

Том 64, Номер 2

ISSN 0044-4669

Февраль 2024



ФИЦ ИУ РАН

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ



НАУКА

— 1727 —

СОДЕРЖАНИЕ

Том 64, номер 2, 2024 год

ОБЩИЕ ЧИСЛЕННЫЕ МЕТОДЫ

- Оценка QTT рангов регулярных функций на равномерной квадратной сетке
А. В. Зыль, Н. Л. Замарашкин 189
- Улучшенная квадратурная формула для прямого значения нормальной производной потенциала простого слоя
П. А. Крутицкий, И. О. Резниченко 200
- Приближение непрерывных функций с помощью классических синков и значений операторов S_L
В. Н. Пасечник 220

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ

- Синтез оптимальной устойчивой аффинной системы
Л. Т. Ащепков 225

ОБЫКНОВЕННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

- Гладкие многообразия Ляпунова для автономных систем нелинейных обыкновенных дифференциальных уравнений и их применение к решению сингулярных краевых задач
Н. Б. Конюхова 232

УРАВНЕНИЯ В ЧАСТНЫХ ПРОИЗВОДНЫХ

- Явные численно реализуемые формулы для операторов Пуанкаре – Стеклова
А. С. Демидов, А. С. Самохин 253
- Метод численного решения нестационарного уравнения Шрёдингера десятого порядка точности
М. А. Захаров 263
- Решение двумерного нелинейного параболического уравнения теплопроводности при краевом режиме, заданном на подвижном многообразии
А. Л. Казаков, О. А. Нефедова, Л. Ф. Спевак 283
- О разрешимости существенно нелинейного эллиптического дифференциального уравнения с нелокальными краевыми условиями
О. В. Солонуха 304

МАТЕМАТИЧЕСКАЯ ФИЗИКА

- Локальная вейвлетная адаптация декартовых сеток в вычислительных задачах газовой динамики
А. Л. Афендииков, В. С. Никитин 322

Новые компьютерно-экономичные аппроксимации случайных функций для решения стохастических задач теории переноса <i>Г. А. Михайлов, И. Н. Медведев</i>	337
Исследование применения явно-итерационной схемы к моделированию дозвуковых реагирующих газовых потоков <i>Е. Е. Пескова, О. С. Язовцева</i>	350
Алгоритм решения четырехволнового кинетического уравнения в задачах волновой турбулентности <i>Б. В. Семисалов, С. Б. Медведев, С. В. Назаренко, М. П. Федорук</i>	364

ОЦЕНКА QTT РАНГОВ РЕГУЛЯРНЫХ ФУНКЦИЙ НА РАВНОМЕРНОЙ КВАДРАТНОЙ СЕТКЕ¹⁾

© 2024 г. А. В. Зыль^{1,2,*}, Н. Л. Замарашкин^{1, **}

¹119333 Москва, ул. Губкина, 8, Ин-т вычисл. математики РАН, Россия

²141701 Долгопрудный, М. о., Институтский переулок, 9,
Московский физико-технический институт, Россия

*e-mail: zyl.av@phystech.edu

** e-mail: nikolai.zamarashkin@gmail.com

Поступила в редакцию 10.07.2023 г.

Переработанный вариант 10.07.2023 г.

Принята к публикации 20.10.2023 г.

В работе доказываются оценки ε -рангов для ТТ-разложений тензоров, полученных путем тензоризации значений регулярной функции одной комплексной переменной на равномерной квадратной сетке на комплексной плоскости. Установлена связь точности приближения и геометрии области регулярности функции. Библ. 8. Фиг. 2.

Ключевые слова: тензорный поезд, QTT, оценка рангов ТТ-аппроксимаций, регулярные функции.

DOI: 10.31857/S0044466924020017, **EDN:** YKVVQO

1. ВВЕДЕНИЕ

В работе получены оценки ε -рангов для ТТ-аппроксимаций тензоров, образующихся после тензоризации значений регулярной функции $f(z)$ одного комплексного переменного на равномерной сетке в квадратной области в \mathbb{C} . Техника доказательства следует работе [1], можно сказать, что результаты из [1] обобщаются на комплексный случай. Похожие вопросы также затрагивались в работах [2], [3], [4]. В разд. 2, 3 будут даны необходимые сведения о тензорах, ТТ-разложении и оптимальных полиномах, доказана лемма 1 об аппроксимации регулярной функции полиномами на квадрате в комплексной плоскости. В разд. 4 сформулирована и доказана оценка ε -рангов.

2. НЕОБХОДИМЫЕ СВЕДЕНИЯ

Тензоры являются естественным обобщением матриц на многомерный случай $\mathbf{T} \in \mathbb{C}^{n^d}$. Число d называется порядком тензора, а n — его размерностью по соответствующей оси. Для обозначения элемента тензора будут использоваться индексы (i_1, i_2, \dots, i_d) . Индексы тензора можно группировать и преобразовывать, в Matlab это соответствует функции `reshape`. В дальнейшем потребуются понятие развертки тензора по измерению k $T_k \in \mathbb{C}^{(n_1 \dots n_k) \times (n_{k+1} \dots n_d)}$, для элементов которого справедливо

$$T_k(i, j) = \mathbf{T}(i_1, \dots, i_d),$$

$$i = i_1 \cdot n^{k-1} + \dots + i_{k-1} \cdot n + i_k,$$

$$j = i_{k+1} \cdot n^{d-k-1} + \dots + i_{d-1} \cdot n + i_d.$$

Определение 2.1. ТТ-разложением d -мерного тензора $\mathbf{T} \in \mathbb{C}^{n^d}$ называется его представление в виде

$$T(i_1, \dots, i_d) = G_1(i_1)G_2(i_2) \dots G_d(i_d). \quad (1)$$

Трехмерные тензоры G_k имеют размеры $r_{k-1} \times n \times r_k$, кроме случаев $k = 1$ и $k = d$, где они вырождаются в матрицы размеров $n_1 \times r_1$ и $r_{d-1} \times n_d$. Элементы $G_k(i_k)$ являются матрицами, взятыми как срезы трехмерных тензоров. Размеры r_k называются рангами ТТ-разложения.

¹⁾ Работа выполнена при финансовой поддержке РФФ (код проекта No 19-11-00338), Московского центра Фундаментальной и Прикладной Математики в ИВМ РАН (Соглашение с Минобрнауки РФ № 075-15-2022-286).

Ключевым вопросом в численных расчетах является поиск ТТ-аппроксимации с малыми рангами, чтобы получить ускорение и выигрыш в памяти. В литературе существуют примеры оценок рангов для тензоров специфического вида или структуры. Однако, на практике часто предположения о структуре тензора весьма общие и не позволяют провести анализ. Для оценки ранга ТТ-аппроксимации введем понятие *ТТ ϵ -ранга тензора \mathbf{T}*

$$\mathcal{R}_\epsilon(\mathbf{T}) := \min_{\mathbf{S}: \|\mathbf{T}-\mathbf{S}\|_F \leq \epsilon} \max_{k \in \{1, d-1\}} r_k(\mathbf{S}). \tag{2}$$

Другими словами мы смотрим на минимально возможный ранг оптимальной аппроксимации тензора с заданной точностью. Полезным результатом, который будет использоваться далее является теорема об аппроксимации из [5]:

Теорема 2.1. *Для любого тензора $\mathbf{A} = [A(i_1, \dots, i_d)]$ существует ТТ-аппроксимация (1) $\mathbf{B} = [B(i_1, \dots, i_d)]$ с рангами r_k , такими что*

$$\|\mathbf{A} - \mathbf{B}\|_F \leq \sqrt{\sum_{k=1}^{d-1} \epsilon_k^2},$$

где ϵ_k определяется, как минимум

$$\epsilon_k = \min_{\text{rank} B \leq r_k} \|A_k - B\|_F.$$

A_k – это матрица развертки тензора \mathbf{A} :

$$A_k = \mathbf{A}[i_1, \dots, i_k; i_{k+1}, \dots, i_d].$$

В статье будут оцениваться ϵ -ранги (2) тензоров, элементы которых являются значениям некоторой комплекснозначной функции. Пусть в некоторой области \mathbb{C} задана квадратная сетка $N \times N$ с равномерным шагом по каждой оси и функция $f(z)$, которая является регулярной в области, содержащей сетку. Сформируем матрицу $A \in \mathbb{C}^{N \times N}$, где $A(i, j) = f(z_{i,j})$ и $z_{i,j}$ это узел сетки в естественной нумерации (i – номер строки сетки, i – номер узла в строке i). Матрицу A можно тензоризовать, превратив в тензор, пример такой процедуры приведен на фиг. 2. Условием тензоризации матрицы является совпадение размеров $n^d = N$. ТТ-формат полученного тензора мы далее будем называть QTT (Quantized Tensor Train).

3. ОПТИМАЛЬНЫЕ ПОЛИНОМЫ

Определение 3.1. Пусть $C \subset \mathbb{C}$ компакт, монический полином с комплексными коэффициентами $P_{n,C}^*(z)$ будем называть *чебышевским полиномом множества C* , если он удовлетворяет равенству:

$$P_{n,C}^* = \underset{P_n}{\operatorname{argmin}} \max_{z \in C} |P_n(z)|. \tag{3}$$

Здесь минимум берется по всевозможным полиномам $P_n(z) = z^n + a_{n-1}z^{n-1} + \dots + a_0$ степени n с комплексными коэффициентами. Другими словами $P_{n,C}^*$ это полином с единичным старшим коэффициентом, наименее уклоняющийся от нуля на C .

Известным является случай $C = [-1; 1] \subset \mathbb{R}$ и *классические полиномы Чебышёва*:

$$T_n(x) = \cos(n \arccos x), \quad -1 \leq x \leq 1, \tag{4}$$

$$T_n(x) = 2^{1-n} T_n(x). \tag{5}$$

После продолжения в комплексную плоскость формула (4) приобретает вид

$$T_n(z)_{\operatorname{arg}} = \frac{1}{2} \left(z + \sqrt{z^2 - 1} \right)^n + \frac{1}{2} \left(z - \sqrt{z^2 - 1} \right)^n. \tag{6}$$

В формуле (6) выбирается одна из регулярных ветвей $\sqrt{z^2 - 1}$. Полином (5) будет единственным решением задачи (3) при $C = [-1; 1]$, согласно теореме об альтернансе. Нам потребуются некоторые свойства этих полиномов.

Определение 3.2. Чебышёвской сеткой назовем множество корней T_n ,

$$c_j = \cos\left(\frac{\pi}{2n} + \frac{\pi}{n} \cdot j\right) \quad j = 0, 1, \dots, n-1.$$

Определение 3.3. Эллипсами Бернштейна называется семейство эллипсов Γ_ρ , $\rho > 0$, с фокусами в ± 1 , которые получаются после отображения точек окружности $w = e^{i\varphi}$ с помощью функции Жуковского

$$z = \frac{1}{2}\left(w + \frac{1}{w}\right) = \frac{1}{2}\left[\left(\rho + \frac{1}{\rho}\right)\cos\varphi + i\left(\rho - \frac{1}{\rho}\right)\sin\varphi\right].$$

Эллипс Бернштейна Γ_ρ имеет полуоси a , b :

$$a = \frac{1}{2}\left(\rho + \frac{1}{\rho}\right), \quad b = \frac{1}{2}\left(\rho - \frac{1}{\rho}\right), \quad a + b = \rho, \\ a^2 - b^2 = 1.$$

Лемма 3.1. Для любого $z \in \Gamma_\rho$ выполнено

$$|T_n(z)| \geq \frac{1}{2}\left(\rho^n - \frac{1}{\rho^n}\right).$$

Доказательство. Заметим, что для любого $z \in \Gamma_\rho$ выполнено

$$z^2 - 1 = \frac{1}{2}\left(\rho e^{i\varphi} - \frac{1}{\rho} e^{-i\varphi}\right)^2.$$

Тогда, согласно формуле (6) имеем

$$T_n(z) = \frac{1}{2}\left(\rho^n e^{i\varphi} + \frac{1}{\rho^n} e^{-i\varphi}\right).$$

Откуда сразу следует оценка из условия леммы.

Лемма 3.2. Длина контура эллипса Γ_ρ удовлетворяет неравенству $|\Gamma_\rho| \leq \pi(\rho + \rho^{-1})$.

Доказательство. Оценим контурный интеграл:

$$|\Gamma_\rho| = \int_{\Gamma_\rho} |dz| = \int_{S_\rho} \left| \frac{1}{2} \left(1 - \frac{1}{y^2} \right) dy \right| = \int_0^{2\pi} \frac{1}{2} \left| 1 - (\rho e^{i\varphi})^{-2} \right| \rho d\varphi \leq \frac{1}{2} (1 + \rho^{-2}) \rho \cdot 2\pi = \pi \left(\rho + \frac{1}{\rho} \right),$$

где S_ρ – это сфера радиуса ρ с центром в нуле.

Далее под Q_1 будет подразумеваться квадрат в \mathbb{C} с центром в нуле и длиной стороны 2 (см. фиг. 1).

$$Q_1 = \{z = x + iy \mid \max(|x|, |y|) \leq 1\} \subset \mathbb{C}. \tag{7}$$

Лемма 3.3. При $\rho_* = \left(1 + \sqrt{5} + \sqrt{2 + 2\sqrt{5}}\right) / 2$ квадрат Q_1 касается вершинами эллипса Γ_{ρ_*} , см. фиг. 1.

Доказательство. Пусть a , b это полуоси Γ_ρ , точки эллипса можно параметризовать в виде:

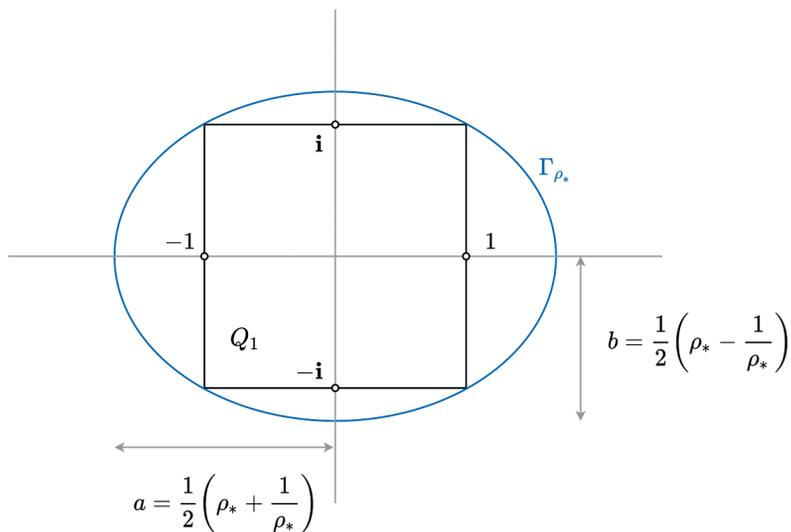
$$x(t) = a(\rho) \cdot \cos t, \\ y(t) = b(\rho) \cdot \sin t, \quad t \in [0; 2\pi].$$

Очевидно, что ρ_* достигается, когда углы квадрата Q_1 принадлежат эллипсу Γ_{ρ_*} . То есть существует $t_* \in [0; 2\pi]$, такое что $x(t_*) = y(t_*) = 1$. Решая систему относительно переменных t_* и ρ_* ,

$$x(t_*) = 1, \\ y(t_*) = 1,$$

получаем

$$\rho_* = \frac{1}{2} \left(1 + \sqrt{5} + \sqrt{2 + 2\sqrt{5}} \right) \approx 2.89.$$



Фиг. 1. Эллипс Γ_{ρ_*} касающийся квадрата Q из леммы 3.3.

Лемма 3.4. Пусть $\rho \geq \rho_{**} = \sqrt{2} + \sqrt{3}$, тогда расстояние между квадратом Q_1 и эллипсом Γ_ρ удовлетворяет неравенству:

$$\text{dist}(Q_1, \Gamma_\rho) \geq \frac{\rho - \rho^{-1}}{2} - \sqrt{2}.$$

Доказательство. Опишем вокруг единичного куба Q_1 окружность диаметром $\sqrt{2}$ с центром в начале координат. Далее грубо оценим расстояние снизу, как разность меньшей полуоси Γ_ρ и описанной вокруг Q_1 окружности,

$$\text{dist}(Q_1, \Gamma_\rho) \geq b(\rho) - \sqrt{2} = \frac{\rho - \rho^{-1}}{2} - \sqrt{2}.$$

Заметим, что чтобы оценка была осмысленной, необходимо требовать неотрицательность правой части неравенства. Легко проверить, что это выполнено при $\rho \geq \rho_{**} = \sqrt{2} + \sqrt{3}$.

В данной работе интерес составляет оценка $|P_{n, \text{int}\Gamma_\rho}^*(z)|$ при $z \in \text{int}\Gamma_\rho$. В [6], [7] исследуются оценки значения (3) для различных компактов. В частности, доказано утверждение леммы 3.1, оно будет использовано далее.

Теорема 3.1 (см. [6]). Пусть $C \subset \mathbb{C}$ это эллипс с фокусами в точках ± 1 и полуосями a, b . Пусть $P_n(z)$ является комплексным полиномом степени n , коэффициент при старшей степени которого равен единице. Тогда выполнено:

$$\max_{z \in C} |P_n(z)| \geq \left(\frac{a+b}{2}\right)^n + \left(\frac{a-b}{2}\right)^n. \tag{8}$$

Причем равенство в (8) достигается лишь в случае $P_n \equiv T_n$ из (5).

Докажем лемму, необходимую для доказательства основной теоремы, отметим, что данная техника использовалась в [1].

Лемма 3.5. Пусть функция $f(z)$ регулярна в эллипсе Γ_ρ при $\rho > \rho_{**} = \sqrt{2} + \sqrt{3}$ (см. 3)), причем $|f(z)| \leq M$ на Γ_ρ . Пусть $P_n(z)$ — полином Лагранжа степени n для $f(z)$ на чебышёвской сетке с $(n+1)$ узлом $\{c_0, \dots, c_n\}$. Тогда для всех точек из квадрата на комплексной плоскости $z \in Q_1 \subset \mathbb{C}$ выполнено следующее:

$$|f(z) - P_n(z)| \leq \frac{M}{2} \cdot \frac{\rho_*^{n+1} + \rho_*^{-n-1}}{\rho_*^{n+1} - \rho_*^{-n-1}} \cdot \frac{\rho + \rho^{-1}}{\frac{\rho - \rho^{-1}}{2} - \sqrt{2}}.$$

Доказательство. Полином Чебышёва $T_{n+1}(z)$ на $[-1;1]$ имеет вид

$$T_{n+1}(z) = \frac{1}{2} \left(z + \sqrt{z^2 - 1} \right)^{n+1} + \frac{1}{2} \left(z - \sqrt{z^2 - 1} \right)^{n+1}. \quad (9)$$

Корнями $T_{n+1}(z)$ являются числа, составляющие чебышёвскую сетку,

$$c_j = \cos \left(\frac{\pi}{2(n+1)} + \frac{\pi}{n+1} \cdot j \right), \quad j = 0, 1, \dots, n.$$

Выберем произвольное $z_* \in Q_1 \setminus \{c_0, \dots, c_n\}$, определим функцию

$$F(z) = \frac{f(z)}{(z - z_*) \cdot \prod_{j=0}^n (z - c_j)}.$$

Заметим, что $F(z)$ имеет полюса первого порядка $\{z_*, c_0, \dots, c_n\}$ в области, ограниченной Γ_ρ . Применим теорему о вычетах для $F(z)$ на Γ_ρ . Тогда

$$\int_{\Gamma_\rho} F(z) dz = 2\pi i \left[\frac{f(z_*)}{\prod_{k=0}^n (z_* - c_k)} + \sum_{i=0}^n \frac{f(c_i)}{(c_i - z_*) \prod_{\substack{j=0 \\ j \neq i}}^n (c_i - c_j)} \right],$$

$$\prod_{k=0}^n (z_* - c_k) \int_{\Gamma_\rho} F(z) dz = 2\pi i \left[f(z_*) - \sum_{i=0}^n \frac{\prod_{k=0}^n (z_* - c_k)}{\prod_{\substack{j=0 \\ j \neq i}}^n (c_i - c_j)} f(c_i) \right]. \quad (10)$$

Заметим, что по определению $\{c_0, \dots, c_n\}$ выполнено

$$T_{n+1}(z) = 2^n \prod_{k=0}^n (z - c_k).$$

Интерполяционный полином Лагранжа функции f в чебышевских узлах $\{c_0, \dots, c_n\}$ имеет вид:

$$P_n(z) = \sum_{i=0}^n \frac{\prod_{k=0}^n (z - c_k)}{\prod_{\substack{j=0 \\ j \neq i}}^n (c_i - c_j)} f(c_i).$$

Тогда уравнение (10) принимает вид

$$\prod_{k=0}^n (z_* - c_k) \int_{\Gamma_\rho} \frac{f(z)}{(z - z_*) \cdot \prod_{j=0}^n (z - c_j)} dz = 2\pi i \left[f(z_*) - \sum_{i=0}^n \frac{\prod_{k=0}^n (z_* - c_k)}{\prod_{\substack{j=0 \\ j \neq i}}^n (c_i - c_j)} f(c_i) \right],$$

$$T_{n+1}(z_*) \cdot \int_{\Gamma_\rho} \frac{f(z)dz}{(z - z_*)T_{n+1}(z)} = 2\pi i [f(z_*) - P_n(z_*)]. \tag{11}$$

Поскольку z_* принадлежит квадрату Q_1 , то z_* принадлежит и внутренности эллипса $\text{int}\Gamma_{\rho_*}$, который касается квадрата согласно лемме 3. Далее применим оценку максимума модуля многочлена на Γ_{ρ_*} , используя теорему 1 из [6] и принцип максимума:

$$|T_{n+1}(z_*)| \leq \max_{z \in Q} |T_{n+1}(z)| \leq \max_{z \in \Gamma_{\rho_*}} |T_{n+1}(z)| = 2^n \cdot \left[\left(\frac{a_* + b_*}{2} \right)^{n+1} + \left(\frac{a_* - b_*}{2} \right)^{n+1} \right].$$

Заметим, что a_* и b_* являются константами и равны соответственно

$$a_* = \frac{1}{2} \left(\rho_* + \frac{1}{\rho_*} \right), \quad b_* = \frac{1}{2} \left(\rho_* - \frac{1}{\rho_*} \right).$$

Используя явно выражение для a_* и b_* , запишем

$$|T_{n+1}(z_*)| \leq \frac{1}{2} \cdot \left[\rho_*^{n+1} + \left(\frac{1}{\rho_*} \right)^{n+1} \right].$$

Из леммы 3.1:

$$|T_{n+1}(z)| \geq \frac{1}{2} \left(\rho^{n+1} - \frac{1}{\rho^{n+1}} \right) \quad z \in \Gamma_\rho.$$

Из леммы 3.2 и 3.4 получим:

$$|\Gamma_\rho| \leq \pi \left(\rho + \frac{1}{\rho} \right), \quad \min_{z_* \in Q} |z - z_*| = \text{dist}(Q, \Gamma_\rho) \geq \frac{\rho - \rho^{-1}}{2} - \sqrt{2} \quad \text{при } z \in \Gamma_\rho.$$

Собирая воедино все полученные выше оценки и подставляя их в уравнение (11), для любого $z_* \in Q \setminus \{c_0, \dots, c_n\}$ имеем

$$\begin{aligned} |f(z_*) - P_n(z_*)| &= \frac{|T_{n+1}(z_*)|}{2\pi} \cdot \left| \int_{\Gamma_\rho} \frac{f(z)dz}{(z - z_*)T_{n+1}(z)} \right| \leq \frac{M}{4\pi} [\rho_*^{n+1} + \rho_*^{-n-1}] \cdot \int_{\Gamma_\rho} \frac{|dz|}{|z - z_*| \cdot |T_{n+1}(z)|} \leq \\ &\leq \frac{M}{4\pi} [\rho_*^{n+1} + \rho_*^{-n-1}] \cdot \frac{2}{\rho^{n+1} - \rho^{-n-1}} \cdot \int_{\Gamma_\rho} \frac{|dz|}{|z - z_*|} \leq \frac{M}{2\pi} \cdot \frac{\rho_*^{n+1} + \rho_*^{-n-1}}{\rho^{n+1} - \rho^{-n-1}} \cdot \frac{|\Gamma_\rho|}{\text{dist}(Q, \Gamma_\rho)} \leq \frac{M}{2} \cdot \frac{\rho_*^{n+1} + \rho_*^{-n-1}}{\rho^{n+1} - \rho^{-n-1}} \cdot \frac{\rho + \rho^{-1}}{\frac{\rho - \rho^{-1}}{2} - \sqrt{2}}, \end{aligned}$$

что и требовалось доказать.

4. QTT ДЛЯ ФУНКЦИЙ ОДНОГОКОМПЛЕКСНОГО ПЕРЕМЕННОГО

Введем равномерную сетку на квадрате $Q_1 \subset \mathbb{C}$, расположим вдоль вещественной оси квадрата m^d точек на равном расстоянии друг от друга, аналогично с мнимой осью, см. фиг. 2. Получаем двумерную сетку с m^{2d} узлами. Далее под обозначением i будет пониматься индекс в тензоре, а под \mathbf{i} — мнимая единица. Все точки, входящие в сетку, можно естественным образом тензоризовать, сопоставив каждой мультииндексу. Узел сетки с мультииндексом $(i_1 j_1, \dots, i_d j_d)$ вычисляется как

$$z(i_1 j_1, \dots, i_d j_d) = -(1 + \mathbf{i}) + \left[\begin{aligned} &(i_1 + \mathbf{i} \cdot j_1) m^{d-1} + \\ &+(i_2 + \mathbf{i} \cdot j_2) m^{d-2} + \dots + (i_d + \mathbf{i} \cdot j_d) \end{aligned} \right] \cdot \frac{2}{m^d}.$$

Заметим, что пара $i_k j_k$ соответствуют горизонтальной и вертикальной части, если отождествлять \mathbb{C} с \mathbb{R}^2 .

Следуя обозначениям работ [8], [1], тензоризацией функции f в узлах равномерной сетки, будем называть тензор $\mathbf{T}_{m,d,Q}(f)$, составленный из значений $f(z(i_1 j_1, \dots, i_d j_d))$. Матрицей развертки тензора $\mathbf{T}_{m,d,Q}(f)$ по оси k будем называть матрицу

$$T_k(i_1 j_1, \dots, i_k j_k; i_{k+1} j_{k+1}, \dots, i_d j_d).$$

Определим открытую область $G = G(l, z_C, L) \subset \mathbb{C}$, см. фиг. 2, как сумму Минковского

$$G = Q(z_C, L) + B(l, 0),$$

где $Q(z_C, L)$ это квадрат, стороны которого параллельны осям \mathbb{C} , левый нижний угол совпадает с z_C , а длина стороны равна L . $B(l, 0)$ это открытый шар с радиусом l и центром в точке 0 .

Теорема 4.1. Пусть на квадрате $Q(z_C, L)$ задана равномерная сетка с m^d отсчетов по каждой оси. Пусть функция $f : G \rightarrow \mathbb{C}$ регулярна в области $G = Q(z_C, L) + B(l, 0)$, где $l > L/m$. Пусть также $|f(z)| \leq M$ при $z \in G$.

Тогда для любого $\varepsilon > 0$, каждой матрицы развертки $T_k \in \mathbb{C}^{m^{2k} \times m^{2(d-k)}}$ тензора $\mathbf{T}_{m,d,Q(z_C,L)}(f)$ существует матрица $B_k \in \mathbb{C}^{m^{2k} \times m^{2(d-k)}}$ такая, что:

$$\|T_k - B_k\|_\infty < \varepsilon, \quad \text{rank}(B_k) \leq s_k + 1, \tag{12}$$

где

$$s_k = \left\lceil \log_{(\rho_k/\rho_{**})} \left(1 + 6 \cdot \frac{M}{\varepsilon} \right) \right\rceil, \quad \rho_k = \left(1 + \frac{2lm^k}{L} \right) + \sqrt{\left(1 + \frac{2lm^k}{L} \right)^2 - 1}. \tag{13}$$

$\rho_{**} = \sqrt{2} + \sqrt{3}$ — константа.

Доказательство. Зафиксируем некоторое $k \in \overline{1, d}$. Определим мультииндексы

$$\begin{aligned} i &= i_1 m^{k-1} + \dots + i_k \in [0; m^k - 1] & \hat{i} &= i_{k+1} m^{d-k-1} + \dots + i_d \in [0; m^{d-k} - 1], \\ j &= j_1 m^{k-1} + \dots + j_k \in [0; m^k - 1] & \hat{j} &= j_{k+1} m^{d-k-1} + \dots + j_d \in [0; m^{d-k} - 1]. \end{aligned}$$

Паре (i, j) можно сопоставить число $i + \mathbf{i} \cdot j$, тогда данный комплексный вектор будет задавать “смещение” в квадрате $Q(z_C, L)$, указывая на подквадрат $Q_{i,j}$ (см. фиг. 2). Вектор $\hat{i} + \mathbf{i} \cdot \hat{j}$, задает смещение в выбранном подквадрате $Q_{i,j}$. Рисунок демонстрирует геометрический смысл матрицы развертки T_k тензора $\mathbf{T}_{m,d,Q(z_C,L)}$: в каждой строке матрицы T_k расположены все значения функции f на узлах сетки соответствующего подквадрата $Q_{i,j}$. Заметим, что длина стороны подквадрата равна $h = L \cdot m^{-k}$.

