43–53 (in Russian).

Информация об авторах

Корзюк Виктор Иванович – академик, профессор, доктор физико-математических наук, Институт математики Национальной академии наук Беларуси (ул. Сурганова, 11, 220072, г. Минск, Республика Беларусь), Белорусский государственный университет (пр. Независимости, 4, 220030, г. Минск, Республика Беларусь). E-mail: korzyuk@bsu.by

Наумовец Светлана Николаевна – старший преподаватель, Брестский государственный технический университет (ул. Московская, 267, 224017, г. Брест, Республика Беларусь). E-mail: e-cveta@tut.by

Севастюк Владимир Александрович – ведущий инженер-программист, Институт математики Национальной академии наук Беларуси (ул. Сурганова, 11, 220072, г. Минск, Республика Беларусь).

Information about the authors

Viktor I. Korzyuk – Academician, Professor, Dr. Sc. (Physics and Mathematics), Institute of Mathematics of the National Academy of Sciences of Belarus (11, Surganov Str., 220072, Minsk, Republic of Belarus); Belarusian State University (4, Nezavisimosti Ave., 220030, Minsk, Republic of Belarus). E-mail: korzyuk@bsu.by

Sviatlana N. Naumavets – Senior Lecturer, Brest State Technical University (267, Moskovskaya Str., 224017, Brest, Republic of Belarus). E-mail: e-cveta@tut.by

Vladimir A. Sevastyuk – Engineer, Institute of Mathematics of the National Academy of Sciences of Belarus (11, Surganov Str., 220072, Minsk, Republic of Belarus).

ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)
УДК 519.63
https://doi.org/10.29235/1561-2430-2019-55-4-413-424

Поступила в редакцию 02.05.2019
Received 02.05.2019

А. Н. Гуревский

Белорусский государственный университет, Минск, Беларусь

ИСПОЛЬЗОВАНИЕ РЕКУРСИВНЫХ ЦИФРОВЫХ ФИЛЬТРОВ ДЛЯ ПОСТРОЕНИЯ РАЗНОСТНЫХ СХЕМ ВЫСОКИХ ПОРЯДКОВ ДЛЯ НЕСТАЦИОНАРНОГО УРАВНЕНИЯ ШРЕДИНГЕРА

Аннотация. Исследованы двухслойные разностные схемы высоких порядков для нестационарного уравнения Шредингера. С использованием методов цифровой обработки сигналов доказан критерий консервативности разностных схем любого порядка для уравнения Шредингера. С помощью достигнутых теоретических результатов вычислены аналитические выражения для коэффициентов разностной схемы восьмого порядка. Получены условия эквивалентности разностных схем восьмого порядка представлению в виде каскада всепропускающих цифровых фильтров первого порядка. На основе численного анализа показано превосходство разностной схемы восьмого порядка при решении линейного уравнения Шредингера над схемой повышенного порядка точности на шеститочечном шаблоне. На примере моделирования двухсолитонного решения нелинейного уравнения Шредингера посредством метода дробных шагов второго порядка точности установлено, что схемы высоких порядков не позволяют радикально улучшить точность полученного решения. Исследован вопрос о вычислительной сложности разностных схем высоких порядков. Полученные результаты могут быть использованы при конструировании эффективных численных алгоритмов численного анализа как линейных, так и нелинейных задач для уравнений шредингеровского типа при применении метода дробных шагов соответствующего порядка точности.

Ключевые слова: разностные схемы, восьмой порядок, уравнение Шредингера, рекурсивный цифровой фильтр, всепропускающий фильтр

Для цитирования. Гуревский, А. Н. Использование рекурсивных цифровых фильтров для построения разностных схем высоких порядков для нестационарного уравнения Шредингера / А. Н. Гуревский // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2019. – Т. 55, № 4. – С. 413–424. https://doi.org/10.29235/1561-2430-2019-55-4-413-424

A. N. Hureuski

Belarusian State University, Minsk, Belarus

USING IIR FILTERS TO BUILD HIGH-ORDER FINITE DIFFERENCE SCHEMES FOR THE UNSTEADY SCHRÖDINGER EQUATION

Abstract. High-order finite difference schemes for the time-dependent Schrödinger equation are investigated. Digital signal processing methods allowed proving the conservativeness of high-order finite difference schemes for the unsteady Schrödinger equation. The eighth-order scheme coefficients were found with the help of the proved theoretical results. The conditions for equivalence between the eighth-order finite difference scheme and the scheme in the form of a cascade of all-pass first-order filters were found. The numerical analysis of the proposed scheme was made. It was shown that the high-order finite difference schemes gave better results on solving the linear Schrödinger equations comparing to the well-known fourth-order scheme on the six-point stencil, however, the high-order schemes in couple with the second-order splitting algorithm to the nonlinear Schrödinger equation do not lead to a radical improvement in the quality of numerical results. Practical issues implementing the proposed numerical technique are considered. The obtained results can be used to construct efficient solvers for linear and nonlinear Schrödinger-type equations by applying the splitting schemes of adequate accuracy order.

Keywords: finite-difference schemes, eighth order, Schrödinger equation, IIR filter, all-pass filter

For citation. Hureuski A. N. Using IIR filters to build high-order finite difference schemes for the unsteady Schrödinger equation. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2019, vol. 55, no. 4, pp. 413–424 (in Russian). https://doi.org/10.29235/1561-2430-2019-55-4-413-424

Введение. Теория разностных схем гласит, что схемы высоких порядков должны обеспечивать лучшую аппроксимацию дифференциальной задачи и, следовательно, давать лучшую точность решения [1]. Очевидно, что в общем случае схемы высоких порядков проигрывают схемам более низких порядков с точки зрения вычислительной сложности, однако их повышенная точность позволяет выбирать намного более крупные шаги сетки по пространственным и эволюционным переменным, что в свою очередь компенсирует некоторый проигрыш в скорости работы. В связи с этим представляет интерес построение компактных разностных схем высоких порядков для нестационарного уравнения Шредингера.

Для уравнения Шредингера классической схемой считается двухслойная схема с весами на шеститочечном шаблоне [1]. При определенных соотношениях на параметры схемы она может обеспечивать четвертый порядок аппроксимации. Под высоким порядком в контексте уравнения Шредингера будем понимать порядок аппроксимации от шестого и выше.

Схемы высоких порядков в основном рассматривались для уравнения теплопроводности [2]. В [3] были рассмотрены компактные разностные схемы высоких порядков для уравнения теплопроводности и уравнения Шредингера. Примечательно, что во всех исследованиях построение схем высоких порядков производится одинаковым способом: аппроксимация невязки дискретной модели на заданном шаблоне методом неопределенных коэффициентов. В [4] было показано, что с не меньшим успехом можно использовать методы теории цифровой обработки сигналов для аппроксимации разностных схем. Также теория фильтров позволяет получать разностные схемы с заданными характеристиками как, например, консервативность [5].

В настоящей работе исследованы критерии консервативности двухслойных разностных схем для нестационарного уравнения Шредингера. Полученные теоретические результаты использованы для построения разностных схем восьмого порядка. Рассмотрена возможность применения схем высоких порядков при решении линейного уравнения Шредингера, а также при вычислении линейной части метода дробных шагов для нелинейного уравнения Шредингера.

Постановка задачи. Рассмотрим нестационарное уравнение Шредингера

$$i\frac{\partial u}{\partial t}+\frac{\partial^2 u}{\partial x^2}=0,\quad x\in[-L,L],\tag{1}$$

с начальными и граничными условиями вида

$$u(x,0)=u_0(x),\quad u(-L,t)=u(L,t)=0.\tag{2}$$

Пусть p – произвольное натуральное число. На прямоугольном $(4p+2)$-точечном шаблоне на равномерной сетке $\omega_x=\{x_k=-L+hk,\ k=\overline{1,N-1},\ h=2L/N,\ t_m=m\tau,\ m=0,1,\ldots\}$ рассмотрим двухслойную разностную схему для задачи (1), (2):

$$i\frac{\hat{U}_k-U_k}{\tau}=-\frac{\sum\limits_{j=-p}^{p}a_j\hat{U}_{k+j}}{h^2}-\frac{\sum\limits_{j=-p}^{p}b_jU_{k+j}}{h^2},\quad \hat{U}_0=\hat{U}_N=0,\tag{3}$$

причем $\forall j\in\overline{1,p}\ a_{-j}=a_j,\ \ b_{-j}=b_j;\ \ a_j,b_j\in\mathbb{C},\ \left|a_p\right|+\left|b_p\right|\neq 0.$

Здесь $U_k=U(x_k,t_m),\ \ \hat{U}_k=U(x_k,t_{m+1})$, набор коэффициентов $\{a_j\}$ используется на верхнем слое разностной схемы, а набор $\{b_j\}$, соответственно, на нижнем. Отметим, что данное определение является обобщением известной двухслойной разностной схемы с весами на шеститочечном шаблоне [1].

В работах [4–6] была показана связь между разностными схемами и рекурсивными цифровыми фильтрами. Так, полагая

$$r=\frac{\tau}{h^2}>0,$$

передаточная функция цифрового фильтра, соответствующего разностной схеме (3), может быть представлена в виде

$$H_{FD}(\omega)=\frac{1+ir\left(b_0+2\sum_{j=1}^{p}b_j\cos j\omega\right)}{1-ir\left(a_0+2\sum_{j=1}^{p}a_j\cos j\omega\right)},\ -\pi\le\omega<\pi. \tag{4}$$

Несложно видеть, что для консервативности схемы (3) требуется, чтобы $\left|H_{FD}(\omega)\right|\equiv 1$, что в свою очередь означает, что цифровой фильтр относится к классу всепропускающих. Покажем, что для выполнения данного условия необходимо и достаточно, чтобы выполнялось следующее утверждение:

$$\forall j\in\overline{0,p}\ b_j=\overline{a_j}.$$

В данной записи горизонтальная черта над соответствующим коэффициентом разностной схемы обозначает комплексное сопряжение. Предварительно докажем две вспомогательные леммы.

Лемма 1. *Рекурсивный цифровой фильтр n-го порядка является всепропускающим тогда и только тогда, когда количество нулей совпадает с количеством полюсов, значения нулей являются обратными, комплексно-сопряженными значениям соответствующих полюсов, модуль коэффициента усиления фильтра равен произведению модулей всех полюсов.*

Доказательство. Передаточная функция рекурсивного цифрового фильтра может быть записана в следующем виде [7]:

$$H(z)=k\frac{\prod_{j=1}^{m}(1-z_jz^{-1})}{\prod_{j=1}^{n}(1-p_jz^{-1})}, \tag{5}$$

где $\{z_j\}$ – нули фильтра, $\{p_j\}$ – полюса фильтра, $k\in\mathbb{C}$ – коэффициент усиления. По определению всепропускающего фильтра $\left|H(e^{i\omega})\right|\equiv 1$, что равносильно следующей цепочке преобразований:

$$\left|H(e^{i\omega})\right|\equiv 1\Leftrightarrow H(e^{i\omega})\overline{H(e^{i\omega})}\equiv 1\Leftrightarrow |k|^2\frac{\prod_{j=1}^{m}(1-z_je^{-i\omega})(1-\overline{z_j}e^{i\omega})}{\prod_{j=1}^{n}(1-p_je^{-i\omega})(1-\overline{p_j}e^{i\omega})}\equiv 1\Leftrightarrow$$

$$\Leftrightarrow |k|^2\frac{\prod_{j=1}^{m}\overline{z_j}}{\prod_{j=1}^{n}\overline{p_j}}\cdot\frac{\prod_{j=1}^{m}(1-z_je^{-i\omega})\left(1-\frac{1}{\overline{z_j}}e^{-i\omega}\right)}{\prod_{j=1}^{n}(1-p_je^{-i\omega})\left(1-\frac{1}{\overline{p_j}}e^{-i\omega}\right)}\equiv 1. \tag{6}$$

Из выражения (6) следует, что комплексный полином, записанный в числителе дроби, должен тождественно равняться комплексному полиному, записанному в ее знаменателе. Применяя основную теорему алгебры, получим, что $n = m$, $|k|=\prod_{j=1}^{m}|p_j|$, а также, с точностью до перестановки, $p_j=\frac{1}{\overline{z_j}}$. Случай $p_j=z_j$ не представляет интереса, так как приводит к вырождению фильтра. Из эквивалентности преобразований следует необходимость и достаточность условий леммы. Лемма доказана.

Следствие 1. *Передаточные функции всепропускающих цифровых фильтров первого и второго порядков имеют соответственно следующий вид:*

$$H(z)=k\frac{1-\frac{1}{\overline{p}}z^{-1}}{1-pz^{-1}},\quad |k|=|p|, \tag{7}$$

$$H(z)=k\frac{\left(1-\frac{1}{\overline{p_1}}z^{-1}\right)\left(1-\frac{1}{\overline{p_2}}z^{-1}\right)}{(1-p_1z^{-1})(1-p_2z^{-1})},\quad |k|=|p_1||p_2|. \tag{8}$$

С л е д с т в и е 2. *Любой всепропускающий фильтр порядка $n \geq 2$ может быть представлен в виде каскада всепропускающих фильтров более низких порядков.*

Л е м м а 2. *Фазово-частотная характеристика всепропускающего фильтра второго порядка представима в виде*

$$H(w)=c\frac{\cos\omega-a}{\cos\omega-b}, \tag{9}$$

где $a,b,c\in\mathbb{C}$ тогда и только тогда, когда

$$p_2=\frac{1}{\overline{p_1}}, \tag{10}$$

где p_1, p_2 – полюсы фильтра.

Д о к а з а т е л ь с т в о. Воспользуемся следствием 1 и формулой (8) и представим передаточную функцию всепропускающего фильтра второго порядка в следующем виде:

$$H(e^{iw})=k\frac{\left(1-\frac{1}{\overline{p_1}}e^{-i\omega}\right)\left(1-\frac{1}{\overline{p_2}}e^{-i\omega}\right)}{(1-p_1e^{-i\omega})(1-p_2e^{-i\omega})}=k\frac{e^{i\omega}+\frac{1}{\overline{p_1p_2}}e^{-i\omega}-\frac{1}{\overline{p_1}}-\frac{1}{\overline{p_2}}}{e^{i\omega}+p_1p_2e^{-i\omega}-p_1-p_2},\quad |k|=|p_1||p_2|. \tag{11}$$

Из условия (11) и того, что $\cos\omega=\frac{e^{-i\omega}+e^{i\omega}}{2}$, и следует выполнение условия (10). Лемма доказана.

С л е д с т в и е 3. *Для формулы* (9) *верны следующие соотношения:*

$$b=\overline{a},\quad |c|=1. \tag{12}$$

Интересно отметить, что при подстановке формулы (10) в формулу (7) получим тот самый сопряженный фильтр, определенный в [5], т. е. такой фильтр, процедура реализации которого состоит в использовании стандартного алгоритма рекурсивного фильтра, реализуемого в обратном направлении дискретных отсчетов. Данный факт следует из следующей цепочки равенств:

$$H_1(z)=k_1\frac{1-\frac{1}{\overline{p}}z^{-1}}{1-pz^{-1}},\quad |k_1|=|p|,$$

$$H_2(z)=k_2\frac{1-\overline{p}z^{-1}}{1-\frac{1}{p}z^{-1}}=k_2\frac{z-\overline{p}}{z-\frac{1}{p}}=k_2|p|^2\frac{1-\frac{1}{\overline{p}}z}{1-pz}=k_1\frac{1-\frac{1}{\overline{p}}z}{1-pz}=H_1(z^{-1}).$$

Применим полученные утверждения для доказательства критерия консервативности разностных схем (3) для решения нестационарного уравнения Шредингера.

Передаточная функция цифрового фильтра, соответствующего дифференциальной задаче (1), (2), имеет вид

$$H(\omega)=e^{-ir\omega^2}. \tag{13}$$

Заметим, что $H(0) = 1$. В связи с этим является естественным выполнение следующего условия для разностных схем:

$$H_{FD}(0) = 1. \tag{14}$$

Сформулируем и докажем следующую теорему.

Т е о р е м а. *Разностная схема* (3), *для которой выполняется условие* (14), *является консервативной тогда и только тогда, когда* $\forall j \in \overline{0, p}\ b_j = \overline{a_j}$.

Д о к а з а т е л ь с т в о. Достаточность данного утверждения очевидна, так как при подстановке значений b_j в передаточную функцию фильтра (4), соответствующего разностной схеме (3), получим дробь, числитель и знаменатель которой являются комплексно-сопряженными функциями. Следовательно, модуль данной функции тождественно равен единице, что в свою очередь означает, что данный фильтр относится к классу всепропускающих, а значит, и разностная схема (3) является консервативной.

Для доказательства необходимости применим полученные ранее леммы и следствия. Для начала отметим, что, рассматривая передаточную функцию (4) как полином относительно cosω, в силу основной теоремы алгебры можно получить следующее представление передаточной функции:

$$H_{FD}(\omega) = k\frac{\prod_{j=1}^{p}\left(\cos\omega - c_j\right)}{\prod_{j=1}^{p}\left(\cos\omega - d_j\right)}. \tag{15}$$

Согласно следствию 3, верны следующие равенства: $d_j = \overline{c_j},\ |k| = 1$. Тогда, подставив выражение (15) в условие (14), получим

$$H_{FD}(0) = k\frac{\prod_{j=1}^{p}\left(1 - c_j\right)}{\prod_{j=1}^{p}\left(1 - \overline{c_j}\right)} = 1 \Leftrightarrow k = \frac{\prod_{j=1}^{p}\left(1 - \overline{c_j}\right)}{\prod_{j=1}^{p}\left(1 - c_j\right)}. \tag{16}$$

Отсюда имеем

$$H_{FD}(\omega) = \frac{\prod_{j=1}^{p}\left(1 - \overline{c_j}\right)}{\prod_{j=1}^{p}\left(1 - c_j\right)} \cdot \frac{\prod_{j=1}^{p}\left(\cos\omega - c_j\right)}{\prod_{j=1}^{p}\left(\cos\omega - \overline{c_j}\right)}. \tag{17}$$

Таким образом, передаточная функция (17) представляет собой отношение двух комплексно-сопряженных функций. Тогда, возвращаясь к исходной записи (4) и приравнивая соответствующие коэффициенты, можем констатировать, что $\forall j \in \overline{0, p}\ b_j = \overline{a_j}$. Теорема доказана.

В частности, из теоремы 1 следует, что двухслойная схема с весами на шеститочечном шаблоне является консервативной тогда и только тогда, когда $\overline{\sigma} = 1 - \sigma \Leftrightarrow \mathrm{Re}\,\sigma = 0{,}5$. Это является широко известным результатом [1], однако методы цифровой обработки сигналов позволили получить его иным способом, отличным от традиционных методов.

Построение разностных схем высоких порядков. Покажем, как применить теорему 1 для построения консервативных разностных схем с локальным порядком аппроксимации $O(\tau^4 + h^8)$. Рассмотрим разностную схему (3) при $p = 2$ на десятиточечном шаблоне. Тогда передаточная функция (4) примет вид

$$H_8(\omega) = \frac{1 + ir(b_0 + 2b_1\cos\omega + 2b_2\cos 2\omega)}{1 - ir(a_0 + 2a_1\cos\omega + 2a_2\cos 2\omega)}, \quad -\pi \le \omega < \pi. \tag{18}$$

Потребуем выполнения условия (14): $H_8(0) = 1$. Значит, верно равенство

$$a_0 + b_0 + 2a_1 + 2b_1 + 2a_2 + 2b_2 = 0. \tag{19}$$

Согласно теореме 1, $b_j = \overline{a_j},\ j = \overline{0,2}$. Представим коэффициенты разностной схемы следующим образом:

$$\begin{aligned} a_0 &= a_0^r + ia_0^i, \\ a_1 &= a_1^r + ia_1^i, \\ a_2 &= a_2^r + ia_2^i, \\ b_0 &= \overline{a_0} = a_0^r - ia_0^i, \\ b_1 &= \overline{a_1} = a_1^r - ia_1^i, \\ b_2 &= \overline{a_2} = a_2^r - ia_2^i. \end{aligned} \tag{20}$$

Тогда условие (19) примет вид

$$a_0^r = -2a_1^r - 2a_2^r. \tag{21}$$

В работе [4] было показано, что из аппроксимации передаточной функции фильтра, соответствующего дифференциальной задаче, передаточной функцией фильтра, соответствующего применению разностной схемы, следует, что и сама разностная схема аппроксимирует дифференциальную задачу. Приравнивая коэффициенты при соответствующих степенях ω разложений в ряд Тейлора функций (13) и (18) и последовательно исключая неизвестные, с учетом равенств (20) и (21) получим

$$\begin{aligned} &\omega^0 : a_0^r = -2a_1^r - 2a_2^r, \\ &\omega^2 : a_1^r = \frac{a_0^i r}{2} - 4a_2^r + a_1^i r + a_2^i r + \frac{1}{2}, \\ &\omega^4 : a_2^r = \frac{5a_1^i r}{12} - \frac{a_0^i r}{24} + \frac{23a_2^i r}{12} - \frac{1}{24}, \\ &\omega^6 : a_0^i = -\frac{1}{r} - \frac{124a_2^i r - 11a_1^i r + 30a_1^i r^3 + 30a_2^i r^3}{15r^3 + 2r}, \\ &\omega^8 : a_2^i = -a_1^i \frac{6300r^4 + 105r^2 - 23}{25200r^4 + 13020r^2 + 688}. \end{aligned} \tag{22}$$

Равенства (22) обеспечивают аппроксимацию передаточной функции (13) до ω^8 включительно. Попытка приравнять коэффициенты при ω^{10} соответствующих разложений в ряд Тейлора приводит к уравнению, зависящему только от r, следовательно, значение a_1^i не влияет на порядок аппроксимации и может быть выбрано произвольным, удобным для нас образом. Отметим также, что данное уравнение не разрешимо относительно r в действительных положительных числах, поэтому можно утверждать, что восьмой порядок аппроксимации – это максимально возможный порядок для данной разностной схемы.

Выражая все неизвестные в равенствах (22) через a_1^i, можно заметить, что в качестве a_1^i удобно положить $a_1^i = 2(25200r^4 + 13020r^2 + 688)$. Отсюда получим следующие выражения для коэффициентов разностной схемы (20):

$$\begin{aligned} a_0 &= -\frac{i}{r} - 75600r^4 i - 113400r^3 + 99540r^2 i - 4770r + 4716i, \\ a_1 &= 50400r^4 i + 50400r^3 + 26040r^2 i + 1920r + 1376i, \\ a_2 &= -12600r^4 i + 6300r^3 - 210r^2 i + 465r + 46i, \end{aligned}$$

$$\begin{aligned} b_0 &= \frac{i}{r} + 75600r^4 i - 113400r^3 - 99540r^2 i - 4770r - 4716i, \\ b_1 &= -50400r^4 i + 50400r^3 - 26040r^2 i + 1920r - 1376i, \\ b_2 &= 12600r^4 i + 6300r^3 + 210r^2 i + 465r - 46i. \end{aligned} \tag{23}$$

Таким образом, коэффициенты (23) обеспечивают разностной схеме (3) восьмой порядок аппроксимации на десятиточечном шаблоне.

Проводя полностью аналогичные рассуждения для построения разностной схемы четвертого порядка на шеститочечном шаблоне, можно получить следующие выражения для коэффициентов схемы, зависящие от произвольного параметра a_1^i:

$$\begin{aligned} a_0 &= -\frac{i}{r} - 12a_1^i r + 10a_1^i i, \\ a_1 &= 6a_1^i r + a_1^i i, \\ b_0 &= \frac{i}{r} - 12a_1^i r - 10a_1^i i, \\ b_1 &= 6a_1^i r - a_1^i i. \end{aligned} \tag{24}$$

Примечательно, что если в равенствах (24) положить $a_1^i = \dfrac{1}{12r}$, то получим известную двухслойную разностную схему с весами на шеститочечном шаблоне. Данное замечание приведено для того, чтобы подчеркнуть – множество консервативных разностных схем четвертого порядка на шеститочечном шаблоне не описывается лишь известной схемой с весами, а содержит схемы и с другими соотношениями коэффициентов.

Эквивалентное представление разностной схемы на основе рекурсивных цифровых фильтров. В процессе применения разностной схемы (3) на десятиточечном шаблоне возникает симметричная пятидиагональная матрица. Существуют вариации метода прогонки для такой матрицы [8], однако соотношения (20), вообще говоря, не гарантируют выполнения необходимых условий для вычислительной устойчивости метода прогонки. В связи с этим представим разностную схему в виде каскада фильтров более низких порядков и используем данное представление для организации вычислений.

С точки зрения цифровых фильтров передаточная функция (18) соответствует всепропускающему фильтру четвертого порядка. Согласно следствию 2 и лемме 2, данный фильтр четвертого порядка можно представить в виде пары всепропускающих фильтров второго порядка, причем каждый из фильтров второго порядка является каскадом пары сопряженных всепропускающих фильтров первого порядка. Таким образом, верно равенство

$$H_8(\omega) = k_1 k_2 \frac{e^{i\omega} + e^{i\omega} - \left(\overline{p_1} + \dfrac{1}{\overline{p_1}}\right)}{e^{i\omega} + e^{i\omega} - \left(p_1 + \dfrac{1}{p_1}\right)} \cdot \frac{e^{i\omega} + e^{i\omega} - \left(\overline{p_2} + \dfrac{1}{\overline{p_2}}\right)}{e^{i\omega} + e^{i\omega} - \left(p_2 + \dfrac{1}{p_2}\right)}, \quad |k_1 k_2| = 1. \tag{25}$$

Для дальнейшего удобства сделаем некоторые замены переменных:

$$\begin{aligned} q_1 &= p_1 + \frac{1}{p_1}, \\ q_2 &= p_2 + \frac{1}{p_2}. \end{aligned} \tag{26}$$

Преобразуя выражение (25) и приравнивая соответствующие коэффициенты в выражениях (18) и (25), а также с учетом замен (26) и равенства $H_8(0) = 1$, получим систему уравнений относительно q_1 и q_2:

$$\begin{cases} \left(4+\overline{q_1}\overline{q_2}-2(\overline{q_1}+\overline{q_2})\right)\left(2+q_1q_2\right)=1-ira_0, \\ \left(4+\overline{q_1}\overline{q_2}-2(\overline{q_1}+\overline{q_2})\right)\left(-2(q_1+q_2)\right)=-2ira_1, \\ \left(4+\overline{q_1}\overline{q_2}-2(\overline{q_1}+\overline{q_2})\right)\cdot 2=-2ira_2. \end{cases} \tag{27}$$

Разделив первое и второе уравнения системы (27) на третье уравнение, имеем

$$\begin{cases} q_1+q_2=-\dfrac{a_1}{a_2}, \\ q_1q_2=-2-\dfrac{1-ira_0}{ira_2}. \end{cases} \tag{28}$$

Следовательно, q_1 и q_2 являются корнями квадратного уравнения с комплексными коэффициентами

$$q^2+\frac{a_1}{a_2}q-\left(2+\frac{1-ira_0}{ira_2}\right)=0. \tag{29}$$

Решая уравнение (29), получим

$$q_{1,2}=-\frac{a_1\pm\sqrt{a_1{}^2+8a_2{}^2-4a_0a_2-\dfrac{4a_2i}{r}}}{2a_2}. \tag{30}$$

Найдем значения p_1 и p_2 из системы (26):

$$\begin{aligned} p_1&=\frac{q_1\pm\sqrt{q_1{}^2-4}}{2}, \\ p_2&=\frac{q_2\pm\sqrt{q_2{}^2-4}}{2}. \end{aligned} \tag{31}$$

Таким образом, с учетом леммы 1 равенства (31) и (30) позволяют представить разностную схему (3) на десятиточечном шаблоне в виде всепропускающего фильтра четвертого порядка с полюсами p_1, $\dfrac{1}{p_1}$, p_2, $\dfrac{1}{p_2}$ и, соответственно, нулями $\dfrac{1}{\overline{p_1}}$, $\overline{p_1}$, $\dfrac{1}{\overline{p_2}}$, $\overline{p_2}$. В силу симметричности всех систем уравнений в выражениях (30) и (31) для простоты можно зафиксировать один из знаков.

Для устойчивости рекурсивного цифрового фильтра необходимо и достаточно, чтобы его полюса лежали внутри единичного круга. Но в силу того, что полюса сопряженных фильтров первого порядка являются взаимообратными величинами, прямой ход будем реализовывать для фильтра с полюсом внутри единичного круга, а обратный – для полюса вне единичного круга. Данный алгоритм прямого-обратного хода широко применяется в теории цифровых фильтров [7]. Также отметим, что из леммы 1 следует невозможность того, что полюс всепропускающего фильтра расположен на единичной окружности, так как в таком случае он будет совпадать с нулем фильтра, и рекурсивный фильтр выродится в константу.

В качестве любопытного факта заметим, что разностную схему на десятиточечном шаблоне можно рассматривать как «каскад» двух разностных схем на шеститочечном шаблоне. Данный процесс выглядит следующим образом: к известному временному слою применяется первая из разностных схем на шеститочечном шаблоне, и отсюда имеем некоторое прогнозное решение. Затем к этому прогнозному решению применяем вторую разностную схему, и полученное решение будем использовать в качестве решения на следующем временном слое. Эта процедура полностью аналогична известной схеме предиктор-корректор. Отметим, что данные разностные схемы на шеститочечном шаблоне не обязаны иметь вид схемы с весами, а в общем виде описываются соотношениями (24).

Результаты численного моделирования. Исследуем возможность применения разностной схемы восьмого порядка для решения уравнения (1), (2) с начальным условием вида

$$u_0(x) = e^{-\frac{x^2}{2}}. \tag{32}$$

Для начального условия (32) известно [9] аналитическое решение уравнения (1), (2):

$$u_a(x,t) = \frac{e^{-\frac{x^2}{2(1+4t^2)}+i\left(\frac{tx^2}{1+4t^2}-\frac{\operatorname{arc\,tg}(2t)}{2}\right)}}{\sqrt[4]{1+4t^2}}. \tag{33}$$

Сравним относительные погрешности приближенных решений уравнения (1), (2), (32), полученных с помощью разностной схемы четвертого порядка на шеститочечном шаблоне, разностной схемы восьмого порядка на десятиточечном шаблоне, дискретного преобразования Фурье, с точным решением (33). Определим погрешность приближенного решения как

$$\delta = \frac{\left\|U - U_a\right\|_{l_2}}{\left\|U_a\right\|_{l_2}}, \tag{34}$$

где U – приближенное решение, U_a – точное решение (33) на сетке ω_x в момент времени t_f. Будем использовать следующие значения параметров: $L = 40$; $t_f = 1$; $r = 0{,}5$; $N = 32 \cdot 2^j$; $j = \overline{1,5}$. Результаты численного моделирования приведены на рис. 1.

Данные результаты показывают превосходство схемы восьмого порядка над схемой четвертого порядка при решении линейного уравнения Шредингера, что полностью согласуется с теоретическими результатами. Полученная линейная зависимость в логарифмическом масштабе

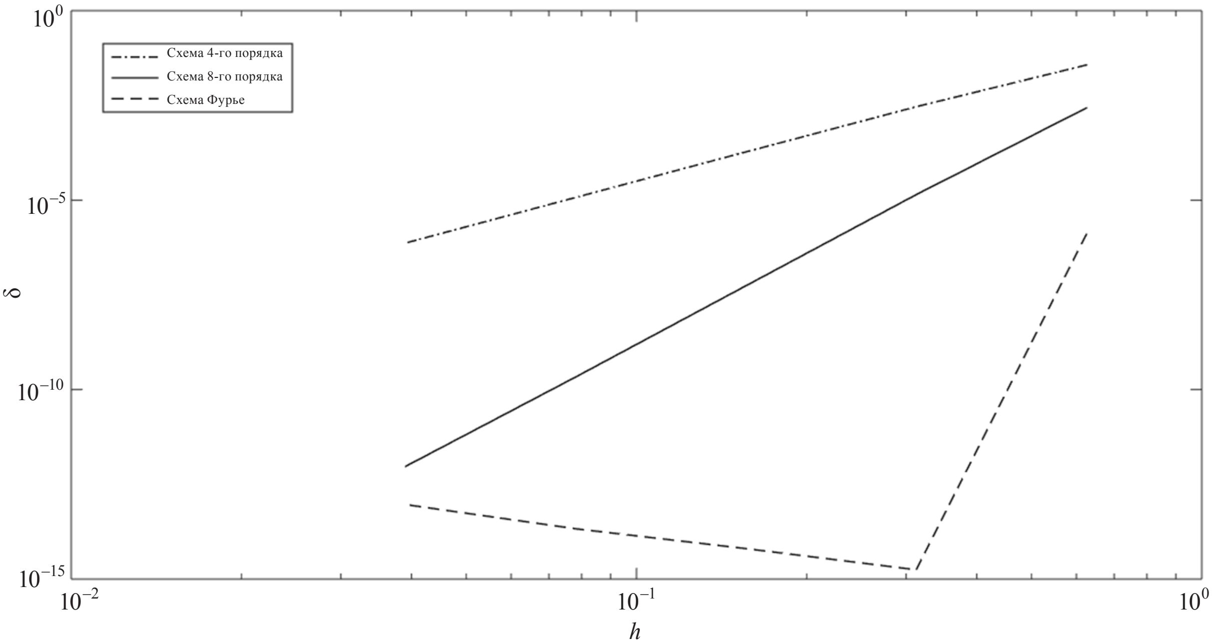


Рис. 1. Зависимости относительной погрешности численных решений линейного уравнения, полученных с помощью разностных схем четвертого и восьмого порядков, схемы Фурье, от размера шага по пространству

Fig. 1. Relative error dependences of the solutions of the linear equation obtained using the fourth- and eighth-order finite difference schemes and the Fourier method on the spatial step size

соответствует восьмому порядку аппроксимации. Отметим также, что при определенном размере шага разностная схема восьмого порядка практически сравнивается по точности с методом Фурье.

С точки зрения вычислительной сложности схема восьмого порядка уступает схеме четвертого порядка примерно в 2 раза, однако для достижения заданного уровня погрешности лучшие характеристики схемы восьмого порядка позволяют использовать более грубые размеры шага сетки по пространству и, следовательно, по эволюционной переменной. Несложно показать, что численное решение с помощью схемы четвертого порядка требует $4\frac{Lt_f}{rh^3}$ операций сложения и $6\frac{Lt_f}{rh^3}$ операций умножения, с помощью схемы восьмого порядка – $8\frac{Lt_f}{rh^3}$ и $10\frac{Lt_f}{rh^3}$ соответственно. Отсюда следует, что увеличение шага по пространству в k раз для схемы восьмого порядка позволяет сократить суммарное количество операций в $\frac{5k^3}{9}$ раз.

Решение линейного уравнения (1), (2) на практике является одним из этапов метода дробных шагов [10] для решения нелинейного уравнения Шредингера, которое моделирует многие физические процессы, например распространение оптических пучков [9]. При применении метода Фурье для решения линейной части метода дробных шагов возникают искусственные эффекты при переходе от дифференциальной задачи к дискретной [11]. Это приводит к падению точности численного решения нелинейного уравнения. Использование разностных схем для решения линейной части в некоторых случаях позволяет получить значительно лучшие результаты по сравнению с методом Фурье [12]. В связи с этим исследуем возможность применения схемы восьмого порядка при решении линейно части метода дробных шагов.

Рассмотрим нелинейное уравнение Шредингера с бикубической нелинейностью:

$$i\frac{\partial u}{\partial t}+\frac{\partial^2 u}{\partial x^2}+2|u|^2 u=0,\quad x\in[-L,L] \tag{35}$$

с начальными и граничными условиями вида (2).

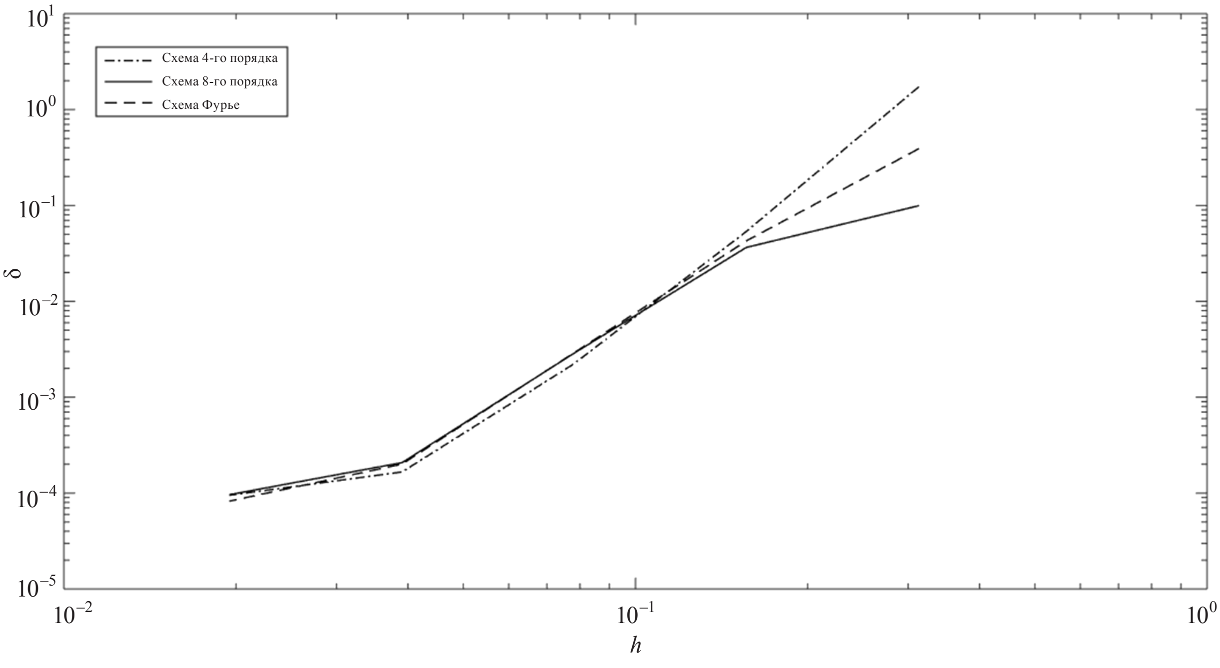


Рис. 2. Зависимости относительной погрешности численных решений нелинейного уравнения, полученных с помощью разностных схем четвертого и восьмого порядков, схемы Фурье, от размера шага по пространству

Fig. 2. Relative error dependences of the solutions of the nonlinear equation obtained using the fourth- and eighth-order finite difference schemes and the Fourier method on the spatial step size

Оценим эффективность разностных схем четвертого и восьмого порядков в сравнении с методом Фурье на примере моделирования двухсолитонных решений уравнения (35) методом дробных шагов. Двухсолитонное решение уравнения (35) имеет вид финитной, периодической по времени, бесконечно дифференцируемой функции [10]:

$$u(t,x)=4\frac{\cosh(3x)+3\exp(8it)\cosh(x)}{\cosh(4x)+4\cosh(2x)+3\cos(8t)}e^{it}. \tag{36}$$

Будем использовать следующие значения параметров: $L=20$; $t_f=1$; $r=0,5$; $N=32\cdot 2^j$; $j=\overline{1,5}$. Результаты численного моделирования приведены на рис. 2.

Из рис. 2 видно, что схема восьмого порядка дает лучшие результаты при достаточно крупных шагах сетки, но при уменьшении размера шага различия в точности всех методов нивелируются, а на определенных участках графика разностная схема четвертого порядка показывает даже лучшую точность. Близость погрешности всех методов можно объяснить тем, что с некоторого значения размера шага по пространству погрешность расщепления метода дробных шагов начинает доминировать над погрешностью решения линейной части метода. Данное обстоятельство не позволяет рассматривать схему восьмого порядка в качестве замены известным методам для решения нелинейного уравнения Шредингера.

Заключение. Представленные выше теоретические результаты показывают, что методы цифровой обработки сигналов могут быть успешно использованы для построения консервативных двухслойных разностных схем сколь угодно высоких порядков для решения линейного нестационарного уравнения Шредингера. В частности, были получены аналитические выражения для разностной схемы восьмого порядка. Также было показано, что представление разностных схем в виде каскада рекурсивных цифровых фильтров низких порядков позволяет организовать вычисления особым способом, который с точки зрения теории алгоритмической сложности превосходит известный метод Фурье.

Что касается нелинейных уравнений Шредингера, установлено, что применение схем высоких порядков не приводит к радикальному улучшению точности полученных решений. В связи с этим при решении нелинейного уравнения Шредингера рекомендуется использовать разностные схемы более низких порядков, которые при меньшей вычислительной сложности дают сравнимые по точности результаты.

Для дальнейших исследований представляет интерес оптимизация параметров разностных схем высоких порядков. В работах [4–5, 12–14] было неоднократно показано, что схемы с более низким порядком аппроксимации способны давать более точные решения при оптимальном наборе своих параметров.

Благодарности. Автор выражает благодарность профессору В. М. Волкову за внимание, проявленное к работе.

Acknowledgements. The author would like to thank Professor V. M. Volkov for attention to the paper.

Список использованных источников

1. Самарский, А. А. Теория разностных схем / А. А. Самарский. – М.: Наука, 1989. – 616 с.

2. Han, F. New higher-order compact finite difference schemes for 1D heat conduction equations / F. Han, W. Dai // Appl. Math. Modell. – 2013. – Vol. 37, № 16/17. – P. 7940–7952. https://doi.org/10.1016/j.apm.2013.03.026

3. Gordin, V. A. Compact differential schemes for the diffusion and Schrödinger equations. Approximation, stability, convergence, effectiveness, monotony / V. A. Gordin, E. A. Tsymbalov // J. Comput. Math. – 2014. – Vol. 32, № 3. – P. 348–370. https://doi.org/10.4208/jcm.1403-cr14

4. Волков, В. М. Спектральная согласованность разностных схем для уравнения теплопроводности / В. М. Волков, А. Н. Гуревский // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2017. – № 3. – С. 7–14.

5. Волков, В. М. Оптимизация компактных разностных схем спектрального разрешения для нестационарного уравнения Шредингера на основе методов цифровой обработки сигналов / В. М. Волков, А. Н. Гуревский, И. В. Жукова // Вестн. БГУ. – 2015. – № 3. – С. 84–89.

6. A time-domain optical transmission system simulation package accounting for nonlinear and polarization-related effects in fiber / A. Carena [et al.] // IEEE J. Selected Areas in Communications. – 1997. – Vol. 15, № 4. – P. 751–765. https://doi.org/10.1109/49.585785

7. Сергиенко, А. Г. Цифровая обработка сигналов / А. Г. Сергиенко. – СПб.: БХВ-Петербург, 2011. – 768 с.

8. Askar, S. S. On Solving Pentadiagonal Linear Systems via Transformations / S. S. Askar, A. A. Karawia // Math. Problems Eng. – 2015. – Vol. 2015. – P. 1–9. https://doi.org/10.1155/2015/232456

9. Ахманов, С. А. Оптика фемтосекундных лазерных импульсов / С. А. Ахманов, В. А. Выслоух, А. С. Чиркин. – М.: Наука, 1988. – 312 с.

10. Агравал, Г. Нелинейная волоконная оптика / Г. Агравал. – М.: Мир, 1996. – 323 с.

11. Optimization of the split-step Fourier method in modeling optical-fiber communications systems / O. V. Sinkin [et al.] // J. Lightwave Technol. – 2003. – Vol. 21, № 1. – P. 61–68. https://doi.org/10.1109/JLT.2003.808628

12. Волков, В. М. Оптимизация компактных разностных схем спектрального разрешения в методе дробных шагов для нелинейного уравнения Шредингера / В. М. Волков, А. Н. Гуревский // Вес. БГПУ. Сер. 3. – 2016. – № 4. – С. 11–17.

13. Lele, S. K. Compact finite difference schemes with spectral-like resolution / S. K. Lele // J. Comput. Phys. – 1992. – Vol. 103, № 1. – P. 16–42. https://doi.org/10.1016/0021-9991(92)90324-r

14. Гуревский, А. Н. Оптимизация спектральных характеристик разностных схем для нестационарного уравнения Шредингера / А. Н. Гуревский //Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2019. – Т. 55, № 1. – С. 62–68. https://doi.org/10.29235/1561-2430-2019-55-1-62-68

References

1. Samarskii A. A. *The theory of Finite Difference schemes*. Moscow, Nauka Publ., 1989. 616 p. (in Russian).

2. Han F., Dai W. New higher-order compact finite difference schemes for 1D heat conduction equations. *Applied Mathematical Modelling*, 2013, vol. 37, no. 16–17, pp. 7940–7952. https://doi.org/10.1016/j.apm.2013.03.026

3. Gordin V. A., Tsymbalov E. A. Compact differential schemes for the diffusion and Schrödinger equations. Approximation, stability, convergence, effectiveness, monotony. *Journal of Computational Mathematics*, 2014, vol. 32, no. 3, pp. 348–370. https://doi.org/10.4208/jcm.1403-cr14

4. Volkov V. M., Hureuski A. N. Spectpal-like resolution of finite-difference schemes for the heat conduction equation. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2017, no. 3, pp. 7–14 (in Russian).

5. Volkov V. M., Gurevskii A. N., Zhukova I. V. Optimization of compact finite difference schemes with spectral-like resolutiona for the non-stationary Schrodinger equation on the base of digital signal processing methods. *Vestnik BGU. Seriya 1. Fizika. Matematika. Informatika = Vestnik BSU. Series 1: Physics. Mathematics. Informatics*, 2015, no. 3, pp. 84–89 (in Russian).

6. Carena A., Curri V., Gaudino R., Poggiolini P., Benedetto S. A time-domain optical transmission system simulation package accounting for nonlinear and polarization-related effects in fiber. *IEEE Journal on Selected Areas in Communications*, 1997, vol. 15, no. 4, pp. 751–765. https://doi.org/10.1109/49.585785

7. Sergienko A. G. *Digital Signal Processing.* Saint Petersburg, BHV-Petersburg Publ., 2011. 768 p. (in Russian).

8. Askar S. S., Karawia A. A. On Solving Pentadiagonal Linear Systems via Transformations. *Mathematical Problems in Engineering,* 2015, vol. 2015, pp. 1–9. https://doi.org/10.1155/2015/232456

9. Akhmanov S. A., Vysloukh V. A., Chirkin A. S. *Optics of Femtosecond Laser Pulses*. Moscow, Nauka Publ., 1988. 312 p. (in Russian).

10. Agrawal G. *Nonlinear Fiber Optics*. 3rd ed. Academic Press, 2001. 481 p.

11. Sinkin O. V, Holzlohner R., Zweck J., Menyuk C. R. Optimization of the split-step Fourier method in modeling optical-fiber communications systems. *Journal of Lightwave Technology,* 2003, vol. 21, no. 1, pp. 61–68. https://doi.org/10.1109/JLT.2003.808628

12. Volkov V. M., Gurevskii A. N. Optimization of compact finite difference schemes with spectral-like resolutions in the split-step method for the nonlinear Schrödinger equation. *Vestsi BDPU. Seryya 3, Fizika. Matematyka. Infarmatyka. Biyalogiya. Geagrafiya* [Bulletin of BSPU. Series 3, Physics. Mathematics. Informatics. Biology. Geography], 2016, no. 4, pp. 11–17 (in Russian).

13. Lele S. K. Compact finite difference schemes with spectral-like resolution. *Journal of Computational Physics*, 1992, vol. 103, no. 1, pp. 16–42. https://doi.org/10.1016/0021-9991(92)90324-r

14. Hureuski A. N. Optimizing the spectral characteristics of the finite-difference schemes for the unsteady Schrödinger equation. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2019, vol. 55, no. 1, pp. 62–68 (in Russian). https://doi.org/10.29235/1561-2430-2019-55-1-62-68

Информация об авторе

Гуревский Алексей Николаевич – старший преподаватель кафедры веб-технологий и компьютерного моделирования, Белорусский государственный университет (пр. Независимости, 4, 220030, г. Минск, Республика Беларусь). E-mail: gurevski@bsu.by

Information about the author

Aliaksei N. Hureuski – Senior Lecturer of the Department Web-Technologies and Computer Modeling, Belarusian State University (4, Nezavisimosti Ave., 220030, Minsk, Republic of Belarus). E-mail: gurevski@bsu.by

ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)
УДК 519.67
https://doi.org/10.29235/1561-2430-2019-55-4-425-434

Поступила в редакцию 23.08.2019
Received 23.08.2019

А. А. Згировский, Н. А. Лиходед

Белорусский государственный университет, Минск, Беларусь

МОДИФИЦИРОВАННЫЙ МЕТОД ПАРАЛЛЕЛЬНОЙ МАТРИЧНОЙ ПРОГОНКИ

Аннотация. Тематика работы относится к области построения параллельных алгоритмов численного решения блочно-трехдиагональных систем линейных алгебраических уравнений. Такие системы часто возникают в приложениях и для ряда задач требуют использования высокопроизводительных многоядерных вычислительных систем. Один из широко применяемых на практике подходов к решению блочно-трехдиагональных систем заключается в использовании оригинальных алгоритмов параллельной матричной прогонки. В настоящей статье рассмотрен метод параллельной матричной прогонки, основанный на разбиении матрицы. Этот метод трехфазный: сначала исходная система разбивается на части и после независимых преобразований каждой из них составляется редуцированная блочно-трехдиагональная система, затем из этой системы находят несколько неизвестных каждой части уравнений, после чего независимо вычисляются остальные неизвестные каждой части. Предложена новая модификация метода; обосновано, что если для исходной системы уравнений справедливы известные (и часто выполненные на практике) условия устойчивости метода матричной прогонки, то вычисления разработанной модификации параллельной матричной прогонки являются устойчивыми.