Построим взаимнооднозначное соответствие между подквадратом $z \in Q_{i,j}$ и единичным квадратом $w \in Q$ из (7):

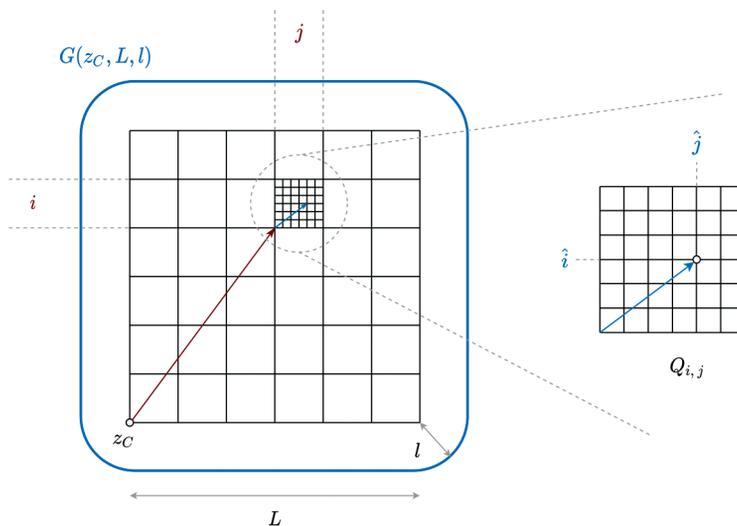
$$z(w) = \beta w + \alpha_{i,j} = \underbrace{\frac{h}{2} \cdot w}_{\text{смещение в } Q_{i,j}} + \underbrace{\left[z_C + (i + \mathbf{i} \cdot j)h + (1 + \mathbf{i})\frac{h}{2} \right]}_{\text{центр квадрата } Q_{i,j}}.$$

На $Q_{i,j}$ определим функцию

$$\begin{aligned} g_{(i,j)}(w) &= f(z(w)) = f(\beta w + \alpha_{i,j}), \\ w &\in B(R_k, 0) \subset \mathbb{C}, \end{aligned} \tag{14}$$

где $B(R, c)$ обозначает шар радиуса R_k с центром в точке $c \in \mathbb{C}$ и

$$R_k = \frac{h/2 + l}{h/2} = \frac{\frac{L}{2m^k} + l}{\frac{L}{2m^k}} = 1 + \frac{2m^k}{L} \cdot l.$$



Фиг. 2. Индекс (i, j) “указывает” на подквадрат $Q_{i,j}$ в тензоризованной сетке. Индекс (\hat{i}, \hat{j}) “указывает” на узел сетке в выбранном подквадрате.

$g_{(i,j)}$ корректно определена, т.е. при всех $w \in B(R_k, 0)$ выполнено $z(w) \in G$. Действительно, в определении R_k заложена половина стороны подквадрата $Q_{i,j}$ и расстояние l до границы области G . Поэтому при отображении $w \mapsto z$ число z останется в G . Геометрически $hR_k / 2$ это максимальный радиус окружности которую можно обвести вокруг центра любого подквадрата (i, j) , т.ч. она не вышла за пределы области G , где регулярна функция f .

Впишем в шар $B(R_k, 0)$ такой эллипс Бернштейна Γ_{ρ_k} , чтобы его большая полуось равнялась R_k ,

$$\frac{\rho_k + \rho_k^{-1}}{2} = R_k, \quad \rho_k = R_k + \sqrt{R_k^2 - 1}. \tag{15}$$

Заметим, что для любого k выполнено $\rho_k > \rho_{**}$. Тогда для $g_{(i,j)}(w)$ при $w \in Q_1$ выполнены все условия леммы 3.5. Действительно, поскольку линейное преобразование сохраняет регулярность, то $g_{(i,j)}(w)$ регулярна на $w \in \text{int}\Gamma_{\rho_k}$ и ограничена константой M . Значит, существует такой полином степени s_k , что для любого $w \in Q_1$ выполнено

$$\left| g_{(i,j)}(w) - \hat{P}_{s_k, (i,j)}(w) \right| \leq \frac{M}{2} \cdot \frac{[\rho_k^{s_k} + \rho_k^{-s_k}]}{\rho_k^{s_k} - \rho_k^{-s_k}} \cdot \frac{\rho_k + \rho_k^{-1}}{\frac{\rho_k - \rho_k^{-1}}{2} - \sqrt{2}}. \tag{16}$$

Оценим слагаемые в выражении (16). Поскольку $l > L / (2m)$, то

$$R_k = 1 + \frac{2m^k}{L} \cdot l > 1 + \frac{2m^k}{L} \cdot \frac{L}{m} = 1 + 2m^{k-1} > 3.$$

Поскольку $k \geq 1$, то выполнено $R_k > 2$. Преобразуем последний множитель в (16) с учетом равенства (15):

$$\begin{aligned} \frac{\rho_k - \rho_k^{-1}}{2} - \sqrt{2} &= \frac{(R_k + \sqrt{R_k^2 - 1})^2 - 1}{2(R_k + \sqrt{R_k^2 - 1})} - \sqrt{2} = \frac{2(R_k^2 + 2R_k\sqrt{R_k^2 - 1} - 2)}{2(R_k + \sqrt{R_k^2 - 1})} - \sqrt{2} = \\ &= R_k - \sqrt{2} - \frac{1}{R_k + \sqrt{R_k^2 - 1}} \geq R_k - \sqrt{2} - \frac{1}{2R_k}. \end{aligned} \tag{17}$$

С учетом $\rho_k + \rho_k^{-1} = 2R_k$ получаем оценку

$$\frac{\rho_k + \rho_k^{-1}}{\frac{\rho_k - \rho_k^{-1}}{2} - \sqrt{2}} \leq \frac{2R_k}{R_k - \sqrt{2} - \frac{1}{2R_k}} = q(R_k) < q(3) = \frac{6}{3 - \sqrt{2} + 1/6} \approx 3.42 < 4. \quad (18)$$

Функция $q(R_k)$ в правой части (18) строго убывает при $R_k > 2$, можем оценить сверху величиной $q(2)$. Теперь оценим первый множитель в правой части (16) с учетом вида степени полинома s_k из условия теоремы

$$\frac{\rho_*^{s_k} + \rho_*^{-s_k}}{\rho_*^{s_k} - \rho_*^{-s_k}} \leq \frac{\rho_*^{s_k} + 1}{\rho_*^{s_k} - 1} \leq \frac{1 + \left(\frac{1}{\rho_*}\right)^{s_k}}{\left(\frac{\rho_k}{\rho_*}\right)^{s_k} - 1} \leq \frac{2}{\left(\frac{\rho_k}{\rho_*}\right)^{s_k} - 1} = \frac{2}{1 + \frac{6M}{\varepsilon} - 1} \leq \frac{\varepsilon}{3M}. \quad (19)$$

Комбинируя результаты (18) и (19) в (16) получаем

$$\left|g_{(i,j)}(w) - \widehat{P}_{s_k,(i,j)}(w)\right| \leq \frac{M}{2} \cdot 4 \cdot \frac{\varepsilon}{3M} < \varepsilon. \quad (20)$$

Поддействуем отображением $w \leftrightarrow z = \beta w + \alpha_{(i,j)}$, тогда для любого $z \in Q_{i,j}$ выполнено

$$\left|f(z) - P_{s_k,(i,j)}(z)\right| < \varepsilon, \quad (21)$$

где

$$P_{s_k,(i,j)}(z) = \widehat{P}_{s_k,(i,j)}(w(z)) = \widehat{P}_{s_k,(i,j)}\left(\frac{1}{\beta}(z - \alpha_{i,j})\right).$$

На основе полиномов для каждого подквадрата Q_{ij} построим матрицу B_k и оценим ее ранг. Как говорилось ранее, каждая строка матрицы развертки $(T_k)_{(i,j)}$, где $i, j \in 0, m^k - 1$ суть значения f в узлах сетки подквадрата Q_{ij} . Для удобства обозначений будем отождествлять мультииндекс (i, j) с индексом v , который пробегает значения от 1 до $V = m^{2k}$, аналогично (\hat{i}, \hat{j}) с индексом b от 1 до $B = m^{2(d-k)}$. В новых обозначениях можно считать $\widehat{P}_{s_k,(i,j)} \equiv \widehat{P}_{s_k,v}$. Пусть w_1, \dots, w_B это все узлы сетки $m^{d-k} \times m^{d-k}$ на единичном квадрате Q_1 . Составим матрицу W :

$$W = \begin{bmatrix} 1 & w_1 & w_1^2 & \dots & w_1^{s_k} \\ 1 & w_2 & w_2^2 & \dots & w_2^{s_k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & w_B & w_B^2 & \dots & w_B^{s_k} \end{bmatrix} \in \mathbb{C}^{m^{2(d-k)} \times (s_k+1)},$$

$$\widehat{P}_{s_k,(i,j)}(w) \equiv \widehat{P}_{s_k,v}(w) = a_{s_k,v} w^{s_k} + a_{s_k-1,v} w^{s_k-1} + \dots + a_{s_k-1,v} w + a_{s_k-1,v}.$$

Тогда значения полинома $\widehat{P}_{s_k,v}$ в узлах сетки выражаются как

$$\begin{bmatrix} \widehat{P}_{s_k,v}(w_1) \\ \widehat{P}_{s_k,v}(w_2) \\ \vdots \\ \widehat{P}_{s_k,v}(w_B) \end{bmatrix} = \begin{bmatrix} 1 & w_1 & w_1^2 & \dots & w_1^{s_k} \\ 1 & w_2 & w_2^2 & \dots & w_2^{s_k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & w_B & w_B^2 & \dots & w_B^{s_k} \end{bmatrix} \cdot \begin{bmatrix} a_{0,v} \\ a_{1,v} \\ \vdots \\ a_{s_k,v} \end{bmatrix} = W \cdot \mathbf{a}_v \in \mathbb{C}^{m^{2(d-k)}}.$$

Значения всевозможных полиномов $\widehat{P}_{s_k, \nu}$ при разных ν в узлах w_b на Q_1 представимы в виде

$$\widehat{P}_k = \begin{bmatrix} \widehat{P}_{s_k,1}(w_1) & \widehat{P}_{s_k,2}(w_1) & \dots & \widehat{P}_{s_k,V}(w_1) \\ \widehat{P}_{s_k,1}(w_2) & \widehat{P}_{s_k,2}(w_2) & \dots & \widehat{P}_{s_k,V}(w_2) \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{P}_{s_k,1}(w_B) & \widehat{P}_{s_k,2}(w_B) & \dots & \widehat{P}_{s_k,V}(w_B) \end{bmatrix} = W \cdot \begin{bmatrix} a_{0,1} & a_{0,2} & \dots & a_{0,V} \\ a_{1,1} & a_{1,2} & \dots & a_{1,V} \\ \vdots & \vdots & \ddots & \vdots \\ a_{s_k,1} & a_{s_k,2} & \dots & a_{s_k,V} \end{bmatrix} = W \cdot A. \tag{22}$$

Где матрица $A \in \mathbb{C}^{(s_k+1) \times m^{2(d-k)}}$ составлена из коэффициентов полинома для каждого индекса $\nu = (i, j)$. Матрица $\widehat{P}_k \in \mathbb{C}^{m^{2(d-k)} \times m^{2k}}$ содержит значения всех полиномов на всех узлах сетки $m^{d-k} \times m^{d-k}$ на Q_1 . По построению, $\widehat{P}_{s_k,(i,j)}(z) = \widehat{P}_{s_k,(i,j)}(w(z))$, поэтому матрица $P_k \in \mathbb{C}^{m^{2(d-k)} \times m^{2k}}$, состоящая из значений всевозможных полиномов $P_{s_k, \nu}(z)$ при различных ν на подквадратах Q_{ij} будет тождественно равна \widehat{P}_k .

Положим $B_k = P_k$, тогда согласно (20) $\|T_k - B_k\|_\infty < \varepsilon$, при этом согласно (22) каждая строка B_k принадлежит линейной оболочке столбцов W , т.е. строчный ранг B_k не превышает $s_k + 1$, а значит, $\text{rank}(B_k) \leq s_k + 1$.

Следствие 4.1. Пусть функция $f : G \rightarrow \mathbb{C}$ регулярна в области $G = Q(z_C, L) + B(0, l)$, где $l > L / m$ и ограничена $|f(z)| \leq M$ при $z \in G$. Тогда для любого $\widehat{\varepsilon} > 0$ выполнено:

$$\mathcal{R}_{\widehat{\varepsilon}}(\mathbf{T}_{m,d,Q(z_C,L)}(f)) \leq \widehat{s} + 1,$$

где

$$\widehat{s} = \left\lceil \log_v \left(1 + 6 \cdot \frac{M m^d \sqrt{d-1}}{\widehat{\varepsilon}} \right) \right\rceil, \quad v = \frac{3 + 2\sqrt{2}}{\sqrt{2} + \sqrt{3}} \approx 1.85.$$

Доказательство. В обозначениях теоремы 4.1, зафиксируем

$$\varepsilon = \frac{\widehat{\varepsilon}}{m^d \sqrt{d-1}}.$$

Заменим в оценке все ρ_k на минимальный из них, т.е. ρ_1 :

$$\rho_1 = \left(1 + \frac{2lm}{L} \right) + \sqrt{\left(1 + \frac{2lm}{L} \right)^2 - 1}.$$

Поскольку $l > L / m$, то $\rho_1 > 3 + 2\sqrt{2}$, совмещая это с явным видом $\rho_{**} = \sqrt{2} + \sqrt{3}$ получаем нижнюю оценку на ρ_k / ρ_{**}

$$\frac{\rho_k}{\rho_{**}} \geq \frac{\rho_1}{\rho_{**}} > \frac{3 + 2\sqrt{2}}{\sqrt{2} + \sqrt{3}} = v \approx 1.85.$$

Тогда ранги из теоремы 4.1 имеют вид

$$s_k = \left\lceil \log_{(\rho_k / \rho_{**})} \left(1 + 6 \cdot \frac{M}{\varepsilon} \right) \right\rceil \leq \left\lceil \log_v \left(1 + 6 \cdot \frac{M}{\varepsilon} \right) \right\rceil = \widehat{s}.$$

Для матрицы развертки T_k тензора $\mathbf{T}_{m,d,Q(z_C,L)}(f)$ имеем:

$$\begin{aligned} \|T_k - B_k\|_\infty &\leq \varepsilon, \\ \|T_k - B_k\|_F &\leq \varepsilon \sqrt{m^d \cdot m^d} = \varepsilon m^d = \frac{\widehat{\varepsilon}}{\sqrt{d-1}}. \end{aligned}$$

Теорема 2.1 (см. [5]) гарантирует существование такого тензорного поезда \mathbf{B} такого, что

$$\left\| \mathbf{T}_{m,d,Q(z_C,L)}(f) - \mathbf{B} \right\|_F \leq \sqrt{\sum_{k=1}^{d-1} \left(\frac{\hat{\varepsilon}}{\sqrt{d-1}} \right)^2} \ll \hat{\varepsilon}.$$

Сравнивая результаты 4.1 с соответствующими в [1], видно, что при похожих предположениях для функции f ее QTT-разложение на двумерной сетке в \mathbb{C} асимптотически будет иметь такие же ранги. Основное отличие в формулировках теорем это область регулярности функции f , в данной работе было показано, что чем “больше” область G (параметр l), тем меньше будут ранги матрицы развертки при той же точности аппроксимации. Применить технику доказательства из [1] во многом помог результат из [6], который позволил оценить значение T_{n+1} на квадрате Q_1 . При больших $l \gg L$, т.е. когда функция f регулярна в намного более обширной области, чем квадрат $Q(z_C, L)$ оценка TT-рангов при фиксированном ε имеет вид

$$\text{rank}(B_k) \sim \mathcal{O}\left(\frac{M(l)}{\ln l}\right).$$

Похожие оценки возможно получить для произвольных областей G , нежели рассмотренный тип в теореме 4.1. В будущем цель авторов получить оценки для случая многомерной сетки в \mathbb{C}^n .

СПИСОК ЛИТЕРАТУРЫ

1. *Высоцкий Л.* О TT-рангах приближенных тензоризаций некоторых гладких функций // Ж. вычисл. матем. и матем. физ. 2021. V. 61. N. 5. P. 750–760.
2. *Vysotsky L., Rakhuba M.* Tensor rank bounds and explicit qtt representations for the inverses of circulant matrices // Numerical Linear Algebra with Applications. 2023. V. 30. N. 3. e2461.
3. *Chertkov A., Oseledets I., Rakhuba M.* Robust discretization in quantized tensor train format for elliptic problems in two dimensions // arXiv: Numerical Analysis. 2016.
4. *Kazeev V., Khoromskij B.* Low-rank explicit qtt representation of the laplace operator and its inverse // SIAM Journal on Matrix Analysis and Applications. 2012. V. 33. N. 3. P. 742–758.
5. *Oseledets I., Tyrtyshnikov E.* TT-cross approximation for multidimensional arrays // Linear Algebra and its Applications. 2010. V. 432. N. 1 P. 70–88.
6. *Munch N. J.* chebyshev theorem for ellipses in the complex plane // The American Mathematical Monthly. 2019. V. 126. N. 5. P. 430–436.
7. *Ransford T.* Capacity // Cambridge University Press. 1995. P. 127–160.
8. *Vysotsky L.* On tensor-train ranks of tensorized polynomials. Large-Scale Scientific Computing: 12th International Conference, LSSC2019, Sozopol, Bulgaria. P. 189–196.

УДК 519.63

УЛУЧШЕННАЯ КВАДРАТУРНАЯ ФОРМУЛА ДЛЯ ПРЯМОГО ЗНАЧЕНИЯ НОРМАЛЬНОЙ ПРОИЗВОДНОЙ ПОТЕНЦИАЛА ПРОСТОГО СЛОЯ¹⁾

© 2024 г. П. А. Крутицкий^{1,*}, И. О. Резниченко^{1,**}

¹ 125047 Москва, Миусская пл., 4, ИПМ им. М. В. Келдыша РАН, Россия

*e-mail: biem@mail.ru

**e-mail: io.reznichenko@physics.msu.ru

Поступила в редакцию 04.09.2023 г.

Переработанный вариант 04.09.2023 г.

Принята к публикации 15.10.2023 г.

Рассматривается потенциал простого слоя для уравнения Гельмгольца в трехмерном случае, а также потенциал простого слоя для уравнения Лапласа. Получена квадратурная формула для прямого значения нормальной производной потенциала простого слоя с непрерывной плотностью, заданной на замкнутой либо разомкнутой поверхности. Квадратурная формула, предложенная в работе, дает значительно более высокую точность, чем известные формулы, что подтверждается численными тестами. Полученная квадратурная формула может использоваться при численном решении краевых задач для уравнений Лапласа и Гельмгольца методом граничных интегральных уравнений. Библ. 17. Табл. 1.

Ключевые слова: потенциал простого слоя, нормальная производная, квадратурная формула.

DOI: 10.31857/S0044466924020021, **EDN:** YKRLKB

1. ВВЕДЕНИЕ

При численном решении краевых задач для уравнений Лапласа и Гельмгольца с использованием потенциала простого слоя необходимо иметь квадратурную формулу, обеспечивающую высокую точность вычисления потенциала в любой точке области, где решается краевая задача. Кроме того, надо иметь квадратурную формулу для прямого значения нормальной производной потенциала, которая используется при численном решении соответствующих граничных интегральных уравнений.

Стандартные квадратурные формулы для потенциала простого слоя (см. [1], гл. 2), основанные на численном интегрировании, не обеспечивают равномерной аппроксимации потенциала вблизи поверхности Γ , где задана плотность потенциала. Эти формулы стремятся к бесконечности, когда точка, в которой вычисляется квадратурная формула, приближается к Γ . При этом сам потенциал простого слоя непрерывен во всем пространстве, в том числе во всех точках на поверхности Γ . Таким образом, стандартные квадратурные формулы не сохраняют важнейшее свойство потенциала, а именно, его ограниченность и непрерывность при переходе через Γ . В [2], [3] предложены квадратурные формулы, которые сохраняют указанное свойство потенциала простого слоя и обеспечивают его равномерную аппроксимацию вблизи Γ .

Квадратурные формулы для вычисления нормальной производной потенциала простого слоя для уравнений Лапласа и Гельмгольца, применяющие численное интегрирование, предлагаются, например, в ([1], гл. 2). Квадратурная формула для нормальной производной потенциала простого слоя, основанная на аналитическом вычислении интегралов и обладающая более высокой точностью, предложена в [4]. В настоящей работе выводится новая квадратурная формула, улучшающая формулу из [4] за счет более точного учета элементов подынтегральной функции при аналитическом вычислении интегралов. Новая формула обеспечивает значительно более низкую погрешность, чем формула из [4], что подтверждается численными тестами.

В двумерном случае квадратурные формулы для потенциалов с плотностью, заданной на разомкнутых кривых и имеющей степенные особенности на концах кривых, предложены в [5]. Эти формулы могут

¹⁾ Работа выполнена при финансовой поддержке Резниченко И.О. Московским центром фундаментальной и прикладной математики, соглашение с Минобрнауки РФ (код проекта 075-15-2022-283).

применяться при численном решении краевых задач на плоскости, когда граница включает разомкнутые кривые. Такие задачи изучались в [6]–[11].

2. ПОСТАНОВКА ЗАДАЧИ

Введем в пространстве декартову систему координат $x = (x_1, x_2, x_3) \in R^3$. Пусть Γ — простая гладкая замкнутая поверхность класса C^2 , ограничивающая объемно-односвязную внутреннюю область, либо простая гладкая ограниченная разомкнутая поверхность класса C^2 , содержащая свои предельные точки (см. [12], гл. 14, § 1). Предположим, что поверхность Γ параметризована так, что на нее отображается прямоугольник:

$$\begin{aligned} y &= (y_1, y_2, y_3) \in \Gamma, \quad y_1 = y_1(u, v), \quad y_2 = y_2(u, v), \quad y_3 = y_3(u, v); \\ u &\in [0, A], \quad v \in [0, B]; \\ y_j(u, v) &\in C^2([0, A] \times [0, B]), \quad j = 1, 2, 3. \end{aligned} \quad (1)$$

Потребуем также, чтобы различным внутренним точкам прямоугольника при указанном отображении соответствовали различные точки поверхности. Сферу, поверхность эллипсоида, гладкие поверхности фигур вращения, поверхность тора и многие другие более сложные поверхности можно параметризовать таким образом. Кроме того, сложные поверхности можно разбить на несколько частей и для каждой части ввести свою параметризацию, тогда дальнейшие рассуждения справедливы для каждой такой части.

Введем N точек u_n с шагом h на отрезке $[0, A]$ и M точек v_m с шагом H на отрезке $[0, B]$ и рассмотрим разбиение прямоугольника $[0, A] \times [0, B]$, который отображается на поверхность Γ :

$$\begin{aligned} A &= Nh, \quad B = MH, \quad u_n = (n + 1/2)h, \\ n &= 0, \dots, N - 1; \quad v_m = (m + 1/2)H, \\ m &= 0, \dots, M - 1. \end{aligned}$$

Тем самым, прямоугольник $[0, A] \times [0, B]$ разбивается на $N \times M$ маленьких прямоугольничков, и через (u_n, v_m) обозначены серединки этих прямоугольничков.

Известно (см. [12], гл. 14, § 1), что компоненты вектора нормали (не единичного) $\eta(y) = (\eta_1(y), \eta_2(y), \eta_3(y))$ в точке поверхности $y = (y_1, y_2, y_3) \in \Gamma$ выражаются через определители второго порядка формулами

$$\eta_1 = \begin{vmatrix} (y_2)_u & (y_3)_u \\ (y_2)_v & (y_3)_v \end{vmatrix}, \quad \eta_2 = \begin{vmatrix} (y_3)_u & (y_1)_u \\ (y_3)_v & (y_1)_v \end{vmatrix}, \quad \eta_3 = \begin{vmatrix} (y_1)_u & (y_2)_u \\ (y_1)_v & (y_2)_v \end{vmatrix}. \quad (2)$$

Положим $|\eta(y)| = \sqrt{(\eta_1(y))^2 + (\eta_2(y))^2 + (\eta_3(y))^2}$. Известно (см. [12], гл. 14, § 1–2), что

$$\int_{\Gamma} F(y) ds_y = \int_0^A du \int_0^B dv F(y(u, v)) |\eta(y(u, v))|.$$

Заметим, что если $|\eta(y(u, v))| = 0$ в некоторой точке, то функция $|\eta(y(u, v))|$ может быть недифференцируемой в этой точке. Поэтому дополнительно потребуем, чтобы

$$|\eta(y(u, v))| \in C^1([0, A] \times [0, B]). \quad (3)$$

Кроме того, потребуем, чтобы

$$|\eta(y(u, v))| > 0 \quad \forall (u, v) \in ((0, A) \times (0, B)). \quad (4)$$

Из условия (4) следует, что $|\eta(y(u, v))| \in C^1((0, A) \times (0, B))$, но условие (3) не следует.

Потенциал простого слоя для уравнения Гельмгольца с заданной на поверхности Γ плотностью $\mu(y) \in C^0(\Gamma)$ имеет вид

$$\mathcal{V}_k[\mu](x) = \frac{1}{4\pi} \int_{\Gamma} \frac{\mu(y) e^{ik|x-y|}}{|x-y|} ds_y = \frac{1}{4\pi} \int_0^A du \int_0^B dv \frac{\mu(y(u, v)) \exp(ik|x-y(u, v)|)}{|x-y(u, v)|} |\eta(y(u, v))|, \quad (5)$$

где $|x - y(u, v)| = \sqrt{(x_1 - y_1(u, v))^2 + (x_2 - y_2(u, v))^2 + (x_3 - y_3(u, v))^2}$, и для простоты константа $k \geq 0$; если же $k = 0$, то потенциал $\mathcal{V}_k[\mu](x)$ переходит в гармонический потенциал $\mathcal{V}_0[\mu](x)$ для уравнения Лапласа.

Пусть $\mathbf{n}_x = \eta(x) / |\eta(x)|$ — вектор единичной нормали в точке $x \in \Gamma$, тогда

$$\begin{aligned} \frac{\partial}{\partial \mathbf{n}_x} \frac{e^{ik|x-y|}}{|x-y|} &= |\eta(x)|^{-1} (\eta(x), \nabla_x) \frac{e^{ik|x-y|}}{|x-y|} = \\ &= \frac{1}{|\eta(x)|} \frac{\exp(ik|x-y|)(ik|x-y|-1)}{|x-y|^2} \sum_{j=1}^3 \frac{\eta_j(x)(x_j - y_j)}{|x-y|}. \end{aligned}$$

Прямое значение нормальной производной потенциала простого слоя $\mathcal{V}_k[\mu](x)$ в точке x имеет вид

$$\begin{aligned} \frac{\partial \mathcal{V}_k[\mu](x)}{\partial \mathbf{n}_x} &= \frac{1}{4\pi} \int_{\Gamma} \mu(y) \frac{\partial}{\partial \mathbf{n}_x} \frac{e^{ik|x-y|}}{|x-y|} ds_y = \frac{1}{4\pi} \int_{\Gamma} \mu(y) \frac{1}{|\eta(x)|} \frac{\exp(ik|x-y|)(ik|x-y|-1)}{|x-y|^2} \times \\ &\times \sum_{j=1}^3 \frac{\eta_j(x)(x_j - y_j)}{|x-y|} ds_y = \frac{1}{4\pi|\eta(x)|} \times \int_0^A du \int_0^B dv \mu(y(u, v)) |\eta(y(u, v))| \times \\ &\times \exp(ik|x-y(u, v)|) \times (ik|x-y(u, v)|-1) \times \sum_{j=1}^3 \frac{\eta_j(x)(x_j - y_j(u, v))}{|x-y(u, v)|^3} = \frac{1}{4\pi|\eta(x)|} \times \\ &\times \sum_{n=0}^{N-1M-1} \sum_{m=0}^{N-1M-1} \int_{u_n-h/2}^{u_n+h/2} du \int_{v_m-H/2}^{v_m+H/2} dv \mu(y(u, v)) \times |\eta(y(u, v))| \times \exp(ik|x-y(u, v)|) \times \\ &\times (ik|x-y(u, v)|-1) \times \sum_{j=1}^3 \frac{\eta_j(x)(x_j - y_j(u, v))}{|x-y(u, v)|^3}. \end{aligned} \tag{6}$$

Известно (см. [13], § 27.5), что прямое значение нормальной производной потенциала простого слоя в наших предположениях является непрерывной на Γ функцией. Пусть $\mu_{nm} = \mu(y(u_n, v_m))$, тогда

$$\mu(y(u, v)) = \mu_{nm} + o(1)$$

для $u \in [u_n - h/2, u_n + h/2]$ и $v \in [v_m - H/2, v_m + H/2]$.

Пусть x лежит в одном из узлов на Γ , т.е. $x = y(u_n^{\wedge}, v_m^{\wedge}) \in \Gamma$. Так же, как и в [2], можно показать, что при $u \in [u_n - h/2, u_n + h/2]$ и $v \in [v_m - H/2, v_m + H/2]$

$$|x - y(u, v)| = |x - y(u_n, v_m)| + O(h + H), \quad \exp(ik|x - y(u, v)|) = \exp(ik|x - y(u_n, v_m)|) + O(h + H).$$

Константы в оценках функций, обозначенных как $O(h + H)$, не зависят от n, m и от расположения x в узлах Γ .

Следовательно, если точка x лежит в узле $y(u_n^{\wedge}, v_m^{\wedge})$ на Γ , то прямое значение нормальной производной потенциала простого слоя можно записать в виде

$$\begin{aligned} \left. \frac{\partial \mathcal{V}_k[\mu](x)}{\partial \mathbf{n}_x} \right|_{x=y(u_n^{\wedge}, v_m^{\wedge}) \in \Gamma} &\approx \frac{1}{4\pi|\eta(x)|} \sum_{n=0}^{N-1M-1} \sum_{m=0}^{N-1M-1} \mu_{nm} \times \exp(ik|x - y(u_n, v_m)|) (1 - ik|x - y(u_n, v_m)|) \times \\ &\times \int_{u_n-h/2}^{u_n+h/2} du \int_{v_m-H/2}^{v_m+H/2} dv |\eta(y(u, v))| \times \sum_{j=1}^3 \frac{\eta_j(x)(y_j(u, v) - x_j)}{|x - y(u, v)|^3}. \end{aligned} \tag{7}$$

Двойной интеграл в (7) будем называть *каноническим* и при его вычислении будем различать 2 случая. В первом случае точка x лежит в области интегрирования, а во втором случае — нет.

3. ВЫЧИСЛЕНИЕ КАНОНИЧЕСКОГО ИНТЕГРАЛА, КОГДА ТОЧКА x ЛЕЖИТ В ОБЛАСТИ ИНТЕГРИРОВАНИЯ

В данном случае интегрирование ведется по прямоугольничку с центром в точке (u_n, v_m) , которой отвечает точка $y(u_n, v_m) = x$ на поверхности Γ . Применяя формулу Тейлора с центром в точке (u_n, v_m) , находим

$$\begin{aligned} |y(u, v) - x|^2 &= |y(u, v) - y(u_n, v_m)|^2 \approx \sum_{j=1}^3 \left((y_j)'_u (u - u_n) + (y_j)'_v (v - v_m) \right)^2 = \\ &= \sum_{j=1}^3 \left(\left((y_j)'_u \right)^2 (u - u_n)^2 + \left((y_j)'_v \right)^2 (v - v_m)^2 + \right. \\ &\quad \left. + 2(y_j)'_u (y_j)'_v (u - u_n)(v - v_m) \right) = \alpha^2 (u - u_n)^2 + \beta^2 (v - v_m)^2 + 2\delta (u - u_n)(v - v_m), \\ \alpha^2 &= \sum_{j=1}^3 \left((y_j)'_u \right)^2, \quad \beta^2 = \sum_{j=1}^3 \left((y_j)'_v \right)^2, \quad \delta = \sum_{j=1}^3 (y_j)'_u (y_j)'_v, \end{aligned}$$

где $(y_j)'_u$ и $(y_j)'_v$ берутся в точке (u_n, v_m) .

Заметим, что $\alpha^2 \beta^2 - \delta^2 = |\eta(y)|^2$, согласно ([12], гл. 14, § 1), поэтому $\alpha^2 > 0$ и $\beta^2 > 0$ в силу условия (4).

С учетом (3) для всех возможных n, m , при $u \in [u_n - h/2, u_n + h/2]$ и $v \in [v_m - H/2, v_m + H/2]$ функция $|\eta(y(u, v))|$ может быть разложена по формуле Тейлора с остаточным членом 1-го порядка

$$|\eta(y(u, v))| = |\eta(y(u_n, v_m))| + O(h + H). \quad (8)$$

В силу (8) канонический интеграл из (7) можно записать в виде

$$\begin{aligned} &\int_{u_n - h/2}^{u_n + h/2} du \int_{v_m - H/2}^{v_m + H/2} dv |\eta(y(u, v))| \times \sum_{j=1}^3 \frac{\eta_j(x)(y_j(u, v) - x_j)}{|x - y(u, v)|^3} \approx \\ &\approx |\eta(y(u_n, v_m))| \int_{u_n - h/2}^{u_n + h/2} du \int_{v_m - H/2}^{v_m + H/2} dv \times \sum_{j=1}^3 \frac{\eta_j(x)(y_j(u, v) - x_j)}{|x - y(u, v)|^3}. \end{aligned} \quad (9)$$

Далее, используя формулу Тейлора в точке (u_n, v_m) с остаточным членом в форме Пеано (см. [12], гл. 10, § 5.3), получаем

$$\begin{aligned} y_j - x_j &= (y_j)'_u (u - u_n) + (y_j)'_v (v - v_m) + \frac{1}{2} (y_j)''_{uu} (u - u_n)^2 + \frac{1}{2} (y_j)''_{vv} (v - v_m)^2 + (y_j)''_{uv} \times \\ &\quad \times (u - u_n)(v - v_m) + o\left((u - u_n)^2 + (v - v_m)^2\right), \end{aligned}$$

производные по u и v берутся в точке (u_n, v_m) .

Легко проверить (см. [12], гл. 14, § 1.2), что

$$\sum_{j=1}^3 \eta_j(x) (y_j)'_u = \sum_{j=1}^3 \eta_j(x) (y_j)'_v = 0,$$

следовательно,

$$\sum_{j=1}^3 \eta_j(x) (y_j - x_j) \approx \xi_1 (u - u_n)^2 + \xi_2 (v - v_m)^2 + \xi_3 (u - u_n)(v - v_m),$$

$$\xi_1 = \frac{1}{2} \sum_{j=1}^3 \eta_j(x)(y_j)''_{uu}, \quad \xi_2 = \frac{1}{2} \sum_{j=1}^3 \eta_j(x)(y_j)''_{vv}, \quad \xi_3 = \sum_{j=1}^3 \eta_j(x)(y_j)''_{uv},$$

производные по u и v берутся в точке (u_n, v_m) .

Из приведенных соотношений вытекает, что в рассматриваемом случае двойной интеграл в (9) приближенно равен следующему интегралу, который обозначим через \mathcal{J}_{nm} :

$$\int_{u_n-h/2}^{u_n+h/2} du \int_{v_m-H/2}^{v_m+H/2} dv \sum_{j=1}^3 \frac{\eta_j(x)(y_j - x_j)}{|x - y|^3} \approx \int_{-h/2}^{h/2} dU \int_{-H/2}^{H/2} dV \frac{dV(\xi_1 U^2 + \xi_2 V^2 + \xi_3 UV)}{(\alpha^2 U^2 + \beta^2 V^2 + 2\delta UV)^{3/2}} = \mathcal{J}_{nm},$$

где $U = u - u_n$, $V = v - v_m$. Вычислим интеграл \mathcal{J}_{nm} в явном виде. Перейдя к полярным координатам $\rho = \sqrt{U^2 + V^2}$, $U = \rho \cos \phi$, $V = \rho \sin \phi$, мы преобразуем выражение под интегралом в сумму двух рациональных дробей. Применяя в получившихся двух интегралах замены $t = \operatorname{tg} \phi$ и $t = \operatorname{ctg} \phi$ соответственно, а затем сделав замену $z = t + \delta / \alpha^2$, мы приходим к табличным интегралам. В итоге получается явное выражение для \mathcal{J}_{nm} . Подробно процесс вывода интеграла \mathcal{J}_{nm} описан в работе [4]. Явное выражение для \mathcal{J}_{nm} имеет вид

$$\mathcal{J}_{nm} = \frac{h}{\beta^3} \left[\begin{aligned} & - \frac{\xi_2 z}{\sqrt{z^2 + (\alpha/\beta)^2 - (\delta/\beta^2)^2}} + \\ & + \xi_2 \ln \left| z + \sqrt{z^2 + (\alpha/\beta)^2 - (\delta/\beta^2)^2} \right| - \\ & - \frac{\xi_3 - 2\xi_2 \delta / \beta^2}{\sqrt{z^2 + (\alpha/\beta)^2 - (\delta/\beta^2)^2}} + \\ & + z \frac{\xi_2 (\delta/\beta^2)^2 + \xi_1 - \xi_3 \delta / \beta^2}{\left[(\alpha/\beta)^2 - (\delta/\beta^2)^2 \right] \sqrt{z^2 + (\alpha/\beta)^2 - (\delta/\beta^2)^2}} \end{aligned} \right]_{-H/h + \delta/\beta^2}^{H/h + \delta/\beta^2} -$$

$$- \frac{H}{\alpha^3} \left[\begin{aligned} & - \frac{\xi_1 z}{\sqrt{z^2 - (\delta/\alpha^2)^2 + (\beta/\alpha)^2}} + \\ & + \xi_1 \ln \left| z + \sqrt{z^2 - (\delta/\alpha^2)^2 + (\beta/\alpha)^2} \right| - \\ & - \frac{\xi_3 - 2\xi_1 \delta / \alpha^2}{\sqrt{z^2 - (\delta/\alpha^2)^2 + (\beta/\alpha)^2}} + \\ & + z \frac{\xi_1 (\delta/\alpha^2)^2 + \xi_2 - \xi_3 \delta / \alpha^2}{\left[-(\delta/\alpha^2)^2 + (\beta/\alpha)^2 \right] \sqrt{z^2 - (\delta/\alpha^2)^2 + (\beta/\alpha)^2}} \end{aligned} \right]_{h/H + \delta/\alpha^2}^{-h/H + \delta/\alpha^2}$$

Таким образом, интеграл \mathcal{J}_{nm} приближенно вычислен в явном виде.

4. ВЫЧИСЛЕНИЕ КАНОНИЧЕСКОГО ИНТЕГРАЛА, КОГДА ТОЧКА x НЕ ЛЕЖИТ В ОБЛАСТИ ИНТЕГРИРОВАНИЯ

Пусть точка x не принадлежит кусочку поверхности Γ , на котором изменяется точка $y = y(u, v)$, когда $(u - u_n) \in [-h/2, h/2]$ и $(v - v_m) \in [-H/2, H/2]$. Разложим $y_j(u, v)$ по формуле Тейлора с центром в точке (u_n, v_m) , тогда для $j = 1, 2, 3$ получим

$$y_j(u, v) = y_j(u_n, v_m) + D_j + O(H^2 + h^2),$$

где

$$D_j = (y_j)'_u (u - u_n) + (y_j)'_v (v - v_m).$$

Здесь и далее все производные по u и v берутся в точке (u_n, v_m) . Положим

$$r^2 = |x - y(u_n, v_m)|^2 = \sum_{j=1}^3 r_j^2 \neq 0, \quad r_j = y_j(u_n, v_m) - x_j, \quad j = 1, 2, 3,$$

тогда

$$y_j(u, v) - x_j = r_j + D_j + O(H^2 + h^2),$$

$$j = 1, 2, 3.$$

Следовательно,

$$\begin{aligned} |x - y(u, v)|^2 &= \sum_{j=1}^3 (x_j - y_j(u, v))^2 \approx \sum_{j=1}^3 (r_j^2 + 2r_j D_j + D_j^2) = r^2 + 2P(u - u_n) + \\ &+ 2Q(v - v_m) + \alpha^2 (u - u_n)^2 + \beta^2 (v - v_m)^2 + 2\delta (u - u_n)(v - v_m) = \\ &= \beta^2 (V + \delta U / \beta^2 + Q / \beta^2)^2 - (\delta U + Q)^2 / \beta^2 + \alpha^2 U^2 + 2PU + r^2, \end{aligned}$$

где $U = u - u_n$, $V = v - v_m$,

$$P = \sum_{j=1}^3 r_j (y_j)'_u, \quad Q = \sum_{j=1}^3 r_j (y_j)'_v, \quad \alpha^2 = \sum_{j=1}^3 \left((y_j)'_u \right)^2,$$

$$\beta^2 = \sum_{j=1}^3 \left((y_j)'_v \right)^2, \quad \delta = \sum_{j=1}^3 (y_j)'_u (y_j)'_v.$$

Производные по u и v берутся в точке $u = u_n$, $v = v_m$.

Можно показать (см. [12], гл. 14, § 1), что

$$\alpha^2 \beta^2 - \delta^2 = |\eta(y(u_n, v_m))|^2. \quad (10)$$

Согласно условию (4), $|\eta(y(u_n, v_m))| > 0$ для всех возможных n, m , поэтому

$$\alpha^2 \beta^2 - \delta^2 > 0. \quad (11)$$

Отсюда следует, что $\alpha^2 > 0$ и $\beta^2 > 0$.

В силу (3), для всех возможных n, m , при $u \in [u_n - h/2, u_n + h/2]$ и $v \in [v_m - H/2, v_m + H/2]$ функция $|\eta(y(u, v))|$ может быть разложена в точке (u_n, v_m) по формуле Тейлора с остаточным членом в форме Пеано (см. [12], гл. 10, § 5.3)

$$\begin{aligned}
|\eta(y(u, v))| &= \eta(y(u_n, v_m)) + |\eta'_u(u - u_n) + \eta'_v(v - v_m) + o\left(\sqrt{(u - u_n)^2 + (v - v_m)^2}\right)| = \\
&= |\eta(y(u_n, v_m))| + |\eta'_u U + \eta'_v V + o\left(\sqrt{(u - u_n)^2 + (v - v_m)^2}\right)|.
\end{aligned} \tag{12}$$

Производные по u и v берутся в точке (u_n, v_m) .

Выражения для $|\eta'_u|$ и $|\eta'_v|$ можно найти с помощью формулы (10), эти выражения имеют вид

$$|\eta'_u| = \frac{(\alpha^2)'_u \beta^2 + \alpha^2 (\beta^2)'_u - 2\delta\delta'_u}{2\sqrt{\alpha^2\beta^2 - \delta^2}}, \quad |\eta'_v| = \frac{(\alpha^2)'_v \beta^2 + \alpha^2 (\beta^2)'_v - 2\delta\delta'_v}{2\sqrt{\alpha^2\beta^2 - \delta^2}}, \tag{13}$$

где

$$\begin{aligned}
(\alpha^2)'_u &= 2\sum_{j=1}^3 (y_j)'_u (y_j)''_{uu}, & (\alpha^2)'_v &= 2\sum_{j=1}^3 (y_j)'_u (y_j)''_{uv}, \\
(\beta^2)'_u &= 2\sum_{j=1}^3 (y_j)'_v (y_j)''_{uv}, & (\beta^2)'_v &= 2\sum_{j=1}^3 (y_j)'_v (y_j)''_{vv}, \\
(\delta)'_u &= \sum_{j=1}^3 (y_j)''_{uu} (y_j)'_v + (y_j)'_u (y_j)''_{uv}, & (\delta)'_v &= \sum_{j=1}^3 (y_j)''_{uv} (y_j)'_v + (y_j)'_u (y_j)''_{vv}.
\end{aligned}$$

Выражения для $|\eta'_u|$ и $|\eta'_v|$ можно также записать в виде

$$|\eta'_u| = \frac{(\eta, \eta'_u)}{|\eta|}, \quad |\eta'_v| = \frac{(\eta, \eta'_v)}{|\eta|}, \tag{14}$$

где (\cdot, \cdot) — скалярное произведение векторов, и

$$\begin{aligned}
(\eta_1)'_u &= (y_2)_{uu} (y_3)_v + (y_2)_u (y_3)_{uv} - (y_3)_{uu} (y_2)_v - (y_3)_u (y_2)_{uv}, \\
(\eta_2)'_u &= (y_3)_{uu} (y_1)_v + (y_3)_u (y_1)_{uv} - (y_1)_{uu} (y_3)_v - (y_1)_u (y_3)_{uv}, \\
(\eta_3)'_u &= (y_1)_{uu} (y_2)_v + (y_1)_u (y_2)_{uv} - (y_2)_{uu} (y_1)_v - (y_2)_u (y_1)_{uv}, \\
(\eta_1)'_v &= (y_2)_{uv} (y_3)_v + (y_2)_u (y_3)_{vv} - (y_3)_{uv} (y_2)_v - (y_3)_u (y_2)_{vv}, \\
(\eta_2)'_v &= (y_3)_{uv} (y_1)_v + (y_3)_u (y_1)_{vv} - (y_1)_{uv} (y_3)_v - (y_1)_u (y_3)_{vv}, \\
(\eta_3)'_v &= (y_1)_{uv} (y_2)_v + (y_1)_u (y_2)_{vv} - (y_2)_{uv} (y_1)_v - (y_2)_u (y_1)_{vv}.
\end{aligned}$$

Воспользуемся разложением по формуле Тейлора в точке (u_n, v_m) с остаточным членом в форме Пеано

$$\begin{aligned}
y_j(u, v) - x_j &= r_j + (y_j)'_u (u - u_n) + (y_j)'_v (v - v_m) + \frac{1}{2}(y_j)''_{uu} (u - u_n)^2 + \\
&+ \frac{1}{2}(y_j)''_{vv} (v - v_m)^2 + (y_j)''_{uv} (u - u_n) \times (v - v_m) + o\left(\sqrt{(u - u_n)^2 + (v - v_m)^2}\right).
\end{aligned}$$

Тогда

$$|\eta(y(u,v))| \sum_{j=1}^3 \eta_j(x) (y_j(u,v) - x_j) \approx R + \xi_4 U + \xi_5 V + \xi_1 U^2 + \xi_2 V^2 + \xi_3 UV,$$

где $U = u - u_n$, $V = v - v_m$ и

$$\begin{aligned} \xi_1 &= \sum_{j=1}^3 \eta_j(x) \left(\frac{1}{2} |\eta(y(u_n, v_m))| (y_j)''_{uu} + |\eta'_u(y_j)'_u| \right), \\ \xi_2 &= \sum_{j=1}^3 \eta_j(x) \left(\frac{1}{2} |\eta(y(u_n, v_m))| (y_j)''_{vv} + |\eta'_v(y_j)'_v| \right), \\ \xi_3 &= \sum_{j=1}^3 \eta_j(x) \left(|\eta(y(u_n, v_m))| (y_j)''_{uv} + \left(|\eta'_u(y_j)'_v| + |\eta'_v(y_j)'_u| \right) \right), \\ \xi_4 &= \sum_{j=1}^3 \eta_j(x) \left(|\eta(y(u_n, v_m))| (y_j)'_u + |\eta'_u r_j| \right), \\ \xi_5 &= \sum_{j=1}^3 \eta_j(x) \left(|\eta(y(u_n, v_m))| (y_j)'_v + |\eta'_v r_j| \right), \\ R &= \sum_{j=1}^3 \eta_j(x) |\eta(y(u_n, v_m))| r_j. \end{aligned}$$

Все производные по u, v берутся в точке (u_n, v_m) . Из приведенных соотношений вытекает, что канонический интеграл из (7) приближенно равен следующему интегралу, который обозначим через $T_{nm}(x)$:

$$\begin{aligned} & \int_{u_n-h/2}^{u_n+h/2} du \int_{v_m-H/2}^{v_m+H/2} dv \frac{|\eta(y(u,v))|}{|x-y(u,v)|^3} \times \sum_{j=1}^3 \eta_j(x) (y_j(u,v) - x_j) \approx \int_{-h/2}^{h/2} dU \int_{-H/2}^{H/2} dV \times \\ & \times \frac{R + \xi_4 U + \xi_5 V + \xi_1 U^2 + \xi_2 V^2 + \xi_3 UV}{\beta^3 \times \left((V + \delta U / \beta^2 + Q / \beta^2)^2 - (\delta U + Q)^2 / \beta^4 + (\alpha^2 U^2 + 2PU + r^2) / \beta^2 \right)^{3/2}} = T_{nm}(x). \end{aligned} \quad (15)$$

4.1. Вычисление интегралов по dV

Введем обозначения

$$\begin{aligned} z &= V + \delta U / \beta^2 + Q / \beta^2 = V + c, \\ c &= \delta U / \beta^2 + Q / \beta^2, \\ a &= -c^2 + (\alpha^2 U^2 + 2PU + r^2) / \beta^2 = -(\delta U + Q)^2 / \beta^4 + (\alpha^2 U^2 + 2PU + r^2) / \beta^2 = \\ &= \frac{1}{\beta^4} \left((\alpha^2 \beta^2 - \delta^2) U^2 + 2(P\beta^2 - \delta Q) U + r^2 \beta^2 - Q^2 \right). \end{aligned} \quad (16)$$

В силу (11) $\alpha^2 > 0$ и $\beta^2 > 0$. Кроме того, отсюда следует, что $a \neq 0$, поскольку a представлено квадратичным полиномом по U , в котором коэффициент при U^2 положителен: $(\alpha^2\beta^2 - \delta^2) / \beta^4 > 0$.

Покажем, что $a \geq 0$. Положим $\tilde{D}_j = (y_j)'_u U - (y_j)'_v c$, где c определено в (16), и рассмотрим преобразования с учетом введенных обозначений

$$\begin{aligned} \sum_{j=1}^3 (r_j + \tilde{D}_j)^2 &= \sum_{j=1}^3 (r_j^2 + 2r_j\tilde{D}_j + \tilde{D}_j^2) = r^2 + 2\sum_{j=1}^3 r_j\tilde{D}_j + \sum_{j=1}^3 \tilde{D}_j^2 = \\ &= r^2 + 2PU - 2Qc + \alpha^2 U^2 + \beta^2 c^2 - 2\delta U c = \beta^2 \left[c^2 - 2\frac{\delta U + Q}{\beta^2} c + \left(\frac{\delta U + Q}{\beta^2} \right)^2 \right] - \\ &\quad - \frac{(\delta U + Q)^2}{\beta^2} + \alpha^2 U^2 + 2PU + r^2 = \beta^2 \left[-c + \frac{\delta U + Q}{\beta^2} \right]^2 - \frac{(\delta U + Q)^2}{\beta^2} + \\ &\quad + \alpha^2 U^2 + 2PU + r^2 = -\frac{(\delta U + Q)^2}{\beta^2} + \alpha^2 U^2 + 2PU + r^2 = a\beta^2 \geq 0. \end{aligned}$$

Как отмечено выше, $\beta^2 > 0$, поэтому, поделив полученное соотношение на β^2 , заключаем, что $a \geq 0$. Следовательно, квадратичный полином, обозначенный через a , неотрицательный.

Положим

$$\begin{aligned} b &= \xi_2 c^2 - \xi_5 c - \xi_3 U c + R + \xi_4 U + \xi_1 U^2 = \\ &= U^2 \left(\xi_2 \frac{\delta^2}{\beta^4} + \xi_1 - \xi_3 \frac{\delta}{\beta^2} \right) + U \left(2\xi_2 \frac{\delta Q}{\beta^4} + \xi_4 - \xi_3 \frac{Q}{\beta^2} - \xi_5 \frac{\delta}{\beta^2} \right) + R - \frac{\xi_5 Q}{\beta^2} + \xi_2 \frac{Q^2}{\beta^4}, \\ z_{\pm} &= \pm H / 2 + c = \pm H / 2 + (\delta U + Q) / \beta^2. \end{aligned}$$

Применяя введенные обозначения, вычислим в (15) интеграл по V , переходя к переменной z :

$$\begin{aligned} &\int_{-H/2}^{H/2} dV \frac{R + \xi_4 U + \xi_5 V + \xi_1 U^2 + \xi_2 V^2 + \xi_3 UV}{\left((V + \delta U / \beta^2 + Q / \beta^2)^2 - (\delta U + Q)^2 / \beta^4 + (\alpha^2 U^2 + 2PU + r^2) / \beta^2 \right)^{3/2}} = \\ &= \int_{-H/2}^{H/2} dV \frac{\xi_2 (V + c - c)^2 + (\xi_3 U + \xi_5)(V + c - c) + R + \xi_4 U + \xi_1 U^2}{\left((V + c)^2 + a \right)^{3/2}} = \\ &= \int_{z_-}^{z_+} dz \frac{\xi_2 z^2 + (\xi_3 U + \xi_5 - 2\xi_2 c)z + b}{(z^2 + a)^{3/2}} = \\ &= \left(\left(\frac{b}{a} - \xi_2 \right) z - \xi_3 U - \xi_5 + 2\xi_2 c \right) \frac{1}{\sqrt{z^2 + a}} \Big|_{z_-}^{z_+} + \xi_2 \ln \left| z + \sqrt{z^2 + a} \right| \Big|_{z_-}^{z_+}, \end{aligned}$$

где использованы интегралы 1.2.43.17–1.2.43.19 из книги [14]. Заметим, что

$$\xi_2 \int_{-h/2}^{h/2} dU \ln \left| z + \sqrt{z^2 + a} \right| \Big|_{z_-}^{z_+} = \xi_2 \beta \theta_{nm}(x),$$

где функция $\theta_{nm}(x)$ выведена в [2], [15] в явном виде. Интеграл в (15) можно записать в виде

$$T_{nm}(x) = \frac{1}{\beta^3} (\xi_2 \beta \theta_{nm}(x) + J(H) - J(-H)).$$

Поскольку функция $\theta_{nm}(x)$ дается в [2], [15], задача сводится к вычислению интеграла

$$\begin{aligned} J(\pm H) &= \int_{-h/2}^{h/2} dU \frac{(b/a - \xi_2)z_{\pm} - \xi_3 U - \xi_5 + 2\xi_2 c}{\sqrt{z_{\pm}^2 + a}} = \\ &= \int_{-h/2}^{h/2} dU \frac{(b/a - \xi_2)(\pm H/2 + c) - \xi_3 U - \xi_5 + 2\xi_2 c}{\sqrt{(\pm H/2 + c)^2 + a}}. \end{aligned}$$

Достаточно вычислить интеграл $J(H)$. Интеграл $J(-H)$ вычисляется по тем же формулам, что и интеграл $J(H)$, в которых H надо заменить на $-H$. Вычислим интеграл $J(H)$. Распишем величины, входящие в подинтегральную функцию, в виде многочленов по U :

$$a = C_2 U^2 + C_1 U + C_0,$$

$$C_2 = (\alpha^2 - \delta^2 / \beta^2) / \beta^2, \quad C_1 = (2P - 2\delta Q / \beta^2) / \beta^2, \quad C_0 = (r^2 - Q^2 / \beta^2) / \beta^2;$$

$$z_+^2 + a = B_2 U^2 + B_1 U + B_0,$$

$$B_2 = \alpha^2 / \beta^2, \quad B_1 = (H\delta + 2P) / \beta^2, \quad B_0 = H^2 / 4 + (HQ + r^2) / \beta^2;$$

$$b = A_2 U^2 + A_1 U + A_0,$$

$$A_2 = \xi_1 - \xi_3 \delta / \beta^2 + \xi_2 \delta^2 / \beta^4, \quad A_1 = \xi_4 - \xi_5 \delta / \beta^2 - \xi_3 Q / \beta^2 + 2\xi_2 \delta Q / \beta^4,$$

$$A_0 = R - \xi_5 Q / \beta^2 + \xi_2 Q^2 / \beta^4;$$

$$bz_+ = (A_2 U^2 + A_1 U + A_0) \times (\delta U / \beta^2 + H/2 + Q / \beta^2) = E_3 U^3 + E_2 U^2 + E_1 U + E_0,$$

$$E_3 = A_2 \delta / \beta^2, \quad E_2 = A_2 (H/2 + Q / \beta^2) + A_1 \delta / \beta^2,$$

$$E_1 = A_1 (H/2 + Q / \beta^2) + A_0 \delta / \beta^2, \quad E_0 = A_0 (H/2 + Q / \beta^2);$$

$$\xi_2 z_+ + \xi_3 U + \xi_5 - 2\xi_2 c = F_1 U + F_0, \quad F_1 = \xi_3 - \xi_2 \delta / \beta^2, \quad F_0 = \xi_5 - \xi_2 Q / \beta^2 + \xi_2 H / 2.$$

Применяя введенные обозначения, запишем интеграл $J(H)$ в виде

$$J(H) = J_1 - J_2,$$

$$J_1 = \int_{-h/2}^{h/2} dU \frac{E_3 U^3 + E_2 U^2 + E_1 U + E_0}{(C_2 U^2 + C_1 U + C_0) \sqrt{B_2 U^2 + B_1 U + B_0}},$$

$$J_2 = \int_{-h/2}^{h/2} dU \frac{F_1 U + F_0}{\sqrt{B_2 U^2 + B_1 U + B_0}}.$$

4.2. Вычисление интегралов по dU

Используя деление многочленов и учитывая, что $C_2 > 0$ в силу (11), приведем интеграл J_1 к виду

$$J_1 = J_{11} + J_{12},$$

$$\begin{aligned}
 J_{11} &= \int_{-h/2}^{h/2} dU \frac{L_1 U + L_0}{\sqrt{B_2 U^2 + B_1 U + B_0}}, \\
 J_{12} &= \int_{-h/2}^{h/2} dU \frac{S_1 U + S_0}{(C_2 U^2 + C_1 U + C_0) \sqrt{B_2 U^2 + B_1 U + B_0}},
 \end{aligned}
 \tag{17}$$

где

$$\begin{aligned}
 L_1 &= \frac{E_3}{C_2}, \quad L_0 = \frac{E_2 C_2 - E_3 C_1}{C_2^2}, \quad S_1 = E_1 - \frac{E_3 C_0}{C_2} - \frac{C_1(E_2 C_2 - E_3 C_1)}{C_2^2}, \\
 S_0 &= E_0 - \frac{C_0(E_2 C_2 - E_3 C_1)}{C_2^2}.
 \end{aligned}$$

Поскольку $a \geq 0$, получаем $z_+^2 + a = B_2 U^2 + B_1 U + B_0 \geq 0$, поэтому $B_1^2 - 4B_2 B_0 \leq 0$. Если $B_1^2 - 4B_2 B_0 = 0$ и $-B_1 / (2B_2) \in [-h / 2, h / 2]$, то в силу ([2], п. 2), ([15], п. 3) точка x лежит на том кусочке касательной плоскости, проведенной в точке $y(u_n, v_m)$, по которому идет интегрирование в каноническом интеграле после линеаризации $y(u, v)$ вблизи узла $y(u_n, v_m)$. Однако точка x не попадает на этот кусочек касательной плоскости, так как лежит в другом узле на Γ . Поэтому если $B_1^2 - 4B_2 B_0 = 0$, то $-B_1 / (2B_2) \notin [-h / 2, h / 2]$.