Ключевые слова: параллельные вычисления, матричная прогонка блочно-трехдиагональных линейных систем, устойчивость параллельной матричной прогонки

Для цитирования. Згировский, А. А. Модифицированный метод параллельной матричной прогонки / А. А. Згировский, Н. А. Лиходед // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2019. – Т. 55, № 4. – С. 425–434. https://doi.org/10.29235/1561-2430-2019-55-4-425-434

A. A. Zgirouski, N. A. Likhoded

Belarusian State University, Minsk, Belarus

MODIFIED METHOD OF PARALLEL MATRIX SWEEP

Abstract. The topic of this paper refers to efficient parallel solvers of block-tridiagonal linear systems of equations. Such systems occur in numerous modeling problems and require usage of high-performance multicore computation systems. One of the widely used methods for solving block-tridiagonal linear systems in parallel is the original block-tridiagonal sweep method. We consider the algorithm based on the partitioning idea. Firstly, the initial matrix is split into parts and transformations are applied to each part independently to obtain equations of a reduced block-tridiagonal system. Secondly, the reduced system is solved sequentially using the classic Thomas algorithm. Finally, all the parts are solved in parallel using the solutions of a reduced system. We propose a modification of this method. It was justified that if known stability conditions for the matrix sweep method are satisfied, then the proposed modification is stable as well.

Keywords: parallel computations, block-tridiagonal linear systems matrix sweep, stability of parallel matrix sweep

For citation. Zgirouski A. A., Likhoded N. A. Modified method of parallel matrix sweep. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2019, vol. 55, no. 4, pp. 425–434 (in Russian). https://doi.org/10.29235/1561-2430-2019-55-4-425-434

Введение. Численные методы решения уравнений в частных производных часто приводят к системам линейных алгебраических уравнений (СЛАУ) с блочно-трехдиагональными матрицами. Одним из широко применяемых на практике для решения таких систем является метод матричной прогонки [1]. Размерность систем и их блочных коэффициентов для ряда задач не позволяет эффективно решать эти задачи с использованием последовательных алгоритмов. Актуальной становится разработка параллельных алгоритмов матричной прогонки для реализации на многоядерных вычислительных системах.

Один подход к реализации матричной прогонки на параллельном компьютере – организовать параллельное выполнение операций алгоритмов перемножения матриц и решения СЛАУ (многократно используемых в матричной прогонке). Если матрицы-блоки плотные и достаточно большие, то такой параллелизм можно эффективно использовать.

Другой подход заключается в использовании специально построенных алгоритмов параллельной матричной прогонки [2–6]. Эти алгоритмы обобщают на случай блочно-трехдиагональных матриц идеи скалярных алгоритмов параллельной матричной прогонки, основанных на разбиении матрицы [7, 8] или на циклической редукции [9]. На блочный случай обобщают скалярные алгоритмы параллельной матричной прогонки, хорошо зарекомендовавшие себя на практике.

Целью настоящей работы является обобщение на блочный случай одного из таких параллельных скалярных алгоритмов [10], основанного на разбиении матрицы. Методы, основанные на разбиении матрицы, являются трехфазными. На первой фазе исходная система разбивается на части и после независимых преобразований каждой из них составляется редуцированная система (блочно-трехдиагональная в блочном случае). На второй – из редуцированной системы находят несколько неизвестных каждой части уравнений. На третьей фазе независимо вычисляются остальные неизвестные каждой части. В данной работе предложена новая модификация параллельной матричной прогонки. Обосновано, что если для исходной системы уравнений справедливы известные (и часто выполненные на практике) условия устойчивости метода матричной прогонки, то вычисления разработанной модификации параллельной матричной прогонки являются устойчивыми.

Матричная прогонка. Пусть $A_1,...,A_N$, $B_0,B_1,..., B_{N-1}$, C_0, $C_1,...,C_N$ – $M{\times}M$-матрицы, $Y_0,Y_1,...,Y_N$, $F_0,F_1,...,F_N$ – M-мерные векторы. Рассмотрим блочно-трехдиагональную систему линейных алгебраических уравнений вида

$$\begin{array}{rr} C_0Y_0 - B_0Y_1 & = F_0, \\ -A_1Y_0 + C_1Y_1 - B_1Y_2 & = F_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ -A_iY_{i-1} + C_iY_i - B_iY_{i+1} & = F_i, \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ & -A_NY_{N-1} + C_NY_N = F_N, \end{array} \tag{1}$$

или, если записывать отдельно по уравнениям,

$$\begin{gathered} C_0Y_0 - B_0Y_1 = F_0, \\ -A_iY_{i-1} + C_iY_i - B_iY_{i+1} = F_i, \quad i = 1,..., N{-}1, \\ -A_iY_{N-1} + C_NY_N = F_N. \end{gathered} \tag{1'}$$

Приведем формулы матричной прогонки для решения системы (1) (см., напр., [1]):

– прямая прогонка – вычисление матриц α_i и векторов β_i по формулам

$$\begin{gathered} \alpha_1 = C_0^{-1}B_0, \quad \beta_1 = C_0^{-1}f_0, \\ \alpha_{i+1} = (C_i - A_i\alpha_i)^{-1}B_i, \;\; \beta_{i+1} = (C_i - A_i\alpha_i)^{-1}(F_i + A_i\beta_i), \; i = 1,2,...,N-1, \\ \beta_{i+1} = (C_i - A_i\alpha_i)^{-1}(F_i + A_i\beta_i); \end{gathered} \tag{2}$$

– обратная прогонка – вычисление решения (вычисление векторов Y_i) по формулам

$$Y_i = \beta_{i+1}, \quad Y_i = \alpha_{i+1}Y_{i+1} + \beta_{i+1}, \quad i = N-1,...,1,0. \tag{3}$$

Говорят, что алгоритм метода матричной прогонки устойчив, если выполнено условие $\|\alpha_i\| \le 1$ для $i = 1,2,...,N$ (предполагается, что в M-мерном пространстве введена какая-либо норма). Имеет место следующее утверждение [1]: если C_i, $i = 0,1,...,N,$ – невырожденные матрицы, а $A_1,...,A_N$, B_0, $B_1,...,B_{N-1}$ – ненулевые матрицы и выполнены условия

$$\left\|C_i^{-1}A_i\right\| + \left\|C_i^{-1}B_i\right\| \le 1, \quad i = 1,2,...,N-1, \tag{4}$$

$$\left\|C_0^{-1}B_0\right\| \le 1, \quad \left\|C_N^{-1}A_N\right\| \le 1, \tag{5}$$

причем хотя бы в одном из условий (4), (5) выполняется строгое неравенство, то алгоритм (2), (3) метода матричной прогонки корректен и устойчив. Под корректностью алгоритма понимается существование обратных матриц в формулах (2).

Параллельная матричная прогонка, основанная на разбиении блочной трехдиагональной матрицы. Пусть дана система – вида (1). Разобьем систему, состоящую из $N + 1$ матрично-векторных уравнений, на K частей, т. е. на K блоков уравнений, где K – параметр алгоритма. Для простоты изложения будем считать, что $N + 1$ нацело делится на K:

$$\frac{N+1}{K} = l,$$

где l – целое число. В k-й блок уравнений, $k = 0,1,\ldots,K - 1$, входят l матрично-векторных уравнений, начиная с уравнения $k \cdot l$ и по уравнение $(k + 1)l - 1$ включительно. Обозначим $s_k = k \cdot l$ через номер первого уравнения в k-м блоке, $f_k = (k + 1)l - 1$ – номер последнего уравнения в k-м блоке. Заметим, что $f_k = s_k + l - 1$, $s_k - 1 = f_{k-1}$, $s_{k+1} - 1 = f_k$. Идея алгоритма заключается в том, чтобы сначала найти граничные неизвестные. Тогда, решая параллельно (независимо друг от друга) блоки уравнений, можно найти оставшиеся неизвестные.

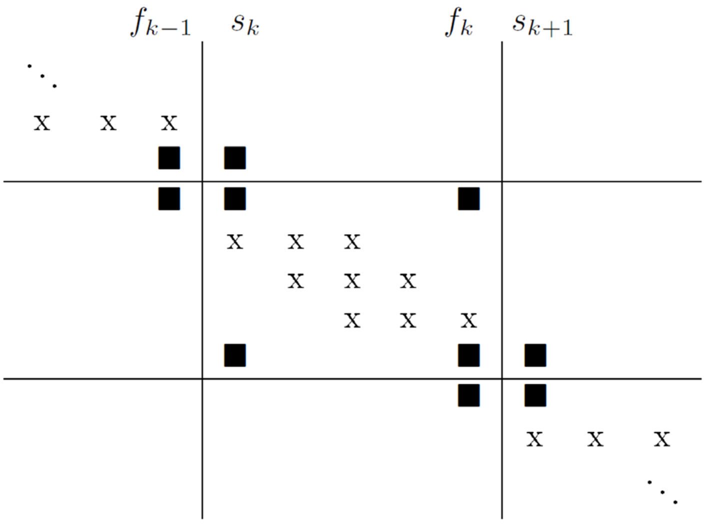
Независимо решить блоки уравнений не представляется возможным, так как первое и последнее уравнения каждого блока содержат переменные, которые входят в соседние блоки (будем называть такие переменные граничными). Известное решение проблемы состоит в том, чтобы считать граничные неизвестные параметрами системы одного блока. А далее в каждом блоке параллельно применить метод матричной прогонки для системы с параметрами. Тогда можно выразить неизвестные исходные системы через значения граничных переменных. Кроме того, можно получить систему блочно-трехдиагональных уравнений относительно граничных неизвестных. Тогда, решив эту систему (последовательно или параллельно), можно параллельно и независимо найти значения неизвестных в каждом блоке.

Таким образом, алгоритм параллельной матричной прогонки состоит из трех фаз. В первой фазе для каждого из K блоков вычисляются коэффициенты матричной прогонки, полагая переменные Y_{s_k} параметрическими. В результате прогонки для каждого неизвестного внутри блока получается выражение относительно граничных неизвестных, а также уравнение, относительно самих граничных неизвестных. Эти уравнения образуют редуцированную блочно-трехдиагональную систему. Во второй фазе указанная система решается последовательным алгоритмом матричной прогонки. После чего в третьей фазе, используя найденные в первой фазе неизвестные и коэффициенты прогонки, независимо вычисляются неизвестные каждого блока.

Модифицированная параллельная матричная прогонка. Отличием предлагаемой модифицированной версии от известного параллельного алгоритма матричной прогонки является отсутствие необходимости хранить после первой фазы для каждого из K блоков новые блочные коэффициенты прогонки для каждого неизвестного.

Пусть произведено разбиение системы вида (1) на K блоков уравнений. В каждом из блоков, кроме нулевого, в первом уравнении есть ненулевые, вообще говоря, коэффициенты при неизвестных с индексами $s_k - 1$ и s_k; эти же неизвестные входят также в уравнения предыдущего блока. Кроме того, в каждом блоке, кроме (K–1)-го, в последнем уравнении есть ненулевые коэффициенты при неизвестных с индексами $s_k + l - 1$ и $s_k + l$, которые входят также в уравнения следующего блока. Это не позволяет решать блоки уравнений независимо.

Неизвестные с индексами $s_k - 1$, s_k и $s_k + l - 1$, $s_k + l$ являются граничными. Всего имеется $2K$ граничных неизвестных. Рассмотрим способ получения редуцированной системы, в которую войдут $2K$ уравнений относительно граничных неизвестных. Для этого в каждом блоке первое уравнение относительно неизвестных с индексами $s_k - 1$, s_k, $s_k + 1$ сведем строчными преобразованиями к уравнению относительно граничных неизвестных с индексами $s_k - 1$, s_k, $s_k + l - 1$,



Схематичное изображение ненулевых блоков матрицы системы уравнений после получения верхних и нижних уравнений; указаны индексы граничных неизвестных k-го блока уравнений

A schematic representation of non-zero blocks of the matrix of the system of equations after obtaining the lower and upper equations; the indexes of the boundary variables of the kth block are indicated

а последнее уравнение относительно неизвестных с индексами $s_k + l - 2$, $s_k + l - 1$, $s_k + l$ сведем к уравнению относительно граничных неизвестных s_k, $s_k + l - 1$, $s_k + l$.

Полученное после преобразований уравнение относительно неизвестных с индексами $s_k - 1$, s_k, $s_k + l - 1$ назовем верхним; уравнение относительно неизвестных s_k, $s_k + l - 1$, $s_k + l$ назовем нижним. На рисунке указаны коэффициенты при граничных неизвестных верхних и нижних уравнений.

Полученная редуцированная система является трехдиагональной. В самом деле, обозначим $Z_{2k-1} = Y_{kl-1}$, $Z_{2k} = Y_{kl}$. Тогда, с учетом $s_k = k \cdot l$, верхние уравнения блоков относительно неизвестных Y_{sk-1}, Y_{s_k}, Y_{s_k+l-1} станут уравнениями относительно Z_{2k-1}, Z_{2k}, Z_{2k+1}, а нижние уравнения относительно Y_{s_k}, Y_{s_k+l-1}, Y_{s_k+l} – уравнениями относительно Z_{2k}, Z_{2k+1}, Z_{2k+2}.

Отметим, что после подстановки граничных неизвестных каждый блок уравнений становится независимой трехдиагональной системой $l - 2$ уравнений относительно $l - 2$ неизвестных.

Рассмотрим подробно первую фазу. Сначала исследуем процесс получения нижних уравнений в k-м блоке уравнений. Текущие коэффициенты-блоки будем хранить в массивах L^A, L^C, L^B, L^F. Получим уравнение относительно неизвестных с индексами s_k, $s_k + l - 1$, $s_k + l$. Для этого понадобится $l - 2$ шага: преобразовать потребуется все, начиная с третьего уравнения блока. Пронумеруем эти шаги числами с $s_k + 2$ по f_k (т. е. с $k \cdot l + 2$ по $k \cdot l + l - 1$).

Второе уравнение в блоке имеет ненулевые коэффициенты при неизвестных с индексами s_k, $s_k + 1$, $s_k + 2$. Оно не преобразуется, но в преобразованиях участвует. На шаге r, $r = s_k + 2,\dots,f_k$, будем преобразовывать уравнение относительно неизвестных с индексами s_k, r, $r + 1$. Запишем коэффициенты-блоки уже преобразованного на предыдущем шаге и этого уравнения:

$$\left.\begin{array}{cccccccc} L^A & 0 & \dots & 0 & L^C & L^B & 0 & \dots \\ 0 & 0 & \dots & 0 & -A_r & C_r & -B_r & \end{array}\right| \begin{array}{c} L^F \\ F_r \end{array}.$$

При начальном значении $r = s_k + 2$ записываются второе и третье уравнения в блоке, причем

$$L^A = -A_{sk+1},\ \ L^C = C_{sk+1},\ \ L^B = -B_{sk+1},\ \ L^F = F_{sk+1}.$$

Выполним преобразования, приводящие во втором уравнении к неизвестным с индексами s_k, $r + 1$, $r + 2$. Сначала умножим первую строку слева на $A_r(L^C)^{-1}$:

$$\cdots\left.\begin{matrix} A_r(L^C)^{-1}L^A & 0 & \ldots & 0 & A_r & A_r(L^C)^{-1}L^B & 0 \\ 0 & 0 & \ldots & 0 & -A_r & C_r & -B_r \end{matrix}\right| \begin{matrix} A_r(L^C)^{-1}L^F \\ F_r \end{matrix}.$$

Затем прибавим ко второй строке первую строку:

$$\cdots\left.\begin{matrix} A_r(L^C)^{-1}L^A & 0 & \ldots & 0 & A_r & A_r(L^C)^{-1}L^B & 0 \\ A_r(L^C)^{-1}L^A & 0 & \ldots & 0 & 0 & C_r + A_r(L^C)^{-1}L^B & -B_r \end{matrix}\right| \begin{matrix} A_r(L^C)^{-1}L^F \\ F_r + A_r(L^C)^{-1}L^F \end{matrix}.$$

Второе уравнение имеет ненулевые коэффициенты при неизвестных с индексами s_k, $r + 1$, $r + 2$.

После шага $r = f_k$ последнее уравнение в блоке становится уравнением относительно неизвестных с индексами s_k, $s_k + l - 1$, $s_k + l$.

Теперь рассмотрим процесс получения верхних уравнений в k-м блоке. Текущие коэффициенты-блоки будем хранить в массивах U^A, U^C, U^B, U^F. Для получения уравнения относительно неизвестных с индексами $s_k - 1$, s_k, $s_k + l - 1$ нужно $l - 2$ шага: преобразовать потребуется все, начиная с третьего снизу и заканчивая первым уравнением блока.

Предпоследнее уравнение в блоке имеет ненулевые коэффициенты при неизвестных с индексами $s_k + l - 3$, $s_k + l - 2$, $s_k + l - 1$. Оно не преобразуется, но в преобразованиях участвует. На шаге r, $r = f_k - 2, f_k - 3, \ldots, s_k$, имеется уравнение относительно неизвестных с индексами $r - 1$, r, $s_k + l - 1$. Запишем коэффициенты-блоки уже преобразованного на предыдущем шаге и этого уравнения:

$$\cdots\left.\begin{matrix} -A_r & C_r & -B_r & 0 & \ldots & 0 & 0 \\ 0 & U^A & U^C & 0 & \ldots & 0 & U^B \end{matrix}\right| \begin{matrix} F_r \\ U^F \end{matrix}.$$

При начальном значении $r = f_k - 1$ записываются предпоследнее и третье снизу уравнения в блоке, в которых полагается

$$U^A = -A_{f_k - 1}, \;\; U^C = C_{f_k - 1}, \;\; U^B = -B_{f_k - 1}, \;\; U^F = F_{f_k - 1}.$$

Выполним преобразования, приводящие в первом уравнении к неизвестным с индексами $r - 2$, $r - 1$, $s_k + l - 1$. Умножим вторую строку слева на $B_r(U^C)^{-1}$, а затем прибавим к первой строке вторую:

$$\cdots\left.\begin{matrix} -A_r & C_r + B_r B_r(U^C)^{-1}U^A & 0 & 0 & \ldots & 0 & B_r B_r(U^C)^{-1}U^B \\ 0 & B_r B_r(U^C)^{-1}U^A & B_r & 0 & \ldots & 0 & B_r B_r(U^C)^{-1}U^B \end{matrix}\right| \begin{matrix} F_r + B_r B_r(U^C)^{-1}U^F \\ B_r B_r(U^C)^{-1}U^F \end{matrix}.$$

Первое уравнение имеет ненулевые коэффициенты при неизвестных с индексами $r - 2$, $r - 1$, $s_k + l - 1$.

После всех преобразований в k-м блоке получены 2 уравнения для редуцированной системы:

$$U^A Z_{2k-1} + U^C Z_{2k} + U^B Z_{2k+1} = U^F,$$

$$L^A Z_{2k} + L^C Z_{2k+1} + L^B Z_{2k+2} = L^F.$$

Здесь для нулевого блока отсутствует слагаемое с U^A, для (K–1)-го блока отсутствует слагаемое с L^B, для всех блоков $L^B = -B_{f_k}$, $U^A = -A_{s_k}$.

Таким образом, получение коэффициентов редуцированной системы уравнений можно представить следующим алгоритмом:

dopar $k = 0, K - 1$

  $s_k = k \cdot l$, $f_k = (k + 1)l - 1$

  \\ Получение коэффициентов нижних уравнений:

$L^A = -A_{s_k+1},\ L^C = C_{s_k+1},\ L^B = -B_{s_k+1},\ L^F = F_{s_k+1}$
do $r = s_k + 2, f_k$
$T^A = A_r (L^C)^{-1}$
$L^A = T^A L^A$
$L^C = C_r + T^A L^B$
$L^B = -B_r$ \\ L^B – нулевая матрица при $r = K \cdot l - 1$
$L^F = F_r + T^A L^F$
enddo
$L_k^A = L^A,\ L_k^C = L^C,\ L_k^B = L^B,\ L_k^F = L^F$
\\ Получение коэффициентов верхних уравнений:
$U^A = -A_{f_k-1},\ U^C = C_{f_k-1},\ U^B = -B_{f_k-1},\ U^F = F_{f_k-1}$
do $r = f_k - 2, s_k$ \\ Цикл с шагом –1
$T^B = B_r (U^C)^{-1}$
$U^A = -A_r$ \\ U^A – нулевая матрица при $r = 0$
$U^C = C_r + T^B U_r^A$
$U^B = T^B U^B$
$U^F = F_r + T^B U^F$
enddo
$U_k^A = U^A,\ U_k^B = U^B,\ U_k^C = U^C,\ U_k^F = U^F$
enddopar

Заметим, что в силу блочно-трехдиагональной структуры исходной системы первый блочный коэффициент первого уравнения A_0 и последний блочный коэффициент последнего уравнения B_N отсутствуют. Аналогично для редуцированной системы отсутствуют первый и последний блочные коэффициенты U_0^A и L_{K-1}^B. При программной реализации удобно хранить произвольные, например нулевые, значения.

На этапе второй фазы решается полученная относительно $Z_0, Z_1, \ldots, Z_{2K-1}$ блочно-трехдиагональная система алгоритмом правой матричной прогонки.

На третьей фазе алгоритма для каждого блока уравнений k, $k = 0, 1, \ldots, K - 1$, после второй фазы найдены граничные неизвестные:

$$Y_{s_k-1} = Z_{2k-1},\ \ Y_{s_k} = Z_{2k},\ \ Y_{s_k+l-1} = Z_{2k+1},\ \ Y_{s_k+l} = Z_{2k+2}.$$

Эти неизвестные входят в первое, второе, предпоследнее и последнее уравнения каждого блока.

Из первого уравнения можно найти неизвестный вектор Y_{s_k+1}, так как уже найдены неизвестные Y_{s_k-1} и Y_{s_k}.

Перенесем во втором уравнении блока (это уравнение с номером $s_k + 1$ исходной системы) известное произведение $(-A_{s_k} Y_{s_k+1})$ в правую часть. Перенесем также в предпоследнем уравнении блока (это уравнение с номером $f_k - 1$) в правую часть известное произведение $(-B_{f_k-1} Y_{f_k})$. Эту систему относительно неизвестных $Y_{s_k+1}, Y_{s_k+2}, \ldots, Y_{s_k+l-2}$ $(Y_{s_k+l-2} = Y_{f_k-1})$ можно решить независимо для каждого из блоков уравнений алгоритмом правой матричной прогонки.

Устойчивость модифицированной параллельной матричной прогонки. На второй и третьей фазах алгоритма выполняется правая матричная прогонка. Если для исходной блочно-трехдиагональной системы выполнены условия устойчивости метода матричной прогонки (4), (5), то они выполняются и для каждой из систем в третьей фазе, так как коэффициенты систем есть коэффициенты исходной системы. Оказывается, аналогичное утверждение справедливо и для второй фазы алгоритма. Предполагается неравенство нулю матриц, встречающихся при выполнении первой фазы алгоритма и, где необходимо, существование обратных матриц.

Т е о р е м а. *Если для исходной блочно-трехдиагональной системы выполнены условия устойчивости метода матричной прогонки* (4), (5), *то они выполняются и для редуцированной системы.*

Д о к а з а т е л ь с т в о. Зафиксируем любой блок уравнений и применим метод математической индукции. Сначала рассмотрим процесс получения коэффициентов нижнего редуцированного уравнения.

Пусть на какой-либо итерации первой фазы алгоритма коэффициенты L^A, L^C, L^B (в итоге входящие в редуцированную систему) удовлетворяют условию

$$\left\|(L^C)^{-1}L^A\right\|+\left\|(L^C)^{-1}L^B\right\|\le 1, \tag{6}$$

причем хотя бы для одного из блоков неравенство (6) является строгим. Для $(K-1)$-го блока отсутствует слагаемое с L^B. Для начальной итерации неравенство (6) справедливо, так как в этом случае L^A, L^C, L^B есть коэффициенты исходной системы.

На каждой из $f_k - s_k - 1$ итераций r происходит вычисление новых значений коэффициентов L^A, L^C, L^B по формулам

$$\overline{L}^A = A_r(L^C)^{-1}L^A,\ \ \overline{L}^C = C_r + A_r(L^C)^{-1}L^B,\ \ \overline{L}^B = -B_r.$$

Требуется показать справедливость неравенства (строгого хотя бы для одного из блоков)

$$\left\|(\overline{L}^C)^{-1}\overline{L}^A\right\|+\left\|(\overline{L}^C)^{-1}\overline{L}^B\right\|\le 1.$$

Имеем

$$\left\|(\overline{L}^C)^{-1}\overline{L}^A\right\|+\left\|(\overline{L}^C)^{-1}\overline{L}^B\right\|=\left\|\left(C_r+A_r(L^C)^{-1}L^B\right)^{-1}A_r(L^C)^{-1}L^A\right\|+\left\|\left(C_r+A_r(L^C)^{-1}L^B\right)^{-1}B_r\right\|=$$

$$=\left\|\left(E+C_r^{-1}A_r(L^C)^{-1}L^B\right)^{-1}C_r^{-1}A_r(L^C)^{-1}L^A\right\|+\left\|\left(E+C_r^{-1}A_r(L^C)^{-1}L^B\right)^{-1}C_r^{-1}B_r\right\|\le$$

$$\le\left\|\left(E+C_r^{-1}A_r(L^C)^{-1}L^B\right)^{-1}\right\|\left\|C_r^{-1}A_r(L^C)^{-1}L^A\right\|+\left\|\left(E+C_r^{-1}A_r(L^C)^{-1}L^B\right)^{-1}\right\|\left\|C_r^{-1}B_r\right\|=$$

$$=\left\|\left(E+C_r^{-1}A_r(L^C)^{-1}L^B\right)^{-1}\right\|\left(\left\|C_r^{-1}A_r(L^C)^{-1}L^A\right\|+\left\|C_r^{-1}B_r\right\|\right),$$

где E – единичная матрица. Таким образом,

$$\left\|(\overline{L}^C)^{-1}\overline{L}^A\right\|+\left\|(\overline{L}^C)^{-1}\overline{L}^B\right\|\le\left\|\left(E+C_r^{-1}A_r(L^C)^{-1}L^B\right)^{-1}\right\|\left(\left\|C_r^{-1}A_r\right\|\left\|(L^C)^{-1}L^A\right\|+\left\|C_r^{-1}B_r\right\|\right). \tag{7}$$

Воспользуемся известным утверждением: если для квадратной матрицы S имеет место оценка $\|S\| < 1$, то существует обратная к $E - S$ матрица, причем $\|(E - S)^{-1}\| \le 1/(1 - \|S\|)$. Положим

$$S=-C_r^{-1}A_r(L^C)^{-1}L^B.$$

Тогда с учетом неравенств (4), (6) и неравенства нулю матрицы $(L^C)^{-1}L^A$ получим

$$\|S\|=\left\|-C_r^{-1}A_r(L^C)^{-1}L^B\right\|\le\left\|C_r^{-1}A_r\right\|\left\|(L^C)^{-1}L^B\right\|\le\left\|C_r^{-1}A_r\right\|\left(1-\left\|(L^C)^{-1}L^A\right\|\right)<1,$$

$$\left\|\left(E-S\right)^{-1}\right\|=\left\|\left(E+C_r^{-1}A_r(L^C)^{-1}L^B\right)^{-1}\right\|\le\frac{1}{1-\|S\|}\le\frac{1}{1-\left\|C_r^{-1}A_r\right\|\left\|(L^C)^{-1}L^B\right\|}\le$$

$$\leq \frac{1}{1-\left\|C_r^{-1}A_r\right\|\left(1-\left\|(L^C)^{-1}L^A\right\|\right)},$$

причем последнее неравенство строгое, если строгим является неравенство (6). Следовательно,

$$\left\|\left(E+C_r^{-1}A_r(L^C)^{-1}L^B\right)^{-1}\right\| \leq \frac{1}{1-\left\|C_r^{-1}A_r\right\|+\left\|C_r^{-1}A_r\right\|\left\|(L^C)^{-1}L^A\right\|} \leq \frac{1}{\left\|C_r^{-1}B_r\right\|+\left\|C_r^{-1}A_r\right\|\left\|(L^C)^{-1}L^A\right\|}.$$

Вернемся к неравенству (7):

$$\left\|(\overline{L}^C)^{-1}\overline{L}^A\right\|+\left\|(\overline{L}^C)^{-1}\overline{L}^B\right\| \leq \frac{1}{\left\|C_r^{-1}B_r\right\|+\left\|C_r^{-1}A_r\right\|\left\|(L^C)^{-1}L^A\right\|}\left(\left\|C_r^{-1}A_r\right\|\left\|(L^C)^{-1}L^A\right\|+\left\|C_r^{-1}B_r\right\|\right)=1.$$

Здесь неравенство строгое, если строгим является неравенство (6).

По аналогии со случаем получения коэффициентов нижнего редуцированного уравнения рассмотрим основные этапы доказательства для случая коэффициентов верхнего уравнения. Пусть справедливо неравенство

$$\left\|(U^C)^{-1}U^A\right\|+\left\|(U^C)^{-1}U^B\right\| \leq 1$$

и новые значения коэффициентов вычисляются по формулам

$$\overline{U}^A=-A_r,\ \ \overline{U}^C=C_r+B_r(U^C)^{-1}U^A,\ \ \overline{U}^B=B_r(U^C)^{-1}U^B.$$

Имеем:

$$\left\|(U^C)^{-1}U^A\right\|+\left\|(U^C)^{-1}U^B\right\|=\left\|\left(C_r+B_r(U^C)^{-1}U^A\right)^{-1}A_r\right\|+\left\|\left(C_r+B_r(U^C)^{-1}U^A\right)^{-1}B_r(U^C)^{-1}U^B\right\| \leq$$

$$\leq \left\|\left(E+C_r^{-1}B_r(U^C)^{-1}U^A\right)^{-1}\right\|\left\|C_r^{-1}A_r\right\|+\left\|\left(E+C_r^{-1}B_r(U^C)^{-1}U^A\right)^{-1}\right\|\left\|C_r^{-1}B_r(U^C)^{-1}U^B\right\| \leq$$

$$\leq \left\|\left(E+C_r^{-1}B_r(U^C)^{-1}U^A\right)^{-1}\right\|\left(\left\|C_r^{-1}A_r\right\|+\left\|C_r^{-1}B_r\right\|\left\|(U^C)^{-1}U^B\right\|\right).$$

Положим $S=-C_r^{-1}B_r(U^C)^{-1}U^A,$ тогда

$$\|S\| \leq \left\|C_r^{-1}B_r\right\|\left\|(U^C)^{-1}U^A\right\| \leq \left\|C_r^{-1}B_r\right\|\left(1-\left\|(U^C)^{-1}U^B\right\|\right) \leq 1,$$

$$\left\|(E-S)^{-1}\right\|=\left\|\left(E+C_r^{-1}B_r(U^C)^{-1}U^A\right)^{-1}\right\| \leq \frac{1}{1-\|S\|} \leq \frac{1}{1-\left\|C_r^{-1}B_r\right\|\left(1-\left\|(U^C)^{-1}U^B\right\|\right)} \leq$$

$$\leq \frac{1}{\left\|C_r^{-1}A_r\right\|+\left\|C_r^{-1}B_r\right\|\left\|(U^C)^{-1}U^B\right\|},$$

$$\left\|(U^C)^{-1}U^A\right\|+\left\|\left\|(U^C)^{-1}U^A\right\|^{-1}U^B\right\| \leq \frac{1}{\left\|C_r^{-1}A_r\right\|+\left\|C_r^{-1}B_r\right\|\left\|(U^C)^{-1}U^B\right\|}\times$$

$$\times\left(\left\|C_r^{-1}A_r\right\|+\left\|C_r^{-1}B_r\right\|\left\|(U^C)^{-1}U^B\right\|\right)=1.$$

Теорема доказана.

На первой фазе алгоритма выполняются вычисления типа исключений Гаусса, в которых в качестве «ведущих элементов» берутся матрицы U^C, L^C. Вследствие этого для устойчивости вычислений первой фазы важно, чтобы норма матрицы L^C превосходила нормы матриц L^A и L^B, а норма матрицы U^C превосходила нормы матриц U^A и U^B. Справедливо следующее утверждение.

С л е д с т в и е. *Если для исходной блочно-трехдиагональной системы выполнены условия устойчивости метода матричной прогонки* (4)*,* (5)*, то на любой итерации r первой фазы имеет место*

$$\left\|U^C\right\| \geq \left\|U^A\right\| + \left\|U^B\right\|, \quad \left\|L^C\right\| \geq \left\|L^A\right\| + \left\|L^B\right\|.$$

Действительно, рассмотрим, например, случай нижних уравнений. Так как

$$\left\|(L^C)^{-1}L^A\right\| + \left\|(L^C)^{-1}L^B\right\| \leq 1,$$

то после умножения левой и правой частей этого неравенства на $\|L^C\|$ и использования свойства мультипликативности матричной нормы ($\|A \times B\| \leq \|A\| \times \|B\|$) получим

$$\left\|L^C\right\|\left\|(L^C)^{-1}L^A\right\| + \left\|L^C\right\|\left\|\left(L^C\right)^{-1}L^B\right\| \leq \left\|L^C\right\|,$$

$$\left\|L^C(L^C)^{-1}L^A\right\| + \left\|L^C(L^C)^{-1}L^B\right\| \leq \left\|L^C\right\|,$$

$$\left\|L^A\right\| + \left\|L^B\right\| \leq \left\|L^C\right\|.$$

Таким образом, нами разработан и исследован с точки зрения устойчивости новый вариант метода параллельной матричной прогонки для решения систем линейных алгебраических уравнений с блочно-трехдиагональными матрицами.

Благодарности. Работа выполнена в рамках государственной программы научных исследований Республики Беларусь «Конвергенция-2020», подпрограмма «Методы математического моделирования сложных систем».

Acknowledgments. The prepared study was sponsored by the Government Program of Scientific Research of the Republic of Belarus "Convergence-2020", the subprogram "Methods of Mathematical Modeling of Complex Systems".

Список использованных источников

1. Самарский, А. А. Методы решения сеточных уравнений / А. А. Самарский, Е. С. Николаев. – М.: Наука, 1978. – 592 с.

2. Heller, D. Some aspects of the cyclic reduction algorithm for block tridiagonal linear systems / D. Heller // SIAM J. Numer. Anal. – 1976. – Vol. 13, №. 4. – P. 484–496. https://doi.org/10.1137/0713042

3. Акимова, Е. Н. Распараллеливание алгоритма матричной прогонки / Е. Н. Акимова // Математическое моделирование. – 1994. – Т. 6, № 9. – С. 61–67.

4. Акимова, Е. Н. Параллельные алгоритмы решения СЛАУ с блочно-трехдиагональными матрицами на многопроцессорных вычислителях / Е. Н. Акимова, Д. В. Белоусов // Вестн. УГАТУ. – 2011. – Т. 15, № 5. – С. 87–93.

5. BCYCLIC: A parallel block tridiagonal matrix cyclic solver / S. P. Hirshman [et al.] // J. Comput. Phys. – 2010. – Vol. 229, №. 18. – P. 6392–6404. https://doi.org/10.1016/j.jcp.2010.04.049

6. Davina, A. Lamas. MPI-CUDA parallel linear solvers for block-tridiagonal matrices in the context of SLEPc'seigensolvers/ A. Lamas Davina, J. E. Roman // Parallel Comput. – 2018. – Vol. 74. – P. 118–135. https://doi.org/10.1016/j.parco.2017.11.006

7. Об организации параллельных вычислений и «распараллеливании» прогонки / Н. Н. Яненко [и др.] // Численные методы механики сплошной среды. – 1978. – Т. 9, № 7. – С. 139–146.

8. Wang, H. H. A parallel method for tridiagonal equations / H. H. Wang / ACM Trans. Math.Software. – 1981. – Vol. 7, № 2. – P. 170–183. https://doi.org/10.1145/355945.355947

9. Buzbee, B. L. On direct methods for solving Poisson's equations / B. L. Buzbee, G. H. Golub, C. W. Nielson // SIAM J. Numer. Anal. – 1970. – Vol. 7, №. 4. – P. 627–656. https://doi.org/10.1137/0707049

10. Austin, T. M. A memory efficient parallel tridiagonal solver: Preprint LA-VR-03-4149 / T. M. Austin, M. Berndt, J. D. Moulton. – 2004. – 13 p.

References

1. Samarskii A. A., Nikolaev E. S. *Numerical Methods for Grid Equations. Vol. 1. Direct Methods*. Birkhauser Verlag, 1989. 242 p.

2. Heller D. Some aspects of the cyclic reduction algorithm for block tridiagonal linear systems. *SIAM Journal on Numerical Analysis*. 1976, vol. 13, no. 4, pp. 484–496. https://doi.org/10.1137/0713042

3. Akimova E. N. Parallelization of the matrix sweep algorithm. *Matematicheskoe modelirovanie = Mathematical Models*, 1994, vol. 6, no. 9, pp. 61–67 (in Russian).

4. Akimova E. N, Belousov D. V. Parallel algorithms for solving the systems of equations with block-three-diagonal matrices on multiprocessors computer systems. *Vestnik UGATU* [Scientific Journal of Ufa State Aviation Technical University], 2011 vol. 15, no. 5, pp. 87–93 (in Russian).

5. Hirshman S. P., Perumalla K. S., Lynch V. E., Sanchez R. BCYCLIC: A parallel block tridiagonal matrix cyclic solver. *Journal of Computational Physics*, 2010, vol. 229, no. 18, pp. 6392–6404. https://doi.org/10.1016/j.jcp.2010.04.049

6. Davina A. Lamas, Roman J. E. MPI-CUDA parallel linear solvers for block-tridiagonal matrices in the context of SLEPc'seigensolvers. *Parallel Computing*, 2018, vol. 74, pp. 118–135. https://doi.org/10.1016/j.parco.2017.11.006

7. Yanenko N. N., Konovalov A. N., Bugrov A. N., Shustov G. V. Organization of Parallel Computing and the Thomas Algorithm Parallelization. *Chislennye metody mekhaniki sploshnoi sredy* [Numerical Methods in Continuum Mechanics], 1978, vol. 9, no. 7, pp. 139–146 (in Russian).

8. Wang H. H. A parallel method for tridiagonal equations. *ACM Transactions on Mathematical Software*, 1981, vol. 7, no. 2, pp. 170–183. https://doi.org/10.1145/355945.355947

9. Buzbee B. L., Golub G. H., Nielson C. W. On direct methods for solving Poisson's equations. *SIAM Journal on Numerical Analysis*, 1970, vol. 7, no. 4, pp. 627–656. https://doi.org/10.1137/0707049

10. Austin T. M., Berndt M., Moulton J. D. *A memory efficient parallel tridiagonal solver.* Preprint LA-VR-03-4149, 2004. 13 p.

Информация об авторах

Згировский Андрей Александрович – магистрант факультета прикладной математики и информатики, Белорусский государственный университет (пр. Независимости, 4, 220030, г. Минск, Республика Беларусь). E-mail: zgirovskya@gmail.com

Лиходед Николай Александрович – доктор физико-математических наук, профессор кафедры вычислительной математики факультета прикладной математики и информатики, Белорусский государственный университет (пр. Независимости, 4, 220030, г. Минск, Республика Беларусь). E-mail: likhoded@bsu.by

Information about the authors

Andrei A. Zgirouski – Undergraduate, Belarusian State University (4, Nezavisimosti Ave., 220030, Minsk, Republic of Belarus). E-mail: zgirovskya@gmail.com

Nikolai A. Likhoded – Dr. Sc. (Physics and Mathematics), Belarusian State University (4, Nezavisimosti Ave., 220030, Minsk, Republic of Belarus). E-mail: likhoded@bsu.by

ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)
УДК 519.6,537.533.7,621.3.032.26
https://doi.org/10.29235/1561-2430-2019-55-4-435-444

Поступила в редакцию 18.02.2019
Received 18.02.2019

А. М. Крот[1], О. Н. Петрович[2], И. С. Русецкий[2]

*[1]Объединеный институт проблем информатики
Национальной академии наук Беларуси, Минск, Беларусь
[2]Полоцкий государственный университет, Новополоцк, Беларусь*

АЛГОРИТМ РАСЧЕТА ТРАЕКТОРИЙ ЭЛЕКТРОНОВ В ЭЛЕКТРОСТАТИЧЕСКОМ И МАГНИТОСТАТИЧЕСКОМ ПОЛЯХ ЭЛЕКТРОННО-ОПТИЧЕСКИХ СИСТЕМ

Аннотация. Предложен алгоритм численного расчета траекторий электронов, эмитированных плазмой, в случае движения пучка в аксиально-симметричных электростатическом и магнитостатическом полях. Данный алгоритм основан на технологии дискретизации пучка заряженных частиц токовыми трубками и методе декомпозиции расчетной области. Моделирование полей и численное решение уравнений движения частиц осуществляются с применением квазиструктурированных сеток.

Ключевые слова: деформируемые трубки тока, электронно-оптические системы, декомпозиция расчетной области, квазиструктурированные сетки, нелинейные самосогласованные задачи, движение электронных пучков, аксиально-симметричные электростатическое и магнитостатическое поля

Для цитирования. Крот, А. М. Алгоритм расчета траекторий электронов в электростатическом и магнитостатическом полях электронно-оптических систем / А. М. Крот, О. Н. Петрович, И. С. Русецкий // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2019. – Т. 55, № 4. – С. 435–444. https://doi.org/10.29235/1561-2430-2019-55-4-435-444

A. M. Krot[1], O. N. Petrovich[2], I. S. Rusetski[2]

*[1]United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus
[2]Polotsk State University, Novopolotsk, Belarus*

CALCULATION ALGORITHM OF THE TRAJECTORIES OF ELECTRONS IN ELECTROSTATIC AND MAGNETOSTATIC FIELDS OF ELECTRON-OPTICAL SYSTEMS

Abstract. An algorithm for numerical calculation of the trajectories of electrons emitted by plasma in the electron beam moving in axially symmetric electrostatic and magnetostatic fields is proposed. This algorithm is based on the technology of charged particle beam discretization by current tubes and the decomposition method of the computational domain. Field simulation and numerical solution of equations for particle motion are carried out with the use of quasi-structured grids.

Keywords: deformable current tubes, electron-optical systems, decomposition of the computational domain, quasi-structured grids, nonlinear self-consistent problems, electron beam motion, axially symmetric electrostatic and magnetostatic fields

For citation. Krot A. M., Petrovich O. N., Rusetski I. S. Calculation algorithm of the trajectories of electrons in electrostatic and magnetostatic fields of electron-optical systems. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2019, vol. 55, no. 4, pp. 435–444 (in Russian). https://doi.org/10.29235/1561-2430-2019-55-4-435-444

Введение. Моделирование формирования электронно-оптическими системами (ЭОС) интенсивных пучков и описание потока электронов в устройствах плазменной эмиссионной электроники требует развития методов расчета движения заряженных частиц в электромагнитных полях и нахождения объемного заряда пучка.

В работе [1] приведено сравнение современных пакетов программ для численного моделирования плазменных эмиссионных систем: PBGUNS [2, 3], KOBRA-3 [4, 5], POISSON-2 [6, 7], ELIS [8]. Вычислительные коды PBGUNS, POISSON-2 и ELIS предназначены для расчетов двумерных задач формирования и транспортировки пучков заряженных частиц в системах с плазменным эмиттером. Код KOBRA-3 позволяет моделировать трехмерные задачи сильноточной электронной и ионной оптики. Расчет электростатических полей в пакетах PBGUNS, KOBRA-3 и ELIS основан на методе конечных разностей для уравнения Пуассона, в пакете POISSON-2 используется метод интегральных уравнений с вычислением потенциала и компонент напряженности поля через поверхностную и объемную плотность зарядов. В программных продуктах PBGUNS, KOBRA-3, POISSON-2 возможен учет магнитостатического поля.

В программных кодах PBGUNS, KOBRA-3, POISSON-2 для описания потока заряженных частиц используется дискретная модель недеформируемых токовых трубок, толщина которых достаточно мала, чтобы трубку можно было задать центральной траекторией. В этом случае плотность тока по сечению такой трубки остается постоянной.

Методом описания движения электронного потока в пакете ELIS служит подход, основанный на применении деформируемых трубок тока [9]. Вместо последовательности частиц, вылетающих одна за другой из данной точки эмиттера в разные моменты времени, рассматривается траектория одной частицы, которая представляет собой граничную траекторию трубки тока. Поверхность эмиттера разбивается на слои, для которых рассчитываются граничные траектории. Частицы потока, вылетевшие с внутренних точек каждого слоя эмиттера, при движении заполняют пространство между граничными траекториями, образуя тем самым деформируемую трубку тока. Поперечное сечение такой токовой трубки изменяется вследствие расхождения или сближения граничных траекторий, и плотность тока такой трубки изменяется по ее сечению.

Указанный подход в пакете ELIS был изначально разработан для расчета траекторий электронов, движущихся в электростатическом поле. В рамках данной статьи предложен алгоритм расчета траекторий электронов, эмитированных плазмой, в совмещенных аксиально-симметричных электро- и магнитостатическом полях. Разработанный алгоритм, с целью построения эффективных численных методов описания потока заряженных частиц, основан на технологии декомпозиции расчетной области, в соответствии с которой моделирование полей и численное решение уравнений движения частиц проводится на квазиструктурированных сетках [10].

Уравнение траектории заряженной частицы в аксиально-симметричных электростатическом и магнитостатическом полях систем формирования электронного потока. Основу метода описания потока заряженных частиц токовыми трубками составляет уравнение движения частицы, которое в статических полях принимает вид [11]

$$m_0 \frac{d\vec{v}}{dt} = -q_0\left(\operatorname{grad} \varphi - \left[\vec{v} \times \vec{B}\right]\right), \tag{1}$$

и *энергетическое* уравнение

$$\frac{m_0 \vec{v}^{\,2}}{2} - \frac{m_0 \vec{v}^{\,2}(0)}{2} = -q_0(\varphi - \varphi(0)), \tag{2}$$

где φ – скалярный потенциал электромагнитного поля, $\vec{B}$ – вектор магнитной индукции, m_0 – масса частицы, q_0 – заряд частицы (например, $q_0 = -e$ – заряд электрона), $\vec{v}$ – скорость частицы, $\vec{v}(0)$ – начальная скорость вылета заряженных частиц (электронов) из плазмы с потенциалом $\varphi(0) = \varphi_{pl}$.

Вектор магнитной индукции $\vec{B}$ может быть представлен через векторный потенциал электромагнитного поля $\vec{A}$ [12]:

$$\vec{B} = \operatorname{rot} \vec{A}. \tag{3}$$

Уравнения (1), (2) являются *базовыми* для расчета траекторий движения частиц (электронов и ионов) в статических электромагнитных полях. При условии достаточной *сплошности* потока [13] заряженных частиц к ним добавляется уравнение непрерывности, записанное в интегральной форме:

$$\oiint \vec{j}\, d\vec{S} = 0, \tag{4}$$

где $\vec{j} = \rho\vec{v}$ – плотность потока заряженных частиц (тока электронного или ионного пучка) в поперечном сечении площадью S, ρ – плотность заряда частиц, $\vec{v}$ – скорость заряженных частиц в заданной точке расчетной области.