Поскольку $B_2 > 0$, интегралы J_{11} и J_2 находятся с помощью табличных интегралов 2.261 и 2.264 из книги [16]:

$$\begin{aligned}
 J_{11} &= \left(\frac{L_1}{B_2} \sqrt{B_2 U^2 + B_1 U + B_0} + \left(L_0 - \frac{L_1 B_1}{2B_2} \right) \frac{1}{\sqrt{B_2}} \ln \left| \frac{2B_2 U + B_1 + 2\sqrt{B_2} \times}{\times \sqrt{B_2 U^2 + B_1 U + B_0}} \right| \right) \Bigg|_{U=-h/2}^{U=h/2}, \\
 J_2 &= \left(\frac{F_1}{B_2} \sqrt{B_2 U^2 + B_1 U + B_0} + \left(F_0 - \frac{F_1 B_1}{2B_2} \right) \frac{1}{\sqrt{B_2}} \ln \left| \frac{2B_2 U + B_1 + 2\sqrt{B_2} \times}{\times \sqrt{B_2 U^2 + B_1 U + B_0}} \right| \right) \Bigg|_{U=-h/2}^{U=h/2},
 \end{aligned}$$

Остается вычислить интеграл J_{12} . Способ вычисления интеграла зависит от знака дискриминанта квадратного трехчлена $C_2 U^2 + C_1 U + C_0$, стоящего в знаменателе подынтегральной функции.

Первый случай: $C_1^2 - 4C_2 C_0 > 0$. Выше показано, что квадратичный полином, обозначенный как a — неотрицательный, следовательно, его дискриминант неположительный: $C_1^2 - 4C_2 C_0 \leq 0$, поэтому первый случай не реализуется.

Второй случай: $C_1^2 - 4C_2 C_0 = 0$. В этом случае $C_2 U^2 + C_1 U + C_0 = C_2 (U - U_1)^2$, где $U_1 = -C_1 / (2C_2)$ — корень многочлена. Применяя соотношение

$$\frac{S_1 U + S_0}{(U - U_1)^2} = \frac{S_1 (U - U_1) + S_0 + U_1 S_1}{(U - U_1)^2} = \frac{S_1}{U - U_1} + \frac{S_0 + U_1 S_1}{(U - U_1)^2},$$

получим

$$J_{12} = \int_{-h/2}^{h/2} \frac{S_1 dU}{C_2 (U - U_1) \sqrt{B_2 U^2 + B_1 U + B_0}} + \frac{S_0 + U_1 S_1}{C_2} \int_{-h/2}^{h/2} \frac{dU}{(U - U_1)^2 \sqrt{B_2 U^2 + B_1 U + B_0}}.$$

Либо $U_1 > h / 2$, либо $U_1 < -h / 2$. Случай $U_1 \in [-h / 2, h / 2]$ не реализуется, так как точка x лежит в узле, отличном от $y(u_n, v_m)$, и не попадает на маленький кусочек касательной плоскости вблизи точки $y(u_n, v_m)$, по которому ведется интегрирование в каноническом интеграле после линеаризации $y(u, v)$ около $y(u_n, v_m)$. Сделав замену $t = 1 / (U - U_1)$, находим (см. [17], ч. 1, гл. 7, § 10, п. 5)

$$t_1 = \frac{2}{h - 2U_1}, \quad t_2 = -\frac{2}{h + 2U_1},$$

$$\begin{aligned}
J_{12} &= -\frac{S_1}{C_2} \int_{t_2}^{t_1} \frac{\operatorname{sgn}(t) dt}{\sqrt{(U_1^2 B_2 + U_1 B_1 + B_0)t^2 + (2B_2 U_1 + B_1)t + B_2}} - \\
&= -\frac{S_0 + U_1 S_1}{C_2} \int_{t_2}^{t_1} \frac{\operatorname{sgn}(t) dt}{\sqrt{(U_1^2 B_2 + U_1 B_1 + B_0)t^2 + (2B_2 U_1 + B_1)t + B_2}} = \\
&= -\frac{S_1}{C_2} \operatorname{sgn}(C_1) \int_{t_2}^{t_1} \frac{dt}{\sqrt{\omega_2 t^2 + \omega_1 t + B_2}} - \frac{S_0 + U_1 S_1}{C_2} \operatorname{sgn}(C_1) \int_{t_2}^{t_1} \frac{t dt}{\sqrt{\omega_2 t^2 + \omega_1 t + B_2}},
\end{aligned}$$

где $\omega_1 = 2B_2 U_1 + B_1$, $\omega_2 = U_1^2 B_2 + U_1 B_1 + B_0$. Как показано выше, $a \geq 0$, поэтому и $\omega_2 \geq 0$. Если $\omega_2 > 0$, то с помощью табличных интегралов 2.261 и 2.264 из книги [16] получаем

$$\begin{aligned}
J_{12}(t_2, t_1) &= \frac{-S_1 \operatorname{sgn}(C_1)}{C_2 \sqrt{\omega_2}} \ln \left| 2\omega_2 t + \omega_1 + 2\sqrt{\omega_2} \sqrt{\omega_2 t^2 + \omega_1 t + B_2} \right| \Big|_{t_2}^{t_1} - \\
&= -\frac{S_0 + U_1 S_1}{C_2} \operatorname{sgn}(C_1) \left(\frac{1}{\omega_2} \sqrt{\omega_2 t^2 + \omega_1 t + B_2} - \right. \\
&\quad \left. - \frac{\omega_1}{2\omega_2} \frac{1}{\sqrt{\omega_2}} \ln \left| 2\omega_2 t + \omega_1 + 2\sqrt{\omega_2} \sqrt{\omega_2 t^2 + \omega_1 t + B_2} \right| \right) \Big|_{t_2}^{t_1}.
\end{aligned}$$

Если $\omega_2 = 0$, а $\omega_1 \neq 0$, то, пользуясь интегралами 1.2.18.5, 1.2.18.6 из книги [14], находим

$$J_{12}(t_2, t_1) = -\frac{S_1}{C_2} \frac{2}{\omega_1} \operatorname{sgn}(C_1) \sqrt{\omega_1 t + B_2} \Big|_{t_2}^{t_1} - \frac{S_0 + U_1 S_1}{C_2} \operatorname{sgn}(C_1) \frac{2(\omega_1 t - 2B_2)}{3\omega_1^2} \sqrt{\omega_1 t + B_2} \Big|_{t_2}^{t_1}.$$

Если $\omega_2 = 0$ и $\omega_1 = 0$, то

$$J_{12}(t_2, t_1) = -\frac{S_1}{C_2} \frac{\operatorname{sgn}(C_1)}{\sqrt{B_2}} t \Big|_{t_2}^{t_1} - \frac{S_0 + U_1 S_1}{C_2} \frac{\operatorname{sgn}(C_1)}{2\sqrt{B_2}} t^2 \Big|_{t_2}^{t_1}.$$

Тем самым, во втором случае интеграл J_{12} вычислен явно.

Третий случай: $C_1^2 - 4C_2 C_0 < 0$. В этом случае многочлен $C_2 U^2 + C_1 U + C_0$ неприводимый. Рассмотрим различные варианты вычисления интеграла J_{12} из (17), воспользовавшись методом, предложенным в ([17], ч. 1, гл. 7, § 10, п. 5, (7.75)) или в ([16], разд. 2.25).

Вариант 1. Если $B_1 = B_2 C_1 / C_2$, то в интеграле J_{12} достаточно сделать замену $U = t - C_1 / (2C_2)$. Пусть, кроме того,

$$t_1 = \frac{h}{2} + \frac{C_1}{2C_2}, \quad t_2 = -\frac{h}{2} + \frac{C_1}{2C_2}, \quad t_2 < t_1,$$

тогда

$$\begin{aligned}
J_{12} &= \int_{t_2}^{t_1} \frac{[S_1 t + (-S_1 C_1 / (2C_2) + S_0)] dt}{C_2 [t^2 + C_0 / C_2 - C_1^2 / (4C_2^2)] \sqrt{B_2 t^2 + B_0 - B_2 C_1^2 / (4C_2^2)}} = \\
&= \frac{1}{C_2} \left(\frac{S_1}{2} \int_{t_2}^{t_1} \frac{dt^2}{[t^2 + \sigma_1] \sqrt{B_2 t^2 + \sigma_2}} + \left(-\frac{S_1 C_1}{2C_2} + S_0 \right) \int_{t_2}^{t_1} \frac{dt}{[t^2 + \sigma_1] \sqrt{B_2 t^2 + \sigma_2}} \right),
\end{aligned}$$

где

$$\sigma_1 = \frac{C_0}{C_2} - \frac{C_1^2}{4C_2^2} > 0, \quad \sigma_2 = B_0 - \frac{B_2 C_1^2}{4C_2^2} \geq 0.$$

Вводя обозначения

$$J_{121} = \frac{S_1}{2C_2} \int_{t=t_2}^{t=t_1} \frac{dt^2}{[t^2 + \sigma_1] \sqrt{B_2 t^2 + \sigma_2}},$$

$$J_{122} = \frac{1}{C_2 \sqrt{B_2}} \left(-\frac{S_1 C_1}{2C_2} + S_0 \right) \int_{t_2}^{t_1} \frac{dt}{[t^2 + \sigma_1] \sqrt{t^2 + \sigma_2 / B_2}},$$

получим

$$J_{12} = J_{121} + J_{122}. \quad (18)$$

Используя табличный интеграл 2.246 из книги [16], находим интеграл J_{121} в явном виде.

1. Если $\sigma_1 B_2 - \sigma_2 > 0$, то $\sigma_2 \geq 0$ и

$$J_{121} = \frac{S_1}{C_2 \sqrt{\sigma_1 B_2 - \sigma_2}} \operatorname{arctg} \left(\frac{\sqrt{B_2 t^2 + \sigma_2}}{\sqrt{\sigma_1 B_2 - \sigma_2}} \right) \Big|_{t_2}^{t_1}.$$

2. Если $\sigma_1 B_2 - \sigma_2 < 0$, то $\sigma_2 > 0$ и

$$J_{121} = \frac{S_1}{2C_2 \sqrt{\sigma_2 - \sigma_1 B_2}} \ln \left| \frac{\sqrt{B_2 t^2 + \sigma_2} - \sqrt{\sigma_2 - \sigma_1 B_2}}{\sqrt{B_2 t^2 + \sigma_2} + \sqrt{\sigma_2 - \sigma_1 B_2}} \right| \Big|_{t_2}^{t_1}.$$

3. Если $\sigma_1 B_2 - \sigma_2 = 0$, то $\sigma_2 > 0$ и

$$J_{121} = -\frac{S_1}{C_2 \sqrt{B_2 t^2 + \sigma_2}} \Big|_{t_2}^{t_1}.$$

Используя табличные интегралы 1.2.45.10, 1.2.43.17 и 1.2.11.10 из книги [14], находим интеграл J_{122} в явном виде.

1. Если $\sigma_2 > 0$ и $\sigma_2 - B_2 \sigma_1 < 0$, то

$$J_{122} = \frac{1}{C_2 \sqrt{B_2}} \left(-\frac{S_1 C_1}{2C_2} + S_0 \right) \frac{1}{\sqrt{\sigma_1 (\sigma_1 - \sigma_2 / B_2)}} \times \ln \left| \frac{t \sqrt{\sigma_1 - \sigma_2 / B_2} + \sqrt{\sigma_1 (t^2 + \sigma_2 / B_2)}}{\sqrt{t^2 + \sigma_1}} \right| \Big|_{t_2}^{t_1}.$$

2. Если $\sigma_2 - B_2 \sigma_1 > 0$, то $\sigma_2 > 0$ и

$$J_{122} = \frac{1}{C_2 \sqrt{B_2}} \left(-\frac{S_1 C_1}{2C_2} + S_0 \right) \times \frac{1}{\sqrt{\sigma_1 (\sigma_2 / B_2 - \sigma_1)}} \operatorname{arctg} \frac{t \sqrt{\sigma_2 / B_2 - \sigma_1}}{\sqrt{\sigma_1 (t^2 + \sigma_2 / B_2)}} \Big|_{t_2}^{t_1}.$$

3. Если $\sigma_1 - \sigma_2 / B_2 = 0$, то $\sigma_2 > 0$ и

$$J_{122} = \frac{1}{C_2 \sqrt{B_2}} \left(-\frac{S_1 C_1}{2C_2} + S_0 \right) \frac{t}{\sigma_1 \sqrt{t^2 + \sigma_1}} \Big|_{t_2}^{t_1}.$$

4. Если $\sigma_2 = 0$, то $B_1^2 - 4B_0B_2 = 0$ и $|B_1 / (2B_2)| > h / 2$. Поскольку $t = U + C_1 / (2C_2)$ и $B_1 = B_2C_1 / C_2$, то $\text{sgn}(t) = \text{sgn}(B_1)$ и

$$J_{122} = \frac{1}{C_2\sqrt{B_2}} \left(-\frac{S_1C_1}{2C_2} + S_0 \right) \text{sgn}(B_1) \frac{1}{2\sigma_1} \ln \frac{t^2}{|t^2 + \sigma_1|} \Bigg|_{t_2}^{t_1}.$$

Итак, в этом варианте интеграл J_{12} из (17) вычисляется явно по формуле (18).

Вариант 2. Если $B_1 \neq B_2C_1 / C_2$, то в интеграле J_{12} из (17) сделаем дробно-линейную подстановку

$$U = \frac{v_2t + v_1}{1+t}, \quad t = \frac{v_1 - U}{U - v_2}, \quad dU = \frac{v_2 - v_1}{(1+t)^2} dt, \quad S_1U + S_0 = \frac{(S_1v_2 + S_0)t + (S_1v_1 + S_0)}{1+t}.$$

Константы v_2 и v_1 выбираются таким образом, чтобы квадратичные полиномы в подынтегральной функции в интеграле J_{12} из (17) не имели членов первой степени относительно t . Можно показать (см. [17], ч. 1, гл. 7, § 10, п. 5), что если $B_1 \neq B_2C_1 / C_2$, то такие константы всегда существуют, различны и являются корнями уравнения

$$z^2 + \frac{2(B_0 - B_2C_0 / C_2)}{B_1 - B_2C_1 / C_2} z + \frac{B_0C_1 / C_2 - B_1C_0 / C_2}{B_1 - B_2C_1 / C_2} = 0.$$

Отсюда следует, что

$$v_1 = \frac{-X_1 + \sqrt{X_1^2 - 4X_2}}{2}, \quad v_2 = \frac{-X_1 - \sqrt{X_1^2 - 4X_2}}{2},$$

где

$$X_1 = \frac{2(B_0 - B_2C_0 / C_2)}{B_1 - B_2C_1 / C_2}, \quad X_2 = \frac{B_0C_1 / C_2 - B_1C_0 / C_2}{B_1 - B_2C_1 / C_2}.$$

Квадратичные полиномы, стоящие в подынтегральной функции в интеграле J_{12} из (17), после перехода к переменной t принимают вид

$$U^2 + \frac{C_1}{C_2}U + \frac{C_0}{C_2} = \frac{v_2^2 + (C_1 / C_2)v_2 + C_0 / C_2}{(1+t)^2} \times \left(t^2 + \frac{v_1^2 + (C_1 / C_2)v_1 + C_0 / C_2}{v_2^2 + (C_1 / C_2)v_2 + C_0 / C_2} \right) = \frac{\psi_1(t^2 + \psi_2 / \psi_1)}{(1+t)^2},$$

$$B_2U^2 + B_1U + B_0 = \frac{(B_2v_2^2 + B_1v_2 + B_0)t^2 + B_2v_1^2 + B_1v_1 + B_0}{(1+t)^2} = \frac{\psi_3t^2 + \psi_4}{(1+t)^2},$$

где $\psi_1 = v_2^2 + (C_1 / C_2)v_2 + C_0 / C_2 > 0$, $\psi_2 = v_1^2 + (C_1 / C_2)v_1 + C_0 / C_2 > 0$, $\psi_3 = B_2v_2^2 + B_1v_2 + B_0$, $\psi_4 = B_2v_1^2 + B_1v_1 + B_0$. Ранее было показано, что $B_2U^2 + B_1U + B_0 \geq 0$ при любом U , следовательно, этот квадратичный полином имеет не более одного вещественного корня, а также $\psi_3 \geq 0$ и $\psi_4 \geq 0$. Поскольку константы v_2 и v_1 всегда существуют и различны, ψ_4 и ψ_3 одновременно не обращаются в нуль.

Переходя к переменной t в интеграле J_{12} из (17), получаем

$$J_{12} = \frac{1}{C_2} \int_{t_2}^{t_1} \frac{(S_1v_2 + S_0)t + S_1v_1 + S_0}{(1+t)} \frac{(v_2 - v_1)}{(1+t)^2} \times \frac{(1+t)^2}{\psi_1(t^2 + \psi_2 / \psi_1)} \frac{\sqrt{(1+t)^2}}{\sqrt{\psi_3t^2 + \psi_4}} dt =$$

$$= \frac{v_2 - v_1}{C_2\psi_1} \int_{t_2}^{t_1} \frac{(S_1v_2 + S_0)t + S_1v_1 + S_0}{(t^2 + \psi_2 / \psi_1)\sqrt{\psi_3t^2 + \psi_4}} \text{sgn}(1+t) dt,$$

где

$$t_1 = \frac{v_1 - h/2}{h/2 - v_2}, \quad t_2 = \frac{v_1 + h/2}{-h/2 - v_2}.$$

Введем обозначение

$$J_{12}^*(\tau_2, \tau_1) = \frac{v_2 - v_1}{C_2 \Psi_1} \int_{\tau_2}^{\tau_1} \frac{(S_1 v_2 + S_0)t + S_1 v_1 + S_0}{(t^2 + \Psi_2 / \Psi_1) \sqrt{\Psi_3 t^2 + \Psi_4}} dt,$$

тогда

$$J_{12} = \begin{cases} -J_{12}^*(t_2, t_1), & t_1 + 1 < 0, t_2 + 1 < 0 \Leftrightarrow t_1 < -1, t_2 < -1, \\ J_{12}^*(t_2, t_1), & t_1 + 1 > 0, t_2 + 1 > 0 \Leftrightarrow t_1 > -1, t_2 > -1, \\ -J_{12}^*(t_2, -1) + J_{12}^*(-1, t_1), & t_1 + 1 > 0, t_2 + 1 < 0 \Leftrightarrow t_1 > -1, t_2 < -1, \\ J_{12}^*(t_2, -1) - J_{12}^*(-1, t_1), & t_1 + 1 < 0, t_2 + 1 > 0 \Leftrightarrow t_1 < -1, t_2 > -1. \end{cases} \quad (19)$$

Интеграл $J_{12}^*(\tau_2, \tau_1)$ можно представить в виде

$$J_{12}^*(\tau_2, \tau_1) = J_{121}^*(\tau_2, \tau_1) + J_{122}^*(\tau_2, \tau_1), \quad (20)$$

где

$$J_{121}^*(\tau_2, \tau_1) = \frac{v_2 - v_1}{C_2 \Psi_1} \frac{S_1 v_2 + S_0}{2} \int_{t=\tau_2}^{t=\tau_1} \frac{dt^2}{(t^2 + \Psi_2 / \Psi_1) \sqrt{\Psi_3 t^2 + \Psi_4}},$$

$$J_{122}^*(\tau_2, \tau_1) = \frac{(v_2 - v_1)(S_1 v_1 + S_0)}{C_2 \Psi_1} \int_{t=\tau_2}^{t=\tau_1} \frac{dt}{(t^2 + \Psi_2 / \Psi_1) \sqrt{\Psi_3 t^2 + \Psi_4}}.$$

Пусть $\Psi_3 \neq 0$. Подынтегральная функция в $J_{121}^*(\tau_2, \tau_1)$ представляет собой линейную иррациональность относительно t^2 . Применяя табличный интеграл 2.246 из книги [16], находим интеграл $J_{121}^*(\tau_2, \tau_1)$ в явном виде.

1. Если $\Psi_2 \Psi_3 / \Psi_1 - \Psi_4 > 0$, то

$$J_{121}^*(\tau_2, \tau_1) = \frac{(v_2 - v_1)(S_1 v_2 + S_0)}{C_2 \Psi_1} \cdot \frac{1}{\sqrt{\Psi_2 \Psi_3 / \Psi_1 - \Psi_4}} \operatorname{arctg} \left(\frac{\sqrt{\Psi_3 t^2 + \Psi_4}}{\sqrt{\Psi_2 \Psi_3 / \Psi_1 - \Psi_4}} \right) \Bigg|_{t_2}^{t_1}.$$

2. Если $\Psi_2 \Psi_3 / \Psi_1 - \Psi_4 < 0$, то $\Psi_4 > 0$ и

$$J_{121}^*(\tau_2, \tau_1) = \frac{(v_2 - v_1)(S_1 v_2 + S_0)}{2C_2 \Psi_1} \frac{1}{\sqrt{\Psi_4 - \Psi_2 \Psi_3 / \Psi_1}} \times \ln \left| \frac{\sqrt{\Psi_3 t^2 + \Psi_4} - \sqrt{\Psi_4 - \Psi_2 \Psi_3 / \Psi_1}}{\sqrt{\Psi_3 t^2 + \Psi_4} + \sqrt{\Psi_4 - \Psi_2 \Psi_3 / \Psi_1}} \right| \Bigg|_{\tau_2}^{\tau_1}.$$

3. Если $\Psi_2 \Psi_3 / \Psi_1 - \Psi_4 = 0$, то $\Psi_4 > 0$ и

$$J_{121}^*(\tau_2, \tau_1) = -\frac{(v_2 - v_1)(S_1 v_2 + S_0)}{C_2 \Psi_1} \frac{1}{\sqrt{\Psi_3 t^2 + \Psi_4}} \Bigg|_{\tau_2}^{\tau_1}.$$

Пусть $\Psi_3 = 0$. Тогда $\Psi_4 \neq 0$ и

$$J_{121}^*(\tau_2, \tau_1) = \frac{(v_2 - v_1)(S_1 v_2 + S_0)}{2C_2 \psi_1 \sqrt{\psi_4}} \times \ln \left| t^2 + \psi_2 / \psi_1 \right|_{\tau_2}^{\tau_1}.$$

Пусть $\psi_3 \neq 0$. Используя табличные интегралы 1.2.43.17, 1.2.45.10 и 1.2.11.10 из книги [14], находим интеграл $J_{122}^*(\tau_2, \tau_1)$ в явном виде.

1. Если $\psi_4 > 0$ и $\psi_4 - \psi_3 \psi_2 / \psi_1 < 0$, то

$$J_{122}^*(\tau_2, \tau_1) = \frac{(v_2 - v_1)(S_1 v_1 + S_0)}{C_2 \psi_1 \sqrt{\psi_3}} \times \frac{1}{\sqrt{\psi_2 / \psi_1 (\psi_2 / \psi_1 - \psi_4 / \psi_3)}} \times \times \ln \left| \frac{t \sqrt{\psi_2 / \psi_1 - \psi_4 / \psi_3} + \sqrt{\psi_2 / \psi_1 (t^2 + \psi_4 / \psi_3)}}{\sqrt{t^2 + \psi_2 / \psi_1}} \right|_{\tau_2}^{\tau_1}.$$

2. Если $\psi_4 - \psi_3 \psi_2 / \psi_1 > 0$, то $\psi_4 > 0$ и

$$J_{122}^*(\tau_2, \tau_1) = \frac{(v_2 - v_1)(S_1 v_1 + S_0)}{C_2 \psi_1 \sqrt{\psi_3}} \times \frac{1}{\sqrt{\psi_2 / \psi_1 (\psi_4 / \psi_3 - \psi_2 / \psi_1)}} \frac{t \sqrt{\psi_4 / \psi_3 - \psi_2 / \psi_1}}{\sqrt{\psi_2 / \psi_1 (t^2 + \psi_4 / \psi_3)}} \Big|_{\tau_2}^{\tau_1}.$$

3. Если $\psi_2 / \psi_1 - \psi_4 / \psi_3 = 0$, то $\psi_4 > 0$ и

$$J_{122}^*(\tau_2, \tau_1) = \frac{(v_2 - v_1)(S_1 v_1 + S_0)}{C_2 \psi_1 \sqrt{\psi_3}} \frac{t}{\psi_2 / \psi_1 \sqrt{t^2 + \psi_2 / \psi_1}} \Big|_{\tau_2}^{\tau_1}.$$

4. Если $\psi_4 = 0$, то

$$J_{122}^*(\tau_2, \tau_1) = \frac{(v_2 - v_1)(S_1 v_1 + S_0)}{C_2 \psi_1 \sqrt{\psi_3}} \times \int_{t=\tau_2}^{t=\tau_1} \operatorname{sgn}(t) \frac{dt}{[t^2 + \psi_2 / \psi_1] t},$$

и точка $t = 0$ на координатной оси t отвечает точке $U = -B_1 / (2B_2)$ на оси U . Как указано в начале п. 4.2, точка $-B_1 / (2B_2)$ не попадает в интервал интегрирования по переменной U , поэтому точка $t = 0$ не попадает в интервал интегрирования по переменной t . Следовательно,

$$J_{122}^*(\tau_2, \tau_1) = \frac{(v_2 - v_1)(S_1 v_1 + S_0)}{C_2 \psi_1 \sqrt{\psi_3}} \times \operatorname{sgn}(\tau_1) \frac{1}{2\psi_2 / \psi_1} \ln \frac{t^2}{|t^2 + \psi_2 / \psi_1|} \Big|_{\tau_2}^{\tau_1}.$$

Пусть $\psi_3 = 0$. Тогда $\psi_4 \neq 0$ и

$$J_{122}^*(\tau_2, \tau_1) = \frac{(v_2 - v_1)(S_1 v_1 + S_0)}{C_2 \psi_1 \sqrt{\psi_4}} \sqrt{\frac{\psi_1}{\psi_2}} \operatorname{arctg} \left(t \sqrt{\frac{\psi_1}{\psi_2}} \right) \Big|_{\tau_2}^{\tau_1}.$$

Итак, в варианте 2 интеграл J_{12} из (17) вычисляется явно по формулам (19), (20).

5. ОСНОВНОЙ РЕЗУЛЬТАТ

Сформулируем основной результат этой работы.

Пусть Γ — простая гладкая замкнутая поверхность класса C^2 , ограничивающая объемно-одно связную внутреннюю область, либо простая гладкая ограниченная разомкнутая поверхность класса C^2 , содержащая свои предельные точки. Пусть Γ допускает параметризацию (1) со свойствами (3), (4); и $\mu(y) \in C^0(\Gamma)$, а точка x расположена в одном из узлов на Γ . Тогда для прямого значения нормальной производной потенциала простого слоя (6) для уравнения Гельмгольца имеет место квадратурная формула

$$\frac{\partial \mathcal{V}_k[\mu](x)}{\partial \mathbf{n}_x} \Big|_{x=y(u_n, v_m) \in \Gamma} \approx \frac{1}{4\pi} \mu_{\hat{n}\hat{m}} \mathcal{J}_{\hat{n}\hat{m}} + \frac{1}{4\pi |\eta(x)|} \sum_{\substack{n=N-1, m=M-1 \\ n=0, m=0 \\ (n=m) \neq (\hat{n}, \hat{m})}} \mu_{nm} \times \\ \times \exp(ik|x-y(u_n, v_m)|) (1 - ik|x-y(u_n, v_m)|) T_{nm}(x), \quad (21)$$

где интеграл $\mathcal{J}_{\hat{n}\hat{m}}$ вычислен в явном виде в п. 3, а интеграл $T_{nm}(x)$ вычислен в явном виде в п. 4.

На рассмотренных ниже тестовых примерах остаточный член в (21) ведет себя как $O(H+h)$ равномерно по всем x , лежащим в узлах сетки.

Результаты для прямого значения нормальной производной гармонического потенциала получаются из приведенных результатов в частном случае $k=0$.

6. СТАНДАРТНАЯ КВАДРАТУРНАЯ ФОРМУЛА

Квадратурная формула (21) является альтернативой стандартной квадратурной формуле для прямого значения нормальной производной потенциала простого слоя, используемой в инженерных расчетах (см. [1], гл. 2):

$$\frac{\partial \mathcal{V}_k[\mu](x)}{\partial \mathbf{n}_x} \Big|_{x=y(u_n, v_m) \in \Gamma} \approx \frac{1}{4\pi |\eta(x)|} \times \sum_{\substack{n=N-1, m=M-1 \\ n=0, m=0 \\ (n=m) \neq (\hat{n}, \hat{m})}} \mu_{nm} \times |\eta(y(u_n, v_m))| \exp(ik|x-y(u_n, v_m)|) \times \\ \times (1 - ik|x-y(u_n, v_m)|) \sum_{j=1}^3 \frac{\eta_j(x)(y_j(u_n, v_m) - x_j)}{|x-y(u_n, v_m)|^3} hH. \quad (22)$$

Стандартная квадратурная формула получается из формулы (21) обнулением $\mathcal{J}_{\hat{n}\hat{m}}$ и заменой канонического интеграла $T_{nm}(x)$ на его следующее приближенное значение:

$$hH \sum_{j=1}^3 \frac{\eta_j(x)(y_j(u_n, v_m) - x_j)}{|x-y(u_n, v_m)|^3}.$$

Обнуление интеграла $\mathcal{J}_{\hat{n}\hat{m}}$ можно обосновать следующим образом. Интеграл

$$\int_{u_n-h/2}^{u_n+h/2} du \int_{v_m-H/2}^{v_m+H/2} dv \sum_{j=1}^3 \frac{\eta_j(x)(y_j(u, v) - x_j)}{|x-y(u, v)|^3} \quad (23)$$

равен интегралу по кусочку поверхности Γ с центром в точке $x = y(u_n, v_m)$, а он, в свою очередь, приближенно равен интегралу от той же функции по кусочку касательной плоскости, проведенной в точке x . Вектор нормали η к поверхности в точке x будет и вектором нормали к касательной плоскости. На касательной плоскости вектор $(y(u, v) - x)$ ортогонален вектору нормали η , поэтому их скалярное произведение, стоящее в числителе в (23), тождественно равно нулю для всех u, v , а значит, и интеграл по кусочку касательной плоскости равен нулю. Поскольку этот интеграл приближенно равен интегралу $\mathcal{J}_{\hat{n}\hat{m}}$, умноженному на взятую со знаком минус длину вектора нормали η в точке x , то можно считать, что и интеграл $\mathcal{J}_{\hat{n}\hat{m}}$ приближенно равен нулю.

7. ЧИСЛЕННЫЕ ТЕСТЫ

Тестирование квадратурных формул для прямого значения нормальной производной потенциала простого слоя для уравнений Лапласа и Гельмгольца проведено в случае, когда поверхность Γ является сферой единичного радиуса, которая задана параметрически уравнениями

$$y_1(u, v) = \sin v \cos u, \quad y_2(u, v) = \sin v \sin u, \quad y_3(u, v) = \cos v, \quad (24)$$

причем $(u, v) \in [0, 2\pi] \times [0, \pi]$. В численных тестах используются явные формулы внутри и вне сферы из [2] для потенциалов простого слоя с заданной плотностью.

Согласно ([13], гл. 5, § 27, п. 7), прямое значение нормальной производной потенциала простого слоя на поверхности Γ можно найти по формуле

$$\left. \frac{\partial \mathcal{V}_k[\mu](x)}{\partial \mathbf{n}_x} \right|_{\Gamma} = \frac{1}{2} \left[\left. \frac{\partial \mathcal{V}_k[\mu](x)}{\partial \mathbf{n}_x} \right|_{\Gamma^+} + \left. \frac{\partial \mathcal{V}_k[\mu](x)}{\partial \mathbf{n}_x} \right|_{\Gamma^-} \right].$$

Здесь поверхность Γ рассматривается как двусторонняя, через Γ^- обозначена сторона, которую мы видим, глядя навстречу вектору нормали \mathbf{n}_x , а через Γ^+ обозначена противоположная сторона. В формуле берутся предельные значения нормальной производной потенциала простого слоя на разных сторонах Γ . Отметим, что направление единичной нормали \mathbf{n}_x совпадает с направлением нормали η , так как вектор \mathbf{n}_x получается из η в результате нормировки. Пусть теперь Γ — единичная сфера, заданная параметризацией (24), тогда формулы (2) для нормали η определяют внутреннюю нормаль на сфере, а значит, Γ^- — внутренняя сторона единичной сферы, а Γ^+ — ее внешняя сторона.

В тестах точное прямое значение нормальной производной потенциала простого слоя в узловых точках сравнивалось с приближенными значениями, вычисленными по квадратурным формулам — по улучшенной формуле (21), по формуле из [4] и по стандартной формуле (22). В каждой узловой точке вычислялась абсолютная погрешность по обеим формулам. Вычисления проводились для разных значений M и N . Значения шагов определяются формулами $h = 2\pi / N$, $H = \pi / M$. Если $N / 2 = M = 25$, то $h = H \approx 0.13$; если $N / 2 = M = 50$, то $h = H \approx 0.063$; если $N / 2 = M = 100$, то $h = H \approx 0.031$. В табл. 1 для каждого теста приводится максимум абсолютной погрешности вычислений по всем узловым точкам сферы. В первой строке таблицы указаны значения N , M , в последующих строках — максимальные погрешности для стандартной формулы (22), для формулы из [4] и для улучшенной квадратурной формулы (21) в каждом тесте.

Тест 1 (для квадратурных формул в случае уравнения Лапласа). В данном тесте использовалась плотность потенциала $\mu(y(u, v)) = 1$, тогда гармонический потенциал простого слоя и прямое значение его нормальной производной имеют вид

Таблица 1. Максимальная абсолютная погрешность квадратурных формул в тестах 1–5

Номер теста	Квадратная формула	$M = N / 2 = 25$	$M = N / 2 = 50$	$M = N / 2 = 100$
1	стандартная формула	0.019	0.0097	0.0062
1	формула из [4]	0.0042	0.0024	0.0013
1	улучшенная формула	0.0028	0.0016	0.00084
2	стандартная формула	0.019	0.0097	0.0049
2	формула из [4]	0.00041	0.0001	2.6E-5
2	улучшенная формула	0.00032	8.3E-5	2.1E-5
3	стандартная формула	0.011	0.0089	0.0062
3	формула из [4]	0.0044	0.0025	0.0013
3	улучшенная формула	0.003	0.0016	0.00085
4	стандартная формула	0.019	0.0097	0.0062
4	формула из [4]	0.0041	0.0024	0.0013
4	улучшенная формула	0.0026	0.0015	0.00083
5	стандартная формула	0.011	0.0089	0.0062
5	формула из [4]	0.0044	0.0025	0.0013
5	улучшенная формула	0.003	0.0016	0.00085

$$\mathcal{V}_0[\mu](x) = \begin{cases} 1 & \text{при } |x| < 1, \\ \frac{1}{|x|} & \text{при } |x| > 1, \end{cases} \quad \left. \frac{\partial \mathcal{V}_0[\mu](x)}{\partial \mathbf{n}_x} \right|_{|x|=1} = \frac{1}{2}.$$

Тест 2 (для квадратурных формул в случае уравнения Лапласа). В данном тесте использовалась плотность потенциала $\mu(y(u, v)) = \cos u \sin v$, тогда гармонический потенциал простого слоя и прямое значение его нормальной производной имеют вид

$$\mathcal{V}_0[\mu](x) = \begin{cases} \frac{|x| \cos \varphi \sin \vartheta}{3}, & \text{если } |x| < 1, \\ \frac{\cos \varphi \sin \vartheta}{3|x|^2}, & \text{если } |x| > 1, \end{cases} \quad \left. \frac{\partial \mathcal{V}_0[\mu](x)}{\partial \mathbf{n}_x} \right|_{|x|=1} = \frac{\cos \varphi \sin \vartheta}{6},$$

где φ и ϑ — азимутальный и зенитный углы в сферических координатах с началом в центре сферы.

Тест 3 (для квадратурных формул в случае уравнения Лапласа). В данном тесте использовалась плотность потенциала $\mu(y(u, v)) = \cos v$, тогда гармонический потенциал простого слоя и прямое значение его нормальной производной имеют вид

$$\mathcal{V}_0[\mu](x) = \begin{cases} \frac{|x| \cos \vartheta}{3} & \text{при } |x| < 1, \\ \frac{\cos \vartheta}{3|x|^2} & \text{при } |x| > 1, \end{cases} \quad \left. \frac{\partial \mathcal{V}_0[\mu](x)}{\partial \mathbf{n}_x} \right|_{|x|=1} = \frac{\cos \vartheta}{6},$$

где ϑ — зенитный угол в сферических координатах с началом в центре сферы.

Тест 4 (для квадратурных формул в случае уравнения Гельмгольца). В данном тесте использовалась плотность потенциала $\mu(y(u, v)) = k$, тогда потенциал простого слоя и прямое значение его нормальной производной имеют вид

$$\mathcal{V}_k[\mu](x) = \begin{cases} e^{ik} \frac{\sin(k|x|)}{|x|}, & \text{если } |x| < 1, \\ \sin k \frac{e^{ik|x|}}{|x|}, & \text{если } |x| > 1, \end{cases}$$

$$\left. \frac{\partial \mathcal{V}_k[\mu](x)}{\partial \mathbf{n}_x} \right|_{|x|=1} = -\frac{e^{ik} \sin k}{2} (k \operatorname{ctg} k - 2 + ik),$$

где $k = 1$.

Тест 5 (для квадратурных формул в случае уравнения Гельмгольца). В данном тесте использовалась плотность потенциала $\mu(y(u, v)) = k^3 \cos v$, тогда потенциал простого слоя и прямое значение его нормальной производной имеют вид

$$\mathcal{V}_k[\mu](x) = \begin{cases} (ik - 1) e^{ik} \frac{k|x| \cos(k|x|) - \sin(k|x|)}{|x|^2} \cos \vartheta, & \text{если } |x| < 1, \\ (k \cos k - \sin k) \frac{(ik|x| - 1) e^{ik|x|}}{|x|^2} \cos \vartheta, & \text{если } |x| > 1, \end{cases}$$

$$\left. \frac{\partial \mathcal{V}_k[\mu](x)}{\partial \mathbf{n}_x} \right|_{|x|=1} = -\frac{e^{ik} \cos \vartheta}{2} \left(\sin k(-ik^3 + 2k^2 + 4ik - 4) + \cos k(-k^3 - 4ik^2 + 4k) \right),$$

где ϑ — зенитный угол в сферических координатах с началом в центре сферы, $k = 1$.

Результаты расчетов в приведенных тестовых примерах показывают, что улучшенная квадратурная формула (21) имеет первый порядок сходимости, как и формула из [4]. При этом погрешность вычислений по улучшенной квадратурной формуле (21) в несколько раз меньше, чем погрешность вычислений по стандартной квадратурной формуле и в полтора раза меньше погрешности вычислений по формуле из работы [4]. Тем самым, улучшенная квадратурная формула обеспечивает значительно более высокую точность вычислений прямого значения нормальной производной потенциала простого слоя.

Отметим, что в тестовых примерах погрешность вычислений по улучшенной квадратурной формуле возрастает к полюсам сферы, которые являются особыми точками в силу выбранной параметризации (24). Вычисления по улучшенной квадратурной формуле (21) в тесте 2 показывают более высокую точность, так как плотность в потенциале простого слоя и прямое значение его нормальной производной обращаются в нуль на полюсах сферы. В то же время погрешность стандартной формулы достигает максимума на экваторе сферы, поэтому во втором тесте такого же эффекта для этой формулы не наблюдается.

Улучшенная квадратурная формула (21) может найти применение при численном решении граничных интегральных уравнений, возникающих на первом этапе решения краевых задач для уравнений Лапласа и Гельмгольца методом потенциалов и граничных интегральных уравнений.

СПИСОК ЛИТЕРАТУРЫ

1. Бреббия К., Теллес Ж., Вроубел Л. Методы граничных элементов. М.: Мир, 1987.
2. Крутицкий П. А., Федотова А. Д., Колыбасова В. В. Квадратурная формула для потенциала простого слоя // Дифференц. ур-ния. 2019. Т. 55. № 9. С. 1269–1284.
3. Крутицкий П. А., Резниченко И. О. Улучшенная квадратурная формула для потенциала простого слоя // Ж. вычисл. матем. и матем. физ. 2023. Т. 63. № 2. С. 44–58.
4. Крутицкий П. А., Резниченко И. О., Колыбасова В. В. Квадратурная формула для прямого значения нормальной производной потенциала простого слоя // Дифференц. ур-ния. 2020. Т. 56. № 9. С. 1270–1288.
5. Крутицкий П. А., Колыбасова В. В. Численный метод решения интегральных уравнений в задаче с наклонной производной для уравнения Лапласа вне разомкнутых кривых // Дифференц. ур-ния. 2016. Т. 52. № 9. С. 1262–1276.
6. Krutitskii P. A. The Helmholtz equation in the exterior of slits in a plane with different impedance boundary conditions on opposite sides of the slits // Quarter. Appl. Math. 2009. V. 67. N 1. P. 73–92.
7. Krutitskii P. A. On the electric current from electrodes in a magnetized semiconductor film // IMA J. Appl. Math. 1998. V. 60. P. 285–297.
8. Krutitskii P. A. The Dirichlet problem for the two-dimensional Laplace equation in a multiply connected domain with cuts // Proc. Edinburgh Math. Soc. 2000. V. 43. N 2. P. 325–341.
9. Крутицкий П. А. Смешанная задача для уравнения Лапласа вне разрезов на плоскости // Дифференц. ур-ния. 1997. Т. 33. № 9. С. 1181–1190.
10. Krutitskii P. A. The Neumann problem on wave propagation in a 2-D external domain with cuts // J. Math. Kyoto Univ. 1998. V. 38. N 3. P. 439–452.
11. Krutitskii P. A. The 2-dimensional Dirichlet problem in an external domain with cuts // Zeitschrift fur Analysis und ihre Anwendungen. 1998. V. 17. № 2. P. 361–378.
12. Бутузов В. Ф., Крутицкая Н. Ч., Медведев Г. Н., Шишкин А. А. Математический анализ в вопросах и задачах. М.: Физматлит, 2000.
13. Владимиров В. С. Уравнения математической физики. М.: Физматлит, 1981.
14. Прудников А. П., Брычков Ю. А., Маричев О. И. Интегралы и ряды. М.: Физматлит, 1981.
15. Крутицкий П. А., Федотова А. Д., Колыбасова В. В. О квадратурной формуле для потенциала простого слоя в трехмерном случае // Препринт ИПМ им. М. В. Келдыша РАН. 2019. № 112. 26 с.
16. Градштейн И. С., Рыжик И. М. Таблицы интегралов, сумм, рядов и произведений. М.: Физматлит, 1963.
17. Ильин В. А., Позняк Э. Г. Основы математического анализа. Часть 1 и 2. М.: Физматлит, 1973.

УДК 519.65

ПРИБЛИЖЕНИЕ НЕПРЕРЫВНЫХ ФУНКЦИЙ С ПОМОЩЬЮ КЛАССИЧЕСКИХ СИНКОВ И ЗНАЧЕНИЙ ОПЕРАТОРОВ C_λ

© 2024 г. В. Н. Пасечник^{1, *}

¹410012 Саратов, ул. Астраханская, 83, Саратовский национальный исследовательский государственный университет им. Н. Г. Чернышевского, Россия

*e-mail: kas.sy@yandex.ru

Поступила в редакцию 05.05.2023 г.

Переработанный вариант 29.09.2023 г.

Принята к публикации 19.10.2023 г.

Рассмотрены свойства синк-приближений. Используемые ранее классические синк-аппроксимации давали плохое приближение, а новый оператор, обобщающий синк-аппроксимации, справляется с приближением этой функции лучше. Приведен график численной реализации эксперимента. Библ. 22. Фиг 2.

Ключевые слова: синк-аппроксимация, оператор интерполяции, приближение непрерывной функции, численный эксперимент, приближение синками, приближение оператором.

DOI: 10.31857/S0044466924020034, EDN: YKPOPD

ВВЕДЕНИЕ

Настоящая статья посвящается изучению свойств синк-приближений. Ранее они использовались в теореме отсчетов Уиттекера–Котельникова–Шеннона (см. [1]–[4]). Для развития теории кодирования сигналов Э. Борель и Э. Т. Уиттекер ввели понятие кардинальной функции, сужение которой с оси на отрезок $[0, \pi]$ выглядит следующим образом:

$$L_n(f, x) = \sum_{k=0}^n \frac{(-1)^k \sin nx}{nx - k\pi} f\left(\frac{k\pi}{n}\right). \quad (1)$$

На сегодняшний день достаточно тщательно изучены свойства синк-аппроксимации аналитической функции на действительной оси, экспоненциально убывающей на бесконечности. В [3] доступен наиболее полный обзор результатов, достигнутых в этом направлении до 1993 г., и несколько важных применений синк-приближений.

Синк-приближения могут широко использоваться при построении различных численных методов в математической физике и при аппроксимациях функций как от одной, так и от нескольких переменных (см. [5], [6]) в теории квадратурных формул (см. [3]) и теории вейвлет-изменений или всплесков (см. [1], [2], [4]). В [7] изучаются изменения в синк-приближениях (1), с помощью которых могут быть аппроксимированы произвольные равномерно непрерывные функции, ограниченные на оси.

Результаты [8] предполагают, что при использовании классических синк-приближений (1) явление Уилбрахама–Гиббса возникает ближе к концу отрезка $[0, \pi]$.

До публикации работ [8], [9], насколько известно, приближения определенных классов аналитических функций (см. [3], [10]) на ограниченном отрезке или интервале с помощью оператора (1) применялась его композиция с конформным отображением. В [9] мы получаем верхнюю оценку наилучшего приближения непрерывных, исчезающих на конце отрезка $[0, \pi]$, функций с помощью линейных комбинаций синков.

Из результатов [11] видно, что при попытке аппроксимации негладких непрерывных функций операторными значениями (1) может возникнуть “резонанс”, приводящий к неограниченному увеличению погрешности аппроксимации по всему интервалу $(0, \pi)$. В той же работе [11] показано отсутствие равносходимости значений операторов (1) и рядов или интегралов Фурье в классе непрерывных функций.

Работы [12], [13] показывают различные изменения синк-аппроксимаций (1), которые позволяют аппроксимировать произвольные непрерывные функции на $[0, \pi]$. Основываясь на исследовании полноты системы синков (1) в [12] в пространствах $C[0, \pi]$ и $C_0[0, \pi] = \{f : f \in C[0, \pi], f(0) = f(\pi) = 0\}$, можно сделать вывод, что попытки построить оператор в форме линейных комбинаций синков, которые допускают равномерную аппроксимацию произвольной непрерывной функции на отрезке, бесполезны. Работы [12], [13], дополнены новыми необходимыми и достаточным условиями равномерной сходимости синк-приближений (1) и некоторыми их изменениям на всем отрезке $[0, \pi]$.

Работа [14] используется для изучения приближенных свойств операторов интерполяции, которые построены по решениям задачи Коши для дифференциальных выражений второго порядка. Операторы, предложенные в [14], являются обобщением классических синк-приближений (1). В работе [15] многие приложения результатов [14] используются для изучения приближенных свойств классических алгебраических интерполяционных многочленов Лагранжа с матрицей интерполяционных узлов, каждая строка которых состоит из нулей многочленов Якоби $P_n^{\alpha_n, \beta_n}$ с параметрами, которые зависят от n .

Начиная с хорошо известной работы Крамера [16], изучаются аналоги теорем отсчетов для операторов интерполяции Лагранжа в узлах из спектра задачи Штурма–Лиувилля.

Процессы интерполяции Лагранжа, построенные с помощью собственных функций задачи Штурма–Лиувилля, тесно связаны с приближениями синк-аппроксимаций. Г. И. Натансон в [17] получил признак равномерной сходимости Дини–Липшица в интервале $(0, \pi)$, т.е. равномерно на любом компакте из $(0, \pi)$, процессов Лагранжа–Штурма–Лиувилля.

В проведенных исследованиях [18], [19] было показано, что при сколь угодно малом изменении параметров задачи Штурма–Лиувилля (потенциала q , или констант h, H), приближенные свойства процессов Лагранжа–Штурма–Лиувилля могут сильно различаться. В [20] установлено существование функции, непрерывной на $[0, \pi]$, у которой процесс интерполяции Лагранжа–Штурма–Лиувилля неограниченно расходится почти всюду на $[0, \pi]$.

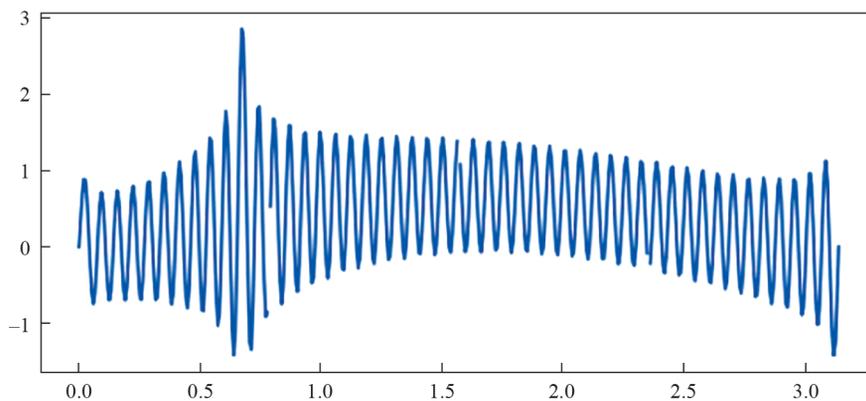
ОСНОВНОЙ РЕЗУЛЬТАТ

Суть эксперимента, приведенного в данной статье, заключается в том, чтобы показать, что для непрерывной функции

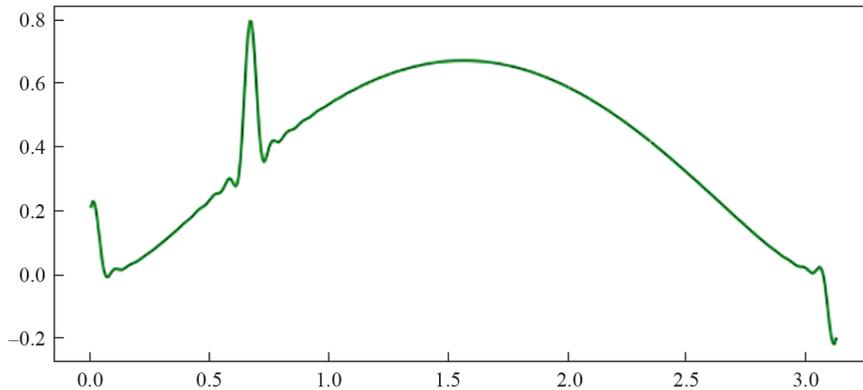
$$f(x) := \sin x e^{\frac{1}{x(x-\pi)}} + \frac{1}{\sqrt{\ln n}} \cos nx \left| \frac{x - \frac{(k_0 + 0.5)\pi}{n}}{x - \frac{(k_0 + 0.5)\pi}{n}} \right| \tag{2}$$

классические синк-аппроксимации дают плохое приближение, а новый оператор

$$C_\lambda(f, x) = \frac{1}{4} \sum_{k=1}^{n-1} \left(\begin{matrix} s_{k+1, \lambda}(x) + 2s_{k, \lambda}(x) + \\ + s_{k-1, \lambda}(x) \end{matrix} \right) f(x_{k, \lambda}), \tag{3}$$



Фиг. 1.



Фиг. 2.

обобщающий синк-аппроксимации, справляется с приближением этой функции лучше. Первое слагаемое в определении функции (2) есть функция, которую приближаем, второе-симуляция помех, которые мешают приближению синками. Но новый оператор (3) позволяет с ними справиться.

Далее приведем график численной реализации эксперимента.

На фиг. 1 изображен график приближения функции (2) с помощью синк-аппроксимаций (1). На фиг. 2 изображен график приближения функции (2) оператором (3).Приведем следующее утверждение.

Теорема 1. Пусть $x_0 \in (0, \pi)$ и $f \in C^1[0, \pi]$, тогда существует последовательность непрерывных на отрезке $[0, \pi]$ функций ψ_n , норма в пространстве Чебышёва $C[0, \pi]$ которых меньше единицы $\|\psi_n\|_{C[0, \pi]} < 1$, такая, что

$$\overline{\lim}_{n \rightarrow \infty} L_n(f + \psi_n, x_0) = \infty.$$

Доказательство. Возьмем $x_0 \in (0, \pi)$, для любого $n \in N$ положим $k_0 = \left\lfloor \frac{nx_0}{\pi} \right\rfloor$,

$$\varphi_n(x) = \cos nx \frac{x - \frac{(k_0 + 0.5)\pi}{n}}{\left| x - \frac{(k_0 + 0.5)\pi}{n} \right|}.$$

Тогда для любого $n \in N$ справедливо соотношение

$$L_n(\varphi_n, x_0) = L_n(x_0),$$

где $L_n(x)$ – функция Лебега [21] интерполяционного процесса (1). Отсюда и [21, Теорема 1, формула 5] следует оценка

$$L_n(\varphi_n, x_0) \geq C_0 |\sin nx_0| \ln n. \tag{4}$$

Решая тригонометрические неравенства, убедимся, что для любого $x_0 \in (0, \pi)$ существует бесконечная последовательность натуральных чисел $\{n_k\}_{k=1}^{\infty}$, для которой будут выполняться соотношения

$$|\sin n_k x_0| \geq \frac{\sqrt{2}}{2}, k = 1, 2, \dots$$

Таким образом, в силу (4), положив $\psi_n = \varphi_n / \sqrt[n]{\ln n}$, для любой точки $x_0 \in (0, \pi)$ получим соотношение

$$\overline{\lim}_{n \rightarrow \infty} L_n(\psi_n, x_0) = \infty.$$

Из [22, Теорема 1] следует, что значения оператора (1) равномерно внутри интервала $(0, \pi)$ приближают любую непрерывно дифференцируемую на отрезке $[0, \pi]$ функцию f (в том числе и функцию $f(x) = \sin x e^{\frac{1}{x(x-\pi)}}$). Действительно, в силу формулы Лагранжа для непрерывно дифференцируемых функций справедлива оценка

$$\begin{aligned} & \max_{0 \leq p \leq n} \left| \sum_{m=1}^{\lfloor \frac{n-1}{2} \rfloor} \frac{f\left(\frac{\pi(2m+1)}{n}\right) - 2f\left(\frac{2\pi m}{n}\right) + f\left(\frac{\pi(2m-1)}{n}\right)}{p-2m} \right| \leq \\ & \leq \max_{0 \leq p \leq n} \left| \sum_{m=1}^{\lfloor \frac{n-1}{2} \rfloor} \frac{\left| f\left(\frac{\pi(2m+1)}{n}\right) - f\left(\frac{2\pi m}{n}\right) \right| + \left| f\left(\frac{\pi(2m-1)}{n}\right) - f\left(\frac{2\pi m}{n}\right) \right|}{p-2m} \right| = \\ & = \frac{\|f'\|_{C[0,\pi]}}{n} O \left(\max_{0 \leq p \leq n} \left| \sum_{m=1}^{\lfloor \frac{n-1}{2} \rfloor} \frac{1}{p-2m} \right| \right) = \|f'\|_{C[0,\pi]} O \left(\frac{\ln n}{n} \right), \end{aligned}$$

где штрих у суммы означает отсутствие в сумме слагаемого со знаменателем равным нулю. Откуда следует условие критерия равномерной сходимости [22, Теорема 1]

$$\lim_{n \rightarrow \infty} \max_{0 \leq p \leq n} \left| \sum_{m=1}^{\lfloor \frac{n-1}{2} \rfloor} \frac{f\left(\frac{\pi(2m+1)}{n}\right) - 2f\left(\frac{2\pi m}{n}\right) + f\left(\frac{\pi(2m-1)}{n}\right)}{p-2m} \right| = 0.$$

Теперь утверждение теоремы 1 вытекает из линейности оператора (1). Теорема 1 доказана.

Замечание 1. Нормы операторов (3) ограничены в совокупности (см. [14] из списка процитированных в статье источников). И, следовательно, по теореме Банаха–Штейнгауза с их помощью можно аппроксимировать равномерно внутри интервала $(0, \pi)$ произвольный элемент чебышёвского пространства $C[0, \pi]$.

Эти утверждения иллюстрируются численным экспериментом.

СПИСОК ЛИТЕРАТУРЫ

1. *Кашин Б. С., Саакян А. А.* Ортогональные ряды. М.: Изд-во АФЦ, 1999.
2. *Новиков И. Я., Стечкин С. Б.* Основы теории всплесков // УМН. 1998. С. 53–128.
3. *Stenger F.* Numerical Methods based on Sinc and analytic functions // Springer Ser. Comput. Math. 20 Springer, 1993.
4. *Добеши И.* Десять лекций по вейвлетам., Ижевск: Регулярная и хаотическая динамика, 2001.
5. *Livne O. E., Brandt A. E.* MuST: The multilevel sine transform // SIAM J. Sci. Comput. 2011. V. 38. N 4. P. 1726–1738.
6. *Marwa M.* Tharwat Sinc approximation of eigenvalues of Sturm – Liouville problems with a Gaussian multiplier Calcolo: a quarterly on numerical analysis and theory of computation. 2014. V. 51. N. 3. P. 465–484.
7. *Kivinukk A.* Tamberg G, Interpolating generalized Shannon sampling operators, their norms and approximation properties // Sampl. Theory Signal Image Process. 2009. V. 8. N 1. P. 77–95.
8. *Trynin A. Yu., Sklyarov V. P.* Error of sine approximation of analytic functions on an interval // Sampling Theory in Signal and Image Processing. 2008. V. 7. N 34. P. 263–270.

9. *Sklyarov V. P.* On the best uniform sinc-approximation on a finite interval // East J. Approximat. 2008. V. 14. N 2. P. 183–192.
10. *Mohsen A., El-Gamel M.* A Sine-Collocation method for the linear Fredholm integro-differential equations // Z. angew. Math. Phys. 2006. P. 1–11, <https://doi.org/10.1007/s00033-006-5124-5>.
11. *Трынин А. Ю.* О расходимости синк-приближений всюду на $(0, \pi)$ // Алгебра и анализ. 2010. Т. 22. N 4. С. 232–256.
12. *Трынин А. Ю.* О необходимых и достаточных условиях сходимости синк-аппроксимаций // Алгебра и анализ. 2015. Т. 27. № 5. С. 170–194.
13. *Трынин А. Ю.* Приближение непрерывных на отрезке функций с помощью линейных комбинаций синков // Изв. высш. уч. заведений. Математика. 2016. № 3. С. 72–81.
14. *Трынин А. Ю.* Обобщение теоремы отсчетов Уиттекера–Котельникова–Шеннона для непрерывных функций на отрезке // Матем. сб. 2009. С. 61–108.
15. *Трынин А. Ю.* Об операторах интерполирования по решениям задачи Коши и многочленах Лагранжа–Якоби // Изв. РАН. Сер. матем. 2011. Т. 75. № 6. С. 129–162.
16. *Kramer H. P.* A generalized sampling theorem // J. Math. Phys. 1959. V. 38. P. 68–72.
17. *Натансон Г. И.* Об одном интерполяционном процессе // Учен. записки Лепинград, пед. ин-та. 1958. Т. 166. С. 213–219.
18. *Трынин А. Ю.* Об отсутствии устойчивости интерполирования по собственным функциям задачи Штурма–Лиувилля // Изв. высш. уч-ых заведений. Математика. 2000. Т. 9. № 460. С. 60–73.
19. *Трынин А. Ю.* Об одной обратной узловых задаче для оператора Штурма–Лиувилля // Уфимск. матем. журн. 2013. Т. 5. № 4. С. 116–129.
20. *Трынин А. Ю.* О расходимости интерполяционных процессов Лагранжа по собственным функциям Задачи Штурма–Лиувилля // Изв. высш. уч-ых заведений. Математика. 2010. № 11. С. 74–85.
21. *Трынин А. Ю.* Оценки функций Лебега и формула Неваи для sinc-приближений непрерывных функций на отрезке // Сиб. матем. журн. 2007. Т. 48. № 5. С. 1155–1166.
22. *Трынин А. Ю.* Критерий равномерной сходимости sinc-приближений на отрезке // Изв. высш. уч-ых заведений. Математика. 2008. № 6. С. 66–78.