При формировании пучков заряженных частиц в *аксиально-симметричных* электромагнитных полях, когда $\varphi = \varphi(r, z)$ и $\vec{A} = \vec{A}(r, z)$, азимутальные компоненты напряженности поля $\vec{E}$ и вектора магнитной индукции $\vec{B}$ равны нулю: $E_\varepsilon = -\operatorname{grad}_\varepsilon \varphi = 0$, $B_\varepsilon = \operatorname{rot}_\varepsilon \vec{A} = 0$. Таким образом, векторный потенциал задается исключительно азимутальной составляющей A_ε, а вектор магнитной индукции в цилиндрической системе координат $\vec{r} = (r, \varepsilon, z)$ определяется только радиальной B_r и осевой B_z компонентами в соответствии с формулами [13]:

$$B_r = \operatorname{rot}_r \vec{A} = \frac{1}{r}\frac{\partial A_z}{\partial \varepsilon} - \frac{\partial A_\varepsilon}{\partial z} = -\frac{\partial A_\varepsilon}{\partial z}; \tag{5а}$$

$$B_z = \operatorname{rot}_z \vec{A} = \frac{1}{r}\left(\frac{\partial (rA_\varepsilon)}{\partial r} - \frac{\partial A_r}{\partial \varepsilon}\right) = \frac{1}{r}\frac{\partial (rA_\varepsilon)}{\partial r}. \tag{5б}$$

Уравнение (1) в цилиндрической системе координат для заряженной частицы, движущейся со скоростью $\vec{v} = (\dot{r}, r\dot{\varepsilon}, \dot{z})$, можно представить в виде системы

$$m_0\ddot{r} = m_0 r\dot{\varepsilon}^2 - q_0\left(\frac{\partial \varphi(r,z)}{\partial r} - r\dot{\varepsilon}B_z\right); \tag{6а}$$

$$m_0\frac{1}{r}\frac{d}{dt}\left(r^2\dot{\varepsilon}\right) = q_0\left(\dot{z}B_r - \dot{r}B_z\right); \tag{6б}$$

$$m_0\ddot{z} = -q_0\left(\frac{\partial \varphi(r,z)}{\partial z} + r\dot{\varepsilon}B_r\right). \tag{6в}$$

Подставляя (5а) и (5б) в формулу (6б) и интегрируя, получим *закон сохранения азимутального обобщенного импульса* [14]:

$$m_0(r^2\dot{\varepsilon} - r_0^2\dot{\varepsilon}_0) = -q_0(rA_\varepsilon - r_0 A_{\varepsilon 0}). \tag{7}$$

Согласно (3) вектор магнитной индукции $\vec{B}$ подчиняется дополнительному условию

$$\mathbf{div}\,\vec{B} = \mathbf{div}\operatorname{rot}\vec{A} = 0, \tag{8}$$

что позволяет выявить *функцию тока* $\psi(r,z)$, связанную с проекциями векторного потенциала на оси цилиндрических координат соотношениями

$$B_r = -\frac{1}{r}\frac{\partial \psi}{\partial z}, \quad B_z = \frac{1}{r}\frac{\partial \psi}{\partial r}. \tag{9}$$

Полагая $\psi = 0$ на оси Oz, с учетом (9) находим величину *магнитного потока* сквозь ортогональное к оси Oz сечение, ограниченное окружностью заданного радиуса r:

$$\Psi = \int_0^r B_z \cdot 2\pi r dr = 2\pi\int_0^r \frac{1}{r}\frac{\partial \psi}{\partial r} \cdot r dr = 2\pi\psi(r). \tag{10}$$

Как известно [13], функцию тока можно рассматривать как *одну из составляющих* векторного потенциала $\vec{A}$, что позволяет выразить азимутальную составляющую векторного потенциала A_ε через магнитный поток Ψ:

$$A_\varepsilon = \frac{\Psi}{2\pi r}. \tag{11}$$

С учетом (11) уравнение (7) можно представить в форме интеграла движения

$$m_0(r^2\dot{\varepsilon} - r_0{}^2\dot{\varepsilon}_0) = -\frac{q_0}{2\pi}(\Psi - \Psi_0). \tag{12}$$

Полагая, что начальное значение азимутальной угловой скорости электронов, эмитированных плазмой, равно нулю $\dot{\varepsilon}_0 = 0$, запишем уравнение (12) в следующей форме:

$$\dot{\varepsilon} = -\frac{q_0}{2\pi m_0}\frac{\Psi - \Psi_0}{r^2}. \tag{13}$$

Подставляя формулы (5б), (11) и (13) в уравнение (6а) и аналогично – формулы (5а), (11) и (13) в (6в), получаем

$$m_0\ddot{r} = -q_0\frac{\partial\varphi(r,z)}{\partial r} - \frac{q_0^2}{8\pi^2 m_0}\frac{\partial}{\partial r}\left[\frac{(\Psi - \Psi_0)^2}{r^2}\right]; \tag{14}$$

$$m_0\ddot{z} = -q_0\frac{\partial\varphi(r,z)}{\partial z} - \frac{q_0^2}{8\pi^2 m_0}\frac{\partial}{\partial z}\left(\frac{(\Psi - \Psi_0)^2}{r^2}\right). \tag{15}$$

Введем *эквивалентный потенциал* Ф [15]:

$$\Phi = \varphi(r,z) + \frac{q_0}{8\pi^2 m_0}\frac{(\Psi - \Psi_0)^2}{r^2} \tag{16}$$

и преобразуем уравнения движения (14), (15) к простому виду:

$$m_0\ddot{r} = -q_0\frac{\partial\Phi}{\partial r}; \tag{17а}$$

$$m_0\ddot{z} = -q_0\frac{\partial\Phi}{\partial z}. \tag{17б}$$

Уравнение траектории заряженной частицы (в частности, электрона) в меридианной (меридиональной) плоскости *rz* находим путем исключения времени из уравнений движения (17а) и (17б) в аксиально-симметричных электростатическом и магнитостатическом полях. С этой целью преобразуем $\ddot{r}$:

$$\ddot{r} = \frac{d^2r}{dt^2} = \frac{d}{dt}\left(\frac{dr}{dz}\cdot\frac{dz}{dt}\right) = \frac{d}{dt}\left(\dot{z}\frac{dr}{dz}\right) = \ddot{z}\frac{dr}{dz} + \dot{z}\frac{d}{dt}\left(\frac{dr}{dz}\right) = \ddot{z}\frac{dr}{dz} + \dot{z}^2\frac{d}{dz}\left(\frac{dr}{dz}\right) = \ddot{z}\frac{dr}{dz} + \dot{z}^2\frac{d^2r}{dz^2}.$$

Учитывая уравнения (17а) и (17б), получим

$$\dot{z}^2\frac{d^2r}{dz^2} = -\frac{q_0}{m_0}\left(\frac{\partial\Phi}{\partial r} - \frac{\partial\Phi}{\partial z}\frac{dr}{dz}\right). \tag{18}$$

Поскольку $\dot{r}^2 + r^2\dot{\varepsilon}^2 + \dot{z}^2 = \vec{v}^2$, то величина квадрата осевой скорости $\dot{z}^2$ может быть выражена из закона сохранения энергии (2):

$$\dot{z}^2 = \frac{v^2(0) - \frac{2q_0}{m_0}\left(\Phi - \varphi_{pl}\right)}{1 + \left(\frac{dr}{dz}\right)^2}. \tag{19}$$

Подставляя (19) в уравнение (18), получаем искомое уравнение траектории в меридианной плоскости в дифференциальной форме [15]:

$$\frac{d^2 r}{dz^2} = \frac{-\dfrac{q_0}{m_0}\left(\dfrac{\partial \Phi}{\partial r} - \dfrac{\partial \Phi}{\partial z}\dfrac{dr}{dz}\right)\left(1 + \left(\dfrac{dr}{dz}\right)^2\right)}{v^2(0) - \dfrac{2q_0}{m_0}\left(\Phi - \varphi_{pl}\right)}. \tag{20}$$

Численное интегрирование уравнения (20) позволяет определить фазовую характеристику пучка заряженных частиц в каждом поперечном сечении потока для заданного значения начальной энергии частиц.

При условии достаточной сплошности пучка заряженных частиц базовые уравнения (1), (2), (4) для расчета потока движущихся заряженных частиц (электронов и ионов) в статических электромагнитных полях соответственно в дифференциальной форме примут вид, аналогичный гидродинамическим уравнениям [13]:

$$\rho\left(\frac{\partial \vec{v}}{\partial t} + (\vec{v}\cdot\nabla)\vec{v}\right) = -\operatorname{grad} p - \rho_{\pm}\left(\operatorname{grad}\varphi - \left[\vec{v}\times\vec{B}\right]\right) + \mu\nabla^2\vec{v}; \tag{21а}$$

$$\rho\frac{d}{dt}\left(\frac{\vec{v}^2}{2}\right) = -\rho_{\pm}\vec{v}\cdot\left(\operatorname{grad}\varphi - \left[\vec{v}\times\vec{B}\right]\right) - \mathbf{div}(p\vec{v}) + N_{in}; \tag{21б}$$

$$\frac{\partial \rho_{\pm}}{\partial t} + \mathbf{div}(\rho_{\pm}\vec{v}) = 0, \tag{21в}$$

где ρ – плотность массы потока частиц; $\rho_{\pm}$ – плотность заряда потока частиц; $\vec{v}$ – скорость потока заряженных частиц в заданной точке расчетной области; p –давление газа, образованного заряженными частицами; φ – скалярный потенциал электромагнитного поля; $\vec{B}$ – вектор магнитной индукции; μ – динамический коэффициент вязкости газа заряженных частиц; N_{in} – плотность распределения мощности внутренних сил [13]. Ввиду незначительной величины давления газа заряженных частиц ($p \approx 0$), первое уравнение (21а) – аналог уравнения Навье – Стокса [13] вырождается в трехмерное уравнения Бюргерса, простейшим решением которого является решение Тейлора в виде ударной волны [16], а дисперсионное обобщение данного нелинейного уравнения имеет решение в виде солитонных волн [17].

Алгоритмическая и программная реализация численных методов расчета электронного потока в устройствах плазменной эмиссионной электроники. Алгоритм компьютерного моделирования потока электронов, движущихся в электро- и магнитостатическом полях систем формирования пучка в устройствах плазменной эмиссионной электроники, основан на поэтапном приближении к самосогласованной задаче. На первом этапе рассчитывается электростатический и эквивалентный потенциалы поля, созданного электронно-оптической системой. На втором – определяются траектории движения электронов в аксиально-симметричных электро- и магнитостатическом полях. Затем вычисляется объемный заряд, вносимый электронным потоком, и характеристики электронного пучка в каждой плоскости поперечного сечения. На следующем этапе проводится расчет потенциалов поля с учетом объемного заряда потока электронов. Далее изменяется положение плазменного эмиттера и форма эмитирующей плазменной поверхности (плазменного мениска). В измененных собственным зарядом электронного пучка электро- и магнитостатическом полях заново рассчитываются траектории частиц и характеристики потока. Этапы расчета в программном цикле продолжаются до тех пор, пока отклонения значений потенциалов поля и траекторий электронов в соседних приближениях не станут меньше заданной точности.

Для повышения точности расчета характеристик электронного потока в устройствах плазменной эмиссионной электроники используется технология построения квазиструктурированных сеток, основанная на методе декомпозиции расчетной области. В расчетной области электронно-оптической системы строится квазиструктурированная сетка (рис. 1), состоящая из макросетки *3* и двух вложенных подсеток: одна из них покрывает подобласть распространения электронного пучка *4*, другая – приэмиттерную подобласть *5*. Каждому сеточному узлу задаются координаты *(i, j)*. Численное моделирование полей проводится на квазиструктурированной сетке. Расчет движения частиц пучка (траектории, объемный заряд) производится на подсетке внутренней области, положение и форма плазменного мениска находится на подсетке приэмиттерной области.

Для описания стационарных полей применяется метод потенциалов. Скалярный и векторный потенциалы удовлетворяют уравнениям Пуассона:

$$\nabla^2\varphi = -\frac{\rho}{\varepsilon_0},\quad \vec{E} = -\operatorname{grad}\varphi = -\nabla\varphi; \tag{22}$$

$$\nabla^2\vec{A} = -\mu_0\vec{j},\quad \vec{B} = \operatorname{rot}\vec{A} = [\nabla\times\vec{A}], \tag{23}$$

где ε_0 и μ_0 – электрическая и магнитная постоянные.

В области, свободной от токов, для описания магнитного поля целесообразно использовать скалярный магнитный потенциал φ_m, который удовлетворяет уравнению Лапласа

$$\nabla^2\varphi_m = 0,\quad \vec{B} = -\nabla\varphi_m. \tag{24}$$

С целью реализации численного расчета электрического поля, создаваемого системой электродов и объемным зарядом потока частиц, а также магнитного поля, уравнения (22)–(24) решаются методом конечных разностей на квадратной сетке с граничными условиями Дирихле.

На первом шаге решение системы (22)–(24) находится на макросетке, заполняющей всю расчетную область в соответствии с конечно-разностными уравнениями:

$$u_{i,j} = \frac{1}{4}\left(u_{i+1,j} + u_{i-1,j} + \frac{2j+1}{2j}u_{i,j+1} + \frac{2j-1}{2j}u_{i,j-1} + h^2 f_{i,j}\right); \tag{25}$$

$$u_{i,0} = \frac{1}{6}\left(u_{i+1,0} + u_{i-1,0} + 4u_{i,1} + h^2 f_{i,0}\right), \tag{26}$$

где потенциалы поля $u = \{\varphi, \vec{A}, \varphi_m\}$, функция $f = \left\{-\frac{\rho}{\varepsilon_0}, -\mu_0\vec{j}, 0\right\}$.

На следующем шаге проводится расчет потенциалов полей на внутренней подсетке с учетом значений потенциалов в общих узлах макросетки и подсетки. Затем вычисляются значения по-

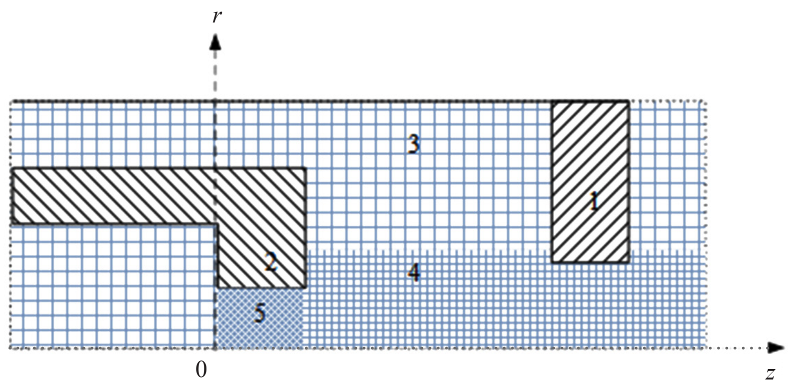


Рис. 1. Квазиструктурированная сетка в расчетной области: *1* – ускоряющий электрод; *2* – фокусирующий электрод; *3* – макросетка; *4* – подсетка в подобласти распространения электронного пучка; *5* – подсетка в приэмиттерной подобласти

Fig. 1. Quasi-structured mesh in the computational domain: *1* – accelerating electrode; *2* – focusing electrode; *3* –macrogrid; *4* – subgrid in the subregion of electron beam propagation; *5* – subgrid in the near emitter domain

тенциалов в приэмиттерной области в дополнительных узлах сетки, вложенной в предыдущую. Таким образом, при расчете полей во внутренней и приэмиттерной подобласти используются значения потенциалов, полученные в узлах более крупных сеток на предыдущих этапах [10].

На границах подобластей применяются формулы Лагранжа для интерполяции значений потенциалов в узлах, расположенных между узлами более крупных сеток [10]:

$$u_{i+\frac{1}{2}} = \frac{1}{16}\left(-u_{i-1} + 9u_i + 9u_{i+1} - u_{i+2}\right); \tag{27}$$

$$u_{i+\frac{1}{2}} = \frac{1}{8}\left(3u_i + 6u_{i+1} - u_{i+2}\right); \tag{28}$$

$$u_{i-\frac{1}{2}} = \frac{1}{8}\left(3u_i + 6u_{i-1} - u_{i-2}\right). \tag{29}$$

В результате решения системы конечно-разностных уравнений (25)–(29) получим скалярный и векторный потенциалы аксиально-симметричных полей $\varphi = \varphi(r,z)$ и $\vec{A} = \vec{A}(r,z)$. Затем вычислим магнитный поток Ψ по формулам (11) или (10) и эквивалентный потенциал Φ по формуле (16).

Алгоритмическая реализация разработанного численного метода расчета траекторий движения электронов в аксиально-симметричных электро- и магнитостатическом полях и характеристик пучка основана на технологии дискретизации электронного потока на токовые слои. Для граничной траектории деформируемой трубки тока в плоскости поперечного сечения путем численного решения уравнения (20) находим значения угла наклона скорости электронов к оси z (расходимость $\theta = \mathrm{arctg}\left(\frac{dr}{dz}\right)$), а также радиус трубки тока и радиус пучка.

Определяя из уравнения (19) осевую скорость $\dot{z}$ и исключая время из уравнений (19) и (13), получим соотношение

$$\frac{d\varepsilon}{dz} = -\frac{q_0}{2\pi m_0}\frac{\Psi - \Psi_0}{r^2}\sqrt{\frac{1 + \left(\frac{dr}{dz}\right)^2}{v^2(0) - \frac{2q_0}{m_0}\left(\Phi - \varphi_{pl}\right)}}. \tag{30}$$

Путем численного решения уравнения (30) рассчитаем азимутальную координату ε электронов, движущихся на границе токовой трубки.

Ток эмиссии I_e для каждой токовой трубки и плотность тока j электронов пучка, вылетевших из эмиттера с начальной скоростью $v(0)$, определяем по формулам

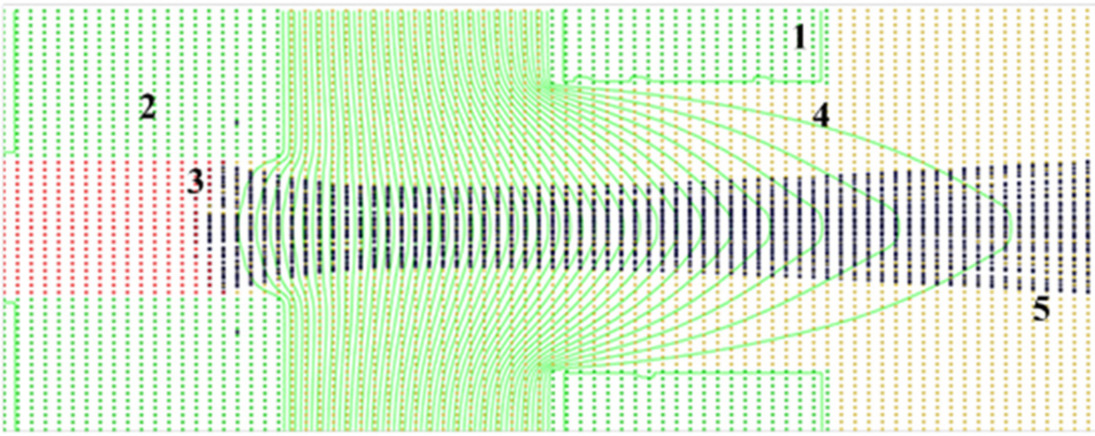
$$I_e = en_{pl}v(0)S_0; \tag{31а}$$

$$j = \frac{I_e}{S_n(z)}, \tag{31б}$$

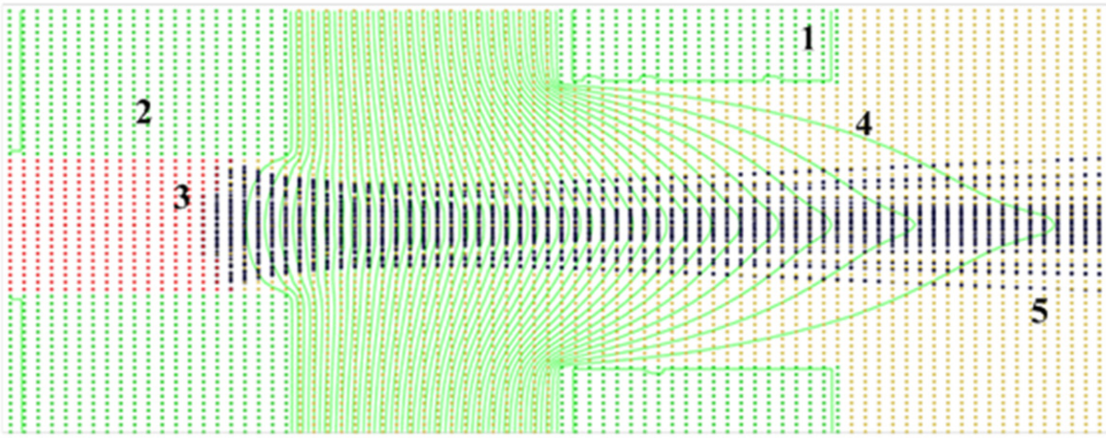
а затем находим объемный заряд ρ:

$$\rho = \frac{I_e}{v \cdot S_n(z)}. \tag{32}$$

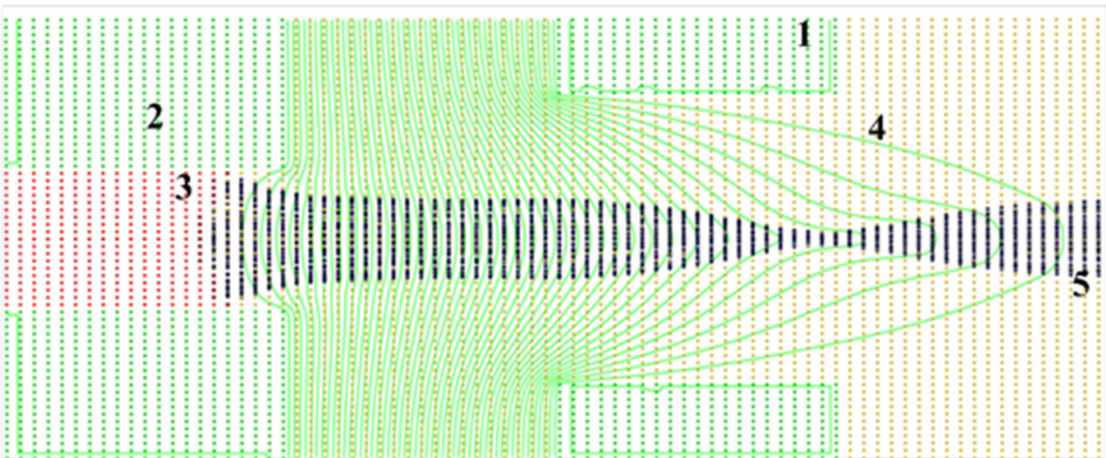
Здесь S_0 – площадь эмитирующей поверхности плазмы для каждой трубки тока, n_{pl} – концентрация электронов плазмы, $S_n(z)$ – площадь поперечного сечения трубки тока пучка плоскостью $z = \mathrm{const}$, v – скорость электронов пучка в точке расчетной области находится из уравнения (2).



a



b



c

Рис. 2. Формирование и фокусировка электронного пучка в электронно-оптической системе с плазменным эмиттером (*1* – ускоряющий электрод; *2* – фокусирующий электрод; *3* – плазменный эмиттер; *4* – эквипотенциальные линии электростатического поля; *5* – траектории электронов): *a* – формирование пучка в электростатическом поле; *b*, *c* – формирование пучка в электро- и магнитостатическом полях

Fig. 2. Formation and focusing of an electron beam in the electron-optical system with a plasma emitter (*1* – accelerating electrode; *2* – focusing electrode; *3* – plasma emitter; *4* – equipotential lines of the electrostatic field; *5* – electron trajectories): *a* – beam formation in the electrostatic field; *b*, *c* – beam formation in electrostatic and magnetostatic fields

На рис. 2 приведен результат моделирования электронно-оптической системы в устройствах плазменной эмиссионной электроники в качестве примера программной реализации изложенного алгоритма. Представлены траектории (*5*) электронов формируемого пучка и эквипотенциальные линии электростатического поля (*4*) в электронно-оптической системе со следующими параметрами: ускоряющая разность потенциалов – 10 кВ; промежуток ускорения между фокусирующим (*2*) и ускоряющим (*1*) электродами – 2 мм; радиус и длина канала в фокусирующем электроде – 1 и 2 мм соответственно; радиус и длина канала в ускоряющем электроде – 2 и 2 мм соответственно; концентрация плазменных электронов – $5 \cdot 10^{18}$ м$^{-3}$. На рис. 2, *a* поток электронов движется в электростатическом поле; на рис. 2, *b* и 2, *c* показано влияние внешнего продольного магнитостатического поля на характеристики пучка. На выходе из электронно-оптической системы диаметр и расходимость электронного пучка равны соответственно 1,8 мм и 8° в электростатическом поле (см. рис. 2, *a*); 1,4 мм и 4° – для электронно-оптической системы с электро- и магнитостатическим (B_z = 0,1 Тл) полями (см. рис. 2, *b*); 1,2 мм и 4° – для электронно-оптической системы с электро- и магнитостатическим (B_z = 0,5 Тл) полями (см. рис. 2, *c*). Ток пучка для случаев, представленных на рис. 2, равен 150 мА.

Заключение. Разработан алгоритм численного решения нелинейных самосогласованных задач по расчету формирования и движения электронных пучков в совмещенных аксиально-симметричных электростатическом и магнитостатическом полях, основанный на технологиях дискретизации электронного потока деформированными трубками тока и декомпозиции расчетной области при расчете полей и уравнений движения частиц.

Представленный алгоритм позволяет проводить сравнительный анализ траекторий электронов и характеристик пучка как в электростатическом поле системы формирования пучка, так и совмещенных электро- и магнитостатическом полях. Данная разработка расширяет возможности программного кода ELIS, так как учитывает влияние внешних и собственных магнитостатических полей на формирование и фокусировку пучка, и может быть использована при проектировании электронно-оптических систем с плазменным эмиттером с целью численного анализа и оптимизации характеристик пучка.

Благодарности. Работа выполнена частично в рамках предоставленного гранта Президента Республики Беларусь в науке на 2019 г., а также при финансовой поддержке Белорусского республиканского фонда фундаментальных исследований (проект № Ф17-122).

Acknowledgements. This work was partially supported by the Grant of the President of the Republic of Belarus in science for 2019 and carried out with financial support of the Belarusian Republican Foundation for Fundamental Research (project No. F17-122).

Список использованных источников

1. Котельников, И. А. Теория плазменного эмиттера положительных ионов / И. А. Котельников, В. Т. Астрелин // Успехи физ. наук. – 2015. – Т. 185, № 7. – С. 753–771. https://doi.org/10.3367/ufnr.0185.201507c.0753
2. Boers, J. E. An interactive version of the PBGUNS program for the simulation of axisymmetric and 2-D, electron and ion beams and guns / J. E. Boers // Proc. Particle Accelerator Conf., 1–5 May 1995. – IEEE, 1996. – P. 2312. https://doi.org/10.1109/pac.1995.505535https://far-tech.com/pbguns.php
3. PBGUNS (Particle Beam GUN Simulations) [Electronic Resource]. Mode of access: https://far-tech.com/pbguns.php
4. KOBRA3-INP, INP, Junkernstr. 99, 65205 Wiesbaden.
5. Spädtke, P. Computer Simulation of High-Current DCIon Beams / P. Spädtke // Proc. 1984 Linear Accelerator Conf. (LINAC’84), Seeheim, Germany, May 1984, paper THP0012. – Seeheim, 1984. – P. 356–358.
6. Астрелин, В. Т. Пакет программ для расчета характеристик интенсивных пучков релятивистских заряженных частиц / В. Т. Астрелин, В. Я. Иванов // Автометрия. – 1980. – № 3. – С. 92–99.
7. Астрелин, В. Т. Особенности решения задач плазменной эмиссионной электроники в пакете прикладных программ POISSON-2 / В. Т. Астрелин // Успехи прикладной физики. – 2013. – Т. 1, № 5. – С. 574–579.
8. Петрович, О. Н. Программный комплекс ELIS для моделирования ЭОС ПИЭЛ / О. Н. Петрович, В. А. Груздев // Прикладная физика. – 2012. – № 2. – С. 79–85.
9. Петрович, О. Н. Моделирование электронно-оптических систем с плазменным эмиттером: дис. … канд. техн. наук: 05.13.18 / О. Н. Петрович. – Новополоцк, 2012. – 199 л.
10. Петрович, О. Н. Численные методы расчета электромагнитных полей на квазиструктурированных сетках в устройствах плазменной эмиссионной электроники / О. Н. Петрович, И. С. Русецкий // Вестн. Полоц. Гос. ун-та. Сер. С, Фундам. науки. – 2018. – № 4. – С. 124–127.
11. Ландау, Л. Д. Теория поля / Л. Д. Ландау, Е. М. Лифшиц. – М.: Наука, 1973. – 504 с.

12. Яворский, Б. М. Справочник по физике для инженеров и студентов ВУЗов / Б. М. Яворский, А. А. Детлаф. – М.: Физматгиз, 1963. – 848 с.

13. Лойцянский, Л. Г. Механика жидкости и газа / Л. Г. Лойцянский. – М.: Наука, 1973. – 840 с.

14. Ландау, Л. Д. Механика / Л. Д. Ландау, Е. М. Лифшиц. – М.: Наука, 1973. – 208 с.

15. Молоковский, С. И. Интенсивные электронные и ионные пучки / С. И. Молоковский, А. Д. Сушков. – М.: Энергоатомиздат, 1991. – 304 с.

16. Солитоны и нелинейные волновые уравнения: пер. с англ. / Р. Додд [и др.]. – М.: Мир, 1988. – 694 с.

17. Krot, A. M. Development of the generalized nonlinear Schrödinger equation of rotating cosmogonical body formation / A. M. Krot // Complex Systems: Theory and Applications, Ch. 3. – New York: Nova Science Publ., 2017. – P. 49–94.

References

1. Kotel'nikov I. A., Astrelin V. T. Theory of plasma emitter of positive ions. *Uspekhi Fizicheskih Nauk = Physics-Uspekhi*, 2015, vol. 185, no. 7, pp. 753–771. https://doi.org/10.3367/ufnr.0185.201507c.0753

2. Boers J. E. An interactive version of the PBGUNS program for the simulation of axisymmetric and 2-D, electron and ion beams and guns. *Proceedings Particle Accelerator Conference*, 1–5 May 1995. IEEE, 1996. P. 2312. https://doi.org/10.1109/pac.1995.505535

3. *PBGUNS (Particle Beam GUN Simulations).* Available at: https://far-tech.com/pbguns.php

4. KOBRA3-INP, INP, Junkernstr. 99, 65205 Wiesbaden.

5. Spädtke P. Computer Simulation of High-Current DC Ion Beams. *Proc. 1984 Linear Accelerator Conf.* (*LINAC'84*), Seeheim, Germany, May 1984, paper THP0012. Seeheim, 1984, pp. 356–358.

6. Astrelin V. T., Ivanov V. Ya. Software package for calculating the characteristics of intense beams of relativistic charged particles. *Avtometriya = Optoelectronics, Instrumentations and Data Processing*, 1980, no. 3, pp. 92–99 (in Russian).

7. Astrelin V. T. Features of solving the plasma emission electronics problems in CAD POISSON-2. *Uspekhi Prikladnoi Fiziki = Advances in Applied Physics*, 2013, vol. 1, no. 5, pp. 574–579 (in Russian).

8. Petrovich O. N., Gruzdev V. A. The software package ELIS for simulation of EOS of PES. *Prikladnaya Fizika = Applied Physics*, 2012, no. 2, pp. 79–85 (in Russian).

9. Petrovich O. N. *Simulation of Electron-Optical Systems with a Plasma Emitter*. Novopolotsk, 2012. 199 p. (in Russian).

10. Petrovich O. N., Rusetskiш I. S. Numerical methods for calculation of electromagnetic fields on quasi-structured grids in devices of plasma emission electronics. *Vestnik Polotskogo gosudarstvennogo universiteta. Seriya S, fundamental'nye nauki = Vestnik of Polotsk State University. Part C, Fundamental Sciences,* 2018, no. 4, pp. 124–127 (in Russian).

11. Landau L. D., Lifschitz E. M. *Classical Theory of Fields*. Reading, Addison-Wesley Publ., 1951.

12. Yavorskii B. M., Detlaf A. A. *A Handbook of Physics for Engineers and Students*. Moscow, Fizmatgiz Publ., 1963. 848 p. (in Russian).

13. Loitsyanskii L. G. *Mechanics of Fluid and Gas*. Moscow, Nauka Publ., 1973. 848 p. (in Russian).

14. Landau L. D., Lifschitz E. M. *Mechanics*. Moscow, Nauka Publ., 1973. 208 p. (in Russian).

15. Molokovskii S. I., Sushkov A. D. *Intensive Electron and IonBeams*. Moscow, Energoatomizdat Publ., 1991. 304 p. (in Russian).

16. Dodd R. K., Eilbeck J. C., Gibbon J. D., Morris H. C. *Solitons and Nonlinear Wave Equations*. London, Academic Press, 1984. 630 p.

17. Krot A. M. Development of the generalized nonlinear Schrödinger equation of rotating cosmogonical body formation. *Complex Systems: Theory and Applications. Ch. 3.* New York, Nova Science Publ., 2017, pp. 49–94.

Информация об авторах

Крот Александр Михайлович – доктор технических наук, профессор, заведующий лабораторией моделирования самоорганизующихся систем, Объединенный институт проблем информатики Национальной академии наук Беларуси (ул. Сурганова, 6, 220012, г. Минск, Республика Беларусь). E-mail: alxkrot@newman.bas-net.by

Петрович Ольга Николаевна – кандидат технических наук, доцент, заведующий кафедрой технологий программирования, Полоцкий государственный университет (ул. Блохина, 29, 211440, г. Новополоцк, Республика Беларусь). E-mail: o.petrovich@psu.by

Русецкий Игорь Станиславович – старший преподаватель кафедры энергетики и электронной техники, Полоцкий государственный университет (ул. Блохина, 29, 211440, Новополоцк, Республика Беларусь). E-mail: i.rusetski@psu.by

Information about the authors

Alexander M. Krot – Dr. Sc. (Engineering), Professor, Head of the Laboratory of Self-Organization Systems Modeling, United Institute of Informatics Problems of the National Academy of Sciences of Belarus (6, Surganov Str., 220012, Minsk, Republic of Belarus). E-mail: alxkrot@newman.bas-net.by

Olga N. Petrovich – Ph. D. (Engineering), Associate Professor, Head of the Department of Programming Technologies, Polotsk State University (29, Blokhin Str., 211440, Novopolotsk, Republic of Belarus). E-mail: o.petrovich@psu.by

Igor S. Rusetski – Senior Lecturer at the Department of Energy and Electronic Engineering, Polotsk State University (29, Blokhin Str., 211440, Novopolotsk, Republic of Belarus). E-mail: i.rusetski@psu.by

ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)
УДК 517.538.52+517.538.53+517.518.84
https://doi.org/10.29235/1561-2430-2019-55-4-445-456

Поступила в редакцию 26.08.2019
Received 26.08.2019

А. П. Старовойтов, Н. В. Рябченко

Гомельский государственный университет им. Ф. Скорины, Гомель, Беларусь

О ЕДИНСТВЕННОСТИ РЕШЕНИЙ ЗАДАЧ ЭРМИТА – ПАДЕ

Аннотация. Введены новые понятия: вполне нормальный индекс и вполне совершенная система функций, с помощью которых доказан критерий единственности решения двух задач Эрмита – Паде, определены явные детерминантные представления многочленов Эрмита – Паде 1-го и 2-го рода для произвольной системы степенных рядов. Полученные результаты дополняют хорошо известные результаты в теории аппроксимаций Эрмита – Паде.

Ключевые слова: задача Эрмита – Паде, многочлены Эрмита – Паде, нормальный индекс, совершенная система функций, определители Адамара

Для цитирования. Старовойтов, А. П. О единственности решений задач Эрмита – Паде / А. П. Старовойтов, Н. В. Рябченко // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2019. – Т. 55, № 4. – С. 445–456. https://doi.org/10.29235/1561-2430-2019-55-4-445-456

A. P. Staravoitov, N. V. Ryabchenko

Francisk Skorina Gomel State University, Gomel, Belarus

UNIQUENESS OF THE SOLUTIONS OF THE HERMITE – PADE PROBLEMS

Abstract. New concepts are introduced in the present work. They are a quite normal index and a quite perfect system of functions. Using these concepts, the uniqueness criterion for solution of two Hermite – Pade problems is proved, the explicit determinant representations of type I and II Hermite – Padé polynomials for an arbitrary system of power series are obtained. The results obtained complement and generalize the well-known result in the theory of Hermite – Padé approximations.

Keywords: problem Hermite – Padé, Hermite – Padé polynomials, normal index, perfect system, Hadamard determinant

For citation. Staravoitov A. P., Ryabchenko N. V. Uniqueness of the solutions of the Hermite – Pade problems. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2019, vol. 55, no. 4, pp. 445–456 (in Russian). https://doi.org/10.29235/1561-2430-2019-55-4-445-456

1. Многочлены Эрмита – Паде 2-го рода. ***Постановка задачи.*** Пусть $f=(f_1,...,f_k)$ – набор степенных рядов

$$f_j(z)=\sum_{i=0}^{\infty} f_i^j z^i,\quad j=1,2,...,k \tag{1}$$

с комплексными коэффициентами. Множество k-мерных мультииндексов (индексов), т. е. упорядоченных k целых неотрицательных чисел, обозначим $\mathbb{Z}_+^k$. Порядок мультииндекса $\vec{m}=(m_1,...,m_k)$ – это сумма $m=m_1+...+m_k$. Зафиксируем индекс $n\in\mathbb{Z}_+^1$ и мультииндекс $\vec{m}=(m_1,\ldots,m_k)\in\mathbb{Z}_+^k$ и рассмотрим следующую задачу Эрмит – Паде (см. [1, гл. 4, § 3; 2–4]).

З а д а ч а А. Найти тождественно не равный нулю многочлен $Q_m(z)=Q_{n,\vec{m}}(z;f)$, $\deg Q_m\leqslant m$ и такие многочлены $P_{n_j}^j(z)=P_{n,\vec{m}}^j(z;f)$, $\deg P_{n_j}^j\leqslant n_j$, $n_j=n+m-m_j$, чтобы при $j=1,...,k$

$$R_{n,\vec{m}}^j(z):=Q_m(z)f_j(z)-P_{n_j}^j(z)=A_j z^{n+m+1}+\ldots\ . \tag{2}$$

Если $k=1$, то f состоит из одной функции $f(z):=f_1(z)$. В этом случае решение поставленной задачи было получено Паде, который нашел явный вид многочленов $Q_m(z)$, $P_n(z):=P_{n,m}^1(z;f)$ (их называют многочленами Паде). Например, если $f_i:=f_i^1$, $i=0,1,...$, то [2, гл. 1, § 1.1, теорема 1.1.1]

$$Q_m(z)=\begin{vmatrix} f_{n-m+1} & f_{n-m+2} & \dots & f_n & f_{n+1} \\ f_{n-m+2} & f_{n-m+3} & \dots & f_{n+1} & f_{n+2} \\ \dots & \dots & \dots & & \dots \\ f_n & f_{n+1} & \dots & f_{n+m-1} & f_{n+m} \\ z^m & z^{m-1} & \dots & z & 1 \end{vmatrix}. \tag{3}$$

Здесь и далее при $i < 0$ считаем, что $f_i^j = 0$.

Когда f состоит из экспонент $\left\{e^{\lambda_j z}\right\}_{j=1}^k$, где λ_j – различные не равные нулю комплексные числа, решение задачи A в явном виде найдено Ш. Эрмитом в его известной работе [5], посвященной доказательству трансцендентности числа e. При этом искомые многочлены представлены им несобственными интегралами Римана.

Хорошо известно [1], что в общем случае решение задачи A существует, а соответствующие многочлены Q_m, $P_{n_j}^j$ находятся с точностью до мультипликативного множителя: если пара (Q_m, P), где $P = (P_{n_1}^1, \dots, P_{n_k}^k)$, удовлетворяет необходимым условиям, то для любого отличного от нуля комплексного числа λ новая пара $(\lambda Q_m, \lambda P)$ также удовлетворяют необходимым условиям. Эта неединственность может быть и более существенной.

П р и м е р 1. Пусть $k = 1$, $n = 2$, $m = 2$, а

$$f(z) = \frac{1}{2-4z} = \frac{1}{2} + z + 2z^2 + 4z^3 + 8z^4 + \dots .$$

Тогда любое решение задачи можно представить в виде (λQ_2, λP_2), где

$$Q_2(z) = a + bz - (4a+2b)z^2,\ P_2(z) = \frac{1}{2} + \left(a + \frac{1}{2}b\right)z,$$

а a и b – произвольные действительные числа.

Принято говорить (см. [1]), что задача A имеет единственное решение, если все решения задачи можно записать в виде $(\lambda Q_m, \lambda P)$, где $\lambda \in \mathbb{C}$, $\lambda \neq 0$, а (Q_m, P) – некоторое одно фиксированное решение.

О п р е д е л е н и е 1. Если пара (Q_m, P), где $P = (P_{n_1}^1, \dots, P_{n_k}^k)$ – решение задачи A с индексом n и мультииндексом $\vec{m} = (m_1, \dots, m_k)$, то многочлены Q_m, $P_{n_1}^1, \dots, P_{n_k}^k$ называют многочленами Эрмита – Паде 2-го рода (German type) для набора (системы) f формальных степенных рядов (1).

Центральными понятиями в теории таких многочленов является понятие нормального индекса и совершенной системы [1, гл. 4, § 1].

О п р е д е л е н и е 2. Индекс $(n, \vec{m}) = (n, m_1, ..., m_k) \in \mathbb{Z}_+^{k+1}$ называется нормальным для системы f относительно задачи A, если для любого решения (Q_m, P) задачи A с индексом n и мультииндексом $\vec{m}$

$$\deg Q_m = m,\ \deg P_{n_j}^j = n_j,\ j = 1, ..., k.$$

Здесь и далее считаем, что степень многочлена T $\deg T = -1$, тогда и только тогда, когда $T(z) \equiv 0$.

О п р е д е л е н и е 3. Систему f назовем совершенной относительно задачи A, если все индексы $(n, \vec{m}) \in \mathbb{Z}_+^{k+1}$ являются нормальными для f относительно задачи A.

При $k = 1$ критерий нормальности индекса (n, m) выражается условием [4]:

$$H_{n,m} \cdot H_{n,m+1} \neq 0, \tag{4}$$

где определители Адамара $H_{n,m}$ определяются равенствами

$$H_{n,m}=\begin{vmatrix} f_{n-m+1} & f_{n-m+2} & \cdots & f_n \\ f_{n-m+2} & f_{n-m+3} & \cdots & f_{n+1} \\ \cdots & \cdots & \cdots & \cdots \\ f_n & f_{n+1} & \cdots & f_{n+m-1} \end{vmatrix}.$$

Если индекс $(n,\vec{m})$ является нормальным, то задача A имеет единственное решение (см. [1]). В этом случае однозначно определяется вектор

$$\pi_{n,\vec{m}}=\left(\pi^1,\ldots,\pi^k\right),\quad \pi^j=\frac{P^j_{n_j}}{Q_m},$$

компоненты которого π^j называют аппроксимациями Эрмита – Паде 2-го рода (совместными аппроксимациями Паде) для системы f. Следующий пример показывает, что уже при $k = 1$ нормальность индекса (n, m) не является необходимым условием единственности решения поставленной задачи.

Пример 2. Пусть $k = 1$, $n = 2$, $m = 2$, а

$$f(z)=\frac{1}{2-z}+\frac{1}{4}z=\frac{1}{z}+\frac{z}{2}+\frac{z^2}{8}+\frac{z^3}{16}+\frac{z^4}{32}+\ldots\,.$$

Тогда любое решение задачи A можно записать в виде $(\lambda Q_2,\lambda P_2)$, где $\lambda\in\mathbb{C},\ \lambda\neq 0,$ а

$$Q_2(z)=2-z,\quad P_2(z)=1+\frac{z}{2}-\frac{z^2}{4}.$$

При этом индекс (2,2) не является нормальным, так как $\deg Q_2=1.$ Нашей ближайшей целью является нахождение необходимых и достаточных условий на индекс $(n,\vec{m})\in\mathbb{Z}_+^{k+1}$ и систему f, при которых решение задачи A единственно.

Критерий единственности решения задачи A. Компоненты вектора $f=\{f_1,\ldots,f_k\}$ являются, вообще говоря, формальными степенными рядами. Уже по этой причине поставленная задача – чисто алгебраическая и, следовательно, имеет алгебраическое решение.

В дальнейшем, не ограничивая общности, будем считать, что радиусы сходимости всех степенных рядов (1) не равны нулю, а $\vec{m}$ – ненулевой мультииндекс. Для нулевого мультииндекса $\vec{m}$ решение задачи A очевидно: $Q_m(z)\equiv 1,$ а P_n^j – многочлены Тейлора функции f_j.

Введем необходимые обозначения. Для каждого $j=1,\ldots,k,$ фиксированных индекса n и мультииндекса $\vec{m}=(m_1,\ldots,m_k)$ в предположении, что $m_j\neq 0,$ определим матрицы-строки порядка $1 \times (m + 1)$

$$F_i^j=\left(f^j_{n-m_j+i}\quad f^j_{n-m_j+i+1}\quad \ldots\quad f^j_{n_j+i}\right),\quad i=1,2,\ldots,$$

матрицу порядка $m_j \times (m + 1)$

$$F^j=\left[F_1^j\quad F_2^j\quad \ldots\quad F_{m_j}^j\right]^T=\begin{bmatrix} F_1^j \\ \vdots \\ F_{m_j}^j \end{bmatrix}, \tag{5}$$

матрицу порядка $m \times (m + 1)$

$$F_{n,\vec{m}}=\left[F^1\quad F^2\quad \ldots\quad F^k\right]^T, \tag{6}$$

и определители (m + 1)-го порядка:

$$d^{j}_{n,\vec{m},i} = \det\begin{bmatrix} F^1 & F^2 & \dots & F^k & F^{j}_{m_j+i} \end{bmatrix}^T,$$

где C^T является матрицей транспонированной к матрице C (транспонирование определяется так же, как и в (5)). В случае, если $m_j = 0$, матрица $F_{n,\vec{m}}$ и определитель $d^{j}_{n,\vec{m},i}$ не содержат блок-матрицу F^j.

При произвольном m_j определим также функциональные матрицы-строки порядка $1\times(m+1)$:

$$E(z) = \begin{pmatrix} z^m & z^{m-1} & \dots & z & 1 \end{pmatrix},$$

$$E_{m_j}(z) = \begin{pmatrix} \sum_{i=0}^{n-m_j} f_i^j z^{m+i} & \sum_{i=0}^{n-m_j+1} f_i^j z^{m+i-1} & \dots & \sum_{i=0}^{n_j} f_i^j z^i \end{pmatrix}.$$

О п р е д е л е н и е 4. Индекс $(n,\vec{m})$ будем называть вполне нормальным для f относительно задачи A, если ранг матрицы $F_{n,\vec{m}}$ равен m.

В примере 1 индекс (2,2) не является нормальным и не является вполне нормальным, а в примере 2 индекс (2,2) не является нормальным, но является вполне нормальным относительно задачи A для рассматриваемых в этих примерах систем функций.

Далее будет установлено, что любой нормальный индекс $(n,\vec{m})$ для f относительно задачи A является также и вполне нормальным индексом для f относительно задачи A. Пример 2 показывает, что обратное утверждение, вообще говоря, неверно.

О п р е д е л е н и е 5. Систему f назовем вполне совершенной относительно задачи A, если все индексы $(n,\vec{m}) \in \mathbb{Z}_+^{k+1}$ являются вполне нормальными для f относительно задачи A.

Отметим, что любая совершенная система f относительно задачи A является также и вполне совершенной системой относительно задачи A.

Сформулируем основную теорему этого раздела.

Т е о р е м а 1. *Для того, чтобы для фиксированного индекса* $(n,\vec{m})$ *задача A имела единственное решение, необходимо и достаточно, чтобы индекс* $(n,\vec{m})$ *был вполне нормальным для* f *относительно задачи A, т. е.* $\operatorname{rang} F_{n,\vec{m}} = m$.

В случае, если $\operatorname{rang} F_{n,\vec{m}} = m$, при определенном выборе мультипликативного множителя для решений задачи (Q_m, P) справедливы следующие представления:

$$Q_m(z) = \det\begin{bmatrix} F^1 & F^2 & \dots & F^k & E(z) \end{bmatrix}^T, \tag{7}$$

$$P^{j}_{n_j}(z) = \det\begin{bmatrix} F^1 & F^2 & \dots & F^k & E_{m_j}(z) \end{bmatrix}^T, \tag{8}$$

$$R^{j}_{n,\vec{m}}(z) = \sum_{i=1}^{\infty} d^{j}_{n,\vec{m},i} z^{n+m+i}. \tag{9}$$

Д о к а з а т е л ь с т в о. Пусть

$$Q_m(z) = b_0 + b_1 z + \dots + b_m z^m.$$

Обозначим через $(g)_k$ коэффициент при z^k степенного ряда $g(k)$. Рассмотрим систему m линейных однородных уравнений относительно m + 1 неизвестных коэффициентов $b_0, b_1, \dots, b_m$:

$$\left(Q_m f_j\right)_p = 0, \tag{10}$$

$$p = n_j + 1,\ n_j + 2, \ldots, n_j + m_j;\ j = 1, 2, \ldots, k.$$

В матричном виде система (10) выглядит так:

$$F_{n,\vec{m}} \times b^T = \theta^T, \tag{11}$$

где $b = (b_0, b_1, \ldots, b_m)$ – матрица-строка, а θ – матрица-строка порядка $1 \times (m + 1)$, все элементы которой нулевые. Поскольку система (11) является однородной и в ней число неизвестных на единицу больше числа уравнений, то из теоремы Кронекера – Капелли следует, что у системы (11) имеется ненулевое решение. Более того, множество всех линейно независимых решений системы (11) состоит из одного фундаментального решения тогда и только тогда, когда $\operatorname{rang} F_{n,\vec{m}} = m$. В этом случае все остальные ненулевые решения получаются домножением этого решения на число $\lambda \neq 0$. Первая часть теоремы 1 доказана.

Докажем теперь равенство (7). Так как ранг матрицы $F_{n,\vec{m}}$ равен m, то при некотором $p \in \{1, \ldots, m+1\}$ определитель, полученный в результате вычеркивания в матрице $F_{n,\vec{m}}$ p-го столбца, отличен от нуля. Для определенности предположим, что $p = m + 1$. Тогда в развернутом виде систему (11) можно переписать в виде

$$\begin{pmatrix} f^1_{n-m_1+1} & f^1_{n-m_1+2} & \cdots & f^1_{n_1} \\ f^1_{n-m_1+2} & f^1_{n-m_1+3} & \cdots & f^1_{n_1+1} \\ \cdots & \cdots & \cdots & \cdots \\ f^1_{n} & f^1_{n+1} & \cdots & f^1_{n+m-1} \\ \cdots & \cdots & \cdots & \cdots \\ f^k_{n-m_k+1} & f^k_{n-m_k+2} & \cdots & f^k_{n_k} \\ f^k_{n-m_k+2} & f^k_{n-m_k+3} & \cdots & f^k_{n_k+1} \\ \cdots & \cdots & \cdots & \cdots \\ f^k_{n} & f^k_{n+1} & \cdots & f^k_{n+m-1} \end{pmatrix} \begin{pmatrix} b_m \\ b_{m-1} \\ . \\ b_{n-m_1+1} \\ . \\ b_{m_k} \\ b_{m_k-1} \\ . \\ b_1 \end{pmatrix} = -b_0 \begin{pmatrix} f^1_{n_1+1} \\ f^1_{n_1+2} \\ . \\ f^1_{n+m} \\ . \\ f^k_{n_k+1} \\ f^k_{n_k+2} \\ . \\ f^k_{n+m} \end{pmatrix}. \tag{12}$$

Обозначим главный определитель системы (12) через $H_{n,\vec{m}}$. По предположению $H_{n,\vec{m}}$ не равен нулю. Если бы $b_0 = 0$, то система (12) имела бы только нулевое решение. Тогда и система (11) имела бы только нулевое решение. Поэтому $b_0 \neq 0$. Учитывая, что мы ищем решение с точностью до мультипликативного множителя, можно считать, что $b_0 = 1$. Решаем систему (12) по правилу Крамера. Пренебрегая числовым множителем, результат можно записать в виде

$$Q_m(z) = \begin{vmatrix} f^1_{n-m_1+1} & f^1_{n-m_1+2} & \cdots & f^1_{n_1+1} \\ \cdots & \cdots & \cdots & \cdots \\ f^1_{n} & f^1_{n+1} & \cdots & f^1_{n+m} \\ \cdots & \cdots & \cdots & \cdots \\ f^k_{n-m_k+1} & f^k_{n-m_k+2} & \cdots & f^k_{n_k+1} \\ \cdots & \cdots & \cdots & \cdots \\ f^k_{n} & f^k_{n+1} & \cdots & f^k_{n+m} \\ z^m & z^{m-1} & \cdots & 1 \end{vmatrix} = \det\begin{bmatrix} F^1 & \ldots & F^k & E(z) \end{bmatrix}^T.$$

Равенство (7) доказано. Справедливость равенств (8), (9) устанавливается непосредственной проверкой. Теорема 1 доказана.