СИНТЕЗ ОПТИМАЛЬНОЙ УСТОЙЧИВОЙ АФФИННОЙ СИСТЕМЫ

© 2024 г. Л. Т. Ащепков^{1,*}¹690922 Владивосток, о. Русский, ДВФУ, Россия

*e-mail: ashchepkov@yahoo.com

Поступила в редакцию 22.08.2023 г.

Переработанный вариант 08.09.2023 г.

Принята к публикации 20.10.2023 г.

Предложен метод построения обратной связи, обеспечивающей притяжение траекторий аффинной системы к состоянию равновесия и заданному многообразию. Искомая обратная связь находится в аналитической форме как решение вспомогательной задачи оптимального управления. Приведены достаточные условия существования оптимального управления. Показано применение метода к некоторым классам линейных и нелинейных систем. Библ. 10.

Ключевые слова: притяжение траекторий, устойчивость, достаточные условия оптимальности, синтез системы, существование оптимального управления.

DOI: 10.31857/S0044466924020046, EDN: YKIQMM

1. ВВЕДЕНИЕ

Настоящая статья посвящена вопросам обеспечения устойчивости аффинных систем, описываемых обыкновенными дифференциальными уравнениями с аффинными по управляющим воздействиям правыми частями. Они широко используются в различных технических областях в качестве моделей управляемых объектов (см. [1], [2]). В фазовых пространствах аффинных систем могут находиться особые многообразия, которые вместе с любой точкой содержат все порождаемые управлениями траектории, проходящие через эту точку. Особые многообразия формируют разную «устойчивость» аффинных систем: в классическом понимании, по части переменных (см. [3]) или в форме притяжения траекторий к состоянию равновесия.

Тематически настоящая статья примыкает к публикациям [4], [5]. Новизну исследования составляют постановка основной задачи, ее формулировка в терминах вспомогательной задачи оптимального управления и полученные выводы. Условия вспомогательной задачи учитывают цель управления, возможность задания особого многообразия и дополнительно обеспечивают гладкость искомого управления на всем фазовом пространстве, за исключением состояния равновесия. В результате упрощается анализ задачи, поскольку исключаются из рассмотрения разрывные управления и сопутствующие им скользящие режимы (см. [6]).

2. ОСНОВНАЯ ЗАДАЧА

Рассмотрим в фазовом пространстве \mathbf{R}^n систему обыкновенных дифференциальных уравнений

$$\dot{x} = f(x) + B(x)u. \quad (1)$$

Здесь $x \in \mathbf{R}^n$ – вектор фазового состояния, $u \in \mathbf{R}^m$ – вектор управляющих воздействий, $\dot{x} = dx/dt$, $f: \mathbf{R}^n \rightarrow \mathbf{R}^n$ и $B: \mathbf{R}^n \rightarrow \mathbf{R}^{n \times m}$ – векторная и матричная функции с гладкими (непрерывными и дифференцируемыми) компонентами, $f(0) = 0$, $m \leq n$. Решение $x = 0$ системы дифференциальных уравнений $\dot{x} = f(x)$ назовем состоянием равновесия.

Пусть в фазовом пространстве векторным уравнением $s(x) = 0$ с гладкой функцией $s: \mathbf{R}^n \rightarrow \mathbf{R}^p$, $p \leq n$, задано многообразие M , содержащее состояние равновесия. Основная задача состоит в нахождении управления $u = u(x)$, обеспечивающего асимптотическое притяжение траекторий замкнутой системы

$$\dot{x} = f(x) + B(x)u(x), \quad x(0) = x_0, \quad (2)$$

к состоянию равновесия и многообразию M для некоторого множества начальных состояний $x_0 \in \mathbb{R}^n$.

Поясним специфику основной задачи на простом примере аффинной системы

$$\dot{x}_1 = x_1 + u, \quad \dot{x}_2 = 0, \quad x_1(0) = x_{10}, \quad x_2(0) = x_{20}$$

с произвольными начальными условиями. Легко видеть, что все траектории системы независимо от выбора управления находятся на прямой $x_2 = x_{20}$ фазовой плоскости. Если $x_{20} \neq 0$, то приблизить с помощью управления траекторию системы сколь угодно близко к состоянию равновесия $(0, 0)$ невозможно. Примем за многообразие M координатную ось $x_2 = 0$ фазовой плоскости. Тогда одним из решений задачи будет управление $u(x_1, x_2) = -2x_1$, обеспечивающее асимптотическое притяжение траектории системы к состоянию равновесия по оси M из любого начального состояния $(x_{10}, 0) \in M$.

В связи с не единственностью решения основной задачи возникает вопрос о выборе наилучшего решения по некоторому заданному критерию. В приведенной ниже вспомогательной задаче в качестве такого критерия используется скорость убывания функции Ляпунова вдоль траектории системы.

3. ВСПОМОГАТЕЛЬНАЯ ЗАДАЧА

Свяжем с основной задачей вспомогательную задачу оптимального управления

$$\begin{aligned} J &= \int_0^{\infty} \alpha V(x(t)) dt \rightarrow \min, \\ \dot{x} &= f(x) + B(x)u, \\ x(0) &= x_0, \quad x(\infty) = 0, \\ \alpha V(x) + \dot{V}(x, u) &\geq 0, \quad t \geq 0. \end{aligned} \quad (3)$$

Здесь α – заданный положительный параметр, $x(\infty)$ – предел функции $x(t)$ при $t \rightarrow \infty$,

$$V(x) = \mu \|x\|^2 + \nu \|s(x)\|^2, \quad x \in \mathbb{R}^n, \quad (4)$$

есть скалярная положительно-определенная функция с постоянными положительными коэффициентами μ, ν ,

$$\dot{V}(x, u) = V_x(x)' (f(x) + B(x)u), \quad x \in \mathbb{R}^n, \quad u \in \mathbb{R}^m, \quad (5)$$

есть производная V в силу системы (1). В формулах (4), (5) и далее все векторы столбцовые, норма векторов евклидова, штрих – знак транспонирования, V_x – градиент функции V .

Обозначим через $x(t), t \geq 0$, произвольное непрерывное кусочно-гладкое решение системы (1), соответствующее некоторому кусочно-непрерывному управлению $u(t), t \geq 0$. Пару $x(t), u(t)$ назовем *процессом*, если она определена на всей полуоси $t \geq 0$ и удовлетворяет краевым условиям и смешанному ограничению (3). Вспомогательная задача состоит в нахождении оптимального процесса с наименьшим значением целевого функционала.

Далее предполагаем

$$\begin{aligned} s(0) &= 0, \quad \text{rank } s_x(x) = p, \quad x \in M, \\ B(x)' V_x(x) &\neq 0, \quad x \neq 0. \end{aligned} \quad (6)$$

Строки матрицы s_x образованы градиентами координат s_1, \dots, s_p векторной функции s .

При $p < n$ первые два условия (6) и теорема о неявной функции (см., например, [7, 5], с. 298) гарантируют существование и единственность гладкого многообразия M в окрестности начала координат. Если $p = n$ и $s(x) = x$, то многообразие M совпадает с состоянием равновесия, и вспомогательная задача не исключает классической асимптотической устойчивости тривиального решения системы (1) при

$u = 0$. Последнее предположение (6) используется для построения управления $u(x)$ и подробно рассматривается в разд. 5.

Остановимся вкратце на смысле условий вспомогательной задачи (3). Функция $V(x)$ играет роль показателя близости фазовой точки x к состоянию равновесия и многообразию M . Целевое требование предполагает уменьшение показателя V на траекториях аффинной системы. Краевые условия отвечают цели управления в основной задаче. Смешанное ограничение определяет множество допустимых скоростей изменения показателя V в силу аффинной системы.

4. РЕШЕНИЕ ВСПОМОГАТЕЛЬНОЙ ЗАДАЧИ

Воспользуемся достаточными условиями оптимальности (см. [8]). Представим (см. [5]) функционал J на произвольном фиксированном процессе $x(t), u(t)$ в виде

$$J = \int_0^\infty \alpha V(x(t)) dt + \int_0^\infty (V(x(t)))^\bullet dt - V(x(t)) \Big|_{t=0}^{t=\infty}. \tag{7}$$

Выполним во втором интеграле (7) дифференцирование по переменной t . Учитывая (3), (5), получим нижнюю оценку функционала

$$J = V(x_0) + \int_0^\infty (\alpha V(x(t)) + \dot{V}(x(t), u(t))) dt \geq V(x_0). \tag{8}$$

На рассматриваемом процессе оценка (8) достижима, если подынтегральная функция равна нулю во всех точках $t \geq 0$ непрерывности управления $u(t)$. В свою очередь, для этого достаточно, чтобы пара $x(t), u(t)$ в том же смысле удовлетворяла уравнению

$$\alpha V(x) + \dot{V}(x, u) = \alpha V(x) + V_x(x)' (f(x) + B(x)u) = 0. \tag{9}$$

Отсюда в силу последнего предположения (6) находим управление

$$u(x) = - \frac{\alpha V(x) + V_x(x)' f(x)}{\|B(x)' V_x(x)\|^2} B(x)' V_x(x), \quad x \neq 0, \tag{10}$$

определенное и гладкое на всем фазовом пространстве \mathbb{R}^n , кроме начала координат. Примем дополнительно $u(0) = 0$.

Пусть задача Коши (2) с управлением (10) имеет решение $x(t)$ на всей полуоси $t \geq 0$. По построению, пара функций $x(t), u(x(t))$ непрерывна при $t \geq 0$ и тождественно удовлетворяет равенству (9) или, равносильно, – дифференциальному уравнению

$$\alpha V(x(t)) + \dot{V}(x(t), u(x(t))) = \alpha V(x(t)) + (V(x(t)))^\bullet = 0. \tag{11}$$

Решая последнее уравнение (11) с начальным условием $V(x(0)) = V(x_0)$, находим

$$V(x(t)) = V(x_0) e^{-\alpha t}, \quad t \geq 0.$$

Отсюда на основании (4) получим экспоненциальные оценки

$$\begin{aligned} \mu \|x(t)\|^2 &\leq V(x_0) e^{-\alpha t}, \\ \nu \|s(x(t))\|^2 &\leq V(x_0) e^{-\alpha t}, \quad t \geq 0. \end{aligned} \tag{12}$$

Из неравенств (12) вытекают предельные соотношения $x(\infty) = 0, s(x(\infty)) = 0$. Следовательно, пара $x(t), u(x(t))$ образует процесс вспомогательной задачи, и траектория $x(t)$ асимптотически притягивается

к состоянию равновесия и многообразию M . Кроме того, оценка (8) и равенство (11) свидетельствуют об оптимальности процесса $x(t), u(x(t))$. Подведем итоги.

Теорема. Если выполнены предположения разд. 1 и 2, то при любом x_0 оптимальное управление (10) вспомогательной задачи (3) существует и обеспечивает с оценками (12) асимптотическое притяжение оптимальной траектории к состоянию равновесия и многообразию M .

Следствие 1. В условиях теоремы формулой (4) задана функция Ляпунова замкнутой системы (2).

Следствие 2. При $s(x) = x$ из оценок (12) вытекает экспоненциальная устойчивость тривиального решения замкнутой системы (2) для любых x_0 .

Пример. Рассмотрим систему дифференциальных уравнений

$$\dot{x} = f(x) + ux, \quad x(0) = x_0,$$

со скалярным управляющим воздействием u . При $B(x) = x, p = n, s(x) = x, V(x) = \frac{1}{2}x'x$ условия теоремы выполнены. По формуле (10) находим оптимальное управление

$$u(x) = -\frac{\alpha}{2} - \frac{x'f(x)}{x'x}, \quad x \neq 0.$$

Примем $u(0) = 0$. На основании следствия 2 теоремы замкнутая оптимальным управлением система имеет экспоненциально устойчивое тривиальное решение при любом x_0 .

5. СУЩЕСТВОВАНИЕ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

Согласно теореме для существования оптимального управления достаточно, чтобы данные основной и вспомогательной задач отвечали перечисленным в разд. 1 и 2 требованиям. Очевидно, наиболее ограничительными из них являются условия (6). Они заведомо выполняются, если многообразие M совпадает с состоянием равновесия, во всех точках фазового пространства матрица B имеет ранг n и градиент V_x равен нулю только в начале координат.

Укажем другие классы троек, образованных дифференциальными уравнениями и функциями s, V , которые удовлетворяют условиям теоремы. Примем

$$p = n - m, \quad V(x) = \frac{1}{2}(\|x\|^2 + \|s(x)\|^2), \quad x \in \mathbb{R}^n.$$

Для упрощения выкладок введем две прямоугольные матрицы $B \in \mathbb{R}^{n \times m}$ и $C \in \mathbb{R}^{n \times (n-m)}$ со свойствами

$$C'B = 0, \quad B'C = 0, \quad B'B = I, \quad C'C = I, \quad (13)$$

где 0 и I – нулевые и единичные матрицы соответствующих размеров. Выберем, далее, произвольно гладкие векторные функции $f_1: \mathbb{R}^m \rightarrow \mathbb{R}^m$ и $f_2: \mathbb{R}^{n-m} \rightarrow \mathbb{R}^{n-m}$ с условием $f_1(0) = 0$ и $f_2(0) = 0$.

Покажем, что для тройки

$$\begin{aligned} \dot{x} &= Bf_1(B'x) + Cf_2(C'x) + Bu, \\ s(x) &= C'x, \quad V(x) = \frac{1}{2}x'(I + CC')x \end{aligned} \quad (14)$$

условия теоремы верны. Действительно, переходя в соотношениях (14) к новым переменным $y \in \mathbb{R}^m, z \in \mathbb{R}^{n-m}$ по формуле

$$x = By + Cz, \quad (15)$$

получим

$$\begin{aligned} \dot{y} &= f_1(y) + u, \quad \dot{z} = f_2(z), \\ \tilde{s}(z) &= z, \quad \tilde{V}(y, z) = \frac{1}{2}y'y + z'z. \end{aligned} \quad (16)$$

В подпространстве $z = 0$ условия теоремы для тройки (16) справедливы. В соответствии с ее заключением найденное по формуле (10) управление

$$\tilde{u}(y) = -\left(\frac{\alpha}{2} + \frac{y f_1(y)}{y' y}\right) y, \quad y \neq 0, \tag{17}$$

оптимально и обеспечивает в подпространстве $z = 0$ притяжение всех траекторий системы дифференциальных уравнений (16) к состоянию равновесия $y = 0, z = 0$. В силу обратимости преобразования (15) аналогичные утверждения имеют место в подпространстве $C'x = 0$ для тройки (14) и соответствующего (17) управления

$$u(x) = -\left(\frac{\alpha}{2} + \frac{x' B f_1(B'x)}{x' B B' x}\right) B'x, \quad B'x \neq 0.$$

Замечание 1. Два последних условия (13) можно без потери общности заменить требованием линейной независимости столбцов матриц B и C .

Замечание 2. Управление (17) обеспечивает асимптотическую y -устойчивость (см. [3, 2]) тривиального решения системы дифференциальных уравнений (16) при $u = 0$, если устойчиво тривиальное решение подсистемы $\dot{z} = f_2(z)$.

6. ПРИМЕНЕНИЯ ТЕОРЕМЫ

В данном разделе рассматриваются линейная и билинейная системы с отмеченными во введении особенностями. Показано, что с помощью теоремы можно организовать притяжение траекторий этих систем к состоянию равновесия на некоторых подпространствах фазового пространства.

6.1. Линейная стационарная система

Рассмотрим систему линейных дифференциальных уравнений

$$\dot{x} = Ax + Bu \tag{18}$$

с постоянной матрицей $A \in \mathbb{R}^{n \times n}$ и постоянной матрицей $B \in \mathbb{R}^{n \times m}$, введенной в разд. 5. Допустим, ранг блочной матрицы $D = (B, AB, \dots, A^{n-1}B)$ равен $m < n$, т.е. система (18) не отвечает критерию полной управляемости Калмана (см. [9], с. 43). Тогда система однородных алгебраических уравнений $c'D = 0$ имеет $n - m$ линейно независимых решений. Построим в линейной оболочке решений методом ортогонализации (см. [10], с. 233) систему $n - m$ ортонормированных векторов и составим из них матрицу $C \in \mathbb{R}^{n \times (n-m)}$. По построению, матрица C имеет ортонормированные столбцы и вместе с матрицей B удовлетворяет условиям (13) и равенствам $C'AB = 0, \dots, C'A^{n-1}B = 0$.

Заменим в тройке (14) дифференциальные уравнения на уравнения (18) и выполним замену переменных (15). Используя свойства матриц B и C , получим

$$\begin{aligned} \dot{y} &= B'ABu + B'ACz + u, \quad \dot{z} = C'ACz, \\ \tilde{s}(z) &= z, \quad \tilde{V}(y, z) = \frac{1}{2}y'y + z'z. \end{aligned} \tag{19}$$

Легко убедиться, что при $z = 0$ для тройка (19) условия теоремы верны. Пользуясь аналогом формулы (10), находим оптимальное управление

$$\tilde{u}(y) = -\left(\frac{\alpha}{2} + \frac{y'B'ABu}{y'y}\right) y, \quad y \neq 0,$$

при котором все траектории системы дифференциальных уравнений (19), лежащие в подпространстве $z = 0$, притягиваются к состоянию равновесия $y = 0, z = 0$. Используя обратное (15) и тождественное преобразования

$$y = B'x, \quad z = C'x, \quad x = (BB' + CC')x$$

легко сформулировать приведенные выводы в исходных координатах.

6.2. Билинейная система

Рассмотрим систему дифференциальных уравнений вида

$$\dot{x} = BAB'x + C(u_1B_1 + \dots + u_mB_m)C'x. \quad (20)$$

Здесь A и B_1, \dots, B_m – постоянные квадратные матрицы порядков n и $n - m$ соответственно, B, C – прямоугольные матрицы со свойствами (13), u_1, \dots, u_m – координаты вектора u управляющих воздействий. Предположим дополнительно, что матрица $B(z) = (B_1z, \dots, B_mz)$, $z \in \mathbb{R}^{n-m}$, удовлетворяет условию

$$B(z)'z \neq 0, z \neq 0. \quad (21)$$

Для выполнения (21) достаточно, например, чтобы при любом $z \neq 0$ нашлся столбец матрицы $B(z)$, не ортогональный z .

Пусть $x(t), u(t), t \geq 0$, – произвольный фиксированный процесс системы (20) с начальным условием $x(0) = x_0$ и $B'x_0 = 0$. Уравнение (20) на рассматриваемом процессе становится тождественным по t равенством. Умножим это равенство слева на матрицу B' . Учитывая (13) и заданные начальные условия, получим

$$(B'x(t))^\bullet = A(B'x(t)), B'x(0) = 0.$$

Данная однородная система дифференциальных уравнений с тривиальным начальным условием имеет единственное решение $B'x(t) = 0, t \geq 0$. В силу произвольности процесса это значит, что независимо от выбора управления каждая траектория билинейной системы, исходящая из любой точки подпространства $B'x = 0$, лежит в том же подпространстве.

Выполнив преобразование (15), приведем систему (20) и функции

$$s(x) = B'x, V(x) = \frac{1}{2}x'(I + BB')x$$

к виду

$$\begin{aligned} \dot{y} &= Ay, \dot{z} = B(z)u, \\ \tilde{s}(y) &= y, \tilde{V}(y, z) = y'y + \frac{1}{2}z'z. \end{aligned} \quad (22)$$

При $y = 0$ тройка (22) в силу (21) соответствует условиям теоремы. По аналогичной (10) формуле находим оптимальное управление

$$\tilde{u}(z) = -\frac{\alpha z'z}{2\|B(z)'z\|^2} B(z)'z, z \neq 0. \quad (23)$$

Согласно теореме в подпространстве $y = 0$ все траектории замкнутой управлением (23) системы дифференциальных уравнений (22) притягиваются к состоянию равновесия $y = 0, z = 0$.

Как было отмечено ранее, в силу обратимости преобразования (15) полученные выводы можно сформулировать в первоначальных обозначениях.

7. ЗАКЛЮЧЕНИЕ

В настоящей статье предложен метод синтеза аффинной системы, траектории которой асимптотически притягиваются к состоянию равновесия и заданному многообразию. Искомое управление типа обратной связи найдено аналитически с помощью решения вспомогательной задачи оптимального управления. Показано применение метода к некоторым классам линейных и нелинейных аффинных систем.

СПИСОК ЛИТЕРАТУРЫ

1. Уткин В. И., Орлов Ю. В. Системы управления с векторным реле // Автоматика и телемехан. 2019. № 9. С. 143–155. Utkin V. I., Orlov Yu. V. Control systems with vector relays // Automat. Remote Control. 2019. V. 80. N 9. P. 1671–1680.

2. *Уткин В. И.* Системы с переменной структурой: состояние, проблемы, перспективы // Автоматика и телемехан. 1983. № 9. С. 5–25.
3. *Воротников В. И.* Частичная устойчивость и управление: состояние проблемы и перспективы развития // Автоматика и телемехан. 2005. № 4. С. 3–59. *Vorotnikov V. I.* Partial stability and control: The state-of-the-art and development prospects // Autom. Remote Control. 2005. V. 66. N 4. P. 511–561.
4. *Ащепков Л. Т.* Аналитическое конструирование регулятора с амплитудным ограничением // Автоматика и телемехан. 2022. № 7. С. 49–58. *Ashchepkov L. T.* Analytical synthesis of an amplitude-constrained controller // Automat. Remote Control. 2022. V. 83. № 7. P. 1050–1058.
5. *Ащепков Л. Т.* Синтез оптимальной системы с устойчивыми режимами скольжения // Ж. вычисл. матем. и матем. физ. 2023. Т. 63. № 5. С. 731–738. *Ashchepkov L. T.* Synthesis of an optimal system with stable sliding modes // Comput. Math. Math. Phys. 2023. V. 63. N 5. P. 743–750.
6. *Филиппов А. Ф.* Дифференциальные уравнения с разрывной правой частью // Матем. сб. 1960. Т. 51(93). № 1. С. 99–128.
7. *Понтрягин Л. С.* Обыкновенные дифференциальные уравнения. М.: Наука, 1965.
8. *Кротов В. Ф., Букреев В. З., Гурман В. И.* Новые методы вариационного исчисления в динамике полета. М.: Машиностр., 1969.
9. *Калман Р., Фалб П., Арбиб М.* Очерки по математической теории систем. М.: Мир, 1971.
10. *Гантмахер Ф. Р.* Теория матриц. М.: Наука, 1967.

ОБЫКНОВЕННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

УДК 519.624

ГЛАДКИЕ МНОГООБРАЗИЯ ЛЯПУНОВА ДЛЯ АВТОНОМНЫХ СИСТЕМ НЕЛИНЕЙНЫХ ОБЫКНОВЕННЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ И ИХ ПРИМЕНЕНИЕ К РЕШЕНИЮ СИНГУЛЯРНЫХ КРАЕВЫХ ЗАДАЧ

© 2024 г. Н. Б. Конюхова^{1, *}¹19333 Москва, ул. Вавилова, 40, ФИЦ ИУ РАН, Россия

*e-mail: n.konyukhova@gmail.com

Поступила в редакцию 20.09.2023 г.

Переработанный вариант 20.09.2023 г.

Принята к публикации 20.10.2023 г.

Для автономной системы N нелинейных обыкновенных дифференциальных уравнений, рассматриваемой на полубесконечном интервале $T_0 \leq t < \infty$ и обладающей точкой покоя (псевдо)гиперболического типа, изучается n -мерное ($0 < n < N$) устойчивое многообразие решений, или многообразие условной устойчивости по Ляпунову, которое для каждого достаточно большого t существует в фазовом пространстве переменных системы в окрестности ее седловой точки. Гладкая сепаратрисная поверхность седла для такой системы описывается с помощью решения сингулярной задачи типа Ляпунова для системы квазилинейных уравнений с частными производными первого порядка с вырождением по начальным данным. Дается применение результатов к правильной постановке граничных условий на бесконечности и их переносу в конечную точку для автономной системы нелинейных уравнений, в том числе с указанием использования этого подхода в некоторых прикладных задачах. Библ. 26.

Ключевые слова: обыкновенные дифференциальные уравнения, автономная система нелинейных уравнений, стационарная точка типа (псевдо)гиперболического седла, граничные условия на бесконечности, устойчивые многообразия решений, сингулярная задача Ляпунова для системы квазилинейных уравнений с частными производными первого порядка.

DOI: 10.31857/S0044466924020053, EDN: YKGYLK

ВВЕДЕНИЕ. ПРЕДВАРИТЕЛЬНЫЕ СВЕДЕНИЯ

Данная работа является итоговой по теме, указанной в названии: излагаются некоторые результаты [1]–[3] в переработанном, дополненном и исправленном виде и дается их применение к построению гладких устойчивых многообразий решений (УМР) для автономных систем нелинейных обыкновенных дифференциальных уравнений (ОДУ), к постановке и аппроксимации для таких систем сингулярных краевых задач (КрЗ); учитывается также малая доступность издания [3].

Система N нелинейных автономных ОДУ рассматривается на полу-бесконечном интервале $T_0 \leq t < \infty$. Исследуется проблема корректной постановки для нее нелинейных сингулярных КрЗ: на бесконечности задается условие типа стремления решения к стационарной точке системы ОДУ; такое граничное условие называем *правильным*, если в фазовом пространстве переменных ОДУ рассматриваемая стационарная точка является (псевдо)гиперболическим седлом (определение многомерного (псевдо)гиперболического седла для автономных систем нелинейных ОДУ см., например, в [4, п. 2.2] и ниже в разд. 1). В этом случае значения решений системы ОДУ, удовлетворяющих заданному условию на бесконечности, образуют в окрестности стационарной точки фазового пространства переменных нелинейное УМР, инвариантное относительно t . Тем самым, понятие правильности граничного условия связано с возможностью сведения сингулярной задачи к регулярной заменой предельного условия на бесконечности требованием: для достаточно больших конечных значений t ($t \geq T$) значения решений принадлежат УМР.

Если система ОДУ обладает стационарной точкой типа многомерного седла, то для гладких функций гладкое УМР задается с помощью решения сингулярной задачи типа Ляпунова для системы квазилинейных уравнений с частными производными первого порядка с вырождением по начальным данным (о задаче Ляпунова см. [5, с. 153–158], [6, гл.VI]). Даются обобщения теорем Ляпунова на системы уравнений с частными производными с неголоморфными заданными функциями и неголоморфными решениями,

а также на системы типа Гамильтона-Якоби, возникающие при описании гладких УМР (или, по терминологии [7, гл. VI], гладких полей) для векторных автономных систем нелинейных ОДУ второго порядка; описание таких нелинейных гладких полей дано в разд. 3. Описывается применение результатов к правильной постановке и аппроксимации сингулярных КрЗ для автономных систем нелинейных ОДУ: перенос из бесконечности граничного условия типа стремления решения к стационарной точке осуществляется требованием: для достаточно больших конечных значений t ($t \geq T$) значения решений принадлежат УМР. В линейном приближении в качестве граничного условия при конечном $t = T$ задается уравнение касательного подпространства к УМР, взятого в стационарной точке. Обобщается понятие допустимого граничного условия на бесконечности, введенное в [8]–[10] для линейных систем ОДУ. Выделяется класс таких условий как стремления решения попасть по направлению в касательное подпространство к УМР, взятое в точке (псевдо)гиперболического равновесия. Частные случаи задачи Ляпунова для систем нелинейных ОДУ представляют самостоятельный интерес для теории сингулярных задач Коши (ЗК).