Замечания и некоторые следствия. Компонента m_j мультииндекса $\vec{m}$ определяет число коэффициентов ряда f_j, которые учитываются при построении многочлена Q_m. В частности, если $m_j = 0$, то матрица $F_{n,\vec{m}}$ и определитель в (7) не содержат блока F^j и, следовательно, при построении многочлена Q_m формальный ряд f_j не участвует, а порядок мультииндекса $\vec{m}$ определяется остальными ненулевыми компонентами.

Например, если $\vec{m} = (m_1, 0, \ldots, 0) \in \mathbb{Z}_+^k$, то $m = m_1$, и тогда, как и в одномерном случае, при нахождении Q_m учитываются только коэффициенты ряда f_1. При этом представление (7) совпадает с (3).

В том случае, если $\vec{m}$ – нулевой индекс, из равенств (7), (8) получаем, что $Q_m(z) \equiv 1$, а P_n^j – многочлен Тейлора функции f_j. Отсюда, в частности, следует, что если система f является совершенной относительно задачи А, то все коэффициенты рядов (1) не равны нулю. Например, если одно из чисел $\{\lambda_p\}_{p=1}^k$ равно нулю, то система экспонент $\left\{e^{\lambda_p z}\right\}_{p=1}^k$ не является совершенной системой относительно задачи А.

Следует также сказать, что если индекс $(n,\vec{m})$ не является вполне нормальным для f относительно задачи А, то многочлены Q_m и $P_{n_j}^j$, определенные равенствами (7) и (8), не являются решениями задачи А. В частности, в примере 1 для индекса (2,2) искомый многочлен $Q_2(z) = a + bz - (4a + 2b)z^2$. Однако, если Q_2 находить по формуле (3), то получим, что $Q_2(z) \equiv 0$. Как уже отмечалось, представление многочлена Паде в виде (3) вытекает из общего представления многочленов Эрмита – Паде (7), поэтому оно также справедливо только в том случае, когда индекс (n, m) является вполне нормальным относительно задачи А. В монографии [2] при доказательстве теоремы 1.1.1 на это обстоятельство внимание не обращено (см. [2, гл. 1, § 1.1, теорема 1.1.1]).

Из (7) и (8) вытекает следующий критерий нормальности индекса $(n,\vec{m})$.

С л е д с т в и е 1. *Индекс $(n,\vec{m}) \in \mathbb{Z}_+^k$ будет нормальным для f относительно задачи А тогда и только тогда, когда*

$$H_{n,\vec{m}} \cdot \prod_{j=1}^{k} H_{n,\vec{m}}^{j} \neq 0, \tag{13}$$

где

$$H_{n,\vec{m}}^{j} = \begin{vmatrix} f_{n-m_1+1}^1 & f_{n-m_1+2}^1 & \cdots & f_{n_1+1}^1 \\ \cdots & \cdots & \cdots & \cdots \\ f_n^1 & f_{n+1}^1 & \cdots & f_{n+m}^1 \\ \cdots & \cdots & \cdots & \cdots \\ f_{n-m_k+1}^k & f_{n-m_k+2}^k & \cdots & f_{n_k+1}^k \\ \cdots & \cdots & \cdots & \cdots \\ f_n^k & f_{n+1}^k & \cdots & f_{n+m}^k \\ f_{n-m_j}^j & f_{n-m_j+1}^j & \cdots & f_{n_j}^j \end{vmatrix}.$$

В частности, при $k = 1$ получим критерий нормальности индекса (n, m), совпадающий с (4). При $m_j = 0$ в (10) предполагается, что определители $H_{n,\vec{m}}$, $H_{n,\vec{m}}^j$ не содержат блок F^j.

Следующее следствие можно рассматривать как некоторый аналог теоремы Кронекера [1].

С л е д с т в и е 2. *Пусть индекс $n = (n,\vec{m})$ является вполне нормальным для $f = \{f_1, \ldots, f_k\}$ относительно задачи А и $m_j \neq 0$. Тогда для того, чтобы функция f_j была рациональной, необходимо и достаточно, чтобы $d_{n,\vec{m},i}^j = 0$ для всех достаточно больших i.*

2. Многочлены Эрмита – Паде 1-го рода. ***Постановка задачи.*** Рассмотрим задачу Эрмита – Паде, двойственную задаче А (см. [1, гл. 4, § 3; 3; 4]).

З а д а ч а В. Для системы f степенных рядов (1) и ненулевого индекса $n=(n_1,...,n_k)\in\mathbb{Z}_+^k$ найти такой набор не равных тождественно нулю одновременно многочленов $A_1=A_n^1,...,A_k=A_n^k$, $\deg A_1\leqslant n_1-1,\ldots,\deg A_k\leqslant n_k-1,$ для которых

$$L_n(z):=\sum_{j=1}^{k}A_j(z)f_j(z)=c_nz^{|n|-1}+\ldots, \tag{14}$$

где по определению $|n|=n_1+\ldots+n_k$.

В случае, когда $k=2$, $(n_1,n_2)=(n+1,m+1)$, $f=(f_1,1)$, задача А совпадает с задачей В. Вследствие этого многочлен A_1 (как и A_2) является многочленом Паде и в предположении, что индекс $(n+1,m+1)$ является вполне нормальным для f относительно задачи А, он, как и Q_m, с точностью до множителя определяется равенством (3).

Если f является набором экспонент $\{e^{\lambda_j z}\}_{j=1}^k$, где λ_j – различные комплексные числа, то решение задачи В в явном виде получено в [6]. В этой работе искомые многочлены представлены Эрмитом в виде интегралов Коши.

В общем случае решение задачи В существует, но не единственно [1, гл. 4, § 1]. В частности, многочлены A_j также находятся с точностью до мультипликативного множителя: если набор $A=(A_1,\ldots,A_k)$ удовлетворяет необходимым условиям, то для любого отличного от нуля комплексного числа λ набор $\lambda A=(\lambda A_1,\ldots,\lambda A_k)$ также удовлетворяет условиям задачи. Эта неединственность может быть и более существенной.

П р и м е р 3. Пусть $k=2$, $n=(3,3)$, а $f=(f_1,1)$, где f_1 – функция из примера 1, тогда любое решение задачи В представимо в виде $(\lambda A_1,\lambda A_2)$, $\lambda\in\mathbb{C}$, $\lambda\neq0$,

$$A_1(z)=a+bz-(4a+2b)z^2,\quad A_2(z)=-\frac{1}{2}a-\left(a+\frac{1}{2}b\right)z,$$

где a, b – произвольные действительные числа, не равные нулю одновременно.

Принято говорить [1], что задача В имеет единственное решение, если все решения задачи можно записать в виде λA, где $\lambda\in\mathbb{C}$, $\lambda\neq0$, а $A=(A_1,\ldots,A_k)$ – некоторое одно фиксированное решение.

О п р е д е л е н и е 6. Если $A=(A_1,...,A_k)$ – решение задачи В с ненулевым индексом $n\in\mathbb{Z}_+^k$, то многочлены $A_1,...,A_k$ называются многочленами Эрмита – Паде 1-го рода (Latin Type) для набора (системы) f степенных рядов (1). Центральными понятиями в теории таких многочленов также являются понятия нормального индекса и совершенной системы [1, гл. 4, § 1].

О п р е д е л е н и е 7. Ненулевой индекс $n=(n_1,...,n_k)\in\mathbb{Z}_+^k$ называется нормальным для f относительно задачи В, если для любого решения задачи В с индексом n

$$\deg A_j=n_j-1,\ \ j=1,...,k.$$

О п р е д е л е н и е 8. Систему f назовем совершенной относительно задачи В, если все ненулевые индексы $n\in\mathbb{Z}_+^k$ являются нормальными для f относительно задачи В.

При $f=(f_1,1)$, $k=2$, критерий нормальности индекса $(n_1,n_2)=(m+1,n+1)$ выражается условием (4).

Известно [1], что если индекс n является нормальным для f относительно задачи В, то последняя имеет единственное решение. Следующий пример показывает, что уже при $k=2$ нормальность индекса n не является необходимым условием единственности решения задачи В.

П р и м е р 4. Пусть $k=2$, $n=(3,3)$, а $f=(f_1,1)$, где f_1 – функция из примера 2. Тогда любое решение задачи В можно записать в виде $(\lambda A_1,\lambda A_2)$, $\lambda\in\mathbb{C}$, $\lambda\neq0$, где

$$A_1(z)=8-4z,\quad A_2(z)=-4-2z+z^2.$$

В этом примере индекс $n=(3,3)$ не является нормальным, так как $\deg A_1=1$.

Нашей ближайшей целью является нахождение необходимых и достаточных условий на индекс $n \in \mathbb{Z}_+^k$ и систему f, определяемую равенствами (1), при которых решение задачи B является единственным.

Критерий единственности решения задачи B. Введем необходимые обозначения. Для каждого $j = 1,...,k$ и индекса $n = (n_1,...,n_k) \in \mathbb{Z}_+^k$ в предположении, что $n_j \neq 0$, определим матрицы-столбцы порядка $(|n|-1)\times 1$

$$G_i^j = \begin{pmatrix} f_{1-i}^j & f_{2-i}^j & \dots & f_{|n|-i-1}^j \end{pmatrix}^T,\ i = 1,...,n_j,$$

матрицы порядка $(|n|-1)\times n_j$

$$G^j = \begin{pmatrix} G_1^j & G_2^j & \dots & G_{n_j}^j \end{pmatrix},$$

и матрицу порядка $(|n|-1)\times |n|$

$$G_n = \begin{pmatrix} G^1 & G^2 & \dots & G^k \end{pmatrix}.$$

В случае, если компонента индекса $n_j = 0$, матрица G_n не содержит блок G^j.

Рассмотрим также функциональные матрицы-строки порядка $1\times |n|$

$$U_1(z) = \begin{pmatrix} 1 & z & \dots & z^{n_1-1} & 0\ 0 \dots 0 & \dots & 0\ 0 \dots 0 \end{pmatrix},$$

$$\dots \quad \dots \quad \dots \quad .. \quad \dots \quad \dots \quad ..$$

$$U_k(z) = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 1 & z & \dots & z^{n_k-1} \end{pmatrix},$$

$$U(z) = U_1(z) + \dots + U_k(z).$$

Если в матрице G_n добавить в качестве последней строки строку $U_j(z)$, то получим квадратную матрицу порядка $|n|\times|n|$. Определитель этой матрицы обозначим через $A_j(z)$. Тогда при $j = 1,2,\dots,k$

$$A_j(z) = \begin{vmatrix}
f_0^1 & \cdot\cdot & 0 & \cdot\cdot & f_0^j & 0 & \cdot\cdot & 0 & \cdot\cdot & f_0^k & \cdot\cdot & 0 \\
f_1^1 & \cdot\cdot & 0 & \cdot\cdot & f_1^j & f_0^j & \cdot\cdot & 0 & \cdot\cdot & f_1^k & \cdot\cdot & 0 \\
. & \cdot\cdot & . & \cdot\cdot & . & . & \cdot\cdot & . & \cdot\cdot & . & \cdot\cdot & . \\
f_{n_j-1}^1 & \cdot\cdot & f_{n_j-n_1}^1 & \cdot\cdot & f_{n_j-1}^j & f_{n_j-2}^j & \cdot\cdot & f_0^j & \cdot\cdot & f_{n_j-1}^k & \cdot\cdot & f_{n_j-n_k}^k \\
f_{n_j}^1 & \cdot\cdot & f_{n_j-n_1+1}^1 & \cdot\cdot & f_{n_j}^j & f_{n_j-1}^j & \cdot\cdot & f_1^j & \cdot\cdot & f_{n_j}^k & \cdot\cdot & f_{n_j-n_k+1}^k \\
. & \cdot\cdot & . & \cdot\cdot & . & . & \cdot\cdot & . & \cdot\cdot & . & \cdot\cdot & . \\
f_{|n|-2}^1 & \cdot\cdot & f_{|n|-n_1-1}^1 & \cdot\cdot & f_{|n|-2}^j & f_{|n|-3}^j & \cdot\cdot & f_{|n|-n_j-1}^j & \cdot\cdot & f_{|n|-2}^k & \cdot\cdot & f_{|n|-n_k-1}^k \\
0 & \cdot\cdot & 0 & \cdot\cdot & 1 & z & \cdot\cdot & z^{n_j-1} & \cdot\cdot & 0 & \cdot\cdot & 0
\end{vmatrix}. \qquad (15)$$

Если в определителе (15) последнюю строку заменить матрицей-строкой

$$\begin{pmatrix} f_{i+|n|-2}^1 & f_{i+|n|-3}^1 & \dots & f_{i+|n|-n_1-1}^1 & \dots & f_{i+|n|-2}^k & f_{i+|n|-3}^k & \dots & f_{i+|n|-n_k-1}^k \end{pmatrix},$$

то полученный определитель обозначим через $\widetilde{d}_{n,i}$.

Определение 9. Ненулевой индекс $n \in \mathbb{Z}_+^k$ назовем вполне нормальным для f относительно задачи B, если ранг матрицы G_n равен $|n|-1$.

В примере 3 индекс $n = (3,3)$ не является нормальным и не является вполне нормальным, а в примере 4 – не является нормальным, но является вполне нормальным относительно задачи B для рассматриваемых в этих примерах систем функций.

О п р е д е л е н и е 10. Систему f назовем вполне совершенной относительно задачи B, если все ненулевые индексы $n \in \mathbb{Z}_+^k$ являются вполне нормальными для f относительно задачи B.

Далее будет установлено, что любой нормальный индекс n для f относительно задачи B является также и вполне нормальным индексом для f относительно задачи B. Вследствие этого любая совершенная система f относительно задачи B является также и вполне совершенной системой относительно задачи B.

Т е о р е м а 2. *Для того чтобы для ненулевого индекса* $n \in \mathbb{Z}_+^k$ *задача B имела единственное решение, необходимо и достаточно, чтобы индекс n был вполне нормальным для f относительно задачи B, т. е.* $\operatorname{rang} G_n = |n| - 1$. *В случае, если* $\operatorname{rang} G_n = |n| - 1$, *при определенном выборе мультипликативного множителя справедливы следующие представления*:

$$A_j(z) = \det\begin{bmatrix} G_n \\ U_j(z) \end{bmatrix}, \quad j = 1,...,k, \tag{16}$$

$$L_n(z) = \sum_{i=1}^{\infty} \widetilde{d}_{n,i}\, z^{|n|+i-2}. \tag{17}$$

Д о к а з а т е л ь с т в о. Пусть $n = (n_1, ..., n_k) \in \mathbb{Z}_+^k$, $n \neq 0$, а

$$A_j(z) = b_0^j + b_1^j z + \ldots + b_{n_j-1}^j z^{n_j-1}, \quad j = 1,\ldots,k.$$

Опираясь на равенство (14), запишем в матричной форме систему уравнений для определения коэффициентов многочлена $A_j(z)$:

$$G_n \times b^T = \theta^T, \tag{18}$$

где $b = (b_0^1\ b_1^1\ \ldots\ b_{n_1-1}^1 \ldots b_0^k\ b_1^k\ \ldots\ b_{n_k-1}^k)$ – матрица-строка порядка $1 \times |n|$ (при $n_j = 0$ матрица b не содержит элементов $b_0^j, \ldots, b_{n_j-1}^j$), а θ – матрица порядка $1 \times |n|$, все элементы которой равны нулю. Система линейных уравнений (18) является однородной, и в ней число неизвестных на единицу больше числа уравнений. Поэтому из теоремы Кронекера – Капелли следует, что система (18) имеет ненулевое решение, а множество всех ее линейно независимых решений состоит из одного фундаментального решения тогда и только тогда, когда $\operatorname{rang} G_n = |n| - 1$. В этом случае все остальные ненулевые решения получаются домножением этого решения на комплексное число $\lambda \neq 0$. Первая часть теоремы 2 доказана.

Докажем теперь равенство (16). Так как $\operatorname{rang} G_n = |n| - 1$, то при некотором $p \in \{1, 2, \ldots, |n|\}$ определитель, полученный из матрицы G_n вычеркиванием в ней p-го столбца , отличен от нуля. Для определенности предположим, что $p = |n|$. Тогда систему (18) можно записать следующим образом:

$$\begin{pmatrix} f_0^1 & 0 & \cdot\cdot & 0 & \cdot\cdot & f_0^k & 0 & \cdot\cdot & 0 \\ f_1^1 & f_0^1 & \cdot\cdot & 0 & \cdot\cdot & f_1^k & f_0^k & \cdot\cdot & 0 \\ \cdot & \cdot & \cdot\cdot & \cdot & \cdot\cdot & \cdot & \cdot & \cdot\cdot & \cdot \\ f_{n_k-1}^1 & f_{n_k-2}^1 & \cdot\cdot & f_{n_k-n_1}^1 & \cdot\cdot & f_{n_k-1}^k & f_{n_k-2}^k & \cdot\cdot & f_1^k \\ f_{n_k}^1 & f_{n_k-1}^1 & \cdot\cdot & f_{n_k-n_1+1}^1 & \cdot\cdot & f_{n_k}^k & f_{n_k-1}^k & \cdot\cdot & f_2^k \\ \cdot & \cdot & \cdot\cdot & \cdot & \cdot\cdot & \cdot & \cdot & \cdot\cdot & \cdot \\ f_{|n|-3}^1 & f_{|n|-4}^1 & \cdot\cdot & f_{|n|-n_k-2}^1 & \cdot\cdot & f_{|n|-3}^k & f_{|n|-4}^k & \cdot\cdot & f_{|n|-n_k-1}^k \\ f_{|n|-2}^1 & f_{|n|-3}^1 & \cdot\cdot & f_{|n|-n_k-1}^1 & \cdot\cdot & f_{|n|-2}^k & f_{|n|-3}^k & \cdot\cdot & f_{|n|-n_k}^k \end{pmatrix} \begin{pmatrix} b_0^1 \\ \vdots \\ b_{n_k-1}^1 \\ \vdots \\ b_0^k \\ \vdots \\ b_{n_k-2}^k \end{pmatrix} = -b_{n_k-1}^k \begin{pmatrix} 0 \\ 0 \\ \vdots \\ f_0^k \\ f_1^k \\ \vdots \\ f_{|n|-n_k-1}^k \end{pmatrix}. \tag{19}$$

Обозначим главный определитель системы (19) через $\widetilde{H}_n^{n_k}$. По предположению $\widetilde{H}_n^{n_k} \neq 0.$ Если бы $b_{n_k-1}^k = 0,$ то система (19) имела бы единственное нулевое решение. Тогда бы и система (18) имела только нулевое решение. Следовательно, $b_{n_k-1}^k \neq 0.$ Решая систему (19) по правилу Крамера, получим решение, которое символически можно записать в виде

$$\det\begin{bmatrix} G_n & U(z) \end{bmatrix}^T = A_1(z) + \ldots + A_k(z), \tag{20}$$

где $A_j(z)$ определяются равенствами (16). В случае, если бы вместо $p = |n|$ вычеркивали столбец матрицы G_n с другим номером, рассуждая аналогичным образом, пришли бы к символической записи решения в виде (20).

Докажем, что многочлены $A_j(z)$, определенные равенствами (16), действительно являются искомыми многочленами. Разложив определитель в (16) по элементам последней строки, получим, что $\deg A_j(z) \leqslant n_j - 1.$ Непосредственная проверка показывает, что для таких многочленов

$$L_n(z) = \begin{vmatrix} f_0^1 & \cdot\cdot & 0 & \cdot\cdot & f_0^k & \cdot\cdot & 0 \\ f_1^1 & \cdot\cdot & 0 & \cdot\cdot & f_1^k & \cdot\cdot & 0 \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\ f_{|n|-2}^1 & & f_{|n|-n_1-1}^1 & \cdot\cdot & f_{|n|-2}^k & \cdot\cdot & f_{|n|-n_k-1}^k \\ \sum\limits_{i=|n|-1}^{\infty} f_i^1 z^i & \cdot\cdot & \sum\limits_{i=|n|-1}^{\infty} f_{i-n_1+1}^1 z^i & \cdot\cdot & \sum\limits_{i=|n|-1}^{\infty} f_i^k z^i & \cdot\cdot & \sum\limits_{i=|n|-1}^{\infty} f_{i-n_k+1}^k z^i \end{vmatrix} = \sum_{i=1}^{\infty} \widetilde{d}_{n,i} z^{|n|+i-2}.$$

Теорема 2 доказана.

Замечания и некоторые следствия. Из теоремы 2 следует, что n_j-я компонента вполне нормального индекса $n = (n_1, \ldots, n_k) \in \mathbb{Z}_+^k$ определяет число коэффициентов ряда f_j, которое учитывается при построении многочленов $\{A_p\}_{p=1}^k$. В частности, если $n_j = 0$, то в матрице G_n отсутствует блок G^j и, следовательно, коэффициенты ряда f_j не учитываются, многочлен $A_j(z) \equiv 0,$ а порядок мультииндекса n определяется остальными ненулевыми компонентами.

Например, если $n = (n_1, n_2, 0, \ldots, 0),$ то при построении многочленов учитываются только коэффициенты рядов f_1, f_2, и если $f_2(z) \equiv 1,$ то, например, многочлен A_1 согласно равенству (16) с точностью до мультипликативного множителя представим в виде

$$A_1(z) = \begin{vmatrix} f_{n_2}^1 & f_{n_2-1}^1 & \cdots & f_{n_2-n_1+1}^1 \\ f_{n_2+1}^1 & f_{n_2}^1 & \cdots & f_{n_2-n_1+2}^1 \\ \cdots & \cdots & \cdots & \cdots \\ f_{n_2+n_1-2}^1 & f_{n_2+n_1-3}^1 & \cdots & f_{n_2-1}^1 \\ 1 & z^2 & \ldots & z^{n_1-1} \end{vmatrix}.$$

Полученное представление многочлена A_1 полностью согласуется с равенством Паде (3). Кроме того, в этом случае $A_j(z) \equiv 0$ при $j = 3, \ldots, k.$

Если же $n = (n_1, 0, \ldots, 0),$ то из (16) получаем, что $A_1(z) = f_0^1 z^{n_1 - 1},$ а $A_j(z) \equiv 0$ при $j = 2, \ldots, k.$ Отсюда, в частности, следует, что если система f совершенна относительно задачи B, то $f_0^j \neq 0,$ $j = 1, \ldots, k.$

Следует также сказать, что если индекс n не является вполне нормальным для f относительно задачи B, то многочлены A_j, определенные равенствами (16), не являются решениями задачи B, так как все они тождественно равны нулю. В частности, в примере 3 $A_1(z) = a + bz - (4a + 2b)z^2.$ Однако если A_1 находить по формуле (16), то получим, что $A_1(z) \equiv 0.$

Из (16) вытекает следующий критерий нормальности индекса *n* для системы *f* относительно задачи B.

С л е д с т в и е 3. *Ненулевой индекс* $n = (n_1, ..., n_k) \in \mathbb{Z}_+^k$ *будет нормальным для f относительно задачи B тогда и только тогда, когда*

$$\prod_{j=1}^{k} \widetilde{H}_n^{n_j} \neq 0, \tag{21}$$

где $\widetilde{H}_n^{n_j}$ *– определитель, полученный из определителя* (15) *вычеркиванием в нем последней строки и столбца, в котором находится элемент* $z^{n_j - 1}$, *при этом, если* $n_j = 0$, *либо* $n_j = 1$, *то в определителе* $\widetilde{H}_n^{n_j}$ *отсутствует блок* G^j.

Если $k = 2$, $f = (f_1, 1)$ и $(n_1, n_2) = (n + 1, m + 1)$, то (21) равносильно условию $H_{n,m} \cdot H_{n,m+1} \neq 0$, поэтому следствие 3 согласуется с критерием нормальности индекса (4).

С л е д с т в и е 4. *Если система* $f = (f_1, ..., f_k)$ *совершенна относительно задачи B, то для любого ненулевого индекса* $n \in \mathbb{Z}_+^k$ *и решения* $A = (A_1, ..., A_k)$ *задачи B с этим индексом в равенстве* (14) *коэффициент* $c_n \neq 0$.

Наряду с системой $f = (f_1, ..., f_k)$ рассмотрим расширенную систему функций $\overline{f} = (1, f_1, ..., f_k)$.

С л е д с т в и е 5. *Система f является совершенной относительно задачи A тогда и только тогда, когда расширенная система* $\overline{f}$ *является совершенной относительно задачи B.*

Из следствия 5 и сделанных ранее замечаний можно сделать вывод о том, что если система $\overline{f}$ является совершенной относительно задачи B, то все коэффициенты степенных рядов (1) не равны нулю.

Следствие 5 не в полной мере является новым. Утверждение, что система $\overline{f}$ является совершенной относительно задачи B, когда система *f* является совершенной относительно задачи A, без подробного доказательства приведено в монографии [1, гл. 4, § 1, утверждение 1.3].

Список использованных источников

1. Никишин, Е. М. Рациональные аппроксимации и ортогональность / Е. М. Никишин, В. Н. Сорокин. – М.: Наука, 1988. – 256 с.
2. Бейкер, Дж. мл. Аппроксимации Паде. 1. Основы теории. 2. Обобщения и приложения / Дж. Бейкер мл., П. Грейвс-Моррис. – М.: Мир, 1986. – 502 с.
3. Stahl, H. Asymptotics for quadratic Hermite-Padé polynomals associated with the exponential function / H. Stahl // Electronic Trans. Num. Anal. – 2002. – № 14. – P. 193–220.
4. Аппроксимации Паде, непрерывные дроби и ортогональные многочлены / А. И. Аптекарев [и др.] // Успехи мат. наук. – 2011. – Т. 66, № 6 (402). – С. 37–122.
5. Hermite, C. Sur la function exponentielle / C. Hermite // C. R. Akad. Sci. – 1873. – Vol. 77. – P. 18–293.
6. Hermite, C. Sur la generalisation des fractions continues algebriques / C. Hermite // Ann. Math. Pura. Appl. – 1893. – Vol. 11, № 1. – P. 289–308. https://doi.org/10.1007/bf02420446

References

1. Nikishin E. M., Sorokin V. N. *Rational Approximations and Orthogonality.* Moscow, Nauka Publ., 1988. 256 p. (in Russian).
2. Baker Jr. G. A., Graves-Morris P. *Pad´e Approximants.* 2nd ed. Cambridge University Press, 1996. https://doi.org/10.1017/CBO9780511530074
3. Stahl H. Asymptotics for quadratic Hermite-Padé polynomals associated with the exponential function. *Electronic Transactions on Numerical Analysis,* 2002, no. 14, pp. 193–220.
4. Aptekarev A. I., Buslaev V. I., Mart´ınez-Finkelshtein A., and Suetin S. P. Pad´e approximants, continued fractions, and orthogonal polynomials. *Russian Mathematical Surveys*, 2011, vol. 66, no. 6, pp. 1049–1131. https://doi.org/10.1070/rm2011v066n06abeh004770
5. Hermite C. Sur la function exponentielle. *Comptes Rendus de l'Académie des Sciences*, 1873, vol. 77, pp. 18–293.
6. Hermite C. Sur la generalisation des fractions continues algebriques. *Annali di Matematica Pura ed Applicata*, 1893, vol. 21, no. 1, pp. 289–308. https://doi.org/10.1007/bf02420446

Информация об авторах

Старовойтов Александр Павлович – доктор физико-математических наук, профессор, профессор кафедры математического анализа и дифференциальных уравнений, Гомельский государственный университет им. Ф. Скорины (ул. Советская, 104, 246019, г. Гомель, Республика Беларусь). E-mail: svoitov@gsu.by; apsvoitov@gmail.com

Рябченко Наталия Валерьевна – старший преподаватель кафедры математического анализа и дифференциальных уравнний, Гомельский государственный университет им. Ф. Скорины (ул. Советская, 104, 246019, г. Гомель, Республика Беларусь). E-mail: nmankevich@tut.by

Information about the authors

Aleksandr P. Staravoitov – Dr. Sc. (Physics and Mathematics), Professor, Francisk Scorina Gomel State University (104, Sovetskaya Str., 246019, Gomel, Republic of Belarus). E-mail: svoitov@gsu.by, apsvoitov@gmail.com

Nataliya V. Ryabchenko – Senior Lecturer, Francisk Scorina Gomel State University (104, Sovetskaya Str., 246019, Gomel, Republic of Belarus). E-mail: nmankevich@tut.by

ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)
UDC 519.237
https://doi.org/10.29235/1561-2430-2019-55-4-457-466

Received 24.10.2019
Поступила в редакцию 24.10.2019

V. S. Mukha, N. F. Kako

Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus

INTEGRALS AND INTEGRAL TRANSFORMATIONS RELATED TO THE VECTOR GAUSSIAN DISTRIBUTION

Abstract. This paper is dedicated to the integrals and integral transformations related to the probability density function of the vector Gaussian distribution and arising in probability applications. Herein, we present three integrals that permit to calculate the moments of the multivariate Gaussian distribution. Moreover, the total probability formula and Bayes formula for the vector Gaussian distribution are given. The obtained results are proven. The deduction of the integrals is performed on the basis of the Gauss elimination method. The total probability formula and Bayes formula are obtained on the basis of the proven integrals. These integrals and integral transformations could be used, for example, in the statistical decision theory, particularly, in the dual control theory, and as table integrals in various areas of research. On the basis of the obtained results, Bayesian estimations of the coefficients of the multiple regression function are calculated.

Keywords: vector Gaussian distribution, multidimensional integrals, total probability formula, Bayes formula, multiple regression function, Bayesian estimations

For citation. Mukha V. S., Kako N. F. Integrals and integral transformations related to the vector Gaussian distribution. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2019, vol. 55, no. 4, pp. 457–466. https://doi.org/10.29235/1561-2430-2019-55-4-457-466

В. С. Муха, Н. Ф. Како

Белорусский государственный университет информатики и радиоэлектроники, Минск, Беларусь

ИНТЕГРАЛЫ И ИНТЕГРАЛЬНЫЕ ПРЕОБРАЗОВАНИЯ, СВЯЗАННЫЕ С ВЕКТОРНЫМ ГАУССОВСКИМ РАСПРЕДЕЛЕНИЕМ

Аннотация. Рассматриваются интегралы и интегральные преобразования, относящиеся к функции плотности вероятности векторного гауссовского распределения и возникающие в вероятностных приложениях. Представлены три интеграла, позволяющие рассчитывать моменты векторного гауссовского распределения, а также формулы полной вероятности и Байеса. Приводятся доказательства полученных результатов. Вывод интегралов выполнен на основе метода исключения Гаусса. Формулы полной вероятности и Байеса получены на основе доказанных интегралов. Представленные интегралы и интегральные преобразования могут быть использованы в различных вероятностных приложениях, например в теории статистических решений, в частности, в теории дуального управления, а также как табличные интегралы в различных областях исследований. На основе полученных результатов рассчитаны байесовские оценки коэффициентов множественной функции регрессии.

Ключевые слова: векторное гауссовское распределение, многомерные интегралы, формула полной вероятности, формула Байеса, множественная функция регрессии, байесовские оценки

Для цитирования. Муха, В. С. Интегралы и интегральные преобразования, связанные с векторным гауссовским распределением / В. С. Муха, Н. Ф. Како // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2019. – Т. 55, № 4. – С. 457–466. https://doi.org/10.29235/1561-2430-2019-55-4-457-466

Introduction. Integrals and integral transformations related to probability distributions have many applications, one of them being the statistical decision theory. The latter attracts much attention due to its capacity to formulate problems in a strict mathematical form. One of the technical problems solved by the statistical decision theory is the problem of dual control [1], requiring the calculation of integrals related to multivariate probability distributions. In this paper, we present three integrals connected with the vector Gaussian distribution and the total probability formula and Bayes formula for the vector Gaussian distribution.

1. Integrals connected with the vector Gaussian distribution. A random vector with k components $\Xi^T = (\Xi_1, \Xi_2, ..., \Xi_k)$ is distributed according to the normal or Gaussian law if its probability density has the form

$$f(\xi) = \frac{1}{\sqrt{(2\pi)^k \,|d_\Xi|}} \exp\left(-\frac{1}{2}(\xi - \nu_\Xi)^T d_\Xi^{-1}(\xi - \nu_\Xi)\right), \quad \xi \in E^k, \tag{1}$$

where $\xi^T = (\xi_1, \xi_2, ..., \xi_k)$ is the vector-row of the arguments of the probability density $f(\xi)$, $\nu_\Xi^T = (\nu_{\Xi,1}, \nu_{\Xi,2}, ..., \nu_{\Xi,k})$ is the vector-row of the parameters of the probability density $f(\xi)$, $d_\Xi = (d_{\Xi,i,j})$, $i, j = \overline{1,k}$, is the symmetric positive-definite matrix of the parameters of the probability density $f(\xi)$, d_Ξ^{-1} is the matrix inverse to the matrix d_Ξ, $|d_\Xi|$ is the determinant of the matrix d_Ξ, E^k is the k-dimensional Euclidean space, and the T symbol stands for *transpose*. The parameters ν_Ξ and d_Ξ of distribution (1) are mathematical expectation and dispersion (variance-covariance) matrices of the random vector Ξ, respectively [2].

The following equalities for function (1) are true:

$$\int_{E^k} \exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi\right) d\xi = \sqrt{(2\pi)^k \,|A^{-1}|} \exp\left(\frac{1}{2}B^T A^{-1} B\right), \tag{2}$$

$$\int_{E^k} C^T\xi \exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi\right) d\xi = \sqrt{(2\pi)^k \,|A^{-1}|} \exp\left(\frac{1}{2}B^T A^{-1} B\right) C^T A^{-1} B, \tag{3}$$

$$\int_{E^k} \xi^T U\xi \exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi\right) d\xi = \sqrt{(2\pi)^k \,|A^{-1}|} \left(\mathrm{Tr}(A^{-1}U) + B^T A^{-1} U A^{-1} B\right) \exp\left(\frac{1}{2}B^T A^{-1} B\right), \tag{4}$$

where $A = (a_{i,j})$, $i, j = \overline{1,k}$, is the symmetric positive-definite matrix, $\xi^T = (\xi_1, \xi_2, ..., \xi_k)$, $B^T = (b_1, b_2, ..., b_k)$, $C^T = (c_1, c_2, ..., c_k)$ are the vector-rows, A^{-1} is the matrix inverse to the matrix A, $|A^{-1}|$ is the determinant of the matrix A^{-1}, $U = (u_{i,j})$, $i, j = \overline{1,k}$, is the symmetric positive-definite matrix, and $\mathrm{Tr}(A^{-1}U)$ is the trace of the matrix $A^{-1}U$.

Integrals (2)–(4) were received in [3], but proof was given only for equality (2). Now we give more detailed proof of equality (2) and prove equalities (3) and (4).

P r o o f o f e q u a l i t y (2). We will use the Gauss elimination method [4], which permits to bring the matrix A to the diagonal form and reduce the calculation of the multiple integral to the calculation of the repeated integral. The conditions for using the Gauss elimination method in this case are held because the matrix A is positive-definite.

We denote the integrand function $F(\xi)$ in (2):

$$F(\xi) = \exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi\right). \tag{5}$$

By applying the Gauss elimination method [4] to the matrix $A = (a_{i,j})$ in (5), we receive the upper triangular matrix

$$G = \begin{pmatrix} a_{1,1}^{(0)} & a_{1,2}^{(0)} & \cdots & a_{1,k}^{(0)} \\ 0 & a_{2,2}^{(1)} & \cdots & a_{2,k}^{(1)} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{k,k}^{(k-1)} \end{pmatrix} = \left(a_{i,j}^{(i-1)}\right), \quad i, j = \overline{1,k}.$$

The determinant of the upper triangular matrix G is equal to the product of its diagonal elements and also to the determinant of the matrix A [4]:

$$|G| = a_{1,1}^{(0)} a_{2,2}^{(1)} \cdots a_{k,k}^{(k-1)} = |A| = \prod_{i=1}^{k} a_{i,i}^{(i-1)}.$$

The mine minors of the matrices A and G are equal, $A_m = G_m$, $m = 1,2,...,k$ [4], where $A_m = \begin{pmatrix} 1,2,...,m \\ 1,2,...,m \end{pmatrix}$ is the main minor of order m of the matrix A and $G_m = \begin{pmatrix} 1,2,...,m \\ 1,2,...,m \end{pmatrix}$ is the main minor of order m of the matrix G. Moreover [4],

$$a_{1,1}^{(0)} = A_1,\ a_{2,2}^{(1)} = \frac{A_2}{A_1},\ a_{3,3}^{(2)} = \frac{A_3}{A_2}, ...,\ a_{k,k}^{(k-1)} = \frac{A_k}{A_{k-1}}. \tag{6}$$

As shown in [4], the matrix A can be represented in the form

$$A = G_l^T \widehat{D}_1 G, \tag{7}$$

where G and G_l are matrices received by the Gauss elimination method from the matrices A and A^T, respectively, and $\widehat{D}_1$ is the diagonal matrix represented in the form

$$\widehat{D}_1 = \text{diag}\left\{\frac{1}{A_1}, \frac{A_1}{A_2}, \frac{A_2}{A_3}, ..., \frac{A_{k-1}}{A_k}\right\}.$$

Since the matrix A is symmetrical, i. e. $A^T = A$, then $G_l = G$, and equation (7) can be written in the form:

$$A = G^T \widehat{D}_1 G.$$

Having expression (6), we get that the matrix A can be represented in the following form:

$$A = G^T \widehat{D} G, \tag{8}$$

where $\widehat{D}$ is a diagonal matrix

$$\widehat{D} = \text{diag}\left\{\left(a_{1,1}^{(0)}\right)^{-1}, \left(a_{2,2}^{(1)}\right)^{-1}, ..., \left(a_{k,k}^{(k-1)}\right)^{-1}\right\}. \tag{9}$$

From (8) it follows that

$$\widehat{D}^{-1} = G A^{-1} G^T. \tag{10}$$

We denote as $\widehat{d}_{i,i} = (a_{i,i}^{(i-1)})^{-1}$, $i = \overline{1,k}$, the diagonal elements of the diagonal matrix $\widehat{D}$ (9). Meanwhile,

$$|\widehat{D}^{-1}| = |G| = |A| = \prod_{i=1}^{k} a_{i,i}^{(i-1)}. \tag{11}$$

Let us introduce in (5) the linear replacement of the variables

$$\xi = G^{-1} z, \tag{12}$$

with which the integrand function (5) is converted to the following function of the z argument:

$$F(z)=\exp\left(-\frac{1}{2}z^T Pz+Dz\right),$$

where

$$P=(G^{-1})^T AG^{-1},$$

$$D=B^T G^{-1}=(d_i),\ \ i=\overline{1,k}. \tag{13}$$

Since, provided (8),

$$P=(G^{-1})^T AG^{-1}=(G^{-1})^T G^T \widehat{D}GG^{-1}=\widehat{D},$$

we have the following function of the z argument:

$$F(z)=\exp\left(-\frac{1}{2}z^T \widehat{D}z+Dz\right). \tag{14}$$

The following equality is obtained after replacing the variables in integral [5]:

$$\int_{E^k} F(\xi)d\xi=\int_{E^k} F(z)\,|\,J\,|\,dz, \tag{15}$$

where $|\,J\,|$ is the absolute value of the Jacobian of transformation (12):

$$|\,J\,|=|\,G^{-1}\,|=|\,\widehat{D}\,|=\prod_{i=1}^{k}(a_{i,i}^{(i-1)})^{-1}.$$

Let us rewrite the function $F(z)$ in (14) as a function of the elements of the matrices $\widehat{D}$ and D, taking into account the notations above:

$$F(z)=\exp\left(-\frac{1}{2}z^T \widehat{D}z+Dz\right)=\exp\left(-\frac{1}{2}\sum_{i=1}^{k}(a_{i,i}^{(i-1)})^{-1}z_i^2+\sum_{i=1}^{k}d_i z_i\right)=\prod_{i=1}^{k}\exp\left(-\frac{1}{2}(a_{i,i}^{(i-1)})^{-1}z_i^2+d_i z_i\right). \tag{16}$$

Substituting (16) into (15) we obtain the following equality:

$$\int_{E^k} F(\xi)d\xi=|\,J\,|\prod_{i=1}^{k}\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}(a_{i,i}^{(i-1)})^{-1}z_i^2+d_i z_i\right)dz_i. \tag{17}$$

The integral in the right part of expression (17) is a table integral [6]:

$$\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}\alpha x^2+\beta x\right)dx=\sqrt{\frac{2\pi}{\alpha}}\exp\left(\frac{\beta^2}{2\alpha}\right). \tag{18}$$

Taking into account integral (18) we receive instead of (17):

$$\int_{E^k} F(\xi)d\xi=|\,J\,|\prod_{i=1}^{k}\sqrt{2\pi a_{i,i}^{(i-1)}}\exp\left(-\frac{1}{2}\sum_{i=1}^{k}a_{i,i}^{(i-1)}d_i^2\right),$$

or in the matrix form

$$\int_{E^k} F(\xi)d\xi = \sqrt{(2\pi)^k \,|\,\widehat{D}\,|}\exp\left(-\frac{1}{2}D\widehat{D}^{-1}D^T\right). \tag{19}$$

Let us turn back from the matrices D and $\widehat{D}$ to the matrices A and B. Since $|\widehat{D}|=|A^{-1}|$ (formula (11)), $\widehat{D}^{-1} = GA^{-1}G^T$ (formula (10)), and $D = B^T G^{-1}$ (formula (13)), then

$$D\widehat{D}^{-1}D^T = B^T G^{-1} G A^{-1} G^T (G^{-1})^T B = B^T A^{-1} B,$$

and instead of (19) we receive (2). Equality (2) is proven.

P r o o f o f e q u a l i t y (3). We calculate the integral

$$I_1 = \frac{1}{\sqrt{(2\pi)^k \,|\,A^{-1}\,|}}\int_{E^k} C^T\xi \exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi\right)d\xi. \tag{20}$$

Let us complete the square of the expression $-\xi^T A\xi/2 + B^T\xi$ in integral (20). We receive the expression

$$-\frac{1}{2}\xi^T A\xi + B^T\xi = -\frac{1}{2}(\xi - A^{-1}B)^T A(\xi - A^{-1}B) + \frac{1}{2}B^T A^{-1}B, \tag{21}$$

and instead of integral (20) we receive the following integral:

$$I_1 = \frac{1}{\sqrt{(2\pi)^k \,|\,A^{-1}\,|}}\int_{E^k} C^T\xi \exp\left(-\frac{1}{2}(\xi - A^{-1}B)^T A(\xi - A^{-1}B)\right)\exp\left(\frac{1}{2}B^T A^{-1}B)\right)d\xi. \tag{22}$$

Since the function

$$f(\xi) = \frac{1}{\sqrt{(2\pi)^k \,|\,A^{-1}\,|}}\exp\left(-\frac{1}{2}(\xi - A^{-1}B)^T A(\xi - A^{-1}B)\right) \tag{23}$$

in (22) is the probability density of the vector Gaussian distribution with the mean value $A^{-1}B$ and the dispersion matrix A^{-1}, then:

$$I_1 = \frac{1}{\sqrt{(2\pi)^k \,|\,A^{-1}\,|}}\int_{E^k} C^T\xi \exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi\right)d\xi = \exp\left(\frac{1}{2}B^T A^{-1}B)\right)C^T A^{-1}B.$$

As a result, we have equality (3).

P r o o f o f e q u a l i t y (4). We will perform it in a way similar to the proof of equality (3). We calculate the integral

$$I_2 = \frac{1}{\sqrt{(2\pi)^k \,|\,A^{-1}\,|}}\int_{E^k} \xi^T U\xi \exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi\right)d\xi. \tag{24}$$

Completing the square of the expression $-\xi^T A\xi/2 + B^T\xi$ in integral (24) gives expression (21). Thus, instead of integral (21) we have:

$$I_2 = \frac{1}{\sqrt{(2\pi)^k \,|\,A^{-1}\,|}}\int_{E^k} \xi^T U\xi \exp\left(-\frac{1}{2}(\xi - A^{-1}B)^T A(\xi - A^{-1}B)\right)\exp\left(\frac{1}{2}B^T A^{-1}B)\right)d\xi. \tag{25}$$

Since the function in (25) of form (23) is the probability density of the vector Gaussian distribution with the mean value $A^{-1}B$ and the dispersion matrix A^{-1}, the value of the integral I_2 is determined by the expression

$$I_2 = E(\Xi^T U\Xi)\exp\left(\frac{1}{2}B^T A^{-1}B)\right), \tag{26}$$

where $E(\Xi^T U\Xi)$ is the mathematical expectation of the quadratic form $\Xi^T U\Xi$ of the random vector Ξ with the Gaussian distribution (23). It is known the equality [2]:

$$\Xi^T U\Xi = \mathrm{Tr}(U\Xi\Xi^T). \tag{27}$$

Then

$$E(\Xi^T U\Xi) = E\left(\mathrm{Tr}(U\Xi\Xi^T)\right) = \mathrm{Tr}\left(UE(\Xi\Xi^T)\right).$$

Further, since for the Gaussian distribution (23) $E(\Xi\Xi^T) = A^{-1} + B^T A^{-1}A^{-1}B,$ then

$$E(\Xi^T U\Xi) = \mathrm{Tr}\left(U(A^{-1} + B^T A^{-1}A^{-1}B)\right) = \mathrm{Tr}(UA^{-1}) + \mathrm{Tr}(UB^T A^{-1}A^{-1}B).$$

If one takes into account the equality of type (27), which can be rewritten as

$$\mathrm{Tr}(UB^T A^{-1}A^{-1}B) = B^T A^{-1}UA^{-1}B,$$

then one receives

$$E(\Xi^T U\Xi) = \mathrm{Tr}(UA^{-1}) + B^T A^{-1}UA^{-1}B,$$

and for the integral I_2:

$$I_2 = \frac{1}{\sqrt{(2\pi)^k\,|A^{-1}|}}\int\limits_{E^k}\xi^T U\xi\exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi\right)d\xi = \left(\mathrm{Tr}(UA^{-1}) + B^T A^{-1}UA^{-1}B\right)\exp\left(\frac{1}{2}B^T A^{-1}B\right).$$

As a result, we have received equality (4).

Let us note that equality (4) is more general then the according equality of work [3], since the matrix U in work [3] is supposed to be diagonal.

2. The total probability formula for vector Gaussian distributions.

Theorem 1 (The total probability formula for vector Gaussian distributions). *Let* $\Xi^T = (\Xi_1, \Xi_2, ..., \Xi_k)$ *be a row random vector with k components,* $X^T = (X_1, X_2, ..., X_n)$ *be a row random vector with n components, f*(ξ) *be the probability density of the vector* Ξ, $f(x/\xi)$ *be the condition probability density of the vector X, and* E^k *be the k-dimensional Euclidean space. If in the total probability formula*

$$f(x) = \int\limits_{E^k} f(x/\xi)f(\xi)d\xi \tag{28}$$

the probability density $f(x/\xi)$ *is represented in the form*

$$f(x/\xi) = \frac{1}{\sqrt{(2\pi)^n\,|d_X|}}\exp\left(-\frac{1}{2}\xi^T S\xi + V^T\xi - \frac{1}{2}W\right),$$

and the probability density f(ξ) *is represented in the form*

$$f(\xi) = \frac{1}{\sqrt{(2\pi)^k \mid d_\Xi \mid}} \exp\left(-\frac{1}{2}\xi^T d_\Xi^{-1}\xi + \nu_\Xi^T d_\Xi^{-1}\xi - \frac{1}{2}\nu_\Xi^T d_\Xi^{-1}\nu_\Xi \right),$$

then integral (28) (*the total probability formula*) *is defined by the following expression*

$$f(x) = \int_{E^k} f(x/\xi) f(\xi) d\xi = \frac{1}{\sqrt{(2\pi)^n \mid d_\Xi A d_X \mid}} \exp\left(\frac{1}{2} B^T A^{-1} B - \frac{1}{2} C \right), \tag{29}$$

where

$$A = d_\Xi^{-1} + S, \tag{30}$$

$$B = d_\Xi^{-1}\nu_\Xi + V, \tag{31}$$

$$C = \nu_\Xi^T d_\Xi^{-1}\nu_\Xi + W. \tag{32}$$

P r o o f. Performing the multiplication under the integral in (28), we receive

$$f(x/\xi) f(\xi) = \frac{1}{\sqrt{(2\pi)^{n+k} \mid d_X \| d_\Xi \mid}} \exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi - \frac{1}{2} C \right), \tag{33}$$

where A, B, and C are defined by formulas (30), (31), and (32). Integration of expression (33) using equality (2) gives expression (29). This completes the proof of theorem 1.

3. The Bayes formula for vector Gaussian distributions.

T h e o r e m 2 (the Bayes formula for vector Gaussian distributions). *Let* $\Xi^T = (\Xi_1, \Xi_2, ..., \Xi_k)$ *be a random vector-row with k components,* $X^T = (X_1, X_2, ..., X_n)$ *be a random vector-row with n components, f*(ξ) *be the probability density of the vector* Ξ, $f(x/\xi)$ *be the condition probability density of the vector X, and* E^k *be the k-dimensional Euclidean space. If in the Bayes formula*

$$f(\xi/x) = \frac{f(\xi) f(x/\xi)}{\int_{E^k} f(\xi) f(x/\xi) d\xi} \tag{34}$$

the probability density $f(x/\xi)$ *is represented in the form*

$$f(x/\xi) = \frac{1}{\sqrt{(2\pi)^n \mid d_X \mid}} \exp\left(-\frac{1}{2}\xi^T S\xi + V^T\xi - \frac{1}{2} W \right),$$

and the probability density f(ξ) *is represented in the form*

$$f(\xi) = \frac{1}{\sqrt{(2\pi)^k \mid d_\Xi \mid}} \exp\left(-\frac{1}{2}\xi^T d_\Xi^{-1}\xi + \nu_\Xi^T d_\Xi^{-1}\xi - \frac{1}{2}\nu_\Xi^T d_\Xi^{-1}\nu_\Xi \right),$$

then the posteriori probability density $f(\xi/x)$ *of the random vector* ξ *defined by the Bayes formula* (34) *has the following form*

$$f(\xi/x) = \frac{1}{\sqrt{(2\pi)^k \mid A^{-1} \mid}} \exp\left(-\frac{1}{2}(\xi - A^{-1}B)^T A(\xi - A^{-1}B) \right), \tag{35}$$

where $A = d_\Xi^{-1} + S$, $B = d_\Xi^{-1}\nu_\Xi + V$.

P r o o f. We note that this theorem is formulated under the same conditions and designations as theorem 1. In this case, the numerator of the Bayes formula (34) is defined by expressions (33), and the denominator is defined by formula (29). Dividing (33) on (29) we receive the formula

$$f(\xi / x) = \frac{1}{\sqrt{(2\pi)^k \,|A^{-1}|}} \exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi - \frac{1}{2}B^T A^{-1}B \right), \tag{36}$$

which can be written as in (35). The equality of expressions (36) and (35) is easily verified by multiplying in expression (35). This completes the proof of theorem 2.

Obviously, the expression $A^{-1}B$ in the Bayes formula (34) is the posteriori mathematical expectation of the random vector Ξ, i. e. $A^{-1}B = E(\Xi / x)$, and the matrix A^{-1} is the posteriori dispersion matrix of the random vector Ξ, i. e. $A^{-1} = E\left((\Xi - A^{-1}B)(\Xi - A^{-1}B)^T / x\right)$.

4. Example. As an example, we consider the problem of the calculation of the Bayesian estimators of the coefficients of the multiple regression function.

Let U and Y be the input and output vectors of the controlled object, respectively, and the object is described by the conditional probability density $f(\overline{y} / \overline{\theta}, \overline{u})$. As a rule, it is the Gaussian (normal) probability density:

$$f(\overline{y} / \overline{\theta}, \overline{u}) \sim N\left(\phi(\overline{\theta}, \overline{u}), d_Y\right), \tag{37}$$

where $\phi(\overline{\theta}, \overline{u}) = \overline{y}$ is the regression function of Y on U, $\overline{u}$ and $\overline{y}$ are the input and output vectors of the regression function, respectively, $\overline{\theta}$ is the vector of the coefficients of the regression function, and d_Y is the constant dispersion matrix of the internal noise of the object. Description (37) could be represented in the form

$$Y = \phi(\overline{\theta}, \overline{u}) + E,$$

where E is the random vector with the Gaussian distribution $N(0, d_E)$.

The multiple regression function is considered most frequently when $\overline{y}$ is scalar (we will denote it y) and $\overline{u}$ is vector.

The class of the functions is represented in the form

$$y = \phi(\overline{\theta}, \overline{u}) = \sum_{j=1}^{m} h_j(\overline{u})\theta_j = \overline{h}^T\overline{\theta},$$

where $h_j(\overline{u})$, $j = 1, 2, ..., m$, are basis functions, $\overline{h}^T = \overline{h}^T(\overline{u}) = \left(h_1(\overline{u}), h_2(\overline{u}), ..., h_m(\overline{u})\right)$ is the vector formed by the basis functions, and $\overline{\theta}^T = (\theta_1, \theta_2, ..., \theta_m)$ is the vector of the coefficients of the regression function, is usually used to describe the multiple regression function. For example, if we want to write in the vector form a function of two variables u_1, u_2 having the form

$$y = \alpha + \beta u_1 + \gamma u_2 + \tau u_1^2,$$

then we have to choose $\overline{h}^T = (1, u_1, u_2, u_1^2)$, $\overline{\theta}^T = (\alpha, \beta, \gamma, \tau)$.

Let $y_{o,i} = \overline{h}_i^T\Theta + \varepsilon_i$, $\Theta^T = (\Theta_1, \Theta_2, ..., \Theta_m)$, be the i-th observed value of the output variable Y of the object on the observed value $\overline{h}_i^T$ of the input vector of the basis functions, $i = 1, 2, ..., n$, and the vector Θ has the normal priory probability density $N(a_\Theta, d_\Theta)$. The task is to find the estimation $\hat{\overline{\theta}}$ of the vector $\overline{\theta}$ on the basis of the observers $(\overline{h}_1, y_{o,1})$, $(\overline{h}_2, y_{o,2})$, ..., $(\overline{h}_n, y_{o,n})$ provided $\Theta = \overline{\theta}$.

In our case we have the following probability density functions:

$$f(\overline{\theta}) = \frac{1}{\sqrt{(2\pi)^m d_\Theta}} \exp\left(-\frac{1}{2}(\overline{\theta} - a_\Theta)^T d_\Theta^{-1}(\overline{\theta} - a_\Theta) \right),$$

$$f(y/\overline{\theta}) = \frac{1}{\sqrt{2\pi d_{\mathrm{E}}}} \exp\left(-\frac{1}{2}(y - \overline{h}^T \overline{\theta})^T d_{\mathrm{E}}^{-1} (y - \overline{h}^T \overline{\theta}) \right).$$

The vector of observations $\vec{y} = (y_{o,1}, y_{o,2}, ..., y_{o,n})$ will have the following probability density function:

$$f(\vec{y}_o / \overline{\theta}) = \prod_{i=1}^{n} f(y_{o,i} / \overline{\theta}) \sim \frac{1}{\sqrt{(2\pi)^n d_{\mathrm{E}}^n}} \exp\left(-\sum_{i=1}^{n} \frac{1}{2} (y_{o,i} - \overline{h}_i^T \overline{\theta})^T d_{\mathrm{E}}^{-1} (y_{o,i} - \overline{h}_i^T \overline{\theta}) \right).$$

We find now the posterior probability density functions $f(\overline{\theta} / \vec{y}_o)$ of the vector coefficient $\overline{\theta}$ by the Bayes formula:

$$f(\overline{\theta} / \vec{y}_o) = \frac{f(\overline{\theta}) f(\vec{y}_o / \overline{\theta})}{\int\limits_{E^m} f(\overline{\theta}) f(\vec{y}_o / \overline{\theta}) d\overline{\theta}}.$$

We will use for this theorem 2. Since

$$f(\overline{\theta}) = \frac{1}{\sqrt{(2\pi)^m d_{\Theta}}} \exp\left(-\frac{1}{2} \overline{\theta}^T d_{\Theta}^{-1} \overline{\theta} + a_{\Theta}^T d_{\Theta}^{-1} \overline{\theta} - \frac{1}{2} a_{\Theta}^T d_{\Theta}^{-1} a_{\Theta} \right),$$

$$f(\vec{y}_o / \overline{\theta}) = \frac{1}{\sqrt{(2\pi)^n d_{\mathrm{E}}^n}} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} y_{o,i}^T d_{\mathrm{E}}^{-1} y_{o,i} + \sum_{i=1}^{n} y_{o,i}^T d_{\mathrm{E}}^{-1} \overline{h}_i^T \overline{\theta} - \frac{1}{2} \sum_{i=1}^{n} \overline{\theta}^T \overline{h}_i d_{\mathrm{E}}^{-1} \overline{h}_i^T \overline{\theta} \right),$$

then, in accordance with theorem 2, we have

$$f(\overline{\theta} / \vec{y}_o) = \frac{1}{\sqrt{(2\pi)^m |A^{-1}|}} \exp\left(-\frac{1}{2} (\overline{\theta} - A^{-1}B)^T A (\overline{\theta} - A^{-1}B) \right),$$

where

$$A = d_{\Theta}^{-1} + \sum_{i=1}^{n} \overline{h}_i d_{\mathrm{E}}^{-1} \overline{h}_i^T, \quad B = d_{\Theta}^{-1} a_{\Theta} + \sum_{i=1}^{n} h_i d_{\mathrm{E}}^{-1} y_{o,i}.$$

Provided the loss function is quadratic, $W(\hat{\overline{\theta}}, \Theta) = (\hat{\overline{\theta}} - \Theta)^T (\hat{\overline{\theta}} - \Theta)$, we get the Bayesian estimation $\hat{\overline{\theta}}$ of the vector $\overline{\theta}$: $\hat{\overline{\theta}} = A^{-1}B$.

Conclusion. The results represented in this article provide a basis for theoretical solutions of the vector problems formulated within the framework of the statistical decision theory. The integrals can also be used as table integrals. The possible generalizations of the obtained results for solving more complicated problems within the framework of the statistical decision theory are of great interest.

References

1. Fel'dbaum A. A. *Optimal Control Systems*. New York, London, Academic Press, 1965. 452 p.

2. Rao C. S. *Linear Statistical Inference and its Applications*, 2nd ed., Wiley, 1973. 648 p.

3. Mukha V. S. Calculation of integrals connected with the multivariate Gaussian distribution. *Proceedings of the LETI*, 1974, vol. 160, pp. 27–30 (in Russian).

4. Gantmacher F. R. *The Theory of Matrices*. New York, Chelsea Publishing Company, 1959. Vol. 1. 374 p.

5. Rudin W. *Principles of Mathematical Analysis*, 3ed ed., New York, McGraw-Hill Inc., 1976. 352 p.

6. Prudnikov A. P., Brychkov Yu. A., Marichev O. I. *Integrals and Series*. Translated from the Russian by N. M. Queen. New York, Gordon and Breach Science Publ., 1986. 753 p.

Information about the authors

Vladimir S. Mukha – Dr. Sc. (Engineering), Professor, Professor of the Department of Information Technologies of Automated Systems, Belarusian State University of Informatics and Radioelectronics (6, P. Brovka Str., 220013, Minsk, Republic of Belarus). E-mail: mukha@bsuir.by

Nancy Farat Kako – Postgraduate Student, Belarusian State University of Informatics and Radioelectronics (6, P. Brovka Str., 220013, Minsk, Republic of Belarus). E-mail: kako.nancy@gmail.com

Информация об авторах

Муха Владимир Степанович – доктор технических наук, профессор, профессор кафедры информационных технологий автоматизированных систем, Белорусский государственный университет информатики и радиоэлектроники (ул. П. Бровки, 6, 220013, г. Минск, Республика Беларусь). E-mail: mukha@bsuir.by

Како Нэнси Фарат – аспирант, Белорусский государственный университет информатики и радиоэлектроники (ул. П. Бровки, 6, 220013, г. Минск, Республика Беларусь). E-mail: kako.nancy@gmail.com

ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)

ФИЗИКА

PHYSICS

УДК 539.12 Поступила в редакцию 02.05.2019
https://doi.org/10.29235/1561-2430-2019-55-4-467-478 Received 02.05.2019

Я. А. Войнова[1], А. Д. Коральков[2], Е. М. Овсиюк[2]

[1]Институт физики им. Б. И. Степанова Национальной академии наук Беларуси, Минск, Беларусь
[2]Мозырский государственный педагогический университет им. И. П. Шамякина, Мозырь, Беларусь

СКАЛЯРНАЯ ЧАСТИЦА СО СТРУКТУРОЙ ДАРВИНА – КОКСА ВО ВНЕШНЕМ КУЛОНОВСКОМ ПОЛЕ

Аннотация. Обобщенное уравнение Клейна – Фока – Гордона для частицы со структурой Дарвина – Кокса, учитывающее распределение заряда частицы по сфере конечного радиуса, исследуется с учетом внешнего кулоновского поля. Проведено разделение переменных, полученное радиальное уравнение сложнее уравнения в случае обычной частицы – оно имеет существенно особые точки $r = 0$ ранга 3, $r = \infty$ ранга 2 и 4 регулярные особые точки. В случае минимального орбитального момента $l = 0$ структура сингулярностей упрощается: есть существенно особые точки $r = 0$, $r = \infty$ ранга 2 и 4 регулярные особые точки. Построены решения Фробениуса этого уравнения, исследована структура рекуррентных соотношений для коэффициентов возникающего 7-членного степенного ряда. В качестве аналитического условия квантования используется обобщенное требование трансцендентности решений, которое позволяет получить алгебраическое уравнение 4-й степени для уровней энергии. Уравнение имеет 4 множества корней, зависящих от орбитального момента l и главного квантового числа $k = 1,2,3,\ldots$. Численный анализ показывает, что одно из множеств корней $0 < \varepsilon_{l,k} < mc^2$ может интерпретироваться как отвечающее некоторым связанным состояниям частицы в кулоновском поле.

Ключевые слова: скалярная частица со структурой Дарвина – Кокса, кулоновское поле, уравнение Клейна – Фока – Гордона, существенно особые точки, решения Фробениуса, связанное состояние

Для цитирования. Войнова, Я. А. Скалярная частица со структурой Дарвина – Кокса во внешнем кулоновском поле / Я. А. Войнова, А. Д. Коральков, Е. М. Овсиюк // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2019. – Т. 55, № 4. – С. 467–478. https://doi.org/10.29235/1561-2430-2019-55-4-467-478

Ya. A. Voynova[1], A. D. Koral'kov[2], E. M. Ovsiyuk[2]

[1]B. I. Stepanov Institute of Physics of the National Academy of Sciences of Belarus, Minsk, Belarus
[2]Mozyr State Pedagogical University named after I. P. Shamyakin, Mozyr, Belarus

SCALAR PARTICLE WITH THE DARWIN – COX INTRINSIC STRUCTURE IN THE EXTERNAL COULOMB FIELD

Abstract. The generalized Klein – Fock – Gordon equation for a particle with the Darwin–Cox structure allowing for a charge distribution of a particle over a sphere of finite radius is studied with regard to the external Coulomb field. The separation of variables is carried out, the obtained radial equation is significantly more complicated than the equation in the case of ordinary particles, it has essentially singular points $r = 0$ of rank 3, $r = \infty$ of rank 2 and 4 regular singular points. In the case of a minimum orbital momentum $l = 0$, the structure of singularities is simplified: there are essentially singular points $r = 0$, $r = \infty$ of rank 2 and 4 regular singular points. Frobenius solutions of this equation are constructed and the structure of the 7-term recurrence relations for the coefficients of the arising power series is investigated. As an analytical quantization condition, the generalized transcendence requirement of solutions is used; it allows one to obtain a fourth-degree algebraic equation for energy levels. The equation has 4 sets of roots depending on the orbital moment l and the main quantum number $k = 1,2,3,\ldots$. The numerical analysis shows that one of the sets of the roots $0 < \varepsilon_{l,k} < mc^2$ can be interpreted as those corresponding to certain bound states of the particle in the Coulomb field.

Keywords: scalar particle with the Darwin – Cox structure, Coulomb field, Klein – Fock – Gordon equation, essentially singular points, Frobenius solutions, bound state

For citation. Voynova Ya. A., Koral'kov A. D., Ovsiyuk E. M. Scalar particle with the Darwin – Cox intrinsic structure in the external Coulomb field. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2019, vol. 55, no. 4, pp. 467–478 (in Russian). https://doi.org/10.29235/1561-2430-2019-55-4-467-478

1. Постановка задачи. Исходим из системы уравнений для скалярной частицы Кокса с распределенным по конечному объему зарядом в тензорной форме Прока (см. [1–8])

$$\frac{mc}{\hbar}\left(\delta_\alpha^\beta + \frac{\lambda}{mc/\hbar}F_\alpha^{\ \beta}\right)\Phi_\beta = D_\alpha\Phi, \quad D^\alpha\Phi_\alpha = \frac{mc}{\hbar}\Phi, \tag{1}$$

где используем обозначение $D_\alpha = i\nabla_\alpha + (e/\hbar c)A_\alpha$ (учитываем отрицательность заряда электрона). Ненулевой параметр λ соответствует наличию у частицы дополнительной структуры Дарвина – Кокса. Представим уравнения (1) в краткой форме:

$$\frac{mc}{\hbar}\Lambda_\alpha^{\ \beta}\Phi_\beta = D_\alpha\Phi, \quad D^\alpha\Phi_\alpha = \frac{mc}{\hbar}\Phi$$

или

$$\frac{mc}{\hbar}\Phi_\rho = (\Lambda^{-1})_\rho^{\ \alpha} D_\alpha\Phi, \quad D^\alpha\Phi_\alpha = \frac{mc}{\hbar}\Phi.$$

Исключая векторную компоненту, получаем обобщенное уравнение для скалярной функции Φ:

$$\left(D_\rho(\Lambda^{-1})^{\rho\alpha}D_\alpha\Phi - \frac{m^2c^2}{\hbar^2}\right)\Phi = 0. \tag{2}$$

В пространстве-времени с метрикой $g_{\alpha\beta}(x)$ уравнение (2) примет вид

$$\left[\left(\frac{i}{\sqrt{-g}}\frac{\partial}{\partial x^\rho}\sqrt{-g} + \frac{e}{c\hbar}A_\rho\right)(\Lambda^{-1})^{\rho\alpha}\left(i\frac{\partial}{\partial x^\alpha} + \frac{e}{c\hbar}A_\alpha\right) - \frac{m^2c^2}{\hbar^2}\right]\Phi = 0. \tag{3}$$

2. Разделение переменных. Будем рассматривать частицу Кокса во внешнем кулоновском поле, используя сферические координаты

$$A_0 = \frac{e}{r}, \quad e > 0, \quad F_{r0} = -\frac{e}{r^2}, \quad dS^2 = c^2dt^2 - dr^2 - r^2d\theta^2 - r^2\sin^2\theta d\phi^2 .$$

Исходная матрица $\Lambda_\alpha^{\ \beta}$ имеет вид

$$\Lambda = (\Lambda_\alpha^{\ \beta}) = \begin{vmatrix} 1 & \frac{\mu}{r^2} & 0 & 0 \\ \frac{\mu}{r^2} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}, \quad \mu = -\frac{\lambda e}{mc/\hbar};$$

обратная матрица задается равенством

$$\Lambda^{-1} = K = (K_\beta^{\ \rho}) = \begin{vmatrix} \frac{r^4}{r^4-\mu^2} & -\frac{\mu r^2}{r^4-\mu^2} & 0 & 0 \\ -\frac{\mu r^2}{r^4-\mu^2} & \frac{r^4}{r^4-\mu^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

или

$$(K^{\beta\rho}) = \begin{vmatrix} \dfrac{r^4}{r^4-\mu^2} & -\dfrac{\mu r^2}{r^4-\mu^2} & 0 & 0 \\ +\dfrac{\mu r^2}{r^4-\mu^2} & -\dfrac{r^4}{r^4-\mu^2} & 0 & 0 \\ 0 & 0 & -1/r^2 & 0 \\ 0 & 0 & 0 & -1/r^2\sin^2\theta \end{vmatrix}.$$

Введем обозначения:

$$\bar{D}_\alpha = \frac{i}{\sqrt{-g}}\frac{\partial}{\partial x^\rho}\sqrt{-g} + \frac{e}{c\hbar}A_\rho, \quad D_\beta = i\frac{\partial}{\partial x^\alpha} + \frac{e}{c\hbar}A_\alpha,$$

тогда уравнение (3) принимает вид

$$\left[\bar{D}_0 K^{00} D_0 + \bar{D}_0 K^{0r} D_r + \bar{D}_r K^{r0} D_0 + \bar{D}_r K^{rr} D_r + \bar{D}_\theta K^{\theta\theta} D_\theta + \bar{D}_\phi K^{\phi\phi} D_\phi - \frac{m^2c^2}{\hbar^2}\right]\Phi = 0.$$

Далее (пусть $\alpha = e^2/\hbar c$)

$$\left\{\left(i\partial_0 + \frac{\alpha}{r}\right)^2 \frac{r^4}{r^4-\mu^2} - \left(i\partial_0 + \frac{\alpha}{r}\right)\frac{\mu r^2}{r^4-\mu^2} i\partial_r + \frac{i}{r^2}\partial_r r^2 \frac{\mu r^2}{r^4-\mu^2}\left(i\partial_0 + \frac{\alpha}{r}\right) - \right.$$

$$\left. - \frac{i}{r^2}\partial_r r^2 \frac{r^4}{r^4-\mu^2} i\partial_r + \frac{1}{r^2}\left(\frac{1}{\sin\theta}\partial_\theta \sin\theta\partial_\theta + \frac{1}{\sin^2\theta}\partial_\phi\partial_\phi\right) - \frac{m^2c^2}{\hbar^2}\right\}\Phi = 0. \tag{4}$$

После разделения переменных на основе подстановки

$$\Phi = e^{-E'x^0/c\hbar}\, Y_{lm}(\theta,\phi)\, R(r), \quad \varepsilon = E'/c\hbar$$

из (4) найдем радиальное уравнение

$$\left\{\left(\varepsilon + \frac{\alpha}{r}\right)^2 \frac{r^4}{r^4+\gamma^2} - \gamma\left(\varepsilon + \frac{\alpha}{r}\right)\frac{r^2}{r^4+\gamma^2}\frac{d}{dr} + \right.$$

$$\left. + \gamma\left(\frac{d}{dr} + \frac{2}{r}\right)\frac{r^2}{r^4+\gamma^2}\left(\varepsilon + \frac{\alpha}{r}\right) + \left(\frac{d}{dr} + \frac{2}{r}\right)\frac{r^4}{r^4+\gamma^2}\frac{d}{dr} - \frac{l(l+1)}{r^2} - M^2\right\}R = 0, \tag{5}$$

где использованы обозначения (учтена вещественность величины $i\mu$)

$$M = \frac{mc}{\hbar}, \quad e^2 = \alpha = \frac{1}{137}, \quad L = l(l+1), \quad i\mu = \gamma, \ \gamma^* = \gamma.$$

Уравнение (5) преобразуется к виду

$$\frac{d^2R}{dr^2} + \left[\frac{2}{r} + \frac{4\gamma^2}{r(r^4+\gamma^2)}\right]\frac{dR}{dr} + \left[\varepsilon^2 - M^2 + \frac{2\alpha\varepsilon}{r} + \frac{\alpha^2 - l(l+1)}{r^2} + \right.$$

$$\left. + \frac{4\gamma\varepsilon}{r^3} + \frac{3\alpha\gamma - \gamma^2M^2}{r^4} - \frac{l(l+1)\gamma^2}{r^6} - \frac{4(\gamma\varepsilon r + \gamma\alpha)}{r^4+\gamma^2}\right]R = 0. \tag{6}$$

Здесь имеем 4 регулярные и 2 нерегулярные особые точки:

$$r^4+\gamma^2=\left(r-e^{+i\pi/4}\sqrt{\gamma}\right)\left(r+e^{+i\pi/4}\sqrt{\gamma}\right)\left(r-e^{-i\pi/4}\sqrt{\gamma}\right)\left(r+e^{-i\pi/4}\sqrt{\gamma}\right),$$

$$r=0,\quad \text{Rang}=3,\quad r=\infty,\quad \text{Rang}=2;$$

справедливо тождество (пусть $\sigma\equiv(-\gamma^2)^{1/4}$)

$$\frac{1}{r^4+\gamma^2}=\frac{1}{4\sigma^3}\left(\frac{1}{r-\sigma}-\frac{1}{r+\sigma}+\frac{i}{r-i\sigma}-\frac{i}{r+i\sigma}\right).$$

Решение около 4 регулярных особых точек имеет простой вид:

$$r\to+\sigma,\quad R\sim(r-\sigma)^{\rho},\quad \rho=0,2;\quad r\to-\sigma,\quad R\sim(r+\sigma)^{\rho},\quad \rho=0,2;$$

$$r\to+i\sigma,\quad R\sim(r-i\sigma)^{\rho},\quad \rho=0,2;\quad r\to-i\sigma,\quad R\sim(r+i\sigma)^{\rho},\quad \rho=0,2.$$

В соответствии с тем, что $r=0$ – особая точка ранга 3, подстановка для локальных решений Фробениуса уравнения (6) около точки $r=0$ должна иметь вид

$$R(r)=r^{C}e^{Ar}e^{B/r}e^{D/r^2}f(r). \tag{7}$$

3. Состояния с нулевым орбитальным моментом l = 0. Ограничимся наиболее простым случаем минимального значения момента, $l=0$, уравнение (6) упрощается:

$$\frac{d^2R}{dr^2}+\left(\frac{6}{r}-\frac{4r^3}{r^4+\gamma^2}\right)\frac{dR}{dr}+$$

$$+\left[\varepsilon^2-M^2+\frac{2\alpha\varepsilon}{r}+\frac{\alpha^2}{r^2}+\frac{4\gamma\varepsilon}{r^3}+\frac{3\alpha\gamma-\gamma^2M^2}{r^4}-\frac{4\gamma(\varepsilon r+\alpha)}{r^4+\gamma^2}\right]R=0. \tag{8}$$

Здесь имеем прежние 4 регулярные особые точки и 2 нерегулярные точки $r=0$, $r=\infty$ ранга 2. Разложим на простые дроби 2 выражения (напоминаем, что $\sqrt[4]{-\gamma^2}=\sigma$):

$$-\frac{4r^3}{(r^4+\gamma^2)}=-\frac{1}{r-\sigma}-\frac{1}{r-i\sigma}-\frac{1}{r+\sigma}-\frac{1}{r+i\sigma},$$

$$\frac{-4\gamma\varepsilon r-4\alpha\gamma}{r^4+\gamma^2}=\frac{1}{\sigma^3}\left\{\frac{-\gamma\varepsilon\sigma-\alpha\gamma}{r-\sigma}+\frac{i(-i\gamma\varepsilon\sigma-\alpha\gamma)}{r-i\sigma}-\frac{\gamma\varepsilon\sigma-\alpha\gamma}{r+\sigma}-\frac{i(i\gamma\varepsilon\sigma-\alpha\gamma)}{r+i\sigma}\right\}.$$

Уравнение (8) примет вид

$$\frac{d^2R}{dr^2}+\left(\frac{6}{r}-\frac{1}{r-\sigma}-\frac{1}{r-i\sigma}-\frac{1}{r+\sigma}-\frac{1}{r+i\sigma}\right)\frac{dR}{dr}+$$

$$+\left[\varepsilon^2-M^2+\frac{2\alpha\varepsilon}{r}+\frac{\alpha^2}{r^2}+\frac{4\gamma\varepsilon}{r^3}+\frac{3\alpha\gamma-\gamma^2M^2}{r^4}-\right.$$

$$\left.-\frac{\gamma(\varepsilon\sigma+\alpha)}{(r-\sigma)\sigma^3}+\frac{\gamma(\varepsilon\sigma+i\alpha)}{(r+i\sigma)\sigma^3}+\frac{\gamma(-\varepsilon\sigma+\alpha)}{(r+\sigma)\sigma^3}-\frac{\gamma(-\varepsilon\sigma+i\alpha)}{(r-i\sigma)\sigma^3}\right]R=0. \tag{9}$$

Специально отметим форму уравнения (9) около точки $r=0$:

$$R'' + \frac{6}{r} R' - \frac{\gamma^2 M^2 - 3\gamma\alpha}{r^4} R = 0. \tag{10}$$

Условие, при котором область около нуля является запрещенной для классического движения частицы, имеет вид

$$\gamma\left(\gamma - \frac{3\alpha}{M^2}\right) > 0 \quad \Rightarrow \quad \gamma < 0;$$

вариант $\gamma > 3\alpha / M^2$ кажется нефизическим, поскольку изначально предполагается, что параметр γ может быть сколь угодно близким к нулю, включая и нуль. Обратное к (10) ограничение (соответствующее разрешенности классического движения в области около нуля) дает

$$\gamma\left(\gamma - \frac{3\alpha}{M^2}\right) < 0 \quad \Rightarrow \quad 0 < \gamma < \frac{3\alpha}{M^2};$$

это условие предполагает ограничение на параметр γ сверху.

Напомним, что для состояний $l = 0$ качественное рассмотрение запрещенности классического движения в области около нуля не является надежной аргументацией и в случае обычной частицы. Действительно, здесь имеем следующее поведение решений:

$$R'' + \frac{2}{r} R' + \frac{\alpha^2}{r^2} R = 0, \quad r \sim r^a, \quad a = -\frac{1}{2} \pm \sqrt{\frac{1}{4} - \alpha^2} < 0;$$

т. е. оба решения около нуля стремятся к бесконечности. С использованием подстановки $R = r^{-1}\bar{R}$ получим обращение решений в нуль в области, разрешенной для классического движения:

$$\bar{R}'' + \frac{\alpha^2}{r^2} \bar{R} = 0, \quad r \sim r^{ba}, \quad b = +\frac{1}{2} \pm \sqrt{\frac{1}{4} - \alpha^2} > 0.$$

Таким образом, пока вопрос с выбором правильного знака для параметра γ остается неясным.

В соответствии со структурой сингулярных точек решения Фробениуса для уравнения (9) ищем в виде (сравн. с (7))

$$R(r) = r^C e^{Ar} e^{B/r} f(r),$$

в результате получаем уравнение для $f(r)$:

$$\frac{d^2 f}{dr^2} + \left(\frac{2C+6}{r} - \frac{2B}{r^2} + 2A - \frac{1}{r+i\sigma} - \frac{1}{r+\sigma} - \frac{1}{r-\sigma} - \frac{1}{r-i\sigma} \right) \frac{df}{dr} +$$

$$+ \left[\frac{2\alpha\varepsilon + 6A + 2AC}{r} + \frac{\alpha^2 - 2AB + 5C + C^2}{r^2} + \frac{4\gamma\varepsilon - 4B - 2BC}{r^3} + \frac{3\alpha\gamma - \gamma^2 M^2 + B^2}{r^4} + \right.$$

$$+ A^2 - M^2 + \varepsilon^2 + \frac{-\gamma\varepsilon\sigma + \alpha\gamma - A\sigma^3 + B\sigma + C\sigma^2}{\sigma^3} \frac{1}{(r+\sigma)} + \frac{-\gamma\varepsilon\sigma - \alpha\gamma - A\sigma^3 + B\sigma - C\sigma^2}{\sigma^3} \frac{1}{(r-\sigma)} +$$

$$\left. + \frac{\gamma\varepsilon\sigma + i\gamma\alpha - A\sigma^3 - B\sigma - iC\sigma^2}{\sigma^3} \frac{1}{(r+i\sigma)} + \frac{\gamma\varepsilon\sigma - i\gamma\alpha - A\sigma^3 - B\sigma + iC\sigma^2}{\sigma^3} \frac{1}{(r-i\sigma)} \right] f = 0.$$

Накладывая 3 ограничения, упрощающие уравнение, на параметры A, B, C, находим 4 набора значений:

$$A^2 - M^2 + \varepsilon^2 = 0 \quad \Rightarrow \quad A = \pm\sqrt{M^2 - \varepsilon^2};$$

$$3\alpha\gamma - \gamma^2 M^2 + B^2 = 0 \quad \Rightarrow \quad B = \pm\sqrt{\gamma(\gamma M^2 - 3\alpha)};$$

$$C = 2\left(\frac{\gamma\varepsilon}{B} - 1\right) = -2 \pm \frac{2\gamma\varepsilon}{\sqrt{\gamma(\gamma M^2 - 3\alpha)}}. \tag{11}$$

Для описания связанных состояний в силу требования обращения решений в нуль на бесконечности будем использовать отрицательное значение параметра

$$A = -\sqrt{M^2 - \varepsilon^2}.$$

Если параметр B вещественный, то требование обращения в нуль решений будет заведомо удовлетворяться только при знаке минус перед корнем:

$$B = -\sqrt{\gamma(\gamma M^2 - 3\alpha)}, \quad \gamma < 0 \quad (R > 0).$$

Теперь предположим, что параметр B чисто мнимый:

$$\gamma\left(\gamma M^2 - 3\alpha\right) < 0, \quad e^{B/r} = e^{\pm iR/r}\left(R > 0\right), \quad \gamma \in \left(0, \frac{3\alpha}{M^2}\right).$$

Следует обратить внимание на то, что при этом имеем очень необычное (расходящееся и осциллирующее) поведение решений в окрестности $r = 0$:

$$r \to 0, \quad R \sim \frac{1}{r^2} e^{\mp i\frac{2\gamma\varepsilon}{R}\ln r} e^{\pm iR/r};$$

едва ли такое поведение совместимо с пониманием связанных состояний в квантовой механике. Вследствие этого в дальнейшем будем предполагать, что $\gamma < 0$.

С учетом ограничений (11) уравнение для $f(r)$ принимает более простой вид

$$\frac{d^2 f}{dr^2} + \left(\frac{2C+6}{r} - \frac{2B}{r^2} + 2A - \frac{1}{r+\sigma} - \frac{1}{r-\sigma} - \frac{1}{r+i\sigma} - \frac{1}{r-i\sigma}\right)\frac{df}{dr} +$$

$$+\left[\frac{2\alpha\varepsilon + 6A + 2AC}{r} + \frac{\alpha^2 - 2AB + 5C + C^2}{r^2} + \right.$$

$$+\frac{-\gamma\varepsilon\sigma + \alpha\gamma - A\sigma^3 + B\sigma + C\sigma^2}{\sigma^3}\frac{1}{(r+\sigma)} + \frac{-\gamma\varepsilon\sigma - \alpha\gamma - A\sigma^3 + B\sigma - C\sigma^2}{\sigma^3}\frac{1}{(r-\sigma)} +$$

$$\left. + \frac{\gamma\varepsilon\sigma + i\gamma\alpha - A\sigma^3 - B\sigma - iC\sigma^2}{\sigma^3}\frac{1}{(r+i\sigma)} + \frac{\gamma\varepsilon\sigma - i\gamma\alpha - A\sigma^3 - B\sigma + iC\sigma^2}{\sigma^3}\frac{1}{(r-i\sigma)}\right] f = 0.$$

Будем использовать сокращенную запись этого уравнения

$$f'' + \left(a + \frac{a_1}{r} + \frac{a_2}{r^2} - \frac{1}{r+\sigma} - \frac{1}{r-\sigma} - \frac{1}{r+i\sigma} - \frac{1}{r-i\sigma}\right) f' +$$

$$+\left(\frac{b_1}{r} + \frac{b_2}{r^2} + \frac{\beta_1}{r+\sigma} + \frac{\beta_2}{r-\sigma} + \frac{\beta_3}{r+i\sigma} + \frac{\beta_4}{r-i\sigma}\right) f = 0.$$

Умножив это уравнение на $r^2(r+\sigma)(r-\sigma)(r+i\sigma)(r-i\sigma)=r^2(r^4-\sigma^4),$ получим

$$\left(r^6-\sigma^4 r^2\right)\frac{d^2 f}{dr^2}+\left[ar^6+\left(a_1-4\right)r^5+a_2 r^4-\sigma^4 ar^2-\sigma^4 a_1 r-\sigma^4 a_2\right]\frac{df}{dr}+$$

$$+\Big[\left(b_1+\beta_1+\beta_2+\beta_3+\beta_4\right)r^5+\left(\left(-\beta_1+\beta_2-i\beta_3+i\beta_4\right)\sigma+b_2\right)r^4+\sigma^2\left(\beta_1+\beta_2-\beta_3-\beta_4\right)r^3+$$

$$+\left(-\beta_1+\beta_2+i\beta_3-i\beta_4\right)\sigma^3 r^2-b_1\sigma^4 r-b_2\sigma^4\Big]f=0.$$

Строим решения в виде степенных рядов:

$$f=\sum_{k=0}^{\infty}c_k r^k,\qquad \frac{df}{dr}=\sum_{k=1}^{\infty}kc_k r^{k-1},\qquad \frac{d^2 f}{dr^2}=\sum_{k=2}^{\infty}k(k-1)c_k r^{k-2}.$$

В результате находим рекуррентные соотношения для коэффициентов ряда

$$k=0,\qquad a_2 c_1+b_2 c_0=0,$$

$$k=1,\qquad 2a_2 c_2+a_1 c_1+b_1 c_0+b_2 c_1=0,$$

$$k=2,\qquad -2\sigma c_2-\sigma a c_1-2\sigma a_1 c_2-3\sigma a_2 c_3+\left(-\beta_1+\beta_2+i\beta_3-i\beta_4\right)c_0-b_1\sigma c_1-b_2\sigma c_2=0,$$

$$k=3,\qquad -6\sigma^2 c_3-2\sigma^2 a c_2-3\sigma^2 a_1 c_3-4\sigma^2 a_2 c_4+\left(\beta_1+\beta_2-\beta_3-\beta_4\right)c_0+$$

$$+\left(-\beta_1+\beta_2+i\beta_3-i\beta_4\right)\sigma c_1-b_1\sigma^2 c_2-b_2\sigma^2 c_3=0,$$

$$k=4,\qquad -12\sigma^4 c_4+a_2 c_1-3\sigma^4 a c_3-\sigma^4 a_1 4c_4-5\sigma^4 a_2 c_5+\left[\left(-\beta_1+\beta_2-i\beta_3+i\beta_4\right)\sigma+b_2\right]c_0+$$

$$+\sigma^2\left(\beta_1+\beta_2-\beta_3-\beta_4\right)c_1+\left(-\beta_1+\beta_2+i\beta_3-i\beta_4\right)\sigma^3 c_2-b_1\sigma^4 c_3-b_2\sigma^4 c_4=0,$$

$$k=5,\qquad -20\sigma^4 c_5+\left(a_1-4\right)c_1+2a_2 c_2-4\sigma^4 a c_4-\sigma^4 a_1 5c_5-6\sigma^4 a_2 c_6+$$

$$+\left(b_1+\beta_1+\beta_2+\beta_3+\beta_4\right)c_0+\left[\left(-\beta_1+\beta_2-i\beta_3+i\beta_4\right)\sigma+b_2\right]c_1+$$

$$+\sigma^2\left(\beta_1+\beta_2-\beta_3-\beta_4\right)c_2+\left(-\beta_1+\beta_2+i\beta_3-i\beta_4\right)\sigma^3 c_3-b_1\sigma^4 c_4-b_2\sigma^4 c_5=0,$$

$$k=6,\qquad 2c_2-30\sigma^4 c_6+ac_1+2\left(a_1-4\right)c_2+3a_2 c_3-5\sigma^4 a c_5-\sigma^4 a_1 6c_6-$$

$$-7\sigma^4 a_2 c_7+\left(b_1+\beta_1+\beta_2+\beta_3+\beta_4\right)c_1+\left[\left(-\beta_1+\beta_2-i\beta_3+i\beta_4\right)\sigma+b_2\right]c_2+$$

$$+\sigma^2\left(\beta_1+\beta_2-\beta_3-\beta_4\right)c_3+\left(-\beta_1+\beta_2+i\beta_3-i\beta_4\right)\sigma^3 c_4-b_1\sigma^4 c_5-b_2\sigma^4 c_6=0,$$

т. е. имеем 7-членное рекуррентное соотношение

$$k=5,6,7,\ldots,\qquad \left[a(k-5)+\left(b_1+\beta_1+\beta_2+\beta_3+\beta_4\right)\right]c_{k-5}+$$

$$+\left[(k-4)(k-5)+\left(a_1-4\right)(k-4)+\left\{\left(-\beta_1+\beta_2-i\beta_3+i\beta_4\right)\sigma+b_2\right\}\right]c_{k-4}+$$

$$+\left[a_2(k-3)+\sigma^2\left(\beta_1+\beta_2-\beta_3-\beta_4\right)\right]c_{k-3}+\left(-\beta_1+\beta_2+i\beta_3-i\beta_4\right)\sigma^3 c_{k-2}+$$

$$+\left[-\sigma^4 a(k-1)-b_1\sigma^4\right]c_{k-1}+\left[-\sigma^4 k(k-1)-\sigma^4 a_1 k-b_2\sigma^4\right]c_k-\sigma^4 a_2(k+1)c_{k+1}=0. \qquad (12)$$

В соответствии с методом Пуанкаре – Перрона разделим последнее соотношение на $k^2 c_{k-5}$:

$$\frac{1}{k^2}\left[a(k-5)+\left(b_1+\beta_1+\beta_2+\beta_3+\beta_4\right)\right]+$$

$$+\frac{1}{k^2}\left[(k-4)(k-5)+\left(a_1-4\right)(k-4)+\left\{\left(-\beta_1+\beta_2-i\beta_3+i\beta_4\right)\sigma+b_2\right\}\right]\frac{c_{k-4}}{c_{k-5}}+$$

$$+\frac{1}{k^2}\left[a_2(k-3)+\sigma^2\left(\beta_1+\beta_2-\beta_3-\beta_4\right)\right]\frac{c_{k-3}}{c_{k-4}}\frac{c_{k-4}}{c_{k-5}}+$$

$$+\frac{1}{k^2}\left(-\beta_1+\beta_2+i\beta_3-i\beta_4\right)\sigma^3\frac{c_{k-2}}{c_{k-3}}\frac{c_{k-3}}{c_{k-4}}\frac{c_{k-4}}{c_{k-5}}+$$

$$+\frac{1}{k^2}\left[-\sigma^4 a(k-1)-b_1\sigma^4\right]\frac{c_{k-1}}{c_{k-2}}\frac{c_{k-2}}{c_{k-3}}\frac{c_{k-3}}{c_{k-4}}\frac{c_{k-4}}{c_{k-5}}+$$

$$+\frac{1}{k^2}\left[-\sigma^4 k(k-1)-\sigma^4 a_1 k-b_2\sigma^4\right]\frac{c_k}{c_{k-1}}\frac{c_{k-1}}{c_{k-2}}\frac{c_{k-2}}{c_{k-3}}\frac{c_{k-3}}{c_{k-4}}\frac{c_{k-4}}{c_{k-5}}-$$

$$-\frac{1}{k^2}\sigma^4 a_2(k+1)\frac{c_{k+1}}{c_k}\frac{c_k}{c_{k-1}}\frac{c_{k-1}}{c_{k-2}}\frac{c_{k-2}}{c_{k-3}}\frac{c_{k-3}}{c_{k-4}}\frac{c_{k-4}}{c_{k-5}}=0$$

и устремим $k \to \infty$. В результате получим алгебраическое уравнение для величины, определяющей возможные радиусы сходимости:

$$R_{\text{conv}}=\frac{1}{|r|},\quad \lim_{k\to\infty}\frac{c_{k-4}}{c_{k-5}}=r,\quad r-\sigma^4 r^5=0,\quad R_{\text{conv}}=\frac{1}{|\sigma|}=|\sqrt{\gamma}|,\infty.$$

Поскольку на границе круга с радиусом $|\gamma|$ поведение решений регулярное, можно полагать, что ряд сходится при всех конечных r.

Будем пробовать в качестве условия квантования использовать ограничение, выделяющее из всех возможных решений Фробениуса так называемые трансцендентные решения (это обобщение условия, применяющегося для выделения трансцендентных функций Гойна [9–11]). Для этого следует обратиться к рекуррентной формуле (12) и потребовать обращения в нуль коэффициента при c_{k-5}:

$$k=5,6,7,\ldots,\qquad P_{k-5}=\left[a(k-5)+\left(b_1+\beta_1+\beta_2+\beta_3+\beta_4\right)\right]=0. \tag{13}$$

Отметим, что если дополнительно к этому условию потребовать в рекуррентном соотношении (12) обращения в нуль множителей при коэффициентах ряда

$$c_{k-4},\quad c_{k-3},\quad c_{k-2},\quad c_{k-1},\quad c_k,$$

это даст еще 5 уравнений:

$$P_{k-4}=0,\quad P_{k-3}=0,\quad P_{k-2}=0,\quad P_{k-1}=0,\quad P_k=0, \tag{14}$$

в силу рекуррентных соотношений все остальные коэффициенты степенного ряда обратятся в нуль, т. е. ряд превратится в полином. Понятно, что при этом уравнения (13), (14) должны иметь совместные решения.

Не следует думать, что не существует примеров, когда условиям полиномиальности можно удовлетворить и попасть при этом в физически интересные области параметров. Однако анализ возможности получения полиномиальных решений для рассматриваемой в работе системы дает отрицательный результат: в области изменения параметра энергии для связанных состояний $\varepsilon\in(0,1)$ совместных решений условий полиномиальности не существует.

4. Условие квантования при минимальном $l = 0$. Прежде чем приступить к анализу условия квантования, удобно в уравнении перейти к безразмерным величинам. Для этого учтем, что в качестве единицы измерения длины можно взять комптоновскую длину волны частицы λ: $M = mc/\hbar = 1/\lambda$, тогда возникают безразмерные величины

$$Mr = x, \quad \frac{\varepsilon}{M} = \frac{E'}{c\hbar}\cdot\frac{\hbar}{mc} = \frac{E'}{mc^2} = E, \quad \gamma M^2 = \Gamma.$$

При этом уравнение при $l = 0$ (см. (8)) преобразуется в следующее:

$$\frac{d^2R}{dx^2} + \left(\frac{6}{x} - \frac{4x^3}{x^4+\Gamma^2}\right)\frac{dR}{dx} +$$

$$+\left[E^2 - 1 + \frac{2\alpha E}{x} + \frac{\alpha^2}{x^2} + \frac{4\Gamma E}{x^3} + \frac{3\alpha\Gamma - \Gamma^2}{x^4} - \frac{4\Gamma(Ex+\alpha)}{x^4+\Gamma^2}\right]R = 0.$$

Переход к безразмерным единицам достигается формальными заменами $r \to x$, $M \to 1$, $\varepsilon \to E$, $\gamma \to \Gamma$.