Всюду далее a_0, T_0 — фиксированные вещественные постоянные, $a_0 > 0$; входящие в уравнения функции многих переменных предполагаются непрерывными по совокупности переменных в заданной области их изменения, то же относится и к их частным производным по соответствующим переменным, если таковые существуют (при определении классов функций это не будет более оговариваться). Используются обозначения: $I_T = [T, \infty)$, $T \geq T_0$; $\mathbb{R}, \mathbb{C}, \mathbb{N}$ — множества вещественных, комплексных и натуральных чисел соответственно, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, $n \in \mathbb{N}$; \mathbb{K}^n — линейное пространство столбцов высоты n , $|\cdot|$ — какая-либо норма в \mathbb{K}^n , для подчиненной нормы матрицы обозначение аналогично; $\mathbb{L}(\mathbb{K}^n)$ ($\mathbb{L}(\mathbb{K}^n, \mathbb{K}^m)$) — линейное пространство $n \times n$ —матриц ($(m \times n)$ —матриц), E_n — единичная матрица порядка n , $\lambda(A)$ — собственное значение (СЗ) матрицы A , $\partial\sigma / \partial x = (\partial\sigma_i / \partial x_j)_{i=1, \dots, m; j=1, \dots, n}$ — матрица Якоби, $\partial\sigma / \partial x \in \mathbb{L}(\mathbb{K}^n, \mathbb{K}^m)$; $\Omega_x(a) = \{x : x \in \mathbb{K}^n, |x| < a\}$, $\bar{\Omega}_x(a) = \{x : x \in \mathbb{K}^n, |x| \leq a\}$, $a > 0$, т.е. $\Omega_x(a)$ ($\bar{\Omega}_x(a)$) — открытый (замкнутый) поликруг в окрестности начала координат x -пространства, $\bar{\Omega}_{x,y}(a) = \bar{\Omega}_x(a) \times \bar{\Omega}_y(a)$; $C_\mu[0, \infty)$ — банахово пространство, элементами которого $z(t)$ являются пары вектор-функций, $z(t) = \{x(t); y(t)\}$, $x : [0, \infty) \rightarrow \mathbb{K}^n$, $y : [0, \infty) \rightarrow \mathbb{K}^m$, $z : [0, \infty) \rightarrow \mathbb{K}^n \times \mathbb{K}^m$, непрерывных на $[0, \infty)$ и ограниченных с весом $\exp(\mu t)$, где $\mu \geq 0$, а норма в пространстве $C_\mu[0, \infty)$ определена формулой

$$|z|_{C_\mu} = |\{x; y\}|_{C_\mu} = \max \left\{ \begin{array}{l} \sup_{t \in [0, \infty)} [\exp(\mu t)|x(t)|]; \\ \sup_{t \in [0, \infty)} [\exp(\mu t)|y(t)|] \end{array} \right\};$$

$C_l(\bar{\Omega}_{x,y}(a_0))$ (или сокращенно C_{l,a_0}) — класс непрерывных по $\{x, y\}$ вектор-функций $\varphi(x, y)$, $\varphi : \bar{\Omega}_x(a_0) \times \bar{\Omega}_y(a_0) \rightarrow \mathbb{K}^l$; $C_l^k(\bar{\Omega}_{x,\rho}(a_0))$ (или сокращенно C_{l,a_0}^k , $k \geq 1$) — класс вектор-функций $\varphi(x, \rho)$, $\varphi : \bar{\Omega}_{x,\rho}(a_0) \rightarrow \mathbb{K}^l$, обладающих в той же области частными производными по $x_1, \dots, x_n, \rho_1, \dots, \rho_m$ до k -го порядка включительно; $H_{l,a_0} = H_l(\bar{\Omega}_{x,\rho}(a_0))$ — класс голоморфных по $\{x, \rho\}$ в некоторой окрестности поликруга $\bar{\Omega}_{x,\rho}(a_0)$ вектор-функций $\varphi(x, \rho)$ со значениями в \mathbb{C}^l ; $[x]_\nu$ — абсолютно сходящийся степенной ряд по координатам x , $x \in \mathbb{K}^n$, начинающийся с членов степени не ниже ν , $0 \leq \nu$ — целое.

Остальные обозначения будут вводиться по мере необходимости.

1. О ГРАНИЧНЫХ УСЛОВИЯХ НА БЕСКОНЕЧНОСТИ И СУЩЕСТВОВАНИИ УМР

Пусть на бесконечном интервале рассматривается автономная система N нелинейных ОДУ

$$z'(\tau) = Q(z(\tau)), \quad -\infty < \tau < \infty, \tag{1.1}$$

где $z \in \mathbb{K}^N$, $Q \in C_N(G_N)$, G_N — некоторая область в \mathbb{K}^N или все пространство. Пусть z_s — стационарная точка этой системы, $z_s : z_s \in G_N$, $Q(z_s) = 0$, и пусть $Q(z)$ в окрестности точки z_s представима в виде

$$Q(z) = A(z - z_s) + P(z), \quad P(z) = o(z - z_s), \quad |z - z_s| \rightarrow 0. \tag{1.2}$$

Здесь A — постоянная матрица, $A \in \mathbb{L}(\mathbb{K}^N)$, $P(z) \in C_N(G_N)$ и удовлетворяет условию Липшица с постоянной $L_p(G_N)$, которая может быть сделана сколь угодно малой за счет выбора в качестве G_N малой окрестности точки z_s . (Стационарные точки системы (1.1) называют также точками равновесия (equilibrium points), или просто равновесиями (equilibriums).)

Определение 1. Стационарную точку $z = z_s$ системы (1.1) назовем седло-узлом (μ, n) -типа при $\mu \geq 0$ ($(\mu, N - n)$ -типа при $\mu \leq 0$), где $0 \leq n \leq N$, n -целое, если для этих μ, n матрица A в (1.2) имеет n СЗ, удовлетворяющих условию $\operatorname{Re} \lambda(A) < -\mu$, и $N - n$ СЗ, удовлетворяющих условию $\operatorname{Re} \lambda(A) > -\mu$. (Если $n(N - n) \neq 0$ и $\mu = 0$ ($\mu \neq 0$), то такую точку называют в литературе гиперболическим (псевдогиперболическим) седлом или равновесием; см., например, [4, п. 2.2].)

Пусть точка $z = z_s$ является для (1.1) седло-узлом (μ, n) -типа при $\mu \geq 0$ и $0 \leq n \leq N$. Поставим для решений системы (1.1) граничное условие на бесконечности в виде

$$\lim_{\tau \rightarrow \infty} \left\{ \exp(\mu\tau) [z(\tau) - z_s] \right\} = 0 \quad (1.3)$$

и рассмотрим задачу (1.1), (1.3) как сингулярную ЗК. Эта задача обладает, как известно, n -параметрическим семейством решений, значения которых порождают в фазовом пространстве системы (1.1) в окрестности точки $z = z_s$ инвариантное относительно τ n -мерное УМР $M_+^{(n)}$ (см. [11, гл. XIII]). Тогда предельное условие (1.3) эквивалентно требованию: для достаточно больших конечных значений τ ($\tau \geq \tau_\infty$) значения решений принадлежат УМР, т.е.

$$z(\tau) \in M_+^{(n)}, \quad \tau \geq \tau_\infty. \quad (1.4)$$

Будем говорить, что (1.4) есть результат перенесения условия (1.3) из бесконечности в конечную точку τ . Для такого переноса надо задать (точно или приближенно) уравнение $M_+^{(n)}$.

Пусть, например, для системы (1.1) рассматривается сингулярная КрЗ на полуоси $0 \leq \tau < \infty$ с каким-либо условием при $\tau = 0$ и предельным условием (1.3) на бесконечности, где $0 \leq \mu$ фиксировано. Возможность сведения сингулярной КрЗ к эквивалентной регулярной, в том числе для исследования корректности постановки исходной задачи по общему числу краевых условий на обоих концах полубесконечного интервала, порождает

Определение 2. Предельное условие (1.3) назовем правильным граничным условием на бесконечности для системы (1.1), если стационарная точка $z = z_s$ является для нее седло-узлом (μ, n) -типа, $0 \leq n \leq N$.

Замечание 1. Очевидно, случай $n = N$ (устойчивый узел) означает, что все решения (1.1), достаточно близкие к z_s , удовлетворяют условию (1.3), а случай $n = 0$ означает, что условию (1.3) удовлетворяет только одно решение $z(\tau) \equiv z_s$. В общем случае при постановке сингулярных КрЗ для (1.1) необходимо учитывать, что количество нелинейных соотношений в конечной точке, эквивалентных в совокупности предельному условию (1.3), равно $N - n$, а сами эти соотношения задаются уравнением УМР. В противном случае возможны следующие ошибки, встречающиеся в литературе при постановке и решении физических задач: неточно задается граничное условие на бесконечности (указываются предельные значения отдельных компонент, а не всего решения); неверно ставится (или неверно трактуется) задача по общему числу краевых условий на обоих концах (полу)бесконечного интервала; неточно аппроксимируется и неудачно численно решается задача на конечном интервале, не эквивалентная исходной (см. примеры в [3] и статьях, перечисленных в заключении).

Пусть система (1.1) рассматривается на всей оси. Пусть $z = z_s$ и $z = \tilde{z}_s$ — точки равновесия системы, причем z_s является седло-узлом (μ, n) -типа при некоторых $\mu \geq 0$ и $n : 0 \leq n \leq N$, а \tilde{z}_s — седло-узлом $(\nu, N - n)$ -типа при некотором $\nu \leq 0$ и том же значении n . Наряду с (1.3) зададим предельное условие на левом конце

$$\lim_{\tau \rightarrow -\infty} \left\{ \exp(\nu\tau) [z(\tau) - \tilde{z}_s] \right\} = 0. \quad (1.5)$$

Аналогично предыдущему, в фазовом пространстве системы (1.1) в окрестности точки $z = \tilde{z}_s$ существует $(N - n)$ -мерное УМР $M_-^{(N - n)}$, инвариантное относительно τ и порожденное значениями $(N - n)$ -параметрического семейства решений сингулярной ЗК (1.1), (1.5). Тогда сингулярная КрЗ (1.1), (1.3), (1.5) правильно поставлена по общему числу краевых условий в окрестностях точек $\tau = -\infty$ и $\tau = \infty$. (Подобные КрЗ встречаются, например, в нелинейной теории поля, в математических моделях химии и биологии, где их решения называют кинк-солитонами, стенками или волновыми фронтами, если $z_s \neq \tilde{z}_s$, и решениями типа колокола, всплеска или волнового вала в противном случае.)

Очевидно, что, если точка равновесия $z = z_s$ системы (1.1) является седло-узлом $(0, n)$ -типа, то она одновременно является и седло-узлом $(0, N - n)$ -типа, так что в ее окрестности в z -пространстве оба многообразия $M_+^{(n)}$ и $M_-^{(N-n)}$ существуют и имеют хотя бы одну общую точку $z = z_s$.

Важным для приложений является случай, когда точка $z = z_s$ в (1.3), будучи седло-узлом (μ, n) -типа, зависит от k параметров $(c_1, \dots, c_k) \in S_k$, где $0 < k < n$, т.е. когда стационарные точки одного типа образуют k -параметрическое семейство — равновесное (или стационарное) множество S_k . Тогда для правильной постановки КрЗ на (полу)оси недостаточно иметь $N - n$ условий на левом конце, требуется задать еще k локальных или нелокальных условий для определения значений параметров c_1, \dots, c_k . (Примеры таких задач в теории пограничного слоя см., например, в [3] и работах, перечисленных здесь в заключении.)

Пусть теперь $Q(z) \in C^1_N(G_N)$, $z_s : z_s \in G_N, Q(z_s) = 0, (\partial Q / \partial z)(z_s) \neq 0$, и пусть на бесконечности задано предельное условие вида

$$\lim_{\tau \rightarrow \infty} \varphi_0 [z(\tau) - z_s] |z(\tau) - z_s| = 0, \tag{1.6}$$

где $\varphi_0 \in \mathbb{L}(\mathbb{K}^N, \mathbb{K}^{N-n})$ и имеет ранг $N - n$, $0 \leq n \leq N$ (решение стремится попасть по направлению в заданное линейное подпространство). При этом полагаем по определению, что $z(\tau) \equiv z_s$ удовлетворяет условию (1.6).

Определение 3. Условие (1.6) назовем *допустимым граничным условием* на бесконечности для системы (1.1), если в окрестности точки $z = z_s$ существует дифференцируемое УМР $M_+^{(n)}$, порожденное значениями решений (1.1), достаточно близких к z_s , а уравнение \mathbb{K}_0^n ,

$$\mathbb{K}_0^n : \varphi_0 [z - z_s] = 0, \quad z \in \mathbb{K}^N, \tag{1.7}$$

задает n -мерное касательное подпространство к $M_+^{(n)}$ в точке $z = z_s$. (Тем самым, если $M_+^{(n)}$ задается уравнением $\varphi(z) = 0$, где $\varphi \in \mathbb{K}^{N-n}$ и $\varphi(z_s) = 0$, то $\varphi_0 = M(\partial\varphi / \partial z)(z_s)$, где M — произвольная неособенная матрица, $M \in \mathbb{L}(\mathbb{K}^{N-n})$, например, $M = E_{N-n}$.)

Класс допустимых φ_0 может быть получен по аналогии с линейным случаем (о допустимых граничных условиях в особой точке или на бесконечности для систем линейных ОДУ и классах таких условий для систем с регулярной или иррегулярной особенностью см., например, в [8]–[10]). А именно, справедливо

Утверждение 1. Пусть \mathbb{H} — пространство строк длины N . Зададим любое вещественное число $\mu \geq 0$ такое, что $A = (\partial Q / \partial z)(z_s)$ не имеет СЗ на прямой $\text{Re}\lambda = -\mu$. Пусть \mathbb{X} — подпространство \mathbb{H} , порожденное всеми левыми корневыми векторами A , для которых $\text{Re}\lambda(A) > -\mu$. В качестве φ_0 возьмем какой-либо базис в \mathbb{X} . Тогда граничное условие (1.6) будет допустимым граничным условием на бесконечности для решений (1.1); оно эквивалентно условию (1.3) и его перенос из бесконечности в конечную точку осуществляется заданием уравнения УМР $M_+^{(n)}$ (в частности, в линейном приближении соотношение (1.7) может рассматриваться как граничное условие в конечной точке $\tau \geq \tau_\infty$).

Замечание 2. Для линейной системы (ЛС) $z' = A(z - z_s)$ с постоянной матрицей $A, A \in \mathbb{L}(\mathbb{K}^N)$, условие (1.6) называется *допустимым граничным условием* на бесконечности, если совокупность решений ЛС, удовлетворяющих (1.6), образует n -мерное подпространство \mathbb{K}_0^n в N -мерном пространстве \mathbb{K}^N всех решений, $0 \leq n \leq N$. Два n -мерных линейных подпространства, задаваемых уравнениями (1.7), называются эквивалентными, если соответствующие им граничные условия (1.6) выделяют одну и ту же совокупность решений ЛС. Тогда класс \mathbb{K}_0^n , порождающих допустимые условия (1.6) для ЛС, можно построить исчерпывающим образом (с точностью до эквивалентных подпространств) по следующей схеме (см. [8]–[10]): зададим любое вещественное число μ ; возьмем все правые корневые векторы матрицы A , соответствующие СЗ $\lambda(A) : \text{Re}\lambda(A) < \mu$; для каждого СЗ $\lambda(A) : \text{Re}\lambda(A) = \mu$, в соответствующем корневом пространстве возьмем какие-либо (например, все или никакие) правые собственные векторы (СВ), не тянущие жордановы цепочки; в качестве \mathbb{K}_0^n возьмем подпространство, порожденное всеми указанными векторами. Из этого результата, однако, виден и недостаток такого определения допустимого условия, так как в этот класс не попадает подпространство всех решений ЛС, ограниченных с весом $\exp(\mu t), t > 0$ (или, в более общем случае, ограниченных с весом $t^{-\alpha} \exp(\mu t), t > 0$, где $0 \leq \alpha \leq \kappa - 1, 1 \leq \kappa \leq N$, κ — наибольший размер жордановой клетки, отвечающей СЗ $\lambda(A) : \text{Re}\lambda(A) = \mu$). Такое подпространство приходится выделять отдельно, например следующим образом: зададим любое вещественное число μ ; возьмем все правые корневые векторы матрицы A , соответствующие СЗ $\lambda(A) : \text{Re}\lambda(A) < \mu$, и все СВ, отвечающие СЗ $\lambda(A) : \text{Re}\lambda(A) = \mu$ (в общем случае берем все корневые векторы высоты не более $\alpha + 1$

); в качестве \mathbb{K}_0^n возьмем подпространство, порожденное всеми указанными векторами, и рассмотрим условие (1.6) в подпространстве решений ЛС, ограниченных с весом $\exp(\mu t)$ (с весом $t^{-\alpha} \exp(\mu t)$ соответственно); тогда на таком классе функций условие (1.6) будет допустимым граничным условием для решений ЛС и эквивалентным соотношению (1.7), рассматриваемому уже в пространстве всех решений.

Суммируя некоторые утверждения [2], [3], приведем теорему существования УМР для случая, когда точка $z = z_s$ является седло-узлом (μ, n) -типа. Без ограничения общности можем положить $z_s = 0$. Если точка $z = 0$ является для системы (1.1) седло-узлом (μ, n) -типа при $\mu \geq 0, 0 \leq n \leq N$, то, как известно (см. [11]), эта система может быть неособенным преобразованием приведена к виду

$$dx/d\tau = \Lambda_1 x + F_1(x, y), \quad dy/d\tau = \Lambda_2 y + F_2(x, y), \quad -\infty < \tau < \infty, \tag{1.8}$$

где $x \in \mathbb{K}^n, y \in \mathbb{K}^m, m + n = N, \Lambda_{1,2}$ — постоянные матрицы, $\Lambda_1 \in \mathbb{L}(\mathbb{K}^n), \Lambda_2 \in \mathbb{L}(\mathbb{K}^m), F_1 : \Omega_{x,y}(a_0) \rightarrow \mathbb{K}^n, F_2 : \Omega_{x,y}(a_0) \rightarrow \mathbb{K}^m$, и выполняются приведенные ниже соотношения.

Соотношение 1. Существует $\mu \geq 0$ такое, что для всех СЗ $\Lambda_{1,2}$ справедливо

$$\operatorname{Re}\lambda(\Lambda_1) < -\mu < \operatorname{Re}\lambda(\Lambda_2). \tag{1.9}$$

Соотношение 2. Пусть $F_1 \in C_{n,a_0}, F_2 \in C_{m,a_0},$

$$F_j(0,0) = 0, \quad j = 1,2, \tag{1.10}$$

и, кроме того, $\forall \varepsilon > 0 \exists \delta_\varepsilon, 0 < \delta_\varepsilon \leq a_0,$ при котором справедливо

$$\begin{aligned} |F_j(x,y) - F_j(\tilde{x},\tilde{y})| &\leq \varepsilon(|x - \tilde{x}| + |y - \tilde{y}|), \\ \forall \{x,y,\tilde{x},\tilde{y}\} &\in \bar{\Omega}_{x,y}(\delta_\varepsilon), \quad j = 1,2. \end{aligned} \tag{1.11}$$

Пусть выполняются соотношения 1, 2. Введем обозначение

$$J_1 = \int_0^\infty |\exp((\Lambda_1 + \mu E_n)\tau)| d\tau, \tag{1.12}$$

$$J_2 = \int_0^\infty |\exp(-(\Lambda_2 + \mu E_m)\tau)| d\tau,$$

$$J = \max_{j=1,2} J_j, \tag{1.13}$$

$$K_1 = \sup_{\tau \geq 0} |\exp((\Lambda_1 + \mu E_n)\tau)|, \tag{1.14}$$

где, в силу соотношения 1, справедливо $0 < J < \infty, 1 \leq K_1 < \infty.$

Зафиксируем $q : 0 < q < 1,$ и возьмем $\varepsilon : 0 < \varepsilon \leq q / (2J);$ по ε выберем $\delta_\varepsilon, 0 < \delta_\varepsilon \leq a_0,$ при котором справедливо (1.11). Зафиксируем $\delta : 0 < \delta \leq \delta_\varepsilon,$ и введем величины

$$c_1(\delta) = (1 - q)\delta / K_1, \quad d(\delta) = q\delta. \tag{1.15}$$

Тогда справедлива следующая теорема о существовании УМР.

Теорема 1. Пусть выполняются соотношения 1, 2 и пусть для фиксированного $q, 0 < q < 1,$ определены $\varepsilon, \delta, c_1(\delta), d(\delta)$ указанным выше способом. Тогда каждому вектору $x^{(0)}, x^{(0)} \in \mathbb{K}^n,$ удовлетворяющему требованию

$$|x^{(0)}| \leq c_1(\delta), \tag{1.16}$$

соответствует единственная траектория системы (1.8), лежащая в шаре

$$\max \left\{ \sup_{\tau \geq 0} [\exp(\mu\tau)|x(\tau)|]; \sup_{\tau \geq 0} [\exp(\mu\tau)|y(\tau)|] \right\} \leq \delta \tag{1.17}$$

и такая, что

$$x(0) = x^{(0)}, \quad x^{(0)} \in \mathbb{K}^n, \tag{1.18}$$

т.е. начальные точки $z(0) = \{x(0); y(0)\}$, порождающие решения системы (1.8), не выходящие из шара (1.17), образуют в “шаре”

$$|x(0)| \leq c_1(\delta), \quad |y(0)| \leq d(\delta) \tag{1.19}$$

n -мерное УМР $M_+^{(n)}$, гомеоморфное окрестности начала координат x -пространства. Уравнение $M_+^{(n)}$ как “поверхности” в декартовом произведении пространств $\mathbb{K}^n \times \mathbb{K}^m$, т.е. связь между $y(0)$ и $x(0)$, неявно содержится в интегральном соотношении

$$y(0) = -\int_0^\infty \exp(-\Lambda_2 \tau) F_2(x(\tau, x(0)), y(\tau, x(0))) d\tau. \tag{1.20}$$

Всякое решение системы (1.8), значения которого в момент $\tau = 0$ принадлежат шару (1.19), но не принадлежат УМР $M_+^{(n)}$, удаляется от $M_+^{(n)}$ и при некотором $\tau = \tau_f > 0$ выходит из шара (1.17) (в частности, при $\mu = 0$ имеет место локальная экспоненциальная расходимость траекторий). Многообразие $M_+^{(n)}$ инвариантно относительно τ и существует, по крайней мере, во всей области фазового пространства, порожденной значениями решений n -параметрического семейства задач (1.8), (1.17), (1.18), где вектор параметров $x^{(0)}$ удовлетворяет ограничению (1.16), так что всякое решение системы (1.8), значения которого в момент $\tau = \tau_0$ принадлежат УМР $M_+^{(n)}$, остается на этом многообразии для всех $\tau \geq \tau_0$.

Доказательство существования и единственности решения задачи (1.8), (1.17), (1.18) основано на применении принципа сжатых отображений к оператору $V : C_\mu[0, \infty) \rightarrow C_\mu[0, \infty)$, заданному выражением

$$(V(\{x; y\}))(\tau) = \begin{cases} \exp(\Lambda_1 \tau) x^{(0)} + \int_0^\tau \exp(\Lambda_1(\tau - s)) F_1(x(s), y(s)) ds; \\ - \int_\tau^\infty \exp(\Lambda_2(\tau - s)) F_2(x(s), y(s)) ds \end{cases}, \quad \tau \geq 0.$$

Остальные утверждения достаточно очевидны (подробнее см. [2]).

Замечание 3. Следуя [11], [12], можно уточнить поведение решений системы (1.8) на УМР $M_+^{(n)}$. Возьмем число $\alpha > 0$ такое, что $\max_{1 \leq j \leq n} \operatorname{Re} \lambda_j(\Lambda_1 + \mu E_n) < -\alpha$. Отсюда и из (1.12) следует, что $\exists K_1, \sigma$, $1 \leq K_1 = \operatorname{const}$, $0 < \sigma = \operatorname{const}$, при которых для $\tau \geq 0$ справедливы неравенства

$$\begin{aligned} |\exp((\Lambda_1 + \mu E_n)\tau)| &\leq K_1 \exp(-(\alpha + \sigma)\tau), \\ |\exp(-(\Lambda_2 + \mu E_m)\tau)| &\leq K_1 \exp(-\sigma\tau). \end{aligned}$$

Из этих оценок в (1.13) получаем $J \leq K_1 / \sigma$. Зафиксируем $q : 0 < q < 1$, и возьмем $\varepsilon : 0 < \varepsilon \leq q\sigma / (2K_1)$; по ε выберем $\delta_\varepsilon, 0 < \delta_\varepsilon \leq a_0$, при котором справедливо (1.11). Зафиксируем $\delta : 0 < \delta \leq \delta_\varepsilon$, и введем величины $c_1(\delta)$, $d(\delta)$ по формулам (1.15). Тогда при указанном выборе δ и фиксированном значении $x^{(0)}$, удовлетворяющем неравенству (1.16), решение задачи (1.8), (1.17), (1.18) удовлетворяет оценкам

$$\begin{aligned} |x(\tau) - \exp(\Lambda_1 \tau) x^{(0)}| &\leq q\delta \exp(-(\alpha + \mu)\tau), \\ |y(\tau)| &\leq q\delta \exp(-(\alpha + \mu)\tau), \quad \tau \geq 0. \end{aligned}$$

Теорема 1 и замечание 3 дополняют и уточняют в современной форме классические результаты [11], [12] о существовании УМР для автономных систем нелинейных ОДУ с экспоненциально дихотомической главной частью (аналогичные теоремы для более общих систем функционально-дифференциальных и операторных уравнений см., например, в [13]–[17]).

Важным является следующее дополнение (см. [2], [3]; некоторые погрешности, допущенные в [2], [3], здесь исправлены). Вместо (1.8) рассмотрим неавтономную систему N нелинейных ОДУ с начальным вырождением по независимой переменной (типа регулярной особенности):

$$tdxdt = \Lambda_1 x + F_1(t, x, y), \quad tdydt = \Lambda_2 y + F_2(t, x, y), \quad 0 < t \leq 1. \tag{1.21}$$

Здесь $x \in \mathbb{K}^n$, $y \in \mathbb{K}^m$, $m + n = N$, $\Lambda_1 \in \mathbb{L}(\mathbb{K}^n)$, $\Lambda_2 \in \mathbb{L}(\mathbb{K}^m)$, $F_1 : [0,1] \times \overline{\Omega}_{x,y}(a_0) \rightarrow \mathbb{K}^n$, $F_2 : [0,1] \times \overline{\Omega}_{x,y}(a_0) \rightarrow \mathbb{K}^m$, причем $\operatorname{Re}\lambda(\Lambda_1) > \mu \quad \forall \lambda(\Lambda_1)$, $\operatorname{Re}\lambda(\Lambda_2) < \mu \quad \forall \lambda(\Lambda_2)$, где μ — фиксированное число из интервала $[0,1]$. Введем новую независимую переменную

$$\tau = -\ln t, \quad t = \exp(-\tau), \quad 0 \leq \tau < \infty, \quad (1.22)$$

и обозначения

$$\hat{x} = \begin{pmatrix} t \\ x \end{pmatrix}, \quad \hat{x} \in \mathbb{K}^{n+1}, \quad \hat{\Lambda}_1 = \begin{pmatrix} -1 & 0 \\ 0 & -\Lambda_1 \end{pmatrix}, \\ \hat{\Lambda}_2 = -\Lambda_2, \quad \hat{F}_j(\hat{x}, y) = -F_j(t, x, y), \quad j = 1, 2.$$

Тогда в новых переменных получим автономную систему $N + 1$ нелинейных ОДУ на полубесконечном интервале

$$d\hat{x}d\tau = \hat{\Lambda}_1 \hat{x} + \hat{F}_1(\hat{x}, y), \quad dyd\tau = \hat{\Lambda}_2 y + \hat{F}_2(\hat{x}, y), \quad 0 \leq \tau < \infty. \quad (1.23)$$

К системе (1.23) уже применима теорема 1 и последующие результаты этого разд. 1 при соответствующих предположениях относительно $F_j(t, x, y)$, $j = 1, 2$. В частности, для системы (1.21) соотношение (1.20), с учетом формул (1.22), переписывается в виде

$$y(t_0) = \int_0^{t_0} \left[(t/t_0)^{-\Lambda_2} F_2 \begin{pmatrix} t, x(t, x(t_0)), \\ y(t, x(t_0)) \end{pmatrix} / t \right] dt. \quad (1.24)$$

Это соотношение задает уравнение УМР, зависящее от t_0 как от параметра.

В заключение этого разд. 1 отметим, что соотношение (1.20) ((1.24)) является нелокальным и зависит от решений системы (1.8) на полуоси (системы (1.21) на отрезке $[0, t_0]$). Локальное соотношение можно получить для гладких УМР в случае гладких функций в правой части системы (1.8) ((1.21)) с помощью решения сингулярной задачи типа Ляпунова для сопутствующей системы квазилинейных уравнений с частными производными первого порядка с вырождением по начальным данным. Эта задача, представляющая самостоятельный интерес, рассматривается далее в разд. 2.

2. О ГЛАДКИХ УМР И СОПУТСТВУЮЩИХ ЗАДАЧАХ ЛЯПУНОВА ДЛЯ УРАВНЕНИЙ С ЧАСТНЫМИ ПРОИЗВОДНЫМИ

Пусть в декартовом произведении пространств $\mathbb{K}^n \times \mathbb{K}^m = \mathbb{K}^N$ заданы система N нелинейных ОДУ

$$dxd\tau = f(x, y), \quad dyd\tau = g(x, y), \quad x \in G_n, \quad y \in G_m, \quad -\infty < \tau < \infty, \quad (2.1)$$

и нелинейное соотношение

$$y = \rho(x), \quad x \in \tilde{G}_n \subseteq G_n, \quad (2.2)$$

где

$$f \in C_n(G_N), \quad g \in C_m(G_N), \quad G_N = G_n \times G_m, \quad \rho(x) \in C_m(\tilde{G}_n).$$

Определение 4. Будем говорить, что соотношение (2.2) задает в N -мерном фазовом пространстве системы (2.1) инвариантное относительно τ n -мерное УМР $M_+^{(n)}$ ($M_-^{(n)}$), если всякое решение системы (2.1), значения которого в момент $\tau = 0$ удовлетворяют соотношению (2.2), обладает тем же свойством для всех $\tau \geq 0$ ($\tau \leq 0$). Такое многообразие назовем *гладким*, если $\rho \in C_m^k(\tilde{G}_n)$, $k \geq 1$.

Следующие утверждения достаточно очевидны.