Подстановка для решений Фробениуса имеет вид

$$l = 0, \quad R(x) = x^C e^{Ax} e^{B/x} f(x), \quad A = \pm\sqrt{1-E^2},$$

$$B = \pm\sqrt{-\Gamma(3\alpha - \Gamma)}, \quad C = 2\left(\frac{\Gamma E}{B} - 1\right).$$

Для исследуемого ниже случая с отрицательным значением Γ имеем следующие выражения для параметров (предполагаем описание связанных состояний):

$$A = -\sqrt{1-E^2}, \quad B = -\sqrt{-\Gamma(3\alpha-\Gamma)}, \quad C = \frac{-2\Gamma E}{\sqrt{-\Gamma(3\alpha-\Gamma)}} - 2$$

и явный вид уравнения

$$\frac{d^2f}{dx^2} + \left(\frac{2C+6}{x} - \frac{2B}{x^2} + 2A - \frac{1}{x+\Sigma} - \frac{1}{x-\Sigma} - \frac{1}{x+i\Sigma} - \frac{1}{x-i\Sigma}\right)\frac{df}{dx} +$$

$$+\left[\frac{2\alpha E + 6A + 2AC}{x} + \frac{\alpha^2 - 2AB + 5C + C^2}{x^2} +\right.$$

$$+\frac{-\Gamma E\Sigma + \alpha\Gamma - A\Sigma^3 + B\Sigma + C\Sigma^2}{\Sigma^3}\cdot\frac{1}{(x+\Sigma)} + \frac{-\Gamma E\Sigma - \alpha\Gamma - A\Sigma^3 + B\Sigma - C\Sigma^2}{\Sigma^3}\cdot\frac{1}{(x-\Sigma)} +$$

$$\left.+\frac{\Gamma E\Sigma + i\Gamma\alpha - A\Sigma^3 - B\Sigma - iC\Sigma^2}{\Sigma^3}\cdot\frac{1}{(x+i\Sigma)} + \frac{\Gamma E\Sigma - i\Gamma\alpha - A\Sigma^3 - B\Sigma + iC\Sigma^2}{\Sigma^3}\cdot\frac{1}{(x-i\Sigma)}\right]f = 0$$

или сокращенно (напоминаем, что $\Gamma^2 = -\Sigma^4$)

$$f'' + \left(a + \frac{a_1}{x} + \frac{a_2}{x^2} - \frac{1}{x+\Sigma} - \frac{1}{x-\Sigma} - \frac{1}{x+i\Sigma} - \frac{1}{x-i\Sigma}\right)f' +$$

$$+\left(\frac{b_1}{x} + \frac{b_2}{x^2} + \frac{\beta_1}{x+\Sigma} + \frac{\beta_2}{x-\Sigma} + \frac{\beta_3}{x+i\Sigma} + \frac{\beta_4}{x-i\Sigma}\right)f = 0.$$

Обратимся к анализу условия трансцендентности решений (13) для случая $l = 0$:

$$k = 5, 6, 7, \ldots, \qquad P_{k-5} = \left[a(k-5) + \left(b_1 + \beta_1 + \beta_2 + \beta_3 + \beta_4\right) \right] = 0;$$

рассматриваем случай отрицательных значений $\Gamma < 0$:

$$a = 2A = -2\sqrt{1 - E^2}, \quad b_1 = 2\alpha E + 6A + 2AC,$$

$$\beta_1 + \beta_2 + \beta_3 + \beta_4 = \frac{-\Gamma E \Sigma + \alpha\Gamma - A\Sigma^3 + B\Sigma + C\Sigma^2}{\Sigma^3} + \frac{-\Gamma E \Sigma - \alpha\Gamma - A\Sigma^3 + B\Sigma - C\Sigma^2}{\Sigma^3} +$$

$$+ \frac{\Gamma E \Sigma + i\Gamma\alpha - A\Sigma^3 - B\Sigma - iC\Sigma^2}{\Sigma^3} + \frac{\Gamma E \Sigma - i\Gamma\alpha - A\Sigma^3 - B\Sigma + iC\Sigma^2}{\Sigma^3} = -4A.$$

Найдем для этого случая аналитическую форму условия квантования

$$a(k-5) + \left(b_1 + \beta_1 + \beta_2 + \beta_3 + \beta_4\right) =$$

$$= 2Ak - 10A + 2\alpha E + 6A + 2AC - 4A = 2Ak - 8A + 2\alpha E + 2AC = 0,$$

далее получаем уравнение относительно энергий E:

$$k = 6, 7, 8, 9, 10, \ldots, \quad 2\alpha E - 2\sqrt{1 - E^2}\left(k - 6 - \frac{2\Gamma E}{\sqrt{-\Gamma(3\alpha - \Gamma)}} \right) = 0.$$

Преобразуем уравнение для E к явному виду уравнения 4-й степени:

$$\frac{4\Gamma^2 E^4}{\Gamma(-3\alpha + \Gamma)} - \frac{4(k-6)\Gamma E^3}{\sqrt{\Gamma(-3\alpha + \Gamma)}} +$$

$$+ \left[(k-6)^2 + \alpha^2 - \frac{4\Gamma^2}{\Gamma(-3\alpha + \Gamma)} \right] E^2 + \frac{4(k-6)\Gamma E}{\sqrt{\Gamma(-3\alpha + \Gamma)}} - (k-6)^2 = 0.$$

Найдем его корни (рассматриваем несколько значений для k) при $\Gamma = -0{,}001$:

$$k = 6, \quad E_1 = 0{,}0, \quad E_2 = 0{,}0, \quad E_3 = 0{,}9998474910, \quad E_4 = -0{,}9998474910;$$

$$k = 7, \quad E_1 = 0{,}9999867507, \quad E_2 = -0{,}9999213828,$$

$$E_3 = -2{,}392615505 + 0{,}01922309598i, \quad E_4 = -2{,}392615505 - 0{,}01922309598i;$$

$$k = 8, \quad E_1 = 0{,}9999954436, \quad E_2 = -0{,}9999893567,$$

$$E_3 = -4{,}785168685 + 0{,}01785830114i, \quad E_4 = -4{,}785168685 - 0{,}01785830114i;$$

$$k = 9, \quad E_1 = 0{,}9999977197, \quad E_2 = -0{,}9999960043,$$

$$E_3 = -7{,}177749321 + 0{,}01763591332i, \quad E_4 = -7{,}177749321 - 0{,}01763591332i;$$

$$k = 10, \quad E_1 = 0{,}9999986350, \quad E_2 = -0{,}9999979237,$$

$$E_3 = -9{,}570331642 + 0{,}01755953433i, \quad E_4 = -9{,}570331642 - 0{,}01755953433i.$$

Уравнение имеет 4 множества корней, зависящих от орбитального момента l и главного квантового числа $k = 1,2,3\ldots$. Помеченное индексом 1 множество значений, одно из множеств корней $0 < E_{1l,k} < mc^2$, может интерпретироваться как отвечающее некоторым связанным состояниям частицы в кулоновском поле.

Анализ радиального уравнения для частицы Кокса в кулоновском поле и условия квантования для уровней энергии при остальных значений $l = 1,2,\ldots$ будут рассмотрены в отдельной работе.

Список использованных источников

1. Cox, W. Higher-rank representations for zero-spin field theories / W. Cox // J. Phys. Math. Gen. – 1982. – Vol. 15, № 2. – P. 627–635. https://doi.org/10.1088/0305-4470/15/2/029

2. Ovsiyuk, E. M. Spin-zero Cox's particle with an intrinsic structure: general analysis in external electromagnetic and gravitational fields / E. M. Ovsiyuk // Ukr. J. Phys. – 2015. – Vol. 60, № 6. – P. 485–496. https://doi.org/10.15407/ujpe60.06.0485

3. Kazmerchuk, K. V. Cox's particle in magnetic and electric field against the background of Euclidean and spherical geometries / K. V. Kazmerchuk, E. M. Ovsiyuk // Ukr. Phys. J. – 2015. – Vol. 60, № 5. – P. 389–400. https://doi.org/10.15407/ujpe60.05.0389

4. Овсиюк, Е. М. Скалярная частица с внутренней структурой в электромагнитном поле в искривленном пространстве-времени / Е. М. Овсиюк, О. В. Веко, К. В. Казмерчук // Проблемы физики, математики и техники. – 2014. – № 3 (20). – С. 32–36.

5. Veko, O. V. Cox's particle in magnetic and electric fields on the background of hyperbolic Lobachevsky geometry / O. V. Veko // Nonlinear Phenomena in Complex Systems. – 2016. – Vol. 19, № 1. – P. 50–61.

6. Частица Кокса во внешнем магнитном поле: анализ в пространстве Лобачевского / О. В. Веко [и др.] // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2017. – № 4. – С. 55–65.

7. Elementary Particles with Internal Structure in External Fields. Vol. 1. General Theory / V. V. Kisel [et al.]. – New York: Nova Science Publishers, Inc., 2018. − 404 p.

8. Elementary Particles with Internal Structure in External Fields. Vol. 2. Physical Problems / V. V. Kisel [et al.]. – New York: Nova Science Publishers, Inc., 2018. − 402 p.

9. Heun, K. Zur Theorie der Riemann'schen Functionen zweiter Ordnung mit vier Verzweigungspunkten / K. Heun // Math. Ann. – 1989. – Bd. 33, № 2. – S. 161–179. https://doi.org/10.1007/bf01443849

10. Ronveaux, A. Heun's Differential Equation / A. Ronveaux. – Oxford: Oxford University Press, 1995. – 354 p.

11. Slavyanov, S. Yu. Special Functions. A Unified Theory Based on Singularities / S. Yu. Slavyanov, W. Lay. – Oxford: Oxford University Press, 2000. – 312 p.

References

1. Cox W. Higher-rank representations for zero-spin field theories. *Journal of Physics A: Mathematical and General*, 1982, vol. 15, no. 2, pp. 627–635. https://doi.org/10.1088/0305-4470/15/2/029

2. Ovsiyuk E. M. Spin-zero Cox's particle with an intrinsic structure: general analysis in external electromagnetic and gravitational fields. *Ukrainian Journal of Physics*, 2015, vol. 60, no. 6, pp. 485–496. https://doi.org/10.15407/ujpe60.06.0485

3. Kazmerchuk K. V., Ovsiyuk E. M. Cox's particle in magnetic and electric field against the background of Euclidean and spherical geometries. *Ukrainian Journal of Physics*, 2015, vol. 60, no. 5, pp. 389–400. https://doi.org/10.15407/ujpe60.05.0389

4. Ovsiyuk E. M., Veko O. V., Kazmerchuk K. V. Scalar particle with internal structure in the electromagnetic field in the curved space-time. *Problemy fiziki, matematiki i tekhniki = Problems of Physics, Mathematics and Technics*, 2014, no. 3 (20), pp. 32–36 (in Russian).

5. Veko O. V. Cox's particle in magnetic and electric fields on the background of hyperbolic Lobachevsky geometry. *Nonlinear Phenomena in Complex Systems*, 2016, vol. 19, no. 1, pp. 50–61.

6. Veko O. V., Voynova Ya. A., Ovsiyuk E. M., Red'kov V. M. Nonrelativistic Cox particle with intrinsic structure in magnetic field, analysis in Lobachevsky space. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2017, no. 4, pp. 55–65 (in Russian).

7. Kisel V. V., Ovsiyuk E. M., Balan V., Veko O. V., Red'kov V. M. *Elementary Particles with Internal Structure in External Field. Vol. 1. General Formalism*. USA, Nova Science Publishers, Inc., 2018. 404 p.

8. Kisel V. V., Ovsiyuk E. M., Balan V., Veko O. V., Red'kov V. M. *Elementary Particles with Internal Structure in External Field. Vol. 2. Physical Problems*. USA, Nova Science Publishers, Inc., 2018. 402 p.

9. Heun K. Zur Theorie der Riemann'schen Functionen zweiter Ordnung mit vier Verzweigungspunkten. *Mathematische Annalen*, 1989, vol. 33, no. 2, pp. 161–179. https://doi.org/10.1007/bf01443849

10. Ronveaux A. *Heun's Differential Equations*. Oxford, Oxford University Press, 1995. 354 p.

11. Slavyanov S. Yu., Lay W. *Special Functions. A Unified Theory Based on Singularities.* Oxford, Oxford University Press, 2000. 312 p.

Информация об авторах

Войнова Янина Александровна – аспирант, Институт физики им. Б. И. Степанова, Национальная академия наук Беларуси (пр. Независимости, 68-2, 220072, г. Минск, Республика Беларусь). E-mail: voinyuschka@mail.ru

Коральков Артем Дмитриевич – стажер младшего научного сотрудника, Мозырский государственный педагогический университет им. И. П. Шамякина (ул. Студенческая, 28, 247760, г. Мозырь, Гомельская обл., Республика Беларусь). E-mail: artemkoralkov@gmail.com

Овсиюк Елена Михайловна – кандидат физико-математических наук, доцент, заведующий кафедрой теоретической физики и прикладной информатики, Мозырский государственный педагогический университет им. И. П. Шамякина (ул. Студенческая, 28, 247760, г. Мозырь, Гомельская обл., Республика Беларусь). E-mail: e.ovsiyuk@mail.ru

Information about the authors

Voynova Yanina Alexandrovna – Postgraduate, B. I. Stepanov Institute of Physics, National Academy of Sciences of Belarus (68-2, Nezavisimosti Ave., 220072, Minsk, Republic of Belarus). E-mail: voinyuschka@mail.ru

Koralkov Artem Dmitrievich – Assistant Junior Researcher, Mozyr State Pedagogical University named after I. P. Shamyakin (28, Studencheskaya Str., 247760, Mozyr, Republic of Belarus). E-mail: artemkoralkov@gmail.com

Ovsiyuk Elena Mikhailovna – Ph. D. (Physics and Mathematics), Assistant Professor, Head of the Department of Theoretical Physics and Applied Informatics, Mozyr State Pedagogical University named after I. P. Shamyakin (28, Studencheskaya Str., 247760, Mozyr, Republic of Belarus). E-mail: e.ovsiyuk@mail.ru

ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)
УДК 534.535
https://doi.org/10.29235/1561-2430-2019-55-4-479-488

Поступила в редакцию 16.11.2019
Received 16.11.2019

В. Н. Белый[1], П. А. Хило[2], Н. С. Казак[1], Н. А. Хило[1]

[1]Институт физики им. Б. И. Степанова Национальной академии наук Беларуси, Минск, Беларусь
[2]Гомельский государственный технический университет им. П. О. Сухого, Гомель, Беларусь

ОСОБЕННОСТИ АКУСТООПТИЧЕСКОГО ВЗАИМОДЕЙСТВИЯ ОПТИЧЕСКИХ И АКУСТИЧЕСКИХ БЕССЕЛЕВЫХ ПУЧКОВ В ПОПЕРЕЧНО ИЗОТРОПНЫХ ОПТИЧЕСКИ ПОЛОЖИТЕЛЬНЫХ КРИСТАЛЛАХ

Аннотация. Исследованы особенности акустооптической (АО) дифракции с участием бесселевых светового и акустического пучков в анизотропных кристаллах. Рассмотрены кристаллы гексагональной симметрии, которые являются оптически положительными, а акустически – поперечно изотропными. Показано, что в отличие от случая АО дифракции плоских волн, переход к бесселевым пучкам позволяет реализовать ряд новых каналов дифракции, имеющих специфические конфигурации волновых векторов взаимодействующих волн при сохранении аксиальной симметрии оптической схемы в целом. Проведена классификация каналов дифракции для анизотропного рассеяния и для них рассчитаны основные параметры рассеянного бесселева светового пучка и параметры бесселева акустического пучка. Обнаружена возможность реализации изотропного типа дифракции, что позволяет повысить эффективность АО преобразования. Определены его характеристики для двух каналов рассеяния, а именно, рассеяния на попутном бесселевом акустическом пучке и на обратном.

Из-за многообразия каналов рассеяния, а также с учетом того, что бесселевы световой и акустический пучки обладают винтовыми дислокациями волнового фронта, а также подавленным дифракционным расплыванием, изучение особенностей АО дифракции таких пучков в оптически положительных кристаллах представляет как научный, так и практический интерес.

Ключевые слова: бесселевы световые пучки, бесселевы акустические пучки, акустооптическая дифракция, оптически положительные кристаллы

Для цитирования. Особенности акустооптического взаимодействия оптических и акустических бесселевых пучков в поперечно изотропных оптически положительных кристаллах / В. Н. Белый [и др.] // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2019. – Т. 55, № 4. – С. 479–488. https://doi.org/10.29235/1561-2430-2019-55-4-479-488

V. N. Belyi[1] , P. A. Khilo[2], N. S. Kazak[1], N. A. Khilo[1]

[1]B. I. Stepanov Institute of Physics of the National Academy of Sciences of Belarus, Minsk, Belarus
[2]Sukhoi State Technical University of Gomel, Gomel, Belarus

SOME FEATURES OF ACOUSTO-OPTIC INTERACTION OF OPTICAL AND ACOUSTIC BESSEL BEAMS IN TRANSVERSELY ISOTROPIC OPTICALLY POSITIVE CRYSTALS

Abstract. Some basic properties of acousto-optical (AO) diffraction involving Bessel light and acoustic beams in anisotropic crystals are investigated. Hexagonal symmetry crystals are considered and are optically uniaxial and positive and acoustically transversely isotropic. It is shown that, unlike the case of AO diffraction of plane waves, the transition to Bessel beams allows one to realize a number of new diffraction channels having specific configurations of the wave vectors of interacting waves while maintaining the axial symmetry of the optical scheme as a whole. The diffraction channels for anisotropic scattering are classified and the main parameters of the scattered Bessel light beam and the parameters of the Bessel acoustic beam are calculated for each of them. The possibility of implementing the isotropic-type diffraction was revealed, which makes it possible to increase the efficiency of AO conversion. The parameters of this-type diffraction are determined for two scattering channels, namely, for scattering by a direct Bessel acoustic beam and by a backward propagating acoustic beam.

Due to the appearance of a set of scattering channels and with regard to the fact that Bessel light and acoustic beams have helical wave front dislocations, as well as suppressed diffraction spreading, the study of the features of AO diffraction of such beams in optically positive crystals has both a scientific and practical interest.

Keywords: Bessel light beams, Bessel acoustic beams, acousto-optic diffraction, optically positive crystals

For citation. Belyi V. N., Khilo P. A., Kazak N. S., Khilo N. A. Some features of acousto-optic interaction of optical and acoustic Bessel beams in transversely isotropic optically positive crystals. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2019, vol. 55, no. 4, pp. 479–488 (in Russian). https://doi.org/10.29235/1561-2430-2019-55-4-479-488

Введение. Актуальной задачей оптики бесселевых световых пучков (БСП) является разработка методов перестройки их параметров, таких как состояние поляризации, угол конуса и порядок фазовой дислокации волнового фронта. Применение методов акустооптики считается перспективным решением данной задачи, так как позволяет одновременно управлять всеми названными параметрами БСП. Особенность акустооптической (АО) дифракции с участием бесселевых пучков (БП) состоит в том, что она предоставляет широкий набор возможностей реализации пространственных синхронизмов. Это связано, во-первых, с наличием конуса волновых векторов, из которых сформированы БП, и, во-вторых, с применением не только плоских акустических волн или пучков гауссова типа, но и бесселевых акустических пучков (БАП).

Акустооптическая дифракция бесселевых световых пучков на плоской акустической волне, распространяющейся вдоль оптической оси одноосного кристалла, была ранее исследована нами в [1–2]. Показано, что дифрагированный и падающий БСП отличаются углом конуса и могут отличаться также состоянием поляризации и порядком m фазовой дислокации волнового фронта. Вследствие распространения вдоль оптической оси переключение поляризации при анизотропной дифракции осуществляется между ТН- и ТЕ-собственными модами одноосного кристалла, являющимися строгими решениями уравнений Максвелла. Изменение порядка фазовой дислокации бесселевых световых пучков происходит из-за азимутальной зависимости параметров АО взаимодействия в кристаллах. В случае обратного рассеяния, когда изменяется направление распространения БСП, возможно пространственное разделение ТН- и ТЕ-поляризованных световых пучков, или, в частном случае $m = 0$, радиально и азимутально поляризованных БСП.

Переход к акустооптической дифракции с участием бесселевых акустических пучков расширяет возможности управления параметрами БСП. Это происходит из-за появления новых схем реализации векторного синхронизма, так как БАП, аналогично БСП, представляет собой конус плоско-волновых компонент. Следствием этого является зависимость эффективности АО дифракции от поперечного согласования трех бесселевых пучков, а не только от выполнения условия продольного фазового синхронизма [3–5]. Отметим, однако, что требование поперечного согласования бесселевых пучков может усложнить структуру акустооптической ячейки, в частности, привести к необходимости использования ультразвуковых преобразователей с конической структурой [4].

Существенно, что волновой фронт бесселевых акустических пучков в общем случае содержит винтовые дислокации волнового фронта, которые будут взаимодействовать с соответствующими дислокациями БСП. Интересно, что это взаимодействие дислокаций сопровождается изменением состояния поляризации бесселевых световых пучков, что представляет собой проявление спин-орбитального взаимодействия в процессе АО дифракции [3].

Исследованию АО дифракции с участием бесселевых акустических пучков способствует наличие развитой к настоящему времени строгой теории БАП в различных средах, включая акустически изотропные, поперечно-изотропные среды, а также кристаллы [6–9]. Также предложен и реализован ряд методов формирования БАП (их краткий обзор см., напр., в [4]).

Проведенные к настоящему времени исследования особенностей АО взаимодействия бесселевых пучков касаются только поперечно-изотропных оптически отрицательных кристаллов. Однако существует широкий набор оптически положительных кристаллов (например, CdS, ZnO, Ti и др.), которые обладают хорошими акустооптическими свойствами и широко используются в различных АО устройствах. Для изучения особенностей акустооптической дифракции в оптически положительных кристаллах необходимо вовлечение в рассмотрение новых конфигураций АО взаимодействия, в частности участия вертикально поляризованной поперечной акустической волны (SV-волны [8]). Таким образом, актуально исследование особенностей акустооптической дифракции БСП на бесселевых акустических пучках SV-типа в оптически положительных кристаллах.

В настоящей работе рассматриваются два типа АО дифракции в таких кристаллах, а именно: анизотропный и изотропный типы. В пределах каждого из них установлен и классифицирован ряд качественно различных каналов дифракции, для которых построены векторные диаграммы, соответствующие синхронному АО рассеянию. Получены формулы, позволяющие рассчитать параметры рассеянного БСП и БАП для каналов дифракции, соответствующих рассеянию вперед. Проведены численные расчеты зависимости этих параметров от угла конуса падающего бесселева светового пучка.

Векторные диаграммы акустооптического взаимодействия. Для акустооптического взаимодействия БСП в одноосных кристаллах оптимальная геометрия, соответствующая цилиндрической симметрии пучков, реализуется при их распространении вдоль оптической оси. В этом случае отсутствуют искажения взаимодействующих пучков, вызванные анизотропией. Что касается БАП, то для уменьшения эффектов анизотропии наиболее простым является случай распространения в акустически поперечно изотропных средах. Далее этот случай и будет рассматриваться.

На рис. 1 показана часть сечения поверхностей волновых векторов одноосного положительного кристалла плоскостью, проходящей через оптическую ось c, вдоль которой направлена декартова ось z. Эллиптическая кривая относится к необыкновенной плоской волне (индекс e), или же ТН-бесселевой моде; круговая расположена внутри эллиптической ($n_o < n_e$) кривой и относится к обыкновенной плоской волне (индекс o), или ТЕ-моде БСП.

В общем случае в рассматриваемой схеме возможны четыре типа акустооптической дифракции: $e \to o$, $e \to e$, $o \to e$ и $o \to o$. Далее мы ограничимся рассмотрением двух из них, а именно: анизотропным $e \to o$ и изотропным $e \to e$ типами. В этом случае падающий БСП является ТН-поляризованным (или e-пучком). Волновой вектор этого пучка, лежащий в плоскости (x, z), обозначен на рис. 1 как $k_{e,in}$. Волновые векторы $k_{o,d}$ и $k_{e,d}$ относятся к дифрагированным БСП, а вектор k_s – к БАП. Световые падающий и дифрагированный бесселевы световые пучки распространяются в положительном направлении оси z, а акустический БП может распространяться как попутно, так и встречно по отношению к БСП. Волновые векторы для светового и акустического БП (см. рис. 1) имеют смысл локальных волновых векторов из бесконечных их суперпозиций, образующих поверхности круговых конусов с осью симметрии, совпадающей с оптической осью. Для локальных волновых векторов БСП и БАП, аналогично плоским волнам, можно ввести в рассмотрение векторные диаграммы АО дифракции, связывающие волновые векторы падающего и дифрагированного БСП с соответствующим волновым вектором БАП. В общем случае эти три вектора могут иметь разные азимутальные координаты, что приводит к большому многообразию типов или каналов АО рассеяния. Далее ограничимся более простым случаем

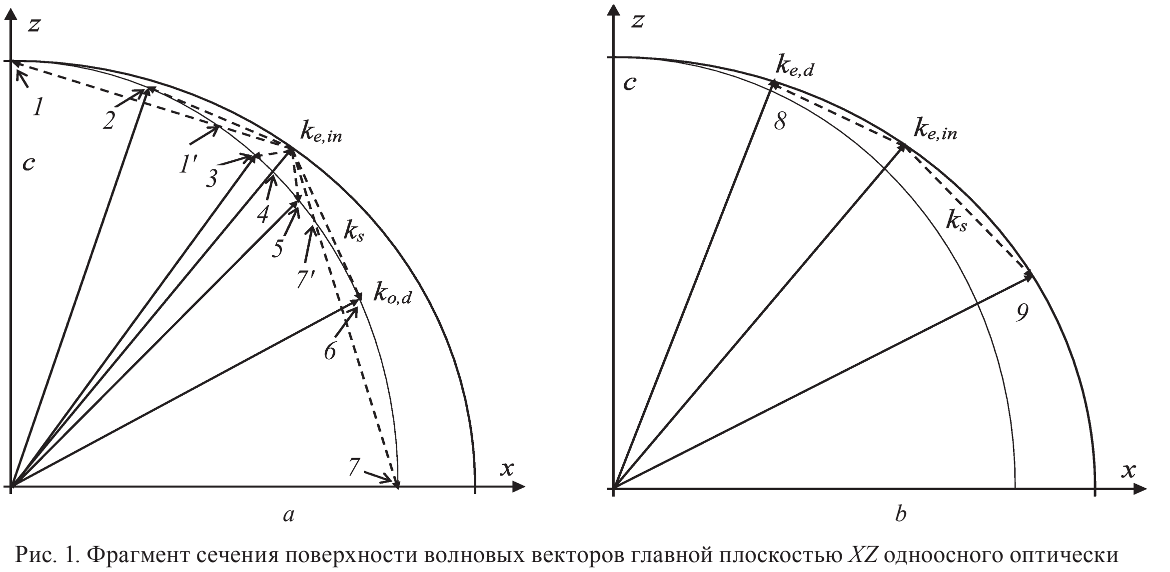


Рис. 1. Фрагмент сечения поверхности волновых векторов главной плоскостью XZ одноосного оптически положительного кристалла (c – оптическая ось) и векторные диаграммы для характерных каналов анизотропной (a) и изотропной (b) АО дифракции БСП и БАП; волновые векторы БАП показаны штриховыми линиями

Fig. 1. Fragment of the cross section of the wave vector surface by the XZ plane of a uniaxial optically positive crystal (c is the optical axis) and vector diagrams for characteristic channels of anisotropic (a) and isotropic (b) AO diffraction of BLB and BAB; BAB wave vectors are shown by the dashed lines

равенства азимутальных координат векторов $\vec{k}_{in}, \vec{k}_d$ и $\vec{k}_s$, когда все возможные векторные диаграммы расположены в плоскости, содержащей оптическую ось (см. рис. 1). Для этого случая можно провести удобную классификацию векторных диаграмм и соответствующих им типов акустооптической дифракции или каналов рассеяния. Было выделено 7 характерных конфигураций волновых векторов или векторных диаграмм для АО дифракции $e \to o$ типа. Соответствующие им каналы дифракции или рассеяния в плоскости (x, z) обозначены цифрами *1–7*. Аналогично каналы *8* и *9* на рис. 1, *b* относятся к $e \to e$ типу дифракции.

При наличии векторного синхронизма локальные волновые векторы падающего и дифрагированного БСП и соответствующий вектор БАП образуют замкнутые треугольники. Отсюда следуют уравнения векторного синхронизма для БП, аналогичные известным уравнениям для плоских волн:

$$\vec{k}_e + \vec{k}_s = \vec{k}_{d,o}, \tag{1а}$$

$$\vec{k}_e + \vec{k}_s = \vec{k}_{d,e}. \tag{1б}$$

Уравнения (1а) и (1б) описывают соответственно анизотропную и изотропную дифракцию.

На рис. 1 для семи каналов анизотропной $e \to o$-дифракции и двух каналов изотропной $e \to e$-дифракции волновые числа падающего и дифрагированных БСП обозначены как $k_{e,in}$ и $k_{o,d}$, $k_{e,d}$, а k_s обозначает волновое число звука.

Некоторые свойства каналов рассеяния *1–7* следуют непосредственно из рис. 1. Так, в канале *1* рассеянное поле распространяется вдоль оптической оси, т. е. является плоской волной; канал *2* характерен тем, что волновой вектор БАП касается o-поверхности, а не пересекает ее; в канале *3* происходит рассеяние на развернутом БАП (угол конуса равен 90°); в канале *4* реализуется коллинеарное рассеяние на обратном БАП (здесь углы конуса трех бесселевых пучков совпадают); канал *5* соответствует дифракции на обратной плоской волне, а канал *6* аналогичен каналу *2*, но для обратного БАП; для канала *7* дифрагированный o-БСП является развернутым, т. е. угол конуса равен 90°; каналы *8* и *9* соответствуют изотропной $e \to e$-дифракции на попутном (или прямом) и обратном БАП.

Для каналов *1–3*, соответствующих дифракции на прямых бесселевых акустических пучках, можно определить минимальный угол конуса $\gamma_{s,\min}$, необходимый для достижения синхронизма. Из рис. 1, *a* видно, что он реализуется для дифракции в канал *2*, т. е. $\gamma_{s,\min} = \gamma_{s,2}$. Поскольку для этого канала парциальный волновой вектор k_s касается o-поверхности, то данный синхронизм может быть назван касательным. Для углов конуса, больших, чем $\gamma_{s,2}$, уравнения синхронизма будут иметь 2 решения, так как акустический волновой вектор в этом случае дважды пересекает o-поверхность. Это видно из рис. 1, *a* для дифракции, например, в каналы *1* и *1′*. Здесь один из каналов (*1′*) является относительно низкочастотным, а второй (*1*) – высокочастотным. При этом по мере увеличения угла конуса γ_s резонансная акустическая частота низкочастотного канала уменьшается, а высокочастотного – возрастает. Далее мы не будем рассматривать малоинтересные с технической точки зрения высокочастотные каналы дифракции, соответствующие точкам пересечения левее точки *1* на рис. 1, *a*.

Минимальное значение акустической частоты требуется в канале рассеяния *3*, который соответствует предельному случаю дифракции на развернутом БАП, когда угол конуса $\gamma_s = 90°$.

Анизотропное рассеяние на прямом бесселевом акустическом пучке. Количественный анализ на основе векторных диаграмм. Для количественного анализа $e \to o$-дифракции обратимся к уравнению синхронизма (1а). Проекция векторов этого уравнения на ось z и на плоскость (x, y) определяет условия так называемых продольного и поперечного синхронизмов при акустооптическом взаимодействии БСП и БАП:

$$\begin{aligned} k_e \cos\gamma_e + k_s \cos\gamma_s &= k_o \cos\gamma_o, \\ k_e \sin\gamma_e - k_s \sin\gamma_s &= k_o \sin\gamma_o. \end{aligned} \tag{2}$$

Отсюда находим волновое число БАП и угол конуса дифрагированного БСП:

$$k_s = -k_e \cos(\gamma_e + \gamma_s) \pm \sqrt{k_o^2 - k_e^2 \sin(\gamma_e + \gamma_s)^2}, \tag{3}$$

$$\operatorname{tg}\gamma_o = \frac{k_e \sin\gamma_e - k_s \sin\gamma_s}{k_e \cos\gamma_e + k_s \cos\gamma_s}. \tag{4}$$

Данные формулы позволяют рассчитать, например, зависимость k_s и γ_o от угла конуса γ_s БАП и угла конуса $\gamma_e = \gamma_{e,in}$ падающего *e*-БСП. По известному волновому числу k_s далее может быть рассчитана акустическая частота $f_s = k_s v(\gamma_s)/2\pi$, где $v(\gamma_o)$ – зависящая от угла конуса γ_s скорость SV – БАП. Эта скорость рассчитывается из уравнения [10]

$$2\rho v^2 = c_{44} + c_{11}\sin^2\gamma_s + c_{33}\cos^2\gamma_s -$$
$$-\left[\left((c_{11} - c_{44})\sin^2\gamma_s + (c_{44} - c_{33})\cos^2\gamma_s\right)^2 + (c_{13} + c_{44})^2 \sin^2 2\gamma_s\right]^{1/2}, \tag{5}$$

где c_{ik} – компоненты тензора упругой жесткости, ρ – плотность кристалла.

Применим полученные формулы для определения угла конуса и частоты бесселева акустического пучка для каналов рассеяния *1*, *2* и *3*. Минимальный угол конуса БАП, при котором возможна синхронная дифракция, соответствует, как указано выше, касательному синхронизму и реализуется для канала *2*. Этот угол можно найти из условия равенства нулю подкоренного выражения в (3), откуда следует

$$\gamma_{s,2} = \cos^{-1}\left(\frac{n_o}{n_e(\gamma_e)}\right) - \gamma_e + \pi/2. \tag{6}$$

Волновое число БАП при этом равно

$$k_s = \sqrt{k_e(\gamma_e)^2 - k_o^2},$$

а угол дифракции γ_o находим из (4).

На рис. 2 показаны зависимости угла конуса $\gamma_{s,2}$, частоты БАП, а также угла конуса γ_o дифрагированного *o*-БСП в канале *2* от угла конуса γ_e падающего *e*-БСП. Графики на рис. 2 и 3–6 построены применительно к кристаллу ZnO со следующими параметрами:

$n_o = 1{,}998$; $n_e = 2{,}0147$; $c_{44} = 4{,}25 \cdot 10^{10}$ Па; $c_{11} = 20{,}97 \cdot 10^{10}$ Па; $c_{33} = 21{,}09 \cdot 10^{10}$ Па; $c_{13} = 10{,}51 \cdot 10^{10}$ Па; $\rho = 5{,}67 \cdot 10^3$ кг/м^3 для длины волны $\lambda = 0{,}63$ мкм.

Из рис. 2, *a* видно, что касательный синхронизм может быть реализован для любого значения угла конуса падающего БСП при соответствующем подборе угла конуса и частоты БАП. Для двух предельных случаев $\gamma_e = 0$ и $\gamma_e = \pi/2$ из (6) находим максимальное и минимальное значения угла конуса бесселева акустического пучка, равные соответственно $\gamma_{s,2} = \pi/2$ и $\gamma_{s,2} = \cos^{-1}(n_o/n_e)$. Для кристалла ZnO минимальный угол $\gamma_{s,2}$ равен 7,38°.

Частота БАП в канале *2* рассчитывается, подстановкой в формулу

$$f_s = k_s v(\gamma_s)/2\pi$$

скорости $v(\gamma_s)$ из (5) и волнового числа БАП, равного

$$k_s = k_0\sqrt{n_e(\gamma_e)^2 - n_o^2}.$$

При $\gamma_e \to 0$ волновое число БАП и его частота также стремятся к нулю (рис. 2, *b*). Максимальная частота БАП требуется для дифракции на развернутом БСП ($\gamma_e = \pi/2$). В этом случае

$$\gamma_{s,2} = \cos^{-1}(n_o/n_e), \quad f_{s,\max} = \lambda^{-1} v(\gamma_{s,2})\sqrt{n_e^2 - n_o^2}.$$

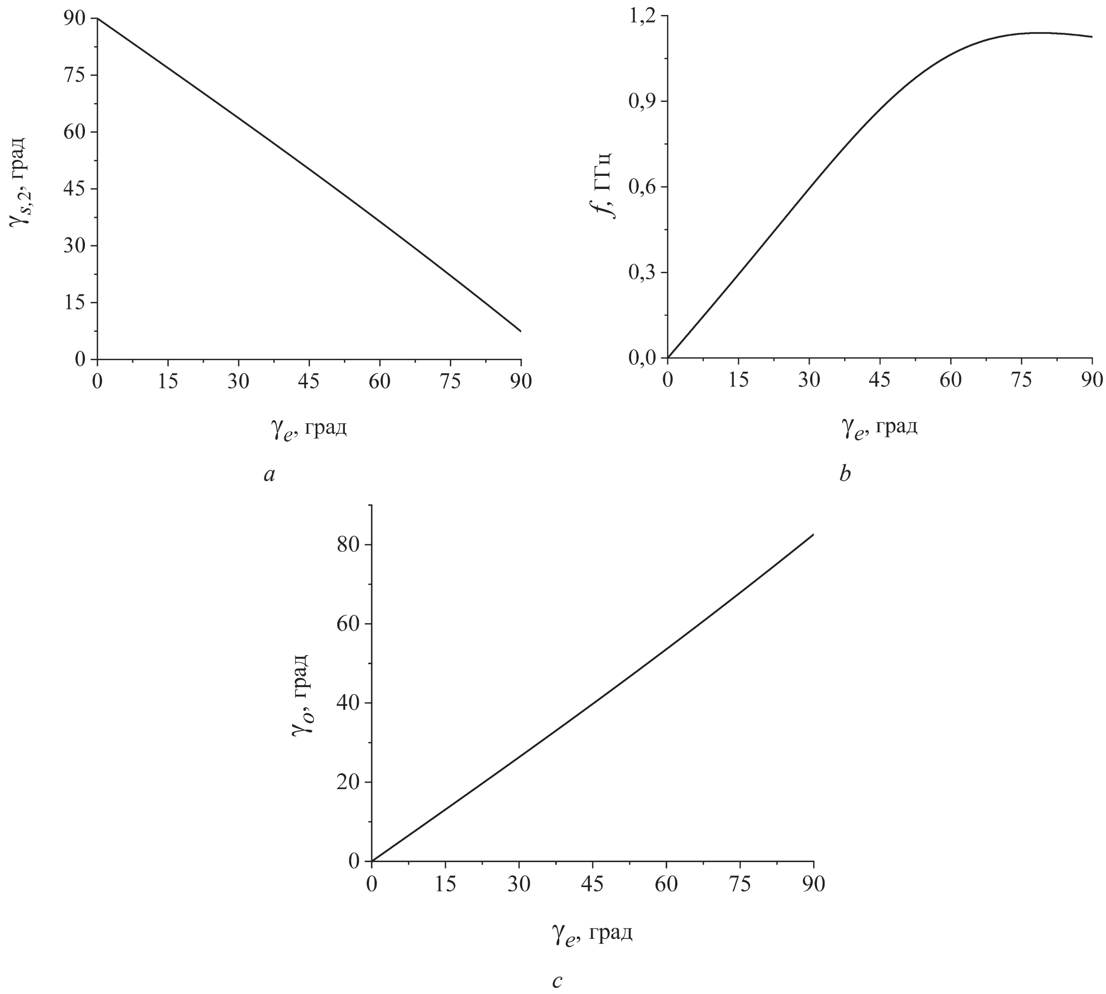


Рис. 2. Зависимость минимального угла конуса бесселева акустического пучка (*a*), частоты звука (*b*) и угла конуса дифрагированного бесселева светового пучка (*c*) от угла конуса падающего БСП для касательного синхронизма

Fig. 2. Dependence of the minimum cone angle of the BAB (*a*), the acoustic frequency (*b*) and the cone angle of the diffracted BLB (*c*) on the cone angle of the incident BLB for tangential synchronism

Расчет для кристалла ZnO дает $f_{s,\max}$ = 1,18 ГГц. В промежутке между этими крайними значениями угла γ_e акустическая частота возрастает нелинейно с насыщением при бóльших значениях γ_e (см. рис. 2, *b*).

Зависимость углов конуса $\gamma_e(\gamma_o)$ падающего и дифрагированного бесселевых световых пучков описывается формулой (4), из которой следует (см. также рис. 2, *c*), что при $\gamma_e = 0$ угол $\gamma_o = 0$, а при $\gamma_e = \pi/2$ угол γ_o находится из уравнения

$$\operatorname{tg}\gamma_o = n_o \Big/ \sqrt{n_e^2 - n_o^2}\,.$$

Для кристалла ZnO этот угол равен 82,62°. Для канала *1* из (2) находим

$$\operatorname{tg}\gamma_s = \frac{k_e(\gamma_e)\sin\gamma_e}{k_o - k_e(\gamma_e)\cos\gamma_e}, \qquad k_s = k_e(\gamma_e)\sin\gamma_e \big/ \sin\gamma_s\,. \tag{7}$$

При $\gamma_e \to 0$ расчет предела первого уравнения (7) дает, что угол $\gamma_s \to \pi/2$. При $\gamma_e = \pi/2$ получим $\gamma_s = \operatorname{tg}^{-1}(n_e/n_o)$. Для кристалла ZnO этот угол равен 45,24°. Зависимость $\gamma_s(\gamma_e)$ во всем диапазоне углов конуса показана на рис. 3, *a*.

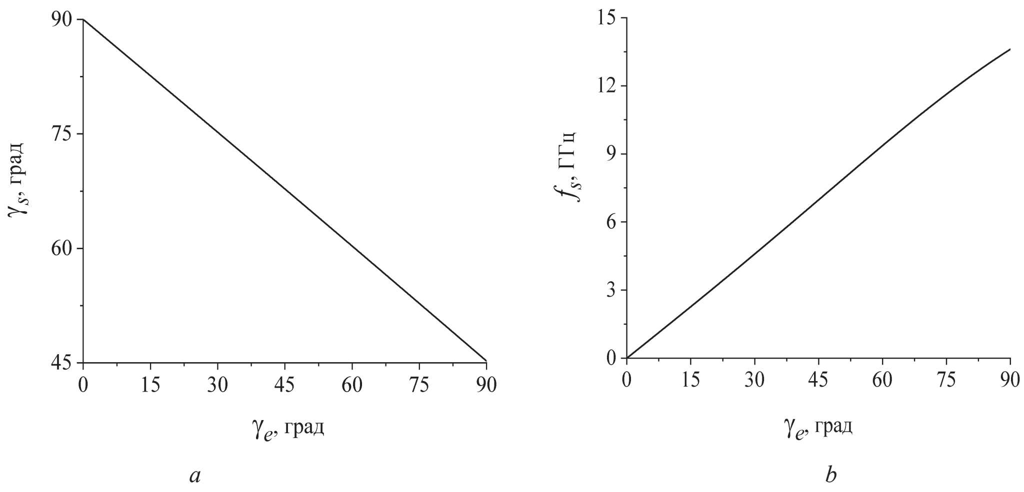


Рис. 3. Зависимость угла конуса бесселева акустического пучка (*a*) и частоты звука (*b*) от угла конуса падающего бесселева светового пучка для канала дифракции *1*

Fig. 3. Dependence of the BAB cone angle (*a*) and the acoustic frequency (*b*) on the cone angle of the incident BLB for diffraction channel *1*

Из второго уравнения (7) определяем зависимость частоты звука от угла γ_e. Видно, что при $\gamma_e \to 0$ частота также стремится к нулю. В другом предельном случае $\gamma_e = \pi/2$ и с учетом того, что в данном случае

$$\sin(\gamma_s) = n_e \Big/ \sqrt{n_o^2 + n_e^2},$$

получим

$$k_s = k_0 \sqrt{n_o^2 + n_e^2}.$$

Как видим, волновое число БАП в данном случае превосходит волновое число БСП, что указывает на предельно высокую акустическую частоту. Согласно численным расчетам (см. рис. 3, *b*), частота здесь равна 13,61 ГГц.

Для канала дифракции *3* (рис. 4) угол $\gamma_s = 90°$ и формула для волнового числа низкочастотного БАП в (3) примет вид

$$k_s = k_e(\gamma_e)\sin\gamma_e - \sqrt{k_o^2 - k_e(\gamma_e)^2 \cos\gamma_e^2}. \tag{8}$$

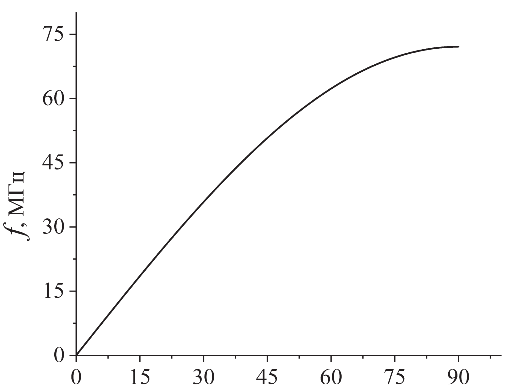


Рис. 4. Зависимость частоты звука от угла конуса падающего бесселева светового пучка для канала дифракции *3*

Fig. 4. Dependence of the acoustic frequency on the cone angle of the incident BLB for diffraction channel *3*

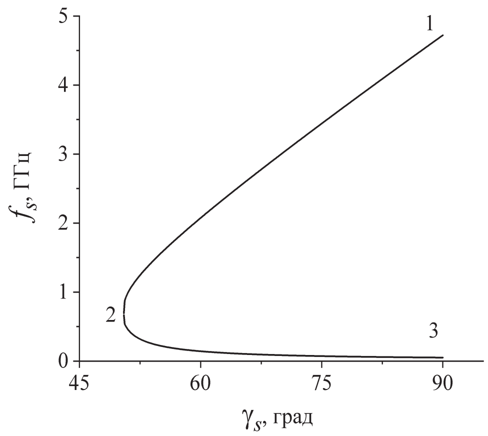


Рис. 5. Зависимость акустической частоты от угла конуса бесселева акустического пучка для промежуточных каналов дифракции между каналами *3* и *1*

Fig. 5. Dependence of the acoustic frequency on the BAB cone angle for the intermediate diffraction channels between *3* and *1*

Из (8) следует, что при $\gamma_e \to 0$ волновое число БАП и его частота также стремятся к нулю. При $\gamma_e = \pi/2$ получаем $k_s = k_0(n_e - n_o)$ или $f_s = v_s(\pi/2)(n_e - n_o)/\lambda$. Для кристалла ZnO эта частота равна 72,1 МГц.

В общем случае канал высокочастотной дифракции располагается в диапазоне между точками *2* и *1*, а низкочастотной – между *2* и *3* (см. рис. 1, *a*). Зависимость между частотой и углом конуса БАП для обоих каналов показана на рис. 5. Угол конуса $\gamma_{e,in}$ принят равным 45°.

Цифры *1–3* на графике соответствуют каналам дифракции, показанным на рис. 1, *a*. Как видим, в окрестности канала *2* имеет место высокая чувствительность акустической частоты по отношению к изменению угла конуса БАП. Ясно, что аналогичная ситуация имеет место и вблизи канала дифракции *6* (см. рис. 1, *a*).

Изотропное рассеяние. Из укороченных уравнений, описывающих акустооптическую дифракцию в кристаллах гексагональной симметрии, следует, что в данных кристаллах индуцированное АО взаимодействием изменение тензора диэлектрической проницаемости для вертикально поляризованной (SV) акустической волны наряду с недиагональными компонентами имеет также и ненулевые диагональные компоненты, т. е. в цилиндрических координатах компоненты $\Delta\varepsilon_{\rho\rho}$, $\Delta\varepsilon_{\varphi\varphi}$ и $\Delta\varepsilon_{zz}$. Как следствие, в оптически положительных кристаллах при возбуждении вертикально поляризованной поперечной волны становится возможным рассеяние типа $e \to e$. Данный тип рассеяния представляет практический интерес в связи с перспективой увеличения эффективности АО преобразования вследствие улучшения пространственного синхронизма при взаимодействиях одинаковых по структуре компонент бесселевых пучков.

Векторная диаграмма для $e{\to}e$-типа рассеяния показана на рис. 1, *b*, из которой для рассеяния на прямом бесселевом акустическом пучке следуют уравнения синхронизма

$$\begin{aligned} k_e \cos\gamma_e + k_s \cos\gamma_s &= k_{e,d} \cos\gamma_{e,d}, \\ k_e \sin\gamma_e - k_s \sin\gamma_s &= k_{e,d} \sin\gamma_{e,d}. \end{aligned} \tag{9}$$

Для рассеяния на обратном БАП в (9) необходимо заменить $k_s \to -k_s$. Для заданного угла конуса падающего бесселева светового пучка из (9) находим зависимости волнового числа k_s и угла конуса γ_s БАП от угла конуса дифрагированного БСП, необходимые для реализации синхронной дифракции

$$\begin{aligned} k_s &= \sqrt{k_{e,d}^2(\gamma_{e,d}) + k_e^2(\gamma_e) - 2k_e k_{e,d}\cos(\gamma_{ed} - \gamma_e)}, \\ \mathrm{tg}\,\gamma_s &= \frac{k_e(\gamma_e)\sin\gamma_e - k_{e,d}(\gamma_{e,d})\sin\gamma_{e,d}}{k_{e,d}(\gamma_{e,d})\cos\gamma_{e,d} - k_e(\gamma_e)\cos\gamma_e}. \end{aligned} \tag{10}$$

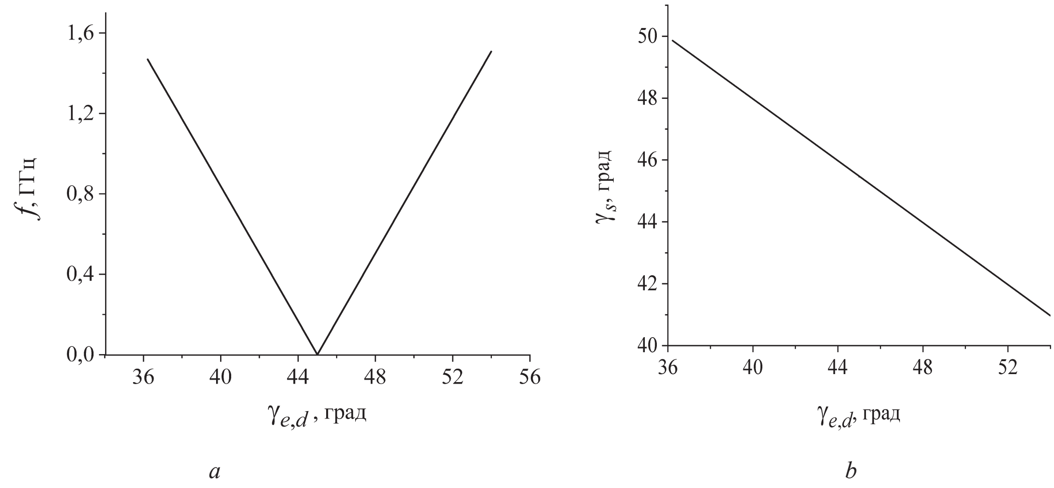


Рис. 6. Зависимость частоты и угла конуса бесселева акустического пучка в изотропном *e*→*e*-канале дифракции

Fig. 6. Dependence of the frequency and the BAB cone angle in the isotropic *e*→*e* diffraction channel

На рис. 6 показаны зависимости частоты и угла конуса бесселева акустического пучка в *e*→*e*-канале дифракции для случая, когда угол конуса дифрагированного бесселева светового пучка расположен в окрестности угла конуса падающего БСП. Угол последнего выбран равным $\gamma_e = 45°$. Из графика на рис. 6, *a* следует, что для изменения угла конуса БСП на 1° необходимо изменение частоты БАП, равное примерно 167 МГц. Из графика на рис. 6, *b* находим, что требуемое при этом изменение угла конуса бесселева акустического пучка равно 0,5º. Тогда, к примеру, для уменьшения угла конуса БСП на 3° до величины $\gamma_{e,d} = 42°$ требуется БАП с частотой ≈ 500 МГц и углом конуса ≈ 43º.

Заключение. Проведено описание акустооптической дифракции бесселевых световых и акустических пучков в поперечно изотропных оптически положительных кристаллах гексагональных классов симметрии. Рассмотрена цилиндрически симметричная геометрия АО рассеяния, когда световые и акустический бесселевы пучки распространяются вдоль оптической оси кристалла. Исследованы анизотропный и изотропный типы дифракции, когда падающий ТН-поляризованный БСП порождает соответственно ТН- и ТЕ-поляризованные БСП. Показано, что в зависимости от частоты и угла конуса бесселева акустического пучка может существовать ряд качественно отличных каналов дифракции. Для этих каналов проведены расчеты параметров БАП, а именно: угла конуса и частоты, необходимых для реализации продольного и поперечного синхронизмов. Также определены параметры рассеянного БСП для случая рассеяния вперед на прямом БАП. Исследованная схема АО дифракции в оптически положительных кристаллах представляет также практический интерес для разработки акустооптических дефлекторов и модуляторов на основе бесселевых пучков.

Список использованных источников

1. Formation of TH- and TE-polarized Bessel light beams at acousto- optic diffraction in anisotropic crystals / V. N. Belyi [et al.] // Proc. SPIE. – 2011. – Vol. 8073. – P. 807327-1-9. https://doi.org/10.1117/12.886443

2. Generation of TH- and TE-polarized Bessel light beams at acousto-optic interaction in anisotropic crystals / P. A. Khilo [et al.] // Opt. Commun. – 2014. – Vol. 325, № 7. – P. 84–91. https://doi.org/10.1016/j. optcom.2014.03.061

3. Transformation of phase dislocations under acousto-optic interaction of optical and acoustical Bessel beams / V. N. Belyi [et al.] // J. Opt. – 2016. – Vol. 18. – 074002 (6 p.). https://doi.org/10.1088/2040-8978/18/7/074002

4. Low-frequency acousto-optic backscattering of Bessel light beams / N. A. Khilo [et al.] // Optics Commun. – 2018. – Vol. 415. – P. 6–12. https://doi.org/10.1016/j.optcom.2018.01.024

5. Низкочастотное обратное акустооптическое рассеяние бесселевых световых пучков / В. Н. Белый [и др.] // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2017. – № 3. – С. 78–86.

6. Nondifracting bulk-acoustic X waves in crystals / J. Salo [et al.] // Phys. Rev. Lett. –1999. – Vol. 83, № 6. – P. 1171–1174. https://doi.org/10.1103/physrevlett.83.1171

7. Salo, J. Nondiffracting waves in anisotropic media / J. Salo, M. M. Salomaa // Phys. Rev. E. – 1999. – Vol. 67, № 5. – 056609 (9 p.). https://doi.org/10.1103/physreve.67.056609

8. Honarvar, F. Acoustic wave scattering from transversely isotropic cylinders / F. Honarvar, A. N. Sinclair // J. Acoust. Soc. Am. – 2003. – Vol. 100, № 1. – P. 57–63. https://doi.org/10.1121/1.415868

9. Ahmad, F. Acoustic scattering by transversely isotropic cylinders / F. Ahmad, A. Rahman // Int. J. Eng. Sci. – 2000. – Vol. 38, № 5. – P. 325–335. https://doi.org/10.1016/s0020-7225(99)00031-2

10. Auld, B. A. Acoustic Fields and Waves in Solids / B. A. Auld. – John Wiley and Sons, Inc., 1973. – Vol. 1. – 430 p.

References

1. Belyi V. N., Khilo P. A., Petrova E. S., Khilo N. A., Kazak N. S. Formation of TH- and TE-polarized Bessel light beams at acousto- optic diffraction in anisotropic crystals. *Proceedings of SPIE*, 2011, vol. 8073, pp. 807327-1-9. https://doi.org/10.1117/12.886443

2. Khilo P. A., Kazak N. S., Khilo N. A. and Belyi V. N. Generation of TH- and TE-polarized Bessel light beams at acousto-optic interaction in anisotropic crystals. *Optics Communications*, 2014, vol. 325, no. 7, pp. 84–91. https://doi.org/10.1016/j.optcom.2014.03.061

3. Belyi V. N., Khilo P. A., Kazak N. S., Khilo N. A. Transformation of phase dislocations under acousto-optic interaction of optical and acoustical Bessel beams. *Journal of Optics*, 2016, vol. 18, 074002 (6 p.). https://doi.org/10.1088/2040-8978/18/7/074002

4. Khilo N. A., Belyi V. N., Khilo P. A., Kazak N. S. Low-frequency acousto-optic backscattering of Bessel light beams. *Optics Communication*, 2018, vol. 415, pp. 6–12. https://doi.org/10.1016/j.optcom.2018.01.024

5. Belyi V. N., Khilo P. A., Kazak N. S., Khilo N. A. Low-frequency backward acousto-optic scattering of Bessel light beams. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2017, no. 3, pp. 78–86 (in Russian).

6. Salo J., Fagerholm J., Friberg A. T., Salomaa M. M. Nondifracting bulk-acoustic X waves in crystals. *Physical Review Letters*, 1999, vol. 83, no. 6, pp. 1171–1174. https://doi.org/10.1103/physrevlett.83.1171

7. Salo J., Salomaa M. M. Nondiffracting waves in anisotropic media. *Physical Review E*, 1999, vol. 67, no. 5, 056609 (9 p.). https://doi.org/10.1103/physreve.67.056609

8. Hanorvar F., Sinclair N. N. Acoustic wave scattering from transversely isotropic cylinders. *The Journal of the Acoustical Society of America,* 2003, vol. 100, no. 1, pp. 57–63. https://doi.org/10.1121/1.415868

9. Ahmad F., Rahman A. Acoustic scattering by transversely isotropic cylinders. *International Journal of Engineering Science*, 2000, vol. 38, no. 5, pp. 325–335. https://doi.org/10.1016/s0020-7225(99)00031-2

10. Auld B. A. *Acoustic Fields and Waves in Solids. Volume 1*. John Wiley and Sons, Inc., 1973. 430 p.

Информация об авторах

Белый Владимир Николаевич – член-корреспондент, доктор физико-математических наук, заведующий центром «Диагностические системы», Институт физики им. Б. И. Степанова Национальной академии наук Беларуси (пр. Независимости, 68-2, 220072, г. Минск, Республика Беларусь). E-mail: v.belyi@dragon.bas-net.by

Хило Петр Анатольевич – доктор физико-математических наук, профессор, заведующий кафедрой общей физики, Гомельский государственный технический университет им. П. О. Сухого (пр. Октября, 48, 246746, г. Гомель, Республика Беларусь). E-mail: khilo_p@tut.by

Казак Николай Станиславович – академик, доктор физико-математических наук, главный научный сотрудник, Институт физики им. Б. И. Степанова Национальной академии наук Беларуси (пр. Независимости, 68, 220072, г. Минск, Республика Беларусь). E-mail: lod@dragon.bas-net.by

Хило Николай Анатольевич – кандидат физико-математических наук, ведущий научный сотрудник, Институт физики им. Б. И. Степанова Национальной академии наук Беларуси (пр. Независимости, 68-2, 220072, г. Минск, Республика Беларусь). E-mail: n.khilo@dragon.bas-net.by

Information about the authors

Vladimir N. Belyi – Corresponding Member, Dr. Sc. (Physics and Mathematics), Professor, Head of the Center "Diagnostic Systems", B. I. Stepanov Institute of Physics of the National Academy of Sciences of Belarus (68-2, Nezavisimosti Ave., 220072, Minsk, Republic of Belarus). E-mail: v.belyi@dragon.bas-net.by

Piotr A. Khilo – Dr. Sc. (Physics and Mathematics), Professor, Head of the Department of General Physics, Sukhoi State Technical University of Gomel (48, Oktyabrya Ave., 48, 246746, Gomel, Republic of Belarus). E-mail: khilo_p@tut.by

Nikolai S. Kazak – Academician, Dr. Sc. (Physics and Mathematics), Chief Scientific Researcher, B. I. Stepanov Institute of Physics of the National Academy of Sciences of Belarus (68-2, Nezavisimosti Ave., 220072, Minsk, Republic of Belarus). E-mail: lod@dragon.bas-net.by

Nikolai A. Khilo – Ph. D. (Physics and Mathematics), Leading Scientific Researcher, B. I. Stepanov Institute of Physics of the National Academy of Sciences of Belarus (68-2, Nezavisimosti Ave., 220072, Minsk, Republic of Belarus). E-mail: n.khilo@dragon.bas-net.by

ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)
УДК 621.315.592
https://doi.org/10.29235/1561-2430-2019-55-4-489-497

Поступила в редакцию 02.09.2019
Received 02.09.2019

Ф. П. Коршунов, Н. Е. Жданович, Д. Н. Жданович

Научно-практический центр Национальной академии наук Беларуси по материаловедению, Минск, Беларусь

ВЛИЯНИЕ ВЫСОКОТЕМПЕРАТУРНОГО ОТЖИГА НА ХАРАКТЕРИСТИКИ ОБЛУЧЕННЫХ БЫСТРЫМИ ЭЛЕКТРОНАМИ *p-n*-СТРУКТУР НА ЯДЕРНО-ЛЕГИРОВАННОМ КРЕМНИИ

Аннотация. Приводятся результаты исследования влияния отжига ($T_{отж}$ = 300–800 ºC) на время жизни неосновных носителей заряда τ_P в *n*-базе *p-n*-структур на базе высокоомного ядерно-легированного кремния (ЯЛК) КОФ300, облученных электронами с E_e = 4 МэВ при комнатной температуре флюенсами $\Phi = 1 \cdot 10^{14}$–$3 \cdot 10^{16}$ см$^{-2}$. Установлено, что при малых флюенсах электронов ($\Phi = 1 \cdot 10^{14}$ см$^{-2}$) отжиг времени жизни неосновных носителей заряда τ_P в *n*-базе структур проходит в две стадии: первая – 320–400 ºC, вторая – 550–650 ºC. При более высоких флюенсах облучения ($\Phi = 5 \cdot 10^{15}$–$2 \cdot 10^{16}$ см$^{-2}$) наблюдается три стадии отжига: первая – 400–450 ºC, вторая – 520–650 ºC и третья – 710–770 ºC. При этом на зависимости барьерной емкости *C* структур от $T_{отж}$ для высоких флюенсов облучения до $T_{отж}$ = 400 ºC измеряется геометрическая емкость. В диапазоне $T_{отж}$ = 420–570 ºC наблюдается рост *C* с максимумом при $T_{отж}$ = 480 ºC и последующий спад до значений геометрической емкости в области $T_{отж}$ = 600–670 ºC, а затем рост в области $T_{отж}$ = 720–770 ºC до значений, соответствующих необлученному образцу с выходом на плато при $T_{отж}$ = 770–800 ºC. Анализ DLTS-спектров исследуемых структур позволил установить образование в процессе отжига глубокого акцепторного уровня E_C – 0,68 эВ при $T_{отж}$ > 400 ºC, глубокого донорного уровня E_C – 0,32 эВ при отжиге в диапазоне $T_{отж}$ = 420–570 ºC и глубокого акцепторного уровня E_C – 0,53 эВ при $T_{отж}$ > 700 ºC, что удовлетворительно объясняет полученные в данной работе зависимости τ_P и *C* от $T_{отж}$.

Ключевые слова: радиационный дефект, радиационно-термический дефект, электронное облучение, высокотемпературный отжиг, время жизни неосновных носителей заряда, DLTS-спектроскопия

Для цитирования. Коршунов, Ф. П. Влияние высокотемпературного отжига на характеристики облученных быстрыми электронами *p-n*-структур на ядерно-легированном кремнии / Ф. П. Коршунов, Н. Е. Жданович, Д. Н. Жданович // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2019. – Т. 55, № 4. – С. 489–497. https://doi.org/10.29235/1561-2430-2019-55-4-489-497

F. P. Korshunov, N. E. Zhdanovich, D. N. Zhdanovich

Scientific-Practical Materials Research Centre of the National Academy of Sciences of Belarus, Minsk, Belarus

INFLUENCE OF HIGH-TEMPERATURE ANNEALING ON THE CHARACTERISTICS OF FAST ELECTRON-IRRADIATED *p-n*-STRUCTURES BASED ON NEUTRON DOPED SILICON

Abstract. The investigation results of the annealing influence (T_{ann} = 300–800 ºC) on the minority charge currier lifetime τ_P in the *n*-base of *p-n*-structures, manufactured on the base of neutron transmutation doped silicon (NTD) КОФ300, irradiated at room temperature by different fluences ($F = 1 \cdot 10^{14} - 3 \cdot 10^{16}$ cm^{-2}) of electrons with the energy of E_e = 4 MeV are presented. It is established that at low electron fluences ($F = 1 \cdot 10^{14}$ cm^{-2}), the annealing of minority charge currier lifetime τ_P in the *n*-base of *p-n*-structures occurs in two stages: the first – 320–400 ºC and the second – 550–650 ºC. At higher electron fluences ($F = 5 \cdot 10^{15}$–$2 \cdot 10^{16}$ cm^{-2}), three annealing stages occur: the first – 400–450 ºC, the second – 520–650 ºC and the third – 710–770 ºC. At this, the structure barrier capacitance *C* dependences on T_{ann} at high electron fluences show the geometry capacitance up to the annealing temperatures T_{ann} = 400 ºC. In the annealing temperature range of T_{ann} = 420–570 ºC, the increase in *C* with maximum is seen at T_{ann} = 480 ºC and a subsequent decrease in the geometry capacitance is seen in the annealing temperature range of T_{ann} = 600–670 ºC, and then again the increase in *C* occurs in the annealing temperature range of T_{ann} = 720–770 ºC reaching the *C* values corresponding to those of the non-irradiated samples in the annealing temperature range of T_{ann} = 770–800 ºC. The analysis of the DLTS-spectra of the investigated structures has allowed establishing the formation in the annealing process of the deep acceptor level E_C – 0.68 eV at T_{ann} > 400 ºC, the deep donor level E_C – 0.32 eV in the annealing temperature range of T_{ann} = 420–570 ºC and the deep acceptor level E_C – 0.53 eV at T_{ann} > 700 ºC, which satisfactorily explains the dependences of τ_P and *C* on T_{ann} obtained in this paper.

Keywords: radiation defect, radiation-thermal defect, electron irradiation, high-temperature annealing, minority charge carrier lifetime, DLTS-spectroscopy

For citation. Korshunov F. P., Zhdanovich N. E., Zhdanovich D. N. Influence of high-temperature annealing on the characteristics of fast electron-irradiated *p-n*-structures based on neutron doped silicon. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2019, vol. 55, no. 4, pp. 489–497 (in Russian). https://doi.org/10.29235/1561-2430-2019-55-4-489-497

Введение. Практика показывает, что качество и надежность полупроводниковых приборов во многом определяется тем, насколько равномерно пролегирован нужными примесями (донорами, акцепторами) исходный полупроводниковый кристалл, по крайней мере, это относится к кремнию. Оказывается, недостаточно получить совершенный кремниевый кристалл, что само по себе является сложной задачей. Для того чтобы найти ему практическое применение в приборостроении, следует ввести в кристалл необходимые примеси в строго дозированных концентрациях, так как именно они определяют его характеристики. Легирование кристалла проводится во время его выращивания и осуществляется при очень высоких температурах (выше 1270 К), при этом неравномерность распределения примесей иногда достигает 15–20 %, а нужно, чтобы она была значительно ниже. В результате параметры многих кремниевых приборов, изготовленных на таких кристаллах, особенно силовых, сильно различаются. По этой причине в производстве силовых приборов используется кремний, полученный с применением другого способа легирования, обеспечивающего более равномерное распределение легирующей примеси, а именно: с использованием нейтронного облучения [1].

Выращенный кристалл кремния состоит в основном из трех его изотопов: Si^{28} (92,21 %), Si^{29} (4,7 %) и Si^{30} (3,09 %). Изотоп кремния Si^{30} захватывает нейтрон и переходит в изотоп Si^{31}, который нестабилен и через 2,6 ч переходит в изотоп фосфора с испусканием отрицательно заряженной β-частицы. Изотопы кремния Si^{28} и Si^{29} при ядерных реакциях не дают других стабильных химических элементов, поэтому не изменяют свойства кремния. При этом ядерная реакция особенно эффективно протекает на медленных нейтронах, т. е. изотоп кремния с высокой вероятностью захватывает именно тепловые медленные нейтроны [2].

В настоящей работе приведены результаты изучения влияния высокоэнергичных электронов на p^+-n-структуры на ядерно-легированном кремнии (ЯЛК) с проведением отжига в широком диапазоне температур и флюенсов облучения. Следует отметить, что ранее такие исследования не проводились, хотя полученные результаты будут представлять интерес как в плане использования в полях излучений приборов на ЯЛК, так и при их производстве для регулирования их параметров, таких как быстродействие, посредством ввода в них вместо золота или платины термостабильных радиационных дефектов с применением облучения ионизирующими излучениями, например быстрыми электронами.

Методика эксперимента. Исследования проводились на специально изготовленных тестовых образцах на пластинах ядерно-легированного кремния КОФ300. Толщина пластин составляла 1000 мкм. *p-n*-Переход формировался диффузией бора. Глубина залегания перехода составляла 200 мкм, а площадь структур – 10 мм^2. Омические контакты создавались путем подлегирования структур со стороны *n*- и *p*-областей фосфором и бором соответственно до концентраций порядка 10^{19} см$^{-3}$. Структуры облучались электронами с энергией 4 МэВ при плотности потока, равной $1 \cdot 10^{12}$ см$^{-2}$·с$^{-1}$. Облучение структур проводилось при комнатной температуре в диапазоне флюенсов $1 \cdot 10^{14}$–$3 \cdot 10^{16}$ см$^{-2}$.

Для исследования влияния отжига на время жизни неосновных носителей заряда были использованы 4 набора образцов, облученных разными флюенсами электронов: 1) $\Phi = 1 \cdot 10^{14}$ см$^{-2}$; 2) $\Phi = 6 \cdot 10^{15}$ см$^{-2}$; 3) $\Phi = 9 \cdot 10^{15}$ см$^{-2}$; 4) $\Phi = 2 \cdot 10^{16}$ см$^{-2}$. Увеличение флюенса электронного облучения позволяло увеличивать концентрацию радиационных дефектов, вводимых в базовую область структур. Облученные и контрольные необлученные образцы отжигались изохронно в интервале температур 300–800 °С в течение 15 мин на воздухе. Время жизни неосновных носителей заряда (ННЗ) измерялось при высоком уровне инжекции по методу Лэкса [3]. До облучения время жизни неосновных носителей заряда составляло ~20–30 мкс. В процессе отжига также контролировалась барьерная емкость *p-n*-структур при двух значениях обратного смещения:

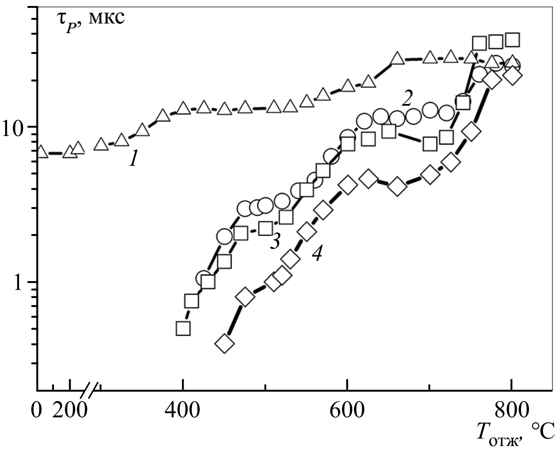


Рис. 1. Зависимости времени жизни неосновных носителей заряда в высокоомной базе диодов τ_P от температуры изохронного отжига (время отжига – 15 мин): *1* – $\Phi = 1 \cdot 10^{14}$ см$^{-2}$; *2* – $\Phi = 6 \cdot 10^{15}$ см$^{-2}$; *3* – $\Phi = 9 \cdot 10^{15}$ см$^{-2}$; *4* – $\Phi = 2 \cdot 10^{16}$ см$^{-2}$

Fig. 1. Dependences of the lifetime of minority charge carriers in the high-resistivity base of diodes τ_P on the isochronous annealing temperature (annealing time – 15 min): *1* – $F = 1 \cdot 10^{14}$ cm^{-2}; *2* – $F = 6 \cdot 10^{15}$ cm^{-2}; *3* – $F = 9 \cdot 10^{15}$ cm^{-2}; *4* – $F = 2 \cdot 10^{16}$ cm^{-2}

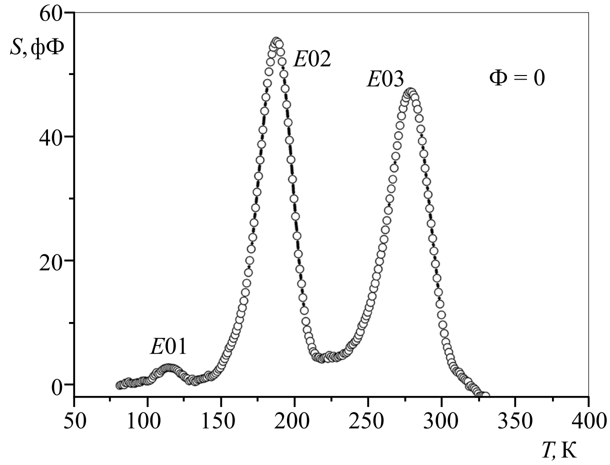
0 и –40 В, что давало возможность отслеживать концентрацию образующихся дефектов как вблизи *p*-*n*-перехода, так и в глубине высокоомной базы. Емкость измерялась с помощью измерителя LCR E7-12 в параллельной схеме замещения на частоте 1 МГц.

Результаты и их обсуждение. Зависимости времени жизни неосновных носителей заряда τ_P в высокоомной *n*-базе диодов от температуры изохронного отжига приведены на рис. 1. Из полученных данных видно, что наблюдается различный характер изменений времени жизни ННЗ в результате термического отжига образцов, облученных разными флюенсами электронов. Для малых флюенсов электронов (кривая *1*, $\Phi = 1 \cdot 10^{14}$ см$^{-2}$ – низкая концентрация радиационных дефектов (РД)) отжиг проходит в две стадии: первая – 320–400 °С, вторая – 550–650 °С. Для высоких флюенсов облучения (кривые *2*–*4*, $\Phi = 5 \cdot 10^{15}$–$2 \cdot 10^{16}$ см$^{-2}$ – высокая концентрация РД) на полученных зависимостях времени жизни неосновных носителей заряда от температуры отжига наблюдается три стадии отжига: первая – 400–450 °С, вторая – 520–650 °С и третья – 710–770 °С.

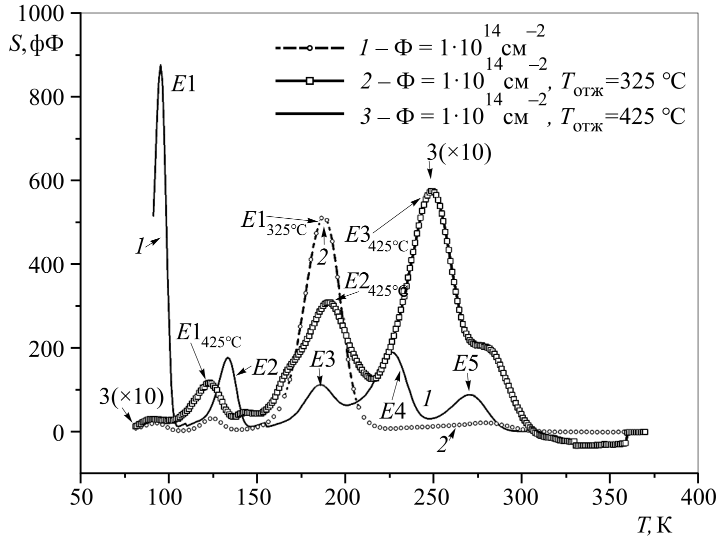
Значения τ_P образцов, облученных флюенсами электронов $5 \cdot 10^{15}$–$2 \cdot 10^{16}$ см$^{-2}$, в интервале температур 20–400 ºС были меньше нижнего предела измерения установки (<100 нс). Вследствие этого зависимости *2*–*4* на рис. 1 начинаются со значений $T_{\text{отж}} \geq 400$ °С и, безусловно, у сильно облученных образцов также присутствует низкотемпературная стадия отжига при $T_{\text{отж}} < 400$ °С.

На рис. 2, *a* показаны *DLTS*-спектры исследуемых структур до электронного облучения. Спектры измерялись в режиме перезарядки глубоких уровней основными носителями заряда. Из полученных результатов видно, что три пика *E*01–*E*03 соответствуют эмиссии в зону проводимости электронов с глубоких уровней дефектов технологического типа, содержащихся в базовой *n*-области исходных *p*-*n*-структур. Пику *E*01 с положением на спектре максимума при 117 К соответствует глубокий уровень акцепторного типа E_C – 0,22 эВ и сечением захвата электронов $\sigma_n = 2{,}44 \cdot 10^{-14}$ см^2, пику *E*02 при 180 К – E_C – 0,27 эВ и $\sigma_n = 2{,}11 \cdot 10^{-17}$ см^2 и пику *E*03 при 280 К – E_C – 0,53 эВ и $\sigma_n = 1{,}55 \cdot 10^{-15}$ см$^{-2}$. Концентрация ловушек в исходном базовом *n*-Si составляет менее одного процента от концентрации легирующей примеси (10^{13} см$^{-3}$) и, вероятнее всего, в их роли выступают неконтролируемые примеси. Данных, полученных методом *DLTS,* недостаточно для однозначной идентификации центров *E*02–*E*03.

На рис. 2, *b* приведены *DLTS*-спектры *p*-*n*-структуры после облучения флюенсом электронов $\Phi = 1 \cdot 10^{14}$ см$^{-2}$ и последующего изохронного отжига при температурах 325 и 425 °С. При такой концентрации введенных РД степень компенсации высокоомной базы структур не превышает 20 %, что позволяет проследить отжиг дефектов методом *DLTS*-спектроскопии. Сразу после элек-



a



b

Рис. 2. *DLTS*-спектры *p*-*n*-структуры до и после облучения флюенсом электронов $\Phi = 1 \cdot 10^{14}$ см$^{-2}$ и последующего изохронного отжига при температурах 325 и 425 °C; *a*: Ф = 0; *b*: *1* – $\Phi = 1 \cdot 10^{14}$ см$^{-2}$; *2* – $\Phi = 1 \cdot 10^{14}$ см$^{-2}$, $T_{\text{отж}}$ =325 °C, $t_{\text{отж}}$ = 15 мин; *3* – $\Phi = 1 \cdot 10^{14}$ см$^{-2}$, $T_{\text{отж}}$ = 425 °C, $t_{\text{отж}}$ = 15 мин. *S* – значение нестационарной емкости, в фемтофарадах; окно скоростей e_m = 191,4 с$^{-1}$

Fig. 2. *DLTS* spectra of the *p*-*n*-structure before and after of irradiation with the electron fluence $F = 1 \cdot 10^{14}$ cm^{-2} and after subsequent annealing at temperatures 325 and 425 °C; *a*: *F* = 0; *b*: *1* – $F = 1 \cdot 10^{14}$ cm^{-2}; *2* – $F = 1 \cdot 10^{14}$ cm^{-2}, T_{ann} = 325 °C, t_{ann} = 15 min; *3* – $F = 1 \cdot 10^{14}$ cm^{-2}, T_{ann} = 425 °C, t_{ann} = 15 min. *S* – value of the non-stationary capacity, in femtofarads; emission rate window e_m = 191.4 s^{-1}

тронного облучения (кривая *1*) на спектре наблюдается пять пиков, связанных с эмиссией электронов с введенных РД глубоких уровней. Значения параметров этих глубоких уровней следующие:

– $E1$ (А-центр): $E = E_C - 0{,}18$ эВ и сечением захвата электронов $\sigma_n = 1{,}86 \cdot 10^{-14}$ см^2;

– $E2$ (мелкий уровень дивакансии): $E = E_C - 0{,}25$ эВ, $\sigma_n = 4{,}73 \cdot 10^{-15}$ см^2;

– $E3$ (природа дефекта не установлена): $E = E_C - 0{,}27$ эВ, $\sigma_n = 2{,}11 \cdot 10^{-17}$ см^2;

– $E4$ (глубокий уровень дивакансии): $E = E_C - 0{,}41$ эВ, $\sigma_n = 8{,}2 \cdot 10^{-16}$ см^2;

– $E5$ (природа дефекта не установлена): $E = E_C - 0{,}53$ эВ, $\sigma_n = 1{,}55 \cdot 10^{-15}$ см^2.

Вид спектра остается неизменным до $T_{\text{отж}}$ = 275 °C. При $T_{\text{отж}} \geq 275$ °C спектр облученной структуры начинает изменяться с одновременным увеличением времени жизни неосновных носителей заряда. При $T_{\text{отж}}$ = 325 °C (кривая *2*) исчезают пики $E1$ (А-центр), а также $E2$ и $E4$ (дивакансия). Отжиг вакансионных комплексов ведет к образованию термически более стабильной ловушки с глубоким уровнем $E_C - 0{,}33$ эВ и сечением захвата основных носителей $\sigma_n = 8{,}87 \times 10^{-16}$ см^2 (пик $E1_{325\,°C}$). С увеличением температуры изохронного отжига до 425 °C вид *DLTS*-спектра продолжает изменяться (кривая *3* на рис. 2, *b*). На этом этапе происходит отжиг ловушки с уровнем E_C – 0,33 эВ и образование новых дефектов с глубокими уровнями E_C – 0,20 эВ и сечением захвата $\sigma_n = 1{,}21 \cdot 10^{-15}$ см^2 (пик $E1_{425\,°C}$), а также E_C – 0,44 эВ и $\sigma_n = 5{,}26 \cdot 10^{-16}$ см^2 (пик $E3_{425\,°C}$). Концентрация образовавшихся ловушек незначительна и соответствующие им амплитуды пиков на рис. 2, *b* для наглядности увеличены в 10 раз. Пик $E2_{425\,°C}$ является суммой двух близко расположенных пиков, которые не удалось разделить и определить параметры соответствующих им глубоких уровней.

На рис. 3 показаны *DLTS*-спектры того же образца, что и на рис. 2, *b*. В этом случае измерения проводились в режиме перезарядки ловушек неосновными носителями заряда, т. е. определялись глубокие уровни радиационных дефектов в нижней половине запрещенной зоны Si.

После облучения образца электронами возникает доминирующий пик *H*1 в результате эмиссии дырок с глубокого уровня E_v + 0,36 эВ и сечением захвата $\sigma_p = 1{,}02 \cdot 10^{-16}$ см^2 комплекса междоузельный углерод – междоузельный кислород C_iO_i (кривая *2*). Данный радиационный дефект обладает высокой термической стабильностью [4] и начинает отжигаться только при $T_{отж}$ = = 450–475 °С (кривые *3* и *4*). При $T_{отж}$ = 450 °С регистрируется *DLTS*-сигнал эмиссии дырок с двух уровней ловушек C_iO_i и вновь образовавшейся C_i–O_{2i} (междоузельный углерод – два междоузельных атома кислорода) с E_v + 0,39 эВ и $\sigma_p = 2{,}32 \cdot 10^{-16}$ см^2. После отжига при 475 °С регистрируется сигнал только от ловушки C_i–O_{2i} (пик *H*2). Комплекс междоузельного типа C_i–O_{2i} отжигается при повышении температуры отжига до 600 °С. При более высоких температурах отжига в нижней половине запрещенной зоны глубоких уровней дефектов для всех флюенсов облучения ($\Phi = 1 \cdot 10^{14}$–$3 \cdot 10^{16}$ см$^{-2}$) не наблюдается.

На рис. 4 приведены зависимости значения барьерной емкости *C* от температуры отжига образцов, облученных флюенсами электронов $\Phi = 1 \cdot 10^{14}$ и $2 \cdot 10^{16}$ см$^{-2}$. Барьерная емкость измерялась при комнатной температуре при двух значениях обратного напряжения $U_{обр}$ = 0 и –40 В.

Каких-либо значимых изменений емкости для флюенса облучения $\Phi = 1 \cdot 10^{14}$ см$^{-2}$ не наблюдается ввиду низкой концентрации радиационных дефектов в исследуемом образце. Это

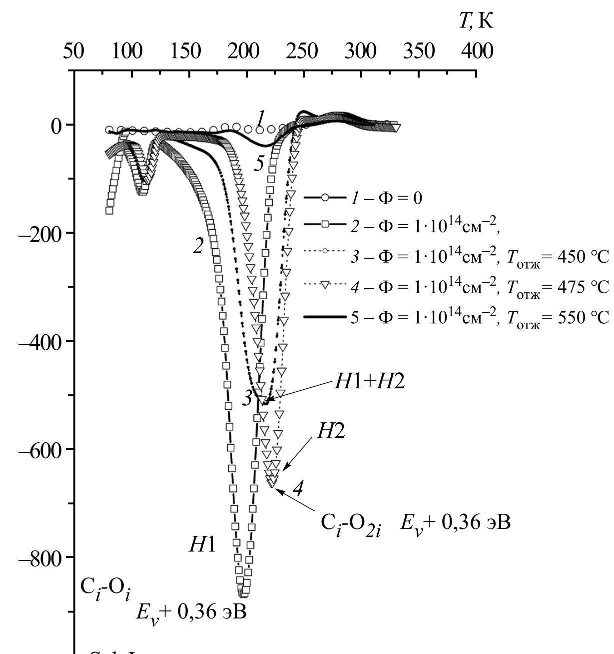


Рис. 3. *DLTS*-спектры *p*-*n*-структуры до и после облучения флюенсом электронов $\Phi = 1 \cdot 10^{14}$ см$^{-2}$ и последующего изохронного отжига при температурах 450–550 °С: *1* – Φ = 0; *2* – $\Phi = 1{\cdot}10^{14}$ см$^{-2}$; *3* – $\Phi = 1 \cdot 10^{14}$ см$^{-2}$, $T_{отж}$ = 450 °С, $t_{отж}$ = 15 мин; *4* – $\Phi = 1 \cdot 10^{14}$ см$^{-2}$, $T_{отж}$ = 475 °С, $t_{отж}$ = 15 мин; *5* – $\Phi = 1 \cdot 10^{14}$ см$^{-2}$, $T_{отж}$ = 550 °С, $t_{отж}$ = 15 мин. *S* – значение нестационарной емкости (фФ); окно скоростей e_m = 191,4 с$^{-1}$

Fig. 3. *DLTS* spectra of the *p*-*n*-structure before and after of irradiation with the electron fluence $F = 1 \cdot 10^{14}$ cm^{-2} and after subsequent annealing at temperatures 450–550 °C: *1* – *F* = 0; *2* – $F = 1 \cdot 10^{14}$ cm^{-2}; *3* – $F = 1 \cdot 10^{14}$ cm^{-2}, T_{ann} = 450 °C, t_{ann} = 15 min; *4* – $F = 1 \cdot 10^{14}$ cm^{-2}, T_{ann} = 475 °C, t_{ann} = 15 min; *5* – $F = 1 \cdot 10^{14}$ cm^{-2}, T_{ann} = 550 °C, t_{ann} = 15 min. *S* – value of the non-stationary capacity (fF); emission rate window e_m = 191.4 s^{-1}

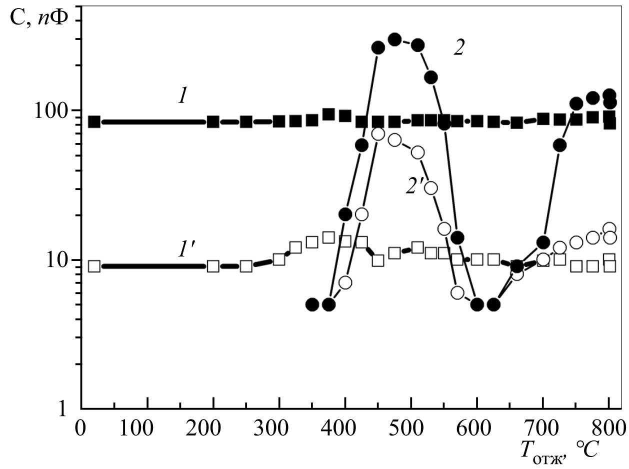


Рис. 4. Зависимости барьерной емкости C образцов с высокой и низкой концентрацией радиационных дефектов от температуры отжига: *1* – $C_{(Ф = 1 \cdot 10^{14}\ см^{-2})}$ – 0 В; *1′* – $C_{(Ф = 1 \cdot 10^{14}\ см^{-2})}$ – 40 В; *2* – $C_{(Ф = 2 \cdot 10^{16}\ см^{-2})}$ – 0 В; *2′* – $C_{(Ф = 2 \cdot 10^{16}\ см^{-2})}$ – 40 В

Fig. 4. Dependences of the barrier capacitance C of samples with high and low RD concentrations on the annealing temperature: $C_{(F = 1 \cdot 10^{14}\ cm^{-2})}$ – 0 V; *1′* – $C_{(F = 1 \cdot 10^{14}\ cm^{-2})}$ – 40 V; *2* – $C_{(F = 2 \cdot 10^{16}\ cm^{-2})}$ – 0 V; *2′* – $C_{(F = 2 \cdot 10^{16}\ cm^{-2})}$ – 40 V

обусловлено тем, что при флюенсе облучения Ф = 1 · 10^{14} см$^{-2}$ суммарная концентрация РД не превышает 20 % от концентрации легирующей примеси и при этом доминирующим является дефект с мелким акцепторным уровнем (E_C – 0,17 эВ – A-центр), который не заполнен носителями при комнатной температуре.

В случае с высокой концентрацией радиационных дефектов (кривые *2, 2′*; Ф = 2 · 10^{16} см$^{-2}$) наблюдается немонотонная зависимость барьерной емкости от температуры отжига. До температур отжига 400 °С измеряется геометрическая емкость структур ввиду полной компенсации высокоомной базы радиационными дефектами [5]. В диапазоне температур отжига 420–570 °С наблюдается рост барьерной емкости с максимумом при $T_{отж}$ = 480 °С и последующий спад до значений геометрической емкости при $T_{отж}$ = 600–670 °С. В области температур 720–800 °С наблюдается восстановление значения барьерной емкости до значений, соответствующих необлученному образцу. Значения барьерной емкости на зависимостях *2, 2′* при $T_{отж}$ = 480 °С многократно превышают аналогичные значения на зависимостях *1* и *1′*. Этот факт позволяет предположить, что в результате отжига радиационных дефектов при $T_{отж}$ = 420–570 °С образуются дефекты с донорным уровнем, концентрация которых больше концентрации легирующей примеси.

Известно, что при длительных термообработках кристаллов кремния в указанной области температур образуются кислородные термодоноры [6]. Их отжиг происходит при $T_{отж}$ = 600–670 °С. В нашем случае время изохронного отжига составляет 15 мин. Это значение меньше, чем указанное в работе [6]. Более того, нами был проведен длительный отжиг необлученных образцов в течение 10 ч при температурах 400 и 480 °С. В результате такой термообработки барьерная емкость структур увеличивалась всего на 10–15 %, что позволяет говорить о маловероятной возможности ускоренной генерации термодоноров и в присутствии радиационных дефектов. Допустимо предположить, что немонотонная зависимость C ($T_{отж}$) при $T_{отж}$ = 420–570 °С связана с образованием более термостабильных дефектов донорного типа при отжиге РД.

Дополнительные данные, которые позволяют проинтерпретировать немонотонный ход зависимости барьерной емкости образцов с высокой концентрацией РД (рис. 4, кривые *2, 2′*) были получены из измерений *DLTS*-спектров этих структур.

На рис. 5 представлены *DLTS*-спектры структуры, облученной флюенсом электронов Ф = = 6 · 10^{15} см$^{-2}$ и отожженной при $T_{отж}$ = 500 °С (спектр *1*), и структуры, облученной флюенсом электронов Ф = 1,2 · 10^{16} см$^{-2}$ и отожженной при температуре 750 °С (спектр *2*). На спектре *1* доминируют 2 пика глубоких уровней: $E1_{500\ °C}$ с энергией активации E_C – 0,32 эВ и сечением захвата основных носителей σ_n = 1,02 · 10^{-16} см^2 и $E2_{500\ °C}$ с E_C – 0,68 эВ и σ_n = 2,15 · 10^{-15} см^2. Анализ температурной зависимости емкости (полученной в ходе *DLTS*-измерений) в интервале температур

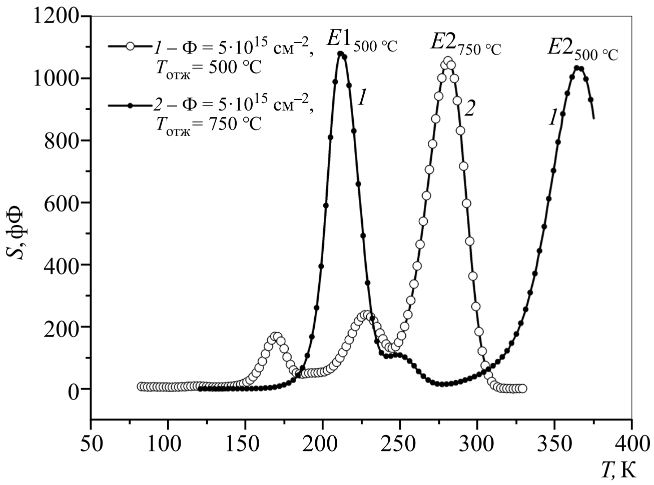


Рис. 5. *DLTS*-спектры исследуемых структур с высокой концентрацией радиационных дефектов после облучения и последующего отжига при различных температурах: *1* – $\Phi = 6 \cdot 10^{15}$ см$^{-2}$, $T_{отж} = 500$ °С, $t_{отж} = 15$ мин; *2* – $\Phi = 1{,}2 \cdot 10^{16}$ см$^{-2}$, $T_{отж} = 750$ °С, $t_{отж} = 15$ мин. Окно скоростей $e_m = 191{,}4$ с$^{-1}$

Fig. 5. *DLTS* spectra of the studied structures with a high concentration of RD after irradiation and subsequent annealing at different temperatures: *1* – $F = 6 \cdot 10^{15}$ cm^{-2}, $T_{ann} = 500$ °C, $t_{ann} = 15$ min; *2* – $F = 1.2 \cdot 10^{16}$ cm^{-2}, $T_{ann} = 750$ °C, $t_{ann} = 15$ min. Emission rate window $e_m = 191.4$ s^{-1}

77–400 K показал, что ее значение уже при температуре измерения, когда на спектре *1* полностью регистрируется пик $E1_{500\,°C}$ ($T > 240$ K), превышают значения емкости на зависимостях, полученных для образца, облученного малым флюенсом электронов (см. рис. 4, кривые *1, 1'*), что позволяет сделать предположение о донорной природе дефекта с уровнем E_C – 0,32 эВ. Также можно предположить, что с определенных флюенсов облучения легирующий вклад этого дефекта в интервале температур отжига 600–670 °С начинает превышать компенсирующий вклад другого регистрируемого на спектре *1* глубокого уровня E_C – 0,68 эВ. Следует отметить, что величина энергии активации уровня $E2_{500\,°C}$ превышает полуширину запрещенной зоны кремния и интерпретация полученного результата требует дополнительных исследований. Возможно, при расчетах энергии активации в данном случае необходимо учитывать влияние последовательного сопротивления базы [7], а также влияние обмена носителями заряда между уровнями и обеими разрешенными зонами [8].

При отжиге в области более высоких температур материал высокоомной базы облученных высоким флюенсом электронов образцов снова становится компенсированным, что связано с отжигом дефекта с глубоким донорным уровнем E_C – 0,32 эВ, и до температур отжига 650 °С запись спектров при флюенсах выше $\Phi = 5 \cdot 10^{14}$ см$^{-2}$ невозможна. Кривая *2* на рис. 5 соответствует образцу, отожженному при температуре 750 °С. В спектре этого образца наблюдается 3 пика, и доминирующим по амплитуде является $E2_{750\,°C}$ с энергией активации глубокого уровня E_C – 0,53 эВ и сечением захвата $\sigma_n = 2{,}15 \cdot 10^{-15}$ см^2. Значение барьерной емкости образца близко к исходному (до облучения). Следовательно, в спектре отсутствуют донорные уровни и нет глубокого компенсирующего уровня (E_C – 0,68 эВ), определявшего барьерную емкость при комнатной температуре при более низких температурах отжига. Также следует обратить внимание на тот факт, что картина отжига и трансформации дефектов в структурах на высокоомном ядерно-легированном кремнии существенно отличается от картины отжига дефектов в структурах на базе кремния, легированного другими способами. Так, согласно [9, 10], уже при температуре отжига 600 °С в структурах на тянутом кремнии наблюдается отжиг практически всех РД при флюенсах облучения вплоть до $1 \cdot 10^{17}$ см$^{-2}$ и отсутствуют дефекты с глубокими уровнями как донорного, так и акцепторного типов.

Заключение. Таким образом, в данной работе приведены результаты исследования по определению электрических характеристик радиационно-индуцированных дефектных комплексов в структурах высоковольтных диодов на ядерно-легированном кремнии. Показано, что в диапазоне температур 400–800 °C методом *DLTS*-спектроскопии можно наблюдать процессы перестройки радиационных дефектов, предварительно введенных электронным облучением в базовую область *p*-*n*-структур при комнатной температуре. Определены основные закономерности формирования термостабильных центров в облученных быстрыми электронами структурах на ядерно-легированном кремнии в широком диапазоне температур отжига.

Список использованных источников

1. Смирнов, Л. С. Легирование полупроводников методом ядерных реакций / Л. С. Смирнов. – Новосибирск: Наука, 1981. – 184 с.
2. Меднис, И. В. Справочные таблицы для нейтронно-активационного анализа / И. В. Меднис. – Рига: Знание, 1974. – 410 с.
3. Lax, B. Transient response of a *p-n* junction / B. Lax, S. F. Neustadter / J. Appl. Phys. – 1984. – Vol. 25, № 9. – P. 1148–1154. https://doi.org/10.1063/1.1721830
4. Defect-impurity complexes with high thermal stability in epi-Si n^+-p diodes irradiated with MeV electrons / F. P. Korshunov [et al.] // Vacuum. – 2009. – Vol. 83. – P. S131–S133. https://doi.org/10.1016/j.vacuum.2009.01.044
5. Берман, Л. С. Емкостная спектроскопия глубоких центров в полупроводниках / Л. С. Берман, А. А. Лебедев. – Л.: Наука, 1981. – 176 с.
6. Claybourn, M. Thermal donor formation and the loss of oxygen from solution in silicon heated at 450 °C / M. Claybourn, R. C. Newman // Appl. Phys. Lett. – 1988. – Vol. 52, № 25. – P. 2139–2141. https://doi.org/10.1063/1.99557
7. Астрова, Е. В. Влияние последовательного сопротивления диода на нестационарное емкостное измерение параметров глубоких уровней / Е. В. Астрова, А. А. Лебедев // Физика и техника полупроводников. – 1985. – Т. 19, № 8. – С. 1382–1385.
8. Емкостная спектроскопия глубоких уровней при обмене носителями заряда между уровнями и обеими разрешенными зонами / Е. А. Татохин [и др.] // Вестн. Воронеж. гос. ун-та. Сер. Физика. Математика. – 2008. – № 2. – С. 60–70.
9. Влияние отжига на перестройку центров рекомбинации в облученных кремниевых структурах / Ф. П. Коршунов [и др.] // Докл. АН БССР. – 1988. – Т. 32, № 9. – С. 781–783.
10. Исследование радиационно-термических дефектов и их влияние на параметры кремниевых диффузионных *p-n*-структур // Ф. П. Коршунов [и др.] // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 1998. – № 3. – С. 64–68.

References

1. Smirnov L. S. *Doping of Semiconductors Using the Method of Nuclear Reactions*. Novosibirsk, Nauka Publ., 1981. 184 p. (in Russian)
2. Mednis I. V. *Reference Tables for Neutron-Activation Analyses*. Riga, Znanie Publ., 1974. 410 p. (in Russian)
3. Lax B., Neustadter S. F. Transient response of a *p-n* junction. *Journal of Applied Physics*, 1984, vol. 25, no. 9, pp. 1148–1154. https://doi.org/10.1063/1.1721830
4. Korshunov F. P., Lastovskii S. B., Markevich V. P., Murin L. I., Bogatyrev Yu. V., Peaker A. R. Defect-impurity complexes with high thermal stability in epi-Si n^+-p diodes irradiated with MeV electrons. *Vacuum*, 2009, vol. 83, pp. S131–S133. https://doi.org/10.1016/j.vacuum.2009.01.044
5. Berman L. S., Lebedev A. A. *Capacity Spectroscopy of Deep Centers in Semiconductors*. Leningrad, Nauka Publ., 1981. 176 p. (in Russian).
6. Claybourn M., Newman R. C. Thermal donor formation and the loss of oxygen from solution in silicon heated at 450 °C. *Applied Physics Letters,* 1988, vol. 52, no. 25, pp. 2139–2141. https://doi.org/10.1063/1.99557
7. Astrova E. V., Lebedev A. A. The influence of diode series resistance on deep levels parameters measuring by DLTS-spectroscopy method. *Fizika i tekhnika poluprovodnikov = Semiconductors. Physics of the Solid State*, 1985, vol. 19, no. 8, pp. 1382–1385 (in Russian).
8. Tatohin E. A., Budanov A. V., Butusov I. Ju., Vasil'eva L. V., Tutov E. A. Capacity spectroscopy of deep levels in case of charge carriers interchange between levels an both allowed energy bands. *Vestnik Voronezhskogo gosudarstvennogo universiteta. Seriya Fizika. Matematika = Proceeding of Voronezh State University. Series Physics. Mathematics*, 2008, no. 2, pp. 60–70 (in Russian).
9. Korshunov F. P., Marchenko I. G., Zdanovich N. E., Troschinskii V. T. Impact of annealing on recombination centers transformation in irradiated silicon structures. *Doklady AN BSSR = Proceedings of the Academy of Sciences of BSSR*, 1988, vol. 32, no. 9, pp. 781–783 (in Russian).
10. Korshunov F. P., Zdanovich N. E., Marchenko I. G. Investigation of radiation-thermal defects and their influence on the characteristics of silicon diffused *p-n*-structures. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 1998, no. 3, pp. 64–68 (in Russian).

Информация об авторах

Коршунов Федор Павлович – член-корреспондент, доктор технических наук, профессор, главный научный сотрудник, лаборатория «Радиационные воздействия», Научно-практический центр Национальной академии наук Беларуси по материаловедению (ул. П. Бровки, 19, 220072, г. Минск, Республика Беларусь). E-mail: korshun@ifttp.bas-net.by

Жданович Николай Евгеньевич – научный сотрудник лаборатории «Радиационные воздействия», Научно-практический центр Национальной академии наук Беларуси по материаловедению (ул. П. Бровки, 19, 220072, г. Минск, Республика Беларусь). E-mail: jdan@ifttp.bas-net.by

Жданович Дмитрий Николаевич – младший научный сотрудник лаборатории «Радиационные воздействия», Научно-практический центр Национальной академии наук Беларуси по материаловедению (ул. П. Бровки, 19, 220072, г. Минск, Республика Беларусь). E-mail: zhdanovich_d@ifttp.bas-net.by

Information about the authors

Fedor P. Korshunov – Corresponding Member, Dr. Sc. (Engineering), Professor, Chief Researcher, Laboratory of the Radiation Effects, Scientific-Practical Materials Research Centre of the National Academy of Sciences of Belarus (19, P. Brovka Str., 220072, Minsk, Republic of Belarus). E-mail: korshun@ifttp.bas-net.by

Nikolai E. Zhdanovich – Researcher of the Laboratory of the Radiation Effects, Scientific-Practical Materials Research Centre of the National Academy of Sciences of Belarus (19, P. Brovka Str., 220072, Minsk, Republic of Belarus). E-mail: jdan@ifttp.bas-net.by

Dmitrij N. Zhdanovich – Junior Researcher of the Laboratory of the Radiation Effects, Scientific-Practical Materials Research Centre of the National Academy of Sciences of Belarus (19, P. Brovka Str., 220072, Minsk, Republic of Belarus). E-mail: zhdanovich_d@ifttp.bas-net.by

ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)
УДК 621.382.3
https://doi.org/10.29235/1561-2430-2019-55-4-498-504

Поступила в редакцию 29.08.2019
Received 29.08.2019

Д. А. Огородников, С. Б. Ластовский, Ю. В. Богатырев

Научно-практический центр Национальной академии наук Беларуси по материаловедению, Минск, Беларусь

МОДЕЛИРОВАНИЕ НАКОПЛЕНИЯ ЗАРЯДА В ОБЛУЧЕННЫХ МОП/КНИ-ТРАНЗИСТОРАХ

Аннотация. С помощью программного комплекса Silvaco проведен расчет накопления встроенного в окисел заряда у границы раздела кремний – скрытый окисел в *n*-канальных МОП/КНИ-транзисторах в зависимости от их геометрических параметров и электрических режимов при воздействии ионизирующего излучения. Показано, что наиболее «жестким» является режим, при котором во время облучения на сток и исток подается напряжение +5 В, а на подложку, затвор и запитку канала – 0 В. При этом величину накопленного заряда удается существенно снизить, прикладывая к подложке отрицательное смещение и уменьшая толщину слоя захороненного окисла. Полученные результаты компьютерного моделирования могут быть использованы испытателями электронных компонентов для предварительной оценки стойкости МОП/КНИ-приборов к накопленной дозе ионизирующего излучения.

Ключевые слова: ионизирующее излучение, МОП/КНИ-транзистор, радиационная стойкость, моделирование, встроенный заряд

Для цитирования. Огородников, Д. А. Моделирование накопления заряда в облученных МОП/КНИ-транзисторах / Д. А. Огородников, С. Б. Ластовский, Ю. В. Богатырев // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2019. – Т. 55, № 4. – С. 498–504. https://doi.org/10.29235/1561-2430-2019-55-4-498-504

D. A. Ogorodnikov, S. B. Lastovskii, Yu. V. Bogatyrev

Scientific-Practical Materials Research Centre of the National Academy of Sciences of Belarus, Minsk, Belarus

SIMULATING OF CHARGE BUILD-UP IN IRRADIATED MOS/SOI TRANSISTORS

Abstract. The charge build-up in the interface of silicon / buried oxide in *n*-channel MOS/SOI transistors depending on their geometric parameters and electrical modes during ionizing irradiation is calculated with the use of the software Silvaco. It is shown that the electrical mode is most “harsh”, when during irradiation the voltage of +5 V is applied to drain and source electrodes and 0 V is applied to substrate, gate and channel feeding. The amount of the built-up charge can be substantially reduced by applying a negative bias to the substrate and by decreasing the thickness of the buried oxide layer.

Keywords: ionizing radiation, MOS/SOI transistor, radiation hardness, simulating, interface charge

For citation. Ogorodnikov D. A., Lastovskii S. B., Bogatyrev Yu. V. Simulating of charge build-up in irradiated MOS/SOI transistors. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2019, vol. 55, no. 4, pp. 498–504 (in Russian). https://doi.org/10.29235/1561-2430-2019-55-4-498-504

Введение. Технология «кремний на изоляторе» (КНИ) перспективна для производства МОП (металл-окисел-полупроводник) интегральных схем с повышенной радиационной стойкостью к воздействию импульсной радиации [1, 2]. Однако в случае стационарного облучения значительную проблему создает наличие у КНИ-структур границы раздела полупроводник – скрытый окисел. При этом образуется паразитный транзистор, у которого подзатворным диэлектриком служит скрытый окисел, а затвором – изолированная подложка. Накопление радиационно-индуцированного заряда в изолирующем окисле ведет к образованию проводящего канала паразитного транзистора на обратной стороне кремниевой пленки. Эффективность накопления заряда в окисле существенно зависит от электрического режима при облучении и параметров транзисторных структур. По величине этого заряда можно судить о стойкости *n*-канальных МОП/КНИ-транзисторов к эффектам накопленной дозы. Ранее в [3–5] экспериментально исследовалось влияние различных электрических режимов на радиационное изменение параметров

МОП/КНИ-транзисторов. Показано, что наихудший электрический режим может по-разному проявляться при облучении приборов с различными конструктивными характеристиками. При этом следует отметить, что в экспериментах не учитывались эффекты отжига радиационно-индуцированного заряда в окисле. Например, в работе [5] измерения параметров МОП/КНИ-транзисторов проводились через несколько часов после облучения, что отразилось на монотонности полученных дозовых зависимостей.

Целью настоящей работы является моделирование распределения встроенного в окисел заряда у границы раздела кремний – скрытый окисел в *n*-канальных МОП/КНИ-транзисторах в зависимости от их геометрических параметров и электрических режимов при воздействии ионизирующего излучения.

Методика исследований. Объектами моделирования являлись тестовые МОП/КНИ-транзисторы с каналом *n*-типа. С помощью программного комплекса Silvaco [6] была создана двумерная модель полностью обедненного МОП/КНИ-транзистора. В первоначальной модели толщина эпитаксиальной пленки кремния составляла d_{Si} = 0,2 мкм, скрытого окисла $d_{\mathrm{SiO_2}}$ = 0,4 мкм, а длина канала L = 0,6 мкм при наличии запитки канала. В ходе дальнейших расчетов толщина пленки кремния варьировалась в пределах d_{Si} = 0,1–0,3 мкм, скрытого окисла $d_{\mathrm{SiO_2}}$ = 0,15–0,5 мкм и длина канала – L = 0,1–1,23 мкм. В программном модуле VictoryDevice было проведено моделирование воздействия на тестовые структуры рентгеновских квантов с энергией 10 кэВ при мощности поглощенной дозы 1 рад/с. После этого из полученных в ходе расчетов структур, подвергшихся воздействию излучения, было извлечено распределение заряда Q вблизи границы раздела кремний – скрытый окисел. Накопленная доза для всех образцов составляла 10^5 рад. Облучение проводилось в четырех электрических режимах:

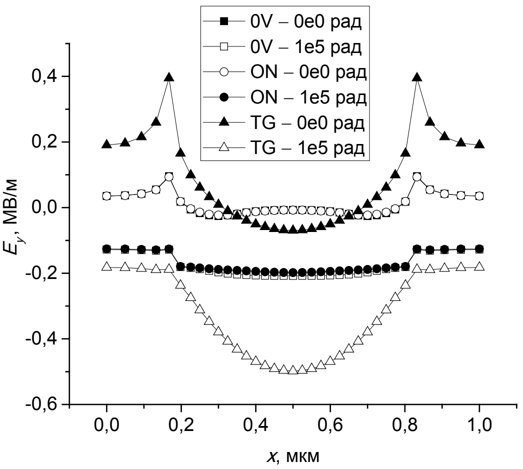
– 0V – напряжение на подложке, стоке, затворе, истоке U_{sub}, U_s, U_g, U_d = 0;
– ON – U_g = 5 В, U_{sub}, U_s, U_d = 0;
– TG – U_s, U_d = 5 В, U_{sub}, U_g = 0;
– TG_Sub – U_s, U_d = 5 В, U_{sub} = 0÷–5 В, U_g = 0.

Во всех исследованных режимах смещение на запитку канала не подавалось.

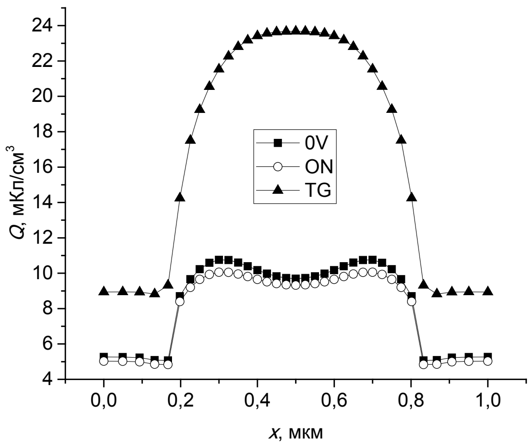
Результаты моделирования. В современных МОП/КНИ-транзисторах подзатворные окислы довольно тонкие (<10 нм), так что при облучении они не испытывают существенной деградации в связи с нейтрализацией положительного накопленного заряда электронами, туннелирующими из внешних слоев [2]. При моделировании будем рассматривать накопление заряда только в слое скрытого окисла у границы раздела со слоем активного кремния, так как там захват дырок на ловушках приводит к снижению порогового напряжения паразитного транзистора и радиационной стойкости.

На рис. 1 приведены результаты моделирования распределения вертикальной составляющей напряженности электрического поля E_y (*a*) и заряда Q (*b*) на глубине 5 нм в захороненном окисле вдоль границы раздела пленка кремния – слой скрытого окисла в моделях МОП/КНИ-транзисторов. Эти транзисторы были облучены рентгеновскими квантами до дозы 10^5 рад в различных электрических режимах. Рассматривались структуры с d_{Si} = 0,2 мкм, $d_{\mathrm{SiO_2}}$ = 0,4 мкм и L = 0,6 мкм.

Данные, представленные на рис. 1, *a*, показывают, что в разных электрических режимах распределение вертикальной составляющей напряженности поля значительно отличается. Распределение этой составляющей напряженности поля в режиме 0V без подачи смещения практически совпадает с распределением в «конденсаторном» режиме ON. Напряженность электрических полей в захороненном окисле в этих случаях является низкой, так как они создаются встроенными потенциалами между стоком или истоком *n*-типа и подложкой *p*-типа. В режиме облучения TG (Transmission Gate) величины вертикальной составляющей напряженности поля в центральной части структуры минимальны. При таких условиях индуцированные излучением дырки перемещаются в область с минимальным потенциалом – в приповерхностную область слоя скрытого окисла, а электрическое поле значительно способствует этому процессу. Дырки в указанной области вызывают сдвиг порогового напряжения паразитного транзистора, что уменьшает стойкость прибора. В режимах 0V и ON напряженности полей значительно выше.



a



b

Рис. 1. Распределение вертикальной составляющей напряженности электрического поля до и после облучения в разных режимах (*a*) и заряда после облучения в разных режимах (*b*)

Fig. 1. Distribution of the vertical component of the electric field intensity before and after irradiation in different modes (*a*) and charge distribution after irradiation in different modes (*b*)

Таким образом, после облучения зависимости не претерпели качественных значительных изменений, несмотря на то, что из-за увеличения количества захваченных зарядов величина напряженности поля значительно снизилась.

На рисунке 1, *b* показаны распределения заряда в разных режимах после облучения до дозы 10^5 рад. Если сравнить эти графики с изображенными на рис. 1, *a*, то можно увидеть, что имеется обратная зависимость между захваченным зарядом и вертикальной составляющей напряженности поля.

Во всех режимах облучения максимальное значение Q достигается в центральной (x = = 0,2–0,8 мкм) области структуры, а под карманами стока (0–0,2 мкм) и истока (0,8–1,0 мкм) оно заметно уменьшается. Также примечателен симметричный вид полученных зависимостей $Q(x)$ относительно прямой, параллельной оси ординат и проходящей через точку x = 0,5 мкм, что обусловлено симметрией как самой транзисторной структуры, так и эквипотенциальных линий электрических полей при исследуемых режимах испытаний.

Полученные результаты согласуются с экспериментальными данными, так как наиболее «жестким» для *n*-канальных МОП/КНИ-транзисторов является режим TG, при котором во время облучения на сток и исток подается напряжение +5 В, а на подложку, затвор и запитку канала – 0 В [4]. При этом электрическое поле в захороненном окисле способствует накоплению избыточного положительного заряда у границы раздела с пленкой кремния и, следовательно, снижению порогового напряжения паразитного транзистора. Уменьшение влияния электрического поля в этом случае достигается подачей на подложку определенного отрицательного смещения U_{sub} [7], что соответствует режиму облучения TG_Sub.

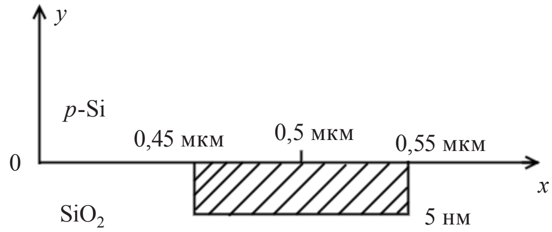


Рис. 2. Схематичное изображение области скрытого окисла, в которой отслеживалось изменение заряда

Fig. 2. Schematic representation of a buried oxide area where the charge alteration was tracked

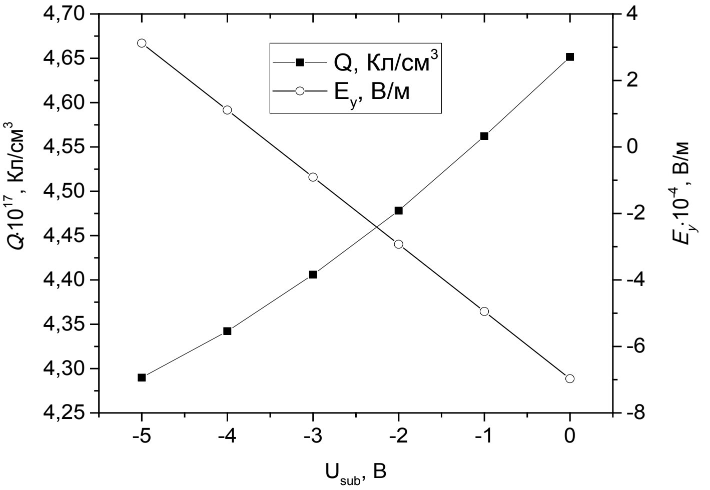


Рис. 3. Зависимости изменения заряда и напряженности поля при разных значениях смещения на подложке в процессе облучения

Fig. 3. Dependences of charge and field intensity alterations at different bias values on the substrate during irradiation

Симметричный вид кривых $Q(x)$ и $E_y(x)$ на рис. 1 позволяет упростить задачу моделирования и рассматривать не распределение заряда вдоль границы раздела пленка кремния – слой скрытого окисла, а только его изменение в приповерхностной центральной области скрытого слоя окисла кремния. На рис. 2 штриховкой обозначена область в центральной приповерхностной части скрытого окисла, в которой накопленный заряд интегрировался по площади.

Зависимости $Q(U_{\text{sub}})$ и $E_y(U_{\text{sub}})$ в режиме облучения TG_Sub при разных смещениях на подложке приведены на рис. 3 для тех же транзисторных структур, что и на рис. 1. В этом случае величина напряженности электрического поля определялась в конкретной точке структуры с координатами $x = 0{,}5$ мкм, $y = 0{,}205$ мкм. Заряд суммировался в области, показанной на рис. 2.

Из рис. 3 видно, что с возрастанием величины отрицательного смещения на подложке значение Q уменьшается, а E_y увеличивается, т. е. эффективность захвата заряда зависит от распределения и величины напряженности электрического поля в транзисторной структуре. Следовательно, величина накопленного заряда должна зависеть и от геометрических размеров транзистора.

Известно [8], что изменение определенных геометрических параметров транзистора приводит к изменению его радиационной стойкости, поэтому далее рассмотрим влияние изменения L, d_{Si} и d_{SiO_2} на накопление заряда в центральной части приповерхностной области структуры в электрических режимах облучения TG и TG_Sub. На рис. 4 показаны зависимости значения Q при разных значениях смещения на подложке в процессе облучения от длины канала L (рис. 4, *a*), толщин пленки кремния d_{Si} (рис. 4, *b*) и скрытого окисла d_{SiO_2} (рис. 4, *c*). В каждом случае изменялся лишь один параметр из трех. Зависимости $Q(L)$ имеют немонотонный вид; на этих кривых (см. рис. 4, *a*) наблюдается максимум при $L \sim 0{,}5$ мкм, где величина захваченного заряда наиболее высока. С уменьшением длины канала от 0,4 до 0,1 мкм снижается не только значение Q, но и положительный эффект отрицательного смещения на подложке транзисторных структур в процессе облучения. Также на стойкости положительно сказывается увеличение длины канала до величин свыше 1 мкм.

При увеличении толщины пленки активного кремния значение накопленного заряда уменьшается (см. рис. 4, *b*). При изменении этого параметра эффективность режима с подачей отрицательного смещения на подложку не снижается. Однако изменение Q выражено слабее, чем в предыдущем случае, и составило около 20 % при уменьшении толщины слоя активного кремния с 0,3 до 0,1 мкм.

Гораздо более значительные изменения в накопленном заряде обнаруживаются при уменьшении толщины скрытого окисла (см. рис. 4, *c*), что согласуется с результатами, полученными в работе [8]. При снижении толщины слоя скрытого окисла с 0,55 до 0,15 мкм Q уменьшается

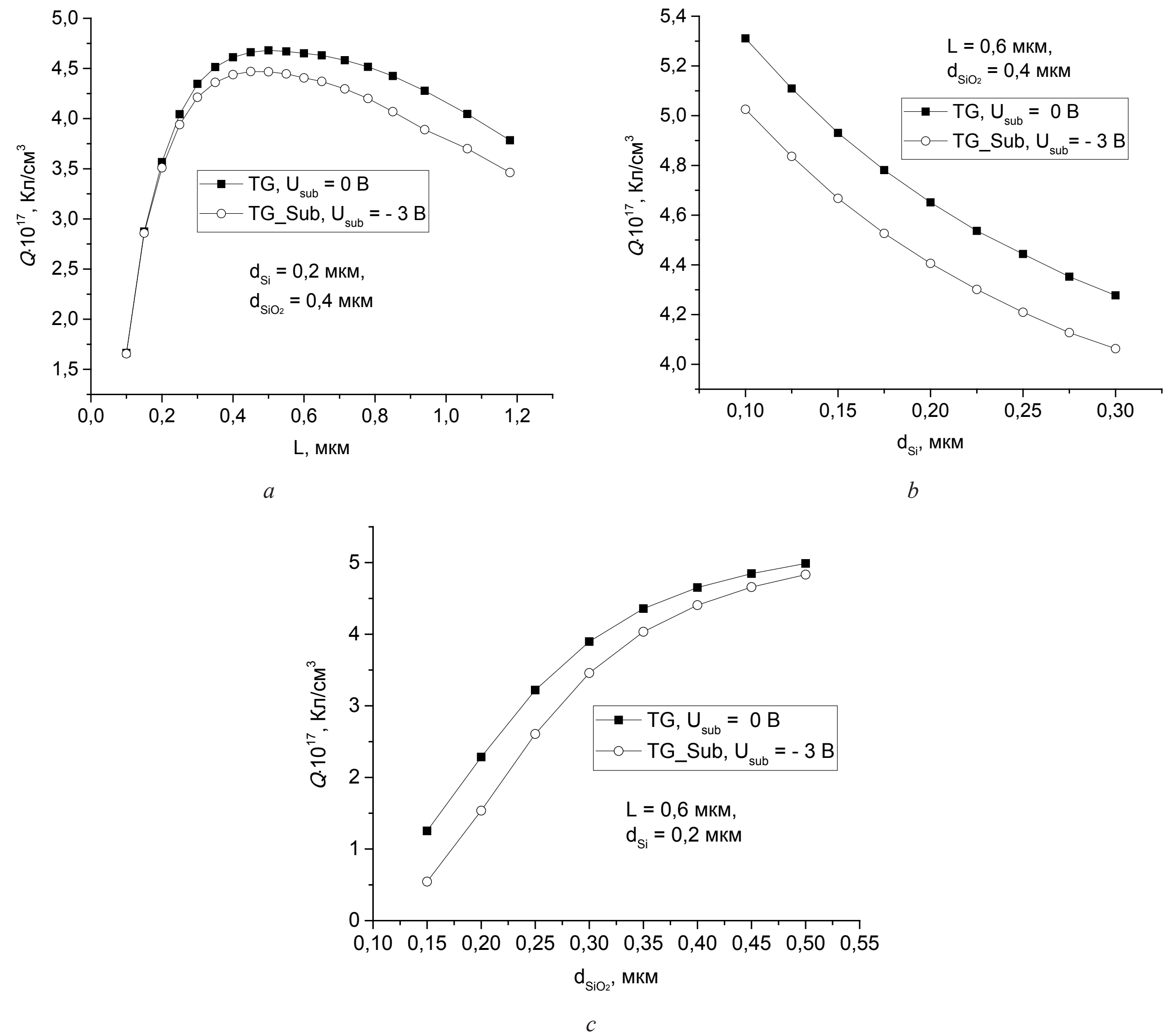


Рис. 4. Зависимость значения поверхностного заряда в точке x = 0,5 мкм при разных значениях смещения на подложке в процессе облучения от длины канала L (*a*), толщины пленки кремния d_{Si} (*b*) и скрытого окисла $d_{\mathrm{SiO_2}}$ (*c*)

Fig. 4. Dependence of the surface charge values at the point x = 0.5 µm at different bias values on the substrate during irradiation from the channel length L (*a*), thickness of the silicon film d_{Si} (*b*) and buried oxide $d_{\mathrm{SiO_2}}$ (*c*)

в ~4 раза для U_{sub} = 0 В и в ~9 раз для U_{sub} = –3 В, что значительно превышает изменения в случае варьирования длины канала и толщины пленки активного кремния.

Обсуждение результатов моделирования. Общая картина радиационного дефектообразования в МОП-приборах описана в различных литературных источниках [1, 2, 9, 10]. При поглощении ионизирующего излучения материалом вещества происходит генерация электронно-дырочных пар. Часть этих образовавшихся электронов и дырок рекомбинирует друг с другом, но количество оставшихся пар зависит от природы источника ионизирующего излучения, а также напряженности электрических полей. Подвижность электронов в SiO_2 (~20 см^2/В · с) намного выше подвижности дырок (~10^{-5} см^2/В · с) [9], поэтому первые быстро покидают окисел, а вторые перемещаются в области окисла с минимальным потенциалом, следуя линиям электрического поля [11]. Дырки под действием поля могут достигать границы раздела пленка кремния – скрытый окисел, где они захватываются ловушками и выступают как дополнительное положительное напряжение, приложенное к затвору паразитного транзистора. Вследствие этого пороговое напряжение паразитного транзистора испытывает существенный сдвиг в сторону отрицательных значений, что в свою очередь приводит к увеличению токов утечки рабочего n-канального транзистора и негативно сказывается на радиационной стойкости.

В силу перечисленных факторов на стойкость МОП/КНИ-транзистора к ионизирующему излучению очень сильное влияние оказывают как электрические режимы во время облучения, так и геометрические размеры структуры транзистора.

Следует отметить, что в настоящей работе при моделировании радиационных эффектов в МОП/КНИ-транзисторах не учитывался заряд быстрых поверхностных состояний на границе раздела кремний – скрытый окисел. Это обусловлено тем, что в приборах современных МОП (в том числе и КНИ) технологий при дозах до 1 Мрад (SiO_2) накопление поверхностных состояний незначительно [12]. В работе [13] также было обнаружено, что генерация дополнительных состояний на границах раздела Si–SiO_2 в КНИ-структурах отсутствует при облучении их электронами (энергия 2,5 МэВ) и гамма-квантами (энергия 662 кэВ).

Заключение. С помощью программного комплекса Silvaco проведены расчеты накопления встроенного в окисел заряда у границы раздела кремний – скрытый окисел в тестовых *n*-канальных МОП/КНИ-транзисторах при разных геометрических параметрах и электрических режимах в ходе облучения. Показано, что наиболее «жестким» является режим TG, при котором во время облучения на сток и исток подается напряжение +5 В, а на подложку, затвор и запитку канала – 0 В. При этом величину накопленного заряда удается существенно снизить, прикладывая к подложке отрицательное смещение и уменьшая толщину слоя захороненного окисла транзисторных структур. Также определенное влияние на стойкость оказывает изменение толщины пленки активного кремния и длины канала, причем зависимости изменения величины накопленного заряда Q от варьирования длины канала имеют немонотонный вид и не оказывают столь значительного влияния на радиационную стойкость, как уменьшение толщины слоя скрытого окисла.

Полученные результаты компьютерного моделирования могут быть использованы испытателями электронных компонентов для предварительной оценки стойкости МОП/КНИ-приборов к накопленной дозе ионизирующего излучения, а также будут полезны для разработчиков радиационно-стойких интегральных МОП-схем на основе КНИ-структур.

Список использованных источников

1. Коршунов, Ф. П. Воздействие радиации на интегральные микросхемы / Ф. П. Коршунов, Ю. В. Богатырев, В. А. Вавилов. – Минск: Наука и техника, 1986. – 254 с.

2. Таперо, К. И. Радиационные эффекты в кремниевых интегральных схемах космического применения / К. И. Таперо, В. Н. Улимов, А. М. Членов. – М.: БИНОМ, Лаб. знаний, 2012. – 304 с.

3. Worst-Case Bias During Total Dose Irradiation of SOI Transistors / V. Ferlet-Cavrois [et al.] // IEEE Trans. Nucl. Sci. – 2000. – Vol. 47, № 6. – P. 2183–2188. https://doi.org/10.1109/23.903751

4. Flament, O. Bias Dependence of FD Transistor Response to TotalDose Irradiation / O. Flament, A. Torres, V. Ferlet-Cavrois// IEEE Trans. Nucl. Sci. – 2003. – Vol. 50, № 6. – P. 2316–2321. https://doi.org/10.1109/tns.2003.822594

5. Сравнение различных вариантов топологии КНИ/МОП-транзисторов для проектирования радиационно-стойких ИС / М. С. Горбунов [и др.] // Вопр. атомной науки и техники. Сер. Физика радиац. воздействия на радиоэлектрон. аппаратуру. – 2010. – Вып. 1. – С. 39–43.

6. SILVACO International. ATLAS User's Manual. Device Simulation Software [Electronic Resource]. – Mode of access: http: // www.silvaco.com. – Date of access: 13.03.2018.

7. Влияние гамма-излучения на МОП/КНИ-транзисторы / Ю. В. Богатырев [и др.] // Докл. БГУИР. – 2016. – № 3 (97). – С. 75–80.

8. Impact of technology scaling in SOI back-channel total dose tolerance. A 2-D numerical study using self-consistent oxide code / J.-L. Leray [et al.] // IEEE Trans. Nucl. Sci. – 2000. – Vol. 47, № 3. – P. 620–627. https://doi.org/10.1109/23.856489

9. Ma, T. P. Ionizing radiation effects in MOS devices and circuits / T. P. Ma, P. V. Dressendorfer. – New York: John Wiley and Sons, 1989. – P. 179.

10. Lacoe, R. CMOS scaling, design principles and hardening-by-design methodologies / R. Lacoe // IEEE NSREC. Short Course. – 2003. – P. II-1–II-142.

11. Correlation Between Co-60 and X-Ray Radiation-Induced Charge Buildup in Silicon-on-Insulator Buried Oxides / J. R. Schwank [et al.] // Trans. Nucl. Sci. – 2000. – Vol. 47, № 6. – P. 2175–2182. https://doi.org/10.1109/23.903750

12. Зебрев, Г. И. Радиационные эффекты в кремниевых интегральных схемах высокой степени интеграции / Г. И. Зебрев. – М.: НИЯУ МИФИ, 2010. – 148 с.

13. Накопление заряда в диэлектрике и состояния на границах структур кремний-на-изоляторе при облучении электронами и гамма-квантами / Д. В. Николаев [и др.] // Физика и техника полупроводников. – 2003. – Т. 37, вып. 4. – С. 443–449.

References

1. Korshunov F. P., Bogatyrev Yu. V., Vavilov V. A. *The Influence of Radiation on Integrated Circuits*. Minsk, Nauka i tekhnika Publ., 1986. 254 p. (in Russian).

2. Tapero K. I., Ulimov V. N., Chlenov A. M. *Radiation Effects in Silicon Integrated Circuits for Space Applications*. Moscow, BINOM, Laboratoriya znanii Publ., 2012. 304 p. (in Russian).

3. Ferlet-Cavrois V., Colladant T., Paillet P., Leray J. L., Musseau O., Schwank J. R., Shaneyfelt M. R., Pelloie J. L., du Port de Poncharra J. Worst-Case Bias During Total Dose Irradiation of SOI Transistors. *IEEE Transactions on Nuclear Science,* 2000, vol. 47, no. 6, pp. 2183–2188. https://doi.org/10.1109/23.903751

4. Flament O., Torres A., Ferlet-Cavrois V. Bias Dependence of FD Transistor Response to TotalDose Irradiation. *IEEE Transactions on Nuclear Science*, 2003, vol. 50, no. 6, pp. 2316–2321. https://doi.org/10.1109/tns.2003.822594

5. Gorbunov M. S., Zebrev G. I., Osipenko P. N., Vasilegin B. V., Il'yaguev V. N. Comparison of different variants of the topology design of MOS transistors for the design of radiation-hardened ICs. *Voprosy atomnoi nauki i tekhniki. Seriya Fizika radiatsionnogo vozdeistviya na radioelektronnuyu apparaturu* [Problems of atomic science and technology, «Physics of radiation effects on electronic equipment»], 2010, no. 1, pp. 39–43 (in Russian).

6. SILVACO International. ATLAS User's Manual. Device Simulation Software. Available at: http: //www.silvaco.com. (accessed 13 March 2018).

7. Bogatyrev Yu. V., Lastovskii S. B., Soroka S. A., Shvedov S. V., Ogorodnikov D. A. The effect of X-ray radiation on MOS / SOI transistors. *Doklady BGUIR*, 2016, no. 3 (97), pp. 75–80 (in Russian).

8. Leray J.-L., Paillet P., Ferlet-Cavrois V., Tavernier C., Belhaddad K., Penzin O. Impact of technology scaling in SOI back-channel total dose tolerance. A 2-D numerical study using self-consistent oxide code. *IEEE Transactions on Nuclear Science*, 2000, vol. 47, no. 3, pp. 620–627. https://doi.org/10.1109/23.856489

9. Ma T. P., Dressendorfer P. V. *Ionizing Radiation Effects in MOS Devices and Circuits*. New York, John Wiley and Sons, 1989, pp. 179.

10. Lacoe R. CMOS scaling, design principles and hardening-by-design methodologies. *IEEE NSREC. Short Course*, 2003, pp. II-1–II-142.

11. Schwank J. R., Shaneyfelt M. R., Dodd P. E., Ferlet-Cavrois V., Loemker R. A., Winokur P. S., Fleetwood D. M., Paillet P., Leray J.-L., Draper B. L., Witczak S. C., Riewe L. C. Correlation Between Co-60 and X-Ray Radiation-Induced Charge Buildup in Silicon-on-Insulator Buried Oxides. *IEEE Transactions on Nuclear Science*, 2000, vol. 47, no. 6, pp. 2175–2182. https://doi.org/10.1109/23.903750

12. Zebrev G. I. *Radiation Effects in High Integration Silicon Integrated Circuits*. Moscow, National Research Nuclear University MEPhI, 2010. 148 p. (in Russian).

13. Nikolaev D. V., Antonova I. V., Naumova O. V., Popov V. P., Smagulova S. A. Charge accumulation in oxide and interface states of silicon-on-insulator structures after irradiation by electrons and γ-rays. *Fizika i tekhnika poluprovodnikov = Semiconductors,* 2003, vol. 37, no. 4. pp. 426–432. https://doi.org/10.1134/1.1568462

Информация об авторах

Огородников Дмитрий Александрович – младший научный сотрудник, ГНПО «НПЦ НАН Беларуси по материаловедению» (ул. П. Бровки, 19, 220072, г. Минск, Республика Беларусь). E-mail: ogorodnikov@ifttp.bas-net.by

Ластовский Станислав Брониславович – кандидат физико-математических наук, заведующий лабораторией радиационный воздействий, ГНПО «НПЦ НАН Беларуси по материаловедению» (ул. П. Бровки, 19, 220072, г. Минск, Республика Беларусь). E-mail: lastov@physics.by

Богатырев Юрий Владимирович – доктор технических наук, главный научный сотрудник, ГНПО «НПЦ НАН Беларуси по материаловедению» (ул. П. Бровки, 19, 220072, г. Минск, Республика Беларусь). E-mail: bogat@ifttp.bas-net.by

Information about the authors

Dmitriy A. Ogorodnikov – Junior Researcher, SSPA "Scientific-Practical Materials Research Centre of NAS of Belarus" (19, P. Brovka Str., 220072, Minsk, Republic of Belarus). E-mail: ogorodnikov@ifttp.bas-net.by

Stanislav B. Lastovskii – Ph. D. (Physics and Mathematics), Head of the Laboratory of Radiation Effects, SSPA "Scientific-Practical Materials Research Centre of NAS of Belarus" (19, P. Brovka Str., 220072, Minsk, Republic of Belarus). E-mail: lastov@physics.by

Yuriy V. Bogatyrev – Dr. Sc. (Engineering), Chief Researcher, SSPA "Scientific-Practical Materials Research Centre of NAS of Belarus" (19, P. Brovka Str., 220072, Minsk, Republic of Belarus). E-mail: bogat@ifttp.bas-net.by

ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)

УЧЕНЫЕ БЕЛАРУСИ

SCIENTISTS OF BELARUS

АРКАДИЙ ПЕТРОВИЧ ИВАНОВ

(К 90-летию со дня рождения)

29 декабря 2019 г. заслуженный деятель науки Республики Беларусь, член-корреспондент Национальной академии наук Беларуси, доктор физико-математических наук, профессор, почетный профессор Винницкого национального технического университета Аркадий Петрович Иванов отмечает юбилей – ему исполняется 90 лет.

Аркадий Петрович родился в 1929 г. в г. Самара (Россия). В 1948 г. поступил в Ленинградский институт точной механики и оптики, после окончания которого в 1953 г. был распределен в Государственный оптический институт им. С. И. Вавилова, где и начал свою научную деятельность.

Первые работы А. П. Иванова были посвящены спектроскопии диспергированных люминесцирующих объектов. По этой теме в 1958 г. им была защищена кандидатская диссертация. В том же году, еще до появления лазеров, он предсказал возможность возникновения отрицательного поглощения (т. е. усиления) света в растворах сложных молекул. В дальнейшем белорусским ученым Б. И. Степанову, А. Н. Рубинову, В. А. Мостовникову впервые удалось получить генерацию света на указанном типе молекул.

В 1959 г. по приглашению академика Б. И. Степанова ученый переехал в г. Минск и приступил к работе в Институте физики АН БССР. С 1959 г. – старший научный сотрудник, с 1964 г. – заведующий лабораторией, с 2008 г. и по настоящее время – главный научный сотрудник Института. В 1967 г. А. П. Иванов защитил докторскую диссертацию, в 1970 г. ему присвоено ученое звание профессора, в 1974 г. избран членом-корреспондентом НАН Беларуси.

Аркадий Петрович стал работать в Академии наук в период бурного развития научной деятельности в республике. Его работы по рассеянию света в дисперсных средах, начатые в Ленинграде, были продолжены в Минске. Развитие этого научного направления диктовалось необходимостью решения широкого круга задач, возникающих в таких областях, как оптика атмосферы и океана, астрономия, спектроскопия, климатология, создание оптических линий связи, спутниковые исследования земных и водных ресурсов, исследование биотканей и биожидкостей, контроль загрязнения атмосферы и воды.

В 1964 г. А. П. Ивановым была создана лаборатория оптики рассеивающих сред, которая быстро стала одной из крупнейших в Институте физики. Выбор основных направлений исследований, предложенный ученым, был весьма необычен с точки зрения того времени: нелинейные явления в рассеивающих средах, закономерности распространения узких направленных световых пучков. Еще не было источников достаточно большой мощности, не был создан лазер, не была разработана базовая теория распространения узких пучков света.

Появление лазеров в 1960-е гг. и последующая лазерная революция в науке и технологиях определили актуальность и перспективность указанных направлений на многие годы и обеспечили Аркадию Петровичу и руководимой им лаборатории обширное поле интересных и актуальных научных проблем, а также контакты и признание мировой научной общественности. При участии и под руководством А. П. Иванова были развиты методы решения прямых и обратных задач рассеяния света и на этой основе выполнены обстоятельные лабораторные и натурные экспериментальные исследования. Благодаря комплексному подходу, удачно сочетающему теоретические и экспериментальные работы, удалось создать большой арсенал методов и приборов для изучения рассеяния света.

Широкое применение нашел предложенный ученым метод экспериментального оптического моделирования, основанный на принципе оптического подобия. Он позволил не только получить ряд принципиально новых результатов, но и резко снизить технические трудности и материальные затраты на проведение натурных экспериментов. Сегодня метод оптического моделирования активно используется во всех научных центрах, занимающихся исследованиями в области оптики атмосферы, океана, научной фотографии, биофизики и т. д.

Аркадий Петрович явился зачинателем многих научных направлений. В частности, под его руководством в Институте физики начались работы по оценке качества оптического изображения. Эти исследования, по результатам которых ученым опубликованы три монографии, сделали возможным решить ряд принципиальных вопросов формирования и передачи оптического изображения фотографическими слоями, люминесцентными и жидкокристаллическими экранами и объективами, послужили теоретической основой для создания систем видения в воде и атмосфере, оптической локации и связи, переноса тепла в средах с внедренными частицами дисперсного вещества.

Особо следует отметить заслуги А. П. Иванова в создании и развитии перспективного научного направления – метода лазерного зондирования воды и атмосферы. Первое обращение лидарного уравнения (восстановление высотного профиля оптических характеристик атмосферы по лидарному сигналу), первые эксперименты по лазерному зондированию воды (на озере Нарочь, 1966 г.) были выполнены именно Аркадием Петровичем и сотрудниками его лаборатории. Ими был предложен, теоретически обоснован и экспериментально проверен ряд новых методик лазерного зондирования. В дальнейшем развитие метода лазерного зондирования атмосферы и воды стало одним из основных направлений научной деятельности А. П. Иванова. В 1960–1990-х гг. под его руководством была создана целая серия лидаров различного типа и назначения, позволивших получить ряд уникальных результатов, в особенности по многочастотному лазерному зондированию атмосферного аэрозоля. Лазерное зондирования атмосферы дает возможность дистанционно определять концентрацию различных газов и аэрозолей в атмосфере, решать многие экологические задачи.

Оригинальность подходов к решению самых разных проблем характерна для Аркадия Петровича, им предложен ряд новых неожиданных постановок и решений многих научных и технических задач. Можно указать на теоретически предсказанное и экспериментально доказанное явление гашения плоской электромагнитной волны при ее прохождении через монослой частиц дисперсного вещества. Исследования, выполненные А. П. Ивановым и его учениками, позволили всесторонне изучить особенности распространения стационарного и импульсного излучения лазерных источников в рассеивающих средах, выяснить условия и закономерности проявления нелинейных эффектов, связанных с большой мощностью излучения, исследовать и использовать для изучения структуры среды явления, обусловленные когерентностью рассеянного излучения, предложить и развить ряд принципиально новых методик спектроскопии светорассеивающих сред.

60-летний период своей плодотворной научной деятельности Аркадий Петрович посвятил развитию физической оптики в Беларуси. Он принадлежит к числу белорусских ученых, которые обогатили физическую науку фундаментальными и прикладными научными трудами. Аркадий Петрович – автор около 800 научных трудов, в том числе 13 монографий, 29 авторских свидетельств. Одной из важнейших заслуг А. П. Иванова стало создание в Беларуси научной школы в области оптики рассеивающих сред, широко известной как в нашей стране, так и далеко за ее пределами. Под его руководством защищено 30 кандидатских и 4 докторские диссертации, один из его учеников избран членом-корреспондентом НАН Беларуси.

Вклад А. П. Иванова в развитие физической науки отмечен правительственными наградами. В 1981 г. за цикл работ по исследованию оптических явлений в атмосфере методом оптического моделирования ему присуждена Премия Совета Министров СССР, в 2002 г. за цикл работ «Диагностика состояния природной среды на основе аэрокосмических, лидарных, наземных и химико-аналитических методов и средств: исследования, разработки, внедрения» – Государственная премия Республики Беларусь. В 2008 г. за организацию лидарной сети стран СНГ и контроля трансграничного переноса загрязнений через Беларусь А. П. Иванов был удостоен Премии НАН Беларуси и Сибирского отделения РАН имени академика В. А. Коптюга.

Для Аркадия Петровича характерны широта научных интересов, глубокое понимание физики исследуемых явлений, высокая требовательность, огромная работоспособность и скромность. И особенно хочется подчеркнуть его редкую научную интуицию и доброжелательную восприимчивость к предлагаемым идеям, черты, столь необходимые в науке. Не ослабевает с годами творческая и общественная активность А. П. Иванова. Свой юбилей Аркадий Петрович встречает полный энергии, являясь примером жизнелюбия и оптимизма.

Сердечно поздравляем Аркадия Петровича с 90-летием, желаем ему крепкого здоровья и дальнейших научных успехов.

Отделение физики, математики и информатики НАН Беларуси,
Институт физики НАН Беларуси,
Белорусское физическое общество

ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)

ЮРИЙ СЕМЕНОВИЧ ХАРИН

(К 70-летию со дня рождения)

17 сентября 2019 г. исполнилось 70 лет известному ученому в области прикладной математики и информатики, члену-корреспонденту Национальной академии наук Беларуси, доктору физико-математических наук, профессору, директору Научно-исследовательского института прикладных проблем математики и информатики БГУ Юрию Семеновичу Харину.

Юрий Семенович родился в 1949 г. в селе Зырянское Томской области. В 1966 г. окончил школу в г. Томске с золотой медалью, в 1971 г. – Томский государственный университет с отличием, в 1974 г. – аспирантуру факультета прикладной математики с досрочной защитой кандидатской диссертации. В 1976 г. в связи с организацией в БГУ кафедры теории вероятностей и математической статистики был приглашен для работы в должности доцента. Здесь, в Белорусском государственном университете, Ю. С. Харин состоялся как ученый и педагог, пройдя путь от доцента до заведующего кафедрой и директора НИИ. В 1986 г. в Институте прикладной математики АН СССР (г. Москва) защитил докторскую диссертацию, в 1988 г. ему присвоено ученое звание профессора, в 2004 г. избран членом-корреспондентом НАН Беларуси.

В 1988 г. по инициативе Юрия Семеновича на факультете прикладной математики и информатики БГУ была организована новая кафедра математического моделирования и анализа данных, которую ученый возглавлял со дня основания до 2018 г., а в настоящее время является ее научным руководителем. В 2000 г. Ю. С. Харин также инициировал создание в БГУ Национального научно-исследовательского центра прикладных проблем математики и информатики, преобразованного в 2008 г. в НИИ прикладных проблем математики и информатики, на который Постановлением Совета Министров Республики Беларусь возложена деятельность по координации в стране научных исследований в области криптографии и анализа данных.

Основные направления научных исследований Ю. С Харина – математическая и прикладная статистика, многомерный статистический анализ, распознавание и прогнозирование сложных случайных процессов, анализ данных, статистическое моделирование, криптография, защита информации.

На основе разработанных ученым теории робастного (устойчивого) статистического распознавания образов, теории робастного статистического анализа данных и прогнозирования построены новые минимаксно устойчивые (к искажениям вероятностной модели) алгоритмы распознавания образов, идентификации и прогнозирования, гарантирующие наименьшее уклонение риска на заданных параметрических и непараметрических семействах искажений, реализованные в пакетах прикладных программ. Они широко используются в Беларуси, России, Германии, Южной Корее и дают возможность решать важные прикладные задачи внедрения информационных технологий в промышленности, медицине и экономике.

Ю. С. Харину принадлежит ведущая роль в создании и развитии Белорусской научной школы криптологии – науки о математических и компьютерных методах защиты информации. С 2001 г. он является научным руководителем Государственной программы по криптологии. Под его личным руководством выполнено более 50 НИР по государственным научным и научно-техническим программам, результаты которых имеют приоритетное значение и внедрены при создании компьютерных систем защиты информации в электронном документообороте республики.

Полученные ученым результаты широко известны как в нашей стране, так и за рубежом. Ю. С. Хариным опубликовано более 600 научных трудов, в том числе пять монографий, две из которых изданы за рубежом. Он является соавтором трех коллективных зарубежных монографий, 27 учебников и учебных пособий, 14 из которых имеют гриф Министерства образования Республики Беларусь, а так же международной энциклопедии «Statistical Sciences». Его статьи опубликованы в таких престижных зарубежных журналах, как «Mathematical Methods of Statistics», «Journal of Statistical Planning and Inference», «Applied Stochastic Models and Data Analysis», «Journal of Classification», «Communications in Statistics», «Pattern Recognition Letters», «Lecture Notes in Computer Science», «Lecture Notes in Statistics», «Теория вероятностей и ее применения», «Дискретная математика». Им сделаны доклады более чем на 90 международных научных конференциях. Как приглашенный профессор читал лекции в Калифорнийском университете (Беркли, США), Байрёйтском и Дрезденском университетах (Германия), Лейстерском (Великобритания) и Венском (Австрия) университетах, а также в других ведущих научных центрах. Под его руководством защищены 1 докторская и 34 кандидатские диссертации.

Ю. С. Харин является организатором новых для Беларуси университетских специальностей – «Компьютерная безопасность», «Прикладная криптография», «Экономическая кибернетика», двух специализаций по «Прикладной математике», а также аспирантуры и докторантуры по специальности «Методы и системы защиты информации, информационная безопасность».

Юрий Семенович – член Президиума Высшей аттестационной комиссии Республики Беларусь, Белорусского математического общества, Белорусской ассоциации распознавания образов, Американского математического общества, Международного института математической статистики, Международной ассоциации вычислительной техники, Европейской ассоциации обработки сигналов, Всемирного общества Бернулли по математической статистике, Белорусской статистической ассоциации (ее организатор, а с 1998 по 2008 гг. – руководитель), шести редколлегий зарубежных и пяти республиканских научных журналов, в том числе издания «Весці Нацыянальнай акадэміі навук Беларусі. Серыя фізіка-матэматычных навук», редактор 25 сборников научных статей, главный редактор «Журнала Белорусского государственного университета. Математика и информатика», заместитель главного редактора периодического межведомственного научного сборника «Проблемы защиты информации», организатор двенадцати международных научных конференций «Computer Data Analysis and Modeling», председатель Программного комитета четырнадцати международных конференций «Комплексная защита информации».

За высокие достижения в области науки Ю. С. Харину присвоено звание Заслуженный деятель науки Республики Беларусь (2010 г.). Он удостоен звания лауреата Государственной премии Республики Беларусь в области науки и техники за цикл работ «Распознавание и анализ стохастических данных и цифровых изображений» (2002 г.), Премии им. А. Н. Севченко (1997 г.). Является Отличником образования Республики Беларусь (1999 г.), Заслуженным работником Белорусского государственного университета (2013 г.), награждался Почетными грамотами Совета Министров Республики Беларусь, Государственного комитета по науке и технологиям Республики Беларусь, Министерства образования Республики Беларусь, НАН Беларуси и др.

Глубокие научные идеи, высокий профессионализм и организаторские способности, культура общения, принципиальность и трудолюбие обеспечили Ю. С. Харину авторитет талантливого ученого и руководителя, чуткого и доброжелательного человека.

Научная общественность и редколлегия журнала сердечно поздравляют Юрия Семеновича с юбилеем, желают ему крепкого здоровья и дальнейших творческих достижений.

Отделение физики, математики и информатики НАН Беларуси,
Институт математики НАН Беларуси,
Объединенный институт проблем информатики НАН Беларуси,
Белорусское математическое общество

ПЕРЕЧЕНЬ СТАТЕЙ, ОПУБЛИКОВАННЫХ В ЖУРНАЛЕ «ВЕСЦІ НАЦЫЯНАЛЬНАЙ АКАДЭМІІ НАВУК БЕЛАРУСІ. СЕРЫЯ ФІЗІКА-МАТЭМАТЫЧНЫХ НАВУК» В 2019 ГОДУ

МАТЕМАТИКА

ФИЗИКА

ИНФОРМАТИКА

УЧЕНЫЕ БЕЛАРУСИ

LIST OF PUBLICATIONS FOR 2019 IN "PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF BELARUS. PHYSICS AND MATHEMATICS SERIES"

MATHEMATICS

PHYSICS

INFORMATICS

SCIENTISTS OF BELARUS