Теорема 2. Соотношение (2.2) тогда, и только тогда, задает инвариантное относительно τ гладкое n -мерное УМР в N -мерном фазовом пространстве системы (2.1), когда $\rho(x)$ есть решение уравнения

$$\partial \rho \partial x f(x, \rho) = g(x, \rho). \tag{2.3}$$

Теорема 3. Пусть $\{x, y\} = \{0, 0\}$ — стационарная точка системы (2.1), т.е. $\{0, 0\} \in G_N$, $f(0, 0) = 0$, $g(0, 0) = 0$. Для того, чтобы это стационарное решение принадлежало гладкому УМР (2.2), необходимо и достаточно, чтобы $\rho(x)$ было решением уравнения (2.3), удовлетворяющим условию

$$\rho(0) = 0. \tag{2.4}$$

Пусть теперь система (2.1) имеет вид (1.8), т.е.

$$f(x, y) = \Lambda_1 x + F_1(x, y), \quad g(x, y) = \Lambda_2 y + F_2(x, y).$$

Тогда получим из (2.3), (2.4) сингулярную задачу, которую будем называть *сингулярной задачей Ляпунова*. Она была поставлена и исследована Ляпуновым в предположениях $F_1 \in H_{n, a_0}, F_2 \in H_{m, a_0}$ и более общих, чем в (1.9), предположениях о взаимном расположении спектров Λ_1 и Λ_2 (см. [5, с. 153–158], [6, гл. VI] и здесь п. 2.1): доказаны теоремы существования и единственности решения в классе голоморфных в нуле функций; различные частные случаи этих теорем использованы для описания и развития метода, известного в настоящее время как второй метод Ляпунова в теории устойчивости.

В предположениях $F_1 \in C_{n, a_0}^1, F_2 \in C_{m, a_0}^1$ и справедливости (1.9) сингулярная задача Ляпунова, ее обобщения и применения изучены в [1]–[3] и цитированных там более ранних работах автора. (До этих работ сингулярная задача в указанных предположениях, по-видимому, не рассматривалась, а ее решение не связывалось с построением гладких УМР для автономных систем нелинейных ОДУ и другими приложениями.)

Основной сингулярной задачей Ляпунова называем задачу вида

$$\partial \rho \partial x [\Lambda_1 x + F_1(x, \rho)] = \Lambda_2 \rho + F_2(x, \rho), \quad x \in \Omega_x(c), \tag{2.5}$$

$$\rho(0) = 0. \tag{2.6}$$

Здесь $x \in \mathbb{K}^n$, $\rho \in \mathbb{K}^m$, $\partial \rho / \partial x \in L(\mathbb{K}^n, \mathbb{K}^m)$, $\Lambda_1 \in L(\mathbb{K}^n)$, $\Lambda_2 \in L(\mathbb{K}^m)$, $F_1 : \bar{\Omega}_{x, \rho}(a_0) \rightarrow \mathbb{K}^n$, $F_2 : \bar{\Omega}_{x, \rho}(a_0) \rightarrow \mathbb{K}^m$.

2.1. Существование и единственность голоморфного решения

Первая теорема Ляпунова. Пусть действительные части всех СВ матрицы Λ_1 отличны от нуля и одного знака и пусть не существует никаких соотношений вида

$$\sum_{j=1}^n l_j \lambda_j(\Lambda_1) = \lambda_s(\Lambda_2), \quad s = 1, \dots, m, \tag{2.7}$$

для целых $l_j \geq 0$, $j = 1, \dots, n$, таких, что

$$\sum_{j=1}^n l_j > 0; \tag{2.8}$$

пусть $F_1 \in H_{n, a_0}$, $F_2 \in H_{m, a_0}$, причем для $(x, \rho) \in \bar{\Omega}_{x, \rho}(a_0)$ справедливо

$$F_1(x, \rho) = [x; \rho]_2, \tag{2.9}$$

$$F_2(0, \rho) = [\rho]_2, \quad F_2(x, 0) = [x]_v, \tag{2.10}$$

где $v \geq 1$.

Тогда $\exists c_0$, $0 < c_0 \leq a_0$, такое, что $\forall c : 0 < c \leq c_0$, в классе $H_m(\Omega_x(c))$ существует единственное решение $\hat{\rho}(x)$ задачи (2.5), (2.6), причем справедливо

$$\hat{\rho}(x) = [x]_v, \quad x \in \Omega_x(c). \tag{2.11}$$

Рассмотрим теперь сингулярную задачу Ляпунова с дополнительным требованием

$$\partial \rho \partial x(0) = 0. \quad (2.12)$$

Вторая теорема Ляпунова. Пусть в предположениях предыдущей теоремы неравенство (2.8) заменено неравенством $\sum_{j=1}^n l_j > 1$, а условия (2.9), (2.10) — требованиями

$$F_1(x, 0) = [x]_2, \quad F_1(0, \rho) = [\rho]_1, \quad F_2(x, \rho) = [x, \rho]_2, \quad (x, \rho) \in \overline{\Omega_{x, \rho}}(a_0).$$

Тогда $\exists c_0, 0 < c_0 \leq a_0$, такое, что $\forall c: 0 < c \leq c_0$, в классе $H_m(\Omega_x(c))$ существует единственное решение $\hat{\rho}(x)$ задачи (2.5), (2.6), (2.12), причем если справедливо (2.10) при $v \geq 2$, то справедливо (2.11).

Доказательства этих утверждений тесно связаны с приведенной ниже леммой, представляющей самостоятельный интерес и следующей из результатов [5, с.101–106].

Обозначим N_p -мерное линейное пространство однородных m -векторных форм p -й степени по x_1, \dots, x_n через $F_{m, n}^{N, p}$, т.е. любой элемент $\rho_{(p)}$ этого пространства имеет вид

$$\rho_{(p)}(x) = \sum_{|l|=p} \rho_l x^l, \quad x \in \mathbb{K}^n, \quad (2.13)$$

где p — фиксировано, $p \geq 1$, $l = (l_1, \dots, l_n)$ — мультииндекс, $0 \leq l_j$ — целые, $|l| = \sum_{j=1}^n l_j$, $x^l = x_1^{l_1} \dots x_n^{l_n}$, $\rho_l \in \mathbb{K}^m$. Размерность N_p этого пространства, как нетрудно убедиться, находится по формуле $N_p = m C_{n+p-1}^p$ (здесь C_{n+p-1}^p — коэффициент при t^p в разложении производящей функции $f(t) = (1-t)^{-n} - 1$).

В пространстве $F_{m, n}^{N, p}$ рассмотрим линейный конечномерный оператор $L_{(p)}: F_{m, n}^{N, p} \rightarrow F_{m, n}^{N, p}$ такой, что

$$L_{(p)}\rho_{(p)} = \frac{\partial \rho_{(p)}}{\partial x} \Lambda_1 x - \Lambda_2 \rho_{(p)}. \quad (2.14)$$

Лемма Ляпунова. Все СЗ оператора $L_{(p)}$ найдутся по формуле

$$\lambda_{s, l_1, \dots, l_n}(L_{(p)}) = \sum_{j=1}^n l_j \lambda_j(\Lambda_1) - \lambda_s(\Lambda_2), \quad s = 1, \dots, m, \quad (2.15)$$

где $0 \leq l_j$ — целые, $|l| = p$, причем при каждом $s: 1 \leq s \leq m$, ни один из наборов $\{l_1, \dots, l_n\}$ не встречается более одного раза.

Отсюда, в частности, следует, что уравнение

$$L_{(p)}\rho_{(p)} = \chi_{(p)}, \quad \chi_{(p)} \in F_{m, n}^{N, p}, \quad (2.16)$$

однозначно разрешимо в пространстве $F_{m, n}^{N, p}$, если не существует никаких соотношений вида (2.15), где $0 \leq l_j$ — целые, $|l| = p$. В частности, при выполнении (1.9) уравнение (2.16) однозначно разрешимо в пространстве $F_{m, n}^{N, p}$.

Следствие 1. Пусть $F_1 \in H_{n, a_0}$, $F_2 \in H_{m, a_0}$, а в остальном выполняются условия первой (или второй) теоремы Ляпунова. Представим голоморфное решение задачи (2.5), (2.6) (или задачи (2.5), (2.6), (2.12)) в виде

$$\hat{\rho}(x) = \sum_{p=v}^{\infty} \rho_{(p)}(x), \quad x \in \Omega_x(c), \quad \rho_{(p)} \in F_{m, n}^{N, p}, \quad (2.17)$$

и положим

$$F_1(x, \hat{\rho}(x)) = \sum_{\substack{|\tau|, |s|=0 \\ |\tau|+|s| \geq 2-\delta_{\tau,0}}}^{\infty} F_{\tau s}^{(1)} x^{\tau} \left(\sum_{|l|=v}^{\infty} \rho_l x^l \right)^s = \sum_{p=2}^{\infty} F_{(p)}^{(1)}(x, \rho_{(v)}, \dots, \rho_{(p-1)}) + \partial F_1 \partial \rho(0,0) \sum_{p=v}^{\infty} \rho_{(p)} \quad (2.18)$$

(второе слагаемое отсутствует для первой теоремы Ляпунова),

$$F_2(x, \hat{\rho}(x)) = \sum_{\substack{|\tau|, |s|=0 \\ |\tau|+|s| \geq v \delta_{|s|,0} + 2(1-\delta_{|s|,0})}}^{\infty} F_{\tau s}^{(2)} x^{\tau} \left(\sum_{|l|=v}^{\infty} \rho_l x^l \right)^s = F_{(v)}^{(2)}(x) + \sum_{p=v+1}^{\infty} F_{(p)}^{(2)}(x, \rho_{(v)}, \dots, \rho_{(p-1)}); \quad (2.19)$$

здесь $\delta_{i,j}$ — символ Кронекера, $F_{(p)}^{(1)} \in F_{n,n}^N$, $F_{(p)}^{(2)} \in F_{m,n}^N$,

$$F_{(v)}^{(2)}(x) = \sum_{|l|=v} F_{l,0}^{(2)} x^l,$$

а коэффициенты разложений $F_{(p)}^{(j)}$ по x^l ($j = 1, 2, |l| = p$), образованы из коэффициентов $F_{\tau s}^{(j)}$ и коэффициентов разложений предыдущих форм $\rho_{(v)}, \dots, \rho_{(p-1)}$ с помощью только действий типа сложения и умножения.

Тогда слагаемые ряда (2.17) удовлетворяют рекуррентной последовательности уравнений

$$L_{(p)} \rho_{(p)} = \chi_{(p)}(x, \rho_{(v)}, \dots, \rho_{(p-1)}), \quad p = v, v+1, \dots, \quad (2.20)$$

где оператор $L_{(p)}$ определен в (2.14),

$$\chi_{(v)} = F_{(v)}^{(2)},$$

$$\chi_{(p)} = F_{(p)}^{(2)} - \sum_{k=v}^{p-1} \partial \rho_{(k)} \partial x \left[\begin{array}{c} F_{(p-k+1)}^{(1)} + \\ + \partial F_1 \partial \rho(0,0) \rho_{(p-k+1)} \end{array} \right], \quad p \geq v+1, \quad (2.21)$$

т.е. при каждом $p \geq v$ из (2.20) получается невырожденная система N_p линейных алгебраических уравнений для определения коэффициентов ρ_l формы (2.13) ($|l| = p$) и тем самым эти коэффициенты могут быть определены и притом однозначно.

2.2. Существование и единственность решения в классе $C_m^1(\Omega_x(c))$, оценки решения

Соотношение 3. Для $F_j(x, \rho), j = 1, 2$, справедливо $F_1 \in C_{n,a}^1, F_2 \in C_{m,a_0}^1$,

$$F_j(0,0) = 0, \quad \partial F_j \partial x(0,0) = 0, \quad \partial F_j \partial \rho(0,0) = 0, \quad j = 1, 2. \quad (2.22)$$

Пусть выполняются соотношения 1, 3. Зафиксируем число $q \in (0,1)$ и выберем $\delta \in (0, a_0)$ из условия

$$\max_{j=1,2} \sup_{(x,\rho) \in \bar{\Omega}_{x,\rho}(\delta)} \left\{ \max \left[\left| \partial F_j \partial x(x,\rho) \right|; \left| \partial F_j \partial \rho(x,\rho) \right| \right] \right\} \leq q / (2J), \quad (2.23)$$

где J определено в (1.12), (1.13). В силу (2.22) выполнения (2.23) можно добиться за счет выбора δ .

Используя обозначение (1.14), введем величины $c_1 = c_1(\delta)$ и $d = d(\delta)$ по формулам (1.15) и приведем обобщение первой теоремы Ляпунова.

Теорема 4. Пусть выполняются соотношения 1, 3; пусть для фиксированного $q \in (0,1)$ определены $\delta, c_1(\delta), d(\delta)$ указанным выше способом.

Тогда справедливы следующие утверждения:

1) $\forall c : 0 < c \leq c_1$, в классе $C_m^1(\Omega_x(c))$ существует единственное решение $\hat{\rho}(x)$ задачи (2.5), (2.6), причем $\hat{\rho}(x)$ удовлетворяет условию (2.12) и справедлива оценка

$$|\hat{\rho}(x)| \leq K_1 q |x| / (1 - q), \quad x \in \Omega_x(c);$$

это решение однозначно продолжается в область $\tilde{\Omega}_x$, $\tilde{\Omega}_x \subseteq \bar{\Omega}_x(\delta)$, порожденную x -характеристиками системы (1.8) — проекциями на x -пространство решений x -параметрического семейства задач (1.8), (1.17), (1.18), где вектор параметров $x^{(0)}$ удовлетворяет требованию (1.16) причем

$$\sup_{x \in \tilde{\Omega}_x} |\hat{\rho}(x)| \leq d;$$

2) если $\exists \beta : 0 < \beta = \text{const}$, и $\forall c : 0 < c \leq a_0, \exists C_1(c) : 0 < C_1(c) = \text{const}$ такие, что выполняется неравенство

$$|\partial F_2 \partial x(x, 0)| \leq C_1(c) |x|^\beta, \quad x \in \Omega_x(c), \tag{2.24}$$

то справедлива оценка

$$|\hat{\rho}(x)| \leq C_1(c) J_2 K_1^{\beta+1} |x|^{\beta+1} / [(1 - q / 2)(1 - q)^{\beta+1}], \quad x \in \Omega_x(c); \tag{2.25}$$

3) если $F_1 \in C_{n,a_0}^k, F_2 \in C_{m,a_0}^k, k \geq 1$, то $\hat{\rho} \in C_m^k(\Omega_x(c))$;

4) если $F_1 \in H_{n,a_0}, F_2 \in H_{m,a_0}$, то $\hat{\rho} \in H_m(\Omega_x(c))$, причем если справедливо (2.10) при $v \geq 2$ и $x \in \bar{\Omega}_x(a_0)$, то $\hat{\rho}(x) = [x]_v, x \in \Omega_x(c)$.

Следствие 2. Пусть не выполнено одно из условий в (2.22), а именно,

$$\partial F_2 \partial x(0, 0) = D \neq 0, \quad D \in \mathbb{L}(\mathbb{K}^n, \mathbb{K}^m),$$

а в остальном пусть выполняются предположения теоремы 4. Пусть $(m \times n)$ -матрица α есть решение уравнения

$$\Lambda_2 \alpha - \alpha \Lambda_1 = -D, \tag{2.26}$$

что влечет представление

$$\alpha = - \int_0^\infty \exp(-\Lambda_2 \tau) D \exp(\Lambda_1 \tau) d\tau. \tag{2.27}$$

Пусть $F_j(x, \rho), j = 1, 2$, определены в области $\bar{\Omega}_x(a_0) \times \{\rho : |\rho| \leq (1 + |\alpha|) a_0\}$. Тогда замена $\rho = \alpha x + \tilde{\rho}$ приводит задачу (2.5), (2.6) к аналогичной задаче для $\tilde{\rho}$ с функциями $\tilde{F}_1(x, \tilde{\rho}) = F_1(x, \alpha x + \tilde{\rho}), \tilde{F}_2(x, \tilde{\rho}) = F_2(x, \alpha x + \tilde{\rho}) - \alpha F_1(x, \alpha x + \tilde{\rho}) - Dx$, для которых уже выполняются условия теоремы 4.

Представление (2.27) можно проверить, следуя подходу [13, § 4, п. 4]: подставляя (2.27) в (2.26), получаем

$$\begin{aligned} -\Lambda_2 \int_0^\infty \exp(-\Lambda_2 \tau) D \exp(\Lambda_1 \tau) d\tau + \int_0^\infty \exp(-\Lambda_2 \tau) D \exp(\Lambda_1 \tau) d\tau \Lambda_1 = \\ = \int_0^\infty d(\exp(-\Lambda_2 \tau) D \exp(\Lambda_1 \tau)) = -D. \end{aligned} \tag{2.28}$$

Замечание 4. Если $F_1(x, \rho) \equiv 0$, то в теореме 4 относительно Λ_1 достаточно предполагать $\text{Re} \lambda(\Lambda_1) \leq -\mu$ для всех СЗ Λ_1 , где СЗ Λ_1 , для которых $\text{Re} \lambda(\Lambda_1) = -\mu$, “не тянут жордановы цепочки”. При этом в (1.15) достаточно положить $c_1(\delta) = \delta / K_1$, а в правой части (2.23) будет $q / (2J_2)$. Если, кроме того, справедливо $F_2(x, \rho) \equiv F_2(x) = \sum_{p=v}^K F_{(p)}^{(2)}(x)$, $F_{(p)}^{(2)} \in F_{m,n}^N, v \geq 1$, т.е. F_2 —векторный многочлен по компонентам x , то решением задачи (2.5), (2.6) также, очевидно, будет многочлен $\hat{\rho}(x) = \sum_{p=v}^K \rho_{(p)}(x)$, $\rho_{(p)} \in F_{m,n}^N$, и в этом простейшем случае решение (2.5), (2.6) определено глобально, т.е. для всех $x \in \mathbb{K}^n$. Например, для скалярного уравнения

$$\rho' \lambda_1 x = \lambda_2 \rho - \sum_{p=v}^K p x^p$$

получаем ответ

$$\hat{\rho}(x) = \sum_{p=v}^K p x^p / (\lambda_2 - p \lambda_1).$$

Замечание 5. В [19] изучается, как вспомогательная, следующая скалярная сингулярная ЗК (см. в [19] задачу (3.2) в п. 3):

$$r h' + b h = r f(r, h) + h^2 g(r, h), \quad h(0) = 0,$$

где $f(r, h), g(r, h)$ — вещественные аналитические в точке $(0, 0)$ функции. Приводится утверждение, что при $b > -1$ существует единственное аналитическое в нуле решение этой задачи (см. в [19] теорему 3.1 и два способа ее доказательства, одно из которых принадлежит Л. Ниренбергу). Нетрудно, однако, заметить, что это утверждение сразу следует из первой теоремы Ляпунова, а из теоремы 4 и следствия 2 следует, что это решение единственно в классе $C^1_1(\Omega_r(r_0))$.

Замечание 6. Легко построить пример, когда при невыполнении соотношения 3 для $F_1(x, \rho)$ задача (2.5), (2.6) имеет не одно решение: скалярное ОДУ

$$\rho' (\lambda_1 x + m \rho) = \lambda_2 \rho,$$

где $m \neq 0$ и $\lambda_1 \neq \lambda_2$, имеет, по крайней мере, два решения, обращающиеся в нуль при $x = 0$:

$$\rho_1(x) \equiv 0, \quad \rho_2(x) = (\lambda_2 - \lambda_1)x / m.$$

При этом для первого из этих решений $\rho_1(x)$, и только для него, выполнены условия второй теоремы Ляпунова, а также условия приведенной здесь далее теоремы 5.

Для обобщения второй теоремы Ляпунова введем

Соотношение 4. Для $F_j(x, \rho), j = 1, 2$, справедливо: $F_1 \in C^1_{n, a_0}, F_2 \in C^1_{m, a_0}, 0 < a_0 \leq 1$,

$$F_j(0, 0) = 0, \quad j = 1, 2, \quad \partial F_1 \partial x(0, 0) = 0, \quad \partial F_2 \partial \rho(0, 0) = 0, \tag{2.29}$$

$$\partial F_1 \partial \rho(0, 0) = M \neq 0, \quad M \in \mathbb{L}(\mathbb{K}^m, \mathbb{K}^n), \tag{2.30}$$

$$|\partial F_2 \partial x(x, \rho)| \leq C (|x|^\gamma + |\rho|^\gamma), \quad (x, \rho) \in \bar{\Omega}_{x, \rho}(a_0), \tag{2.31}$$

где $0 < C = \text{const}, 0 < \gamma = \text{const}$.

Пусть выполняются соотношения 1, 4. Зафиксируем числа q и $\Delta, 0 < q < 1, 0 < \Delta < \gamma$, и выберем $\delta \in (0, a_0)$, при котором справедливы неравенства

$$\sup_{(x, \rho) \in \Omega_{x, \rho}(\delta)} \left\{ \max \left[2|\partial F_1 \partial x(x, \rho)|; 2|\partial F_1 \partial \rho(x, \rho) - M|; |\partial F_2 \partial \rho(x, \rho)| \right] \right\} \leq q / (2J), \tag{2.32}$$

$$|M| \delta^\Delta \leq q / (2J_1), \quad 2C \delta^{\gamma - \Delta} \leq q / (2J_2), \tag{2.33}$$

где J, J_1, J_2 определены в (1.12), (1.13) (здесь $\delta = \delta(\Delta) \rightarrow 0$ при $\Delta \rightarrow 0$ и при $\Delta \rightarrow \gamma - 0$). Для выбранного δ определим величины $c_1(\delta)$ и $d(\delta)$ по формулам (1.15) и приведем обобщение второй теоремы Ляпунова.

Теорема 5. Пусть выполняются соотношения 1, 4; пусть для фиксированных $q \in (0, 1)$ и $\Delta \in (0, \gamma)$ определены $\delta, c_1(\delta), d(\delta)$ указанным выше способом. Тогда справедливы следующие утверждения:

1) $\forall c : 0 < c \leq c_1$, в классе $C^1_m(\Omega_x(c))$ существует единственное решение $\hat{\rho}(x)$ задачи (2.5), (2.6), удовлетворяющее дополнительному требованию

$$\rho(x) = O(|x|^{1+\Delta}), \quad |x| \rightarrow 0, \tag{2.34}$$

причем при выполнении (2.24), где $\beta \geq \gamma$, справедлива оценка (2.25);

2) если $F_1 \in C_{n,a_0}^k$, $F_2 \in C_{m,a_0}^k$, $k \geq 1$, то $\hat{\rho} \in C_m^k(\Omega_x(c))$;

3) если $F_1 \in H_{n,a_0}$, $F_2 \in H_{m,a_0}$, то $\hat{\rho} \in H_m(\Omega_x(c))$, причем если справедливо (2.10) при $v \geq 2$ и $x \in \Omega_x(a_0)$, то $\hat{\rho}(x) \stackrel{\text{гл}}{=} [x]_v$, $x \in \Omega_x(c)$, и справедлива оценка (2.25) при $\beta = v - 1$.

Учитывая следствие 1, получаем

Утверждение 2. Пусть $F_1 \in H_{n,a_0}$, $F_2 \in H_{m,a_0}$, а в остальном выполняются условия теоремы 4 (или теоремы 5). Тогда для голоморфного решения основной сингулярной задачи Ляпунова (2.5), (2.6) (или задачи (2.5), (2.6), (2.12)) слагаемые ряда (2.17) последовательно определяются по формулам

$$\rho_{(p)}(x) = -\int_0^\infty \exp(-\Lambda_2 \tau) \chi_{(p)}(\exp(\Lambda_1 \tau)x, \rho_{(v)}(\exp(\Lambda_1 \tau)x), \dots, \rho_{(p-1)}(\exp(\Lambda_1 \tau)x)) d\tau, \quad (2.35)$$

где $p = v, v+1, \dots$, а однородные формы $\chi_{(p)}(x, \rho_{(v)}, \dots, \rho_{(p-1)})$ образованы по формулам (2.18), (2.19), (2.21).

Замечание 7. Как уже отмечалось, для примера в замечании 6 условия теоремы 5 выполнены для решения $\rho_1(x) \equiv 0$. Следующий пример показывает, что нарушение условий обеих теорем 4, 5 для $F_j(x, \rho)$, $j = 1, 2$, приводит, например, к неединственности решения задачи (2.5), (2.6) и несуществованию решения задачи (2.5), (2.6), (2.12) (или задачи (2.5), (2.6), (2.34)). Например, скалярное ОДУ

$$\rho'(\lambda x + \rho) = ax$$

при $a \neq -\lambda^2 / 4$ имеет два решения, обращающиеся в нуле в нуль,

$$\rho_j = \alpha_j x, \quad j = 1, 2,$$

где

$$\alpha_{1,2} = -\lambda / 2 \pm \sqrt{\lambda^2 / 4 + a},$$

и не имеет ни одного решения, удовлетворяющего условиям $\rho(0) = (d\rho / dx)(0) = 0$.

Приведем еще одно ОДУ, удобное тем, что оно также полностью интегрируется явно:

$$d\rho dx (\lambda x + m\rho) = -\lambda\rho + ax^k,$$

где $m \neq 0, \lambda \neq 0, k \geq 1, a \neq 0$ при $k > 1$ и $a > -\lambda^2 / m^2$ при $k = 1$. Это уравнение записывается в виде

$$\left(\lambda x\rho + m\rho^2 / 2 \right)' = ax^k,$$

т.е.

$$\lambda x\rho + m\rho^2 / 2 = ax^{k+1} / (k+1) + \text{const.}$$

Для решений, удовлетворяющих условию $\rho(0) = 0$, имеем

$$\rho_\pm(x) = -\lambda x m \pm \sqrt{\lambda^2 x^2 m^2 + 2ax^{k+1}m(k+1)},$$

т.е. при $k \geq 1$, когда условия теоремы 4 не выполнены, имеем два решения, причем при $k > 1$ ровно одно из них $\rho_+(x)$ удовлетворяет дополнительному требованию типа (2.34) (для этого решения выполнены условия теоремы 5), а при $k = 1$, когда условия обеих теорем 4, 5 не выполняются, таких решений нет.

Замечание 8. Частным случаем задачи (2.5), (2.6) является сингулярная задача Коши для ОДУ

$$\begin{aligned} dy dt [\lambda_1 t + f(t, y)] &= \Lambda_2 y + F_2(t, y), \\ |t| &\leq t_0, \quad t_0 > 0, \quad y(0) = 0, \end{aligned}$$

где λ_1, f — скалярные величины. При $f(t, y) \equiv 0, \lambda_1 = -1$ эта задача, представляющая самостоятельный интерес (см., в частности, замечание 5), подробно изучена при достаточно общих предположениях как для некоторых классов систем нелинейных ОДУ (см. [18]), так и для функционально-дифференциальных и операторных уравнений (см., например, [13]–[17] и цитированные там работы). В частности, во втором примере замечания 7 такая задача отвечает случаю $m = 0$. Тогда выполнены условия теоремы 4

(при $k > 1$) или следствия 2 (при $k = 1$), и единственное решение основной сингулярной задачи Ляпунова имеет вид

$$\rho(x) = ax^k / [\lambda(k + 1)]$$

(параметр m является параметром бифуркации для указанной задачи с критическим значением $m = 0$). Это решение имеет смысл и при $0 < k < 1$, но уже в классе функций, которые не дифференцируемы в нуле; его существование и единственность не укладываются в рамки теорем данной работы, но следуют из более общих результатов для функционально-дифференциальных уравнений.

3. О ГЛАДКИХ ПОЛЯХ ДЛЯ ВЕКТОРНЫХ СИСТЕМ ОДУ ВТОРОГО ПОРЯДКА И СОПУТСТВУЮЩИХ СИСТЕМАХ УРАВНЕНИЙ ТИПА ГАМИЛЬТОНА-ЯКОБИ

Пусть на бесконечном интервале рассматривается автономная система N нелинейных ОДУ второго порядка

$$y'' = Q(y, y'), \quad -\infty < \tau < \infty, \quad y, y' \in G_N, \tag{3.1}$$

где $Q \in C_N(G_N \times G_N)$, и пусть в $2N$ -мерном фазовом пространстве этой системы задано разрешенное относительно производной нелинейное соотношение

$$y' = \alpha(y), \quad y \in \tilde{G}_N \subseteq G_N, \tag{3.2}$$

где $\alpha \in C_N(\tilde{G}_N)$.

Определение 5. Пусть соотношение (3.2) задает инвариантное относительно τ N -мерное УМР в $2N$ -мерном фазовом пространстве системы (3.1) для всех $\tau \geq 0$ ($\tau \leq 0$). Такое УМР будем называть N -мерным полем, отвечающим данной системе. Поле назовем гладким, если $\alpha \in C_N^k(\tilde{G}_N)$, $k \geq 1$.

Теорема 6. Соотношение (3.2) тогда, и только тогда, задает гладкое N -мерное поле для решений системы (3.1), когда $\alpha(y)$ есть решение уравнения

$$\partial \alpha d y \alpha = Q(y, \alpha). \tag{3.3}$$

Теорема 7. Пусть $\{y, y'\} = \{0, 0\}$ — стационарная точка системы (3.1), т.е. $\{0, 0\} \in G_N \times G_N$, $Q(0, 0) = 0$. Для того, чтобы это стационарное решение принадлежало гладкому полю (3.2), необходимо и достаточно, чтобы $\alpha(y)$ было решением уравнения (3.3), удовлетворяющим условию

$$\alpha(0) = 0. \tag{3.4}$$

Пусть теперь система (3.1) имеет вид

$$y'' = Ay' + By + F(y, y'), \quad -\infty < \tau < \infty, \tag{3.5}$$

где A и B — постоянные матрицы, $A, B \in \mathbb{L}(\mathbb{K}^N)$, $F : \bar{\Omega}_{y, y'}(a_0) \rightarrow \mathbb{K}^N$, $F(y, y') = o(|y| + |y'|)$, $|y| + |y'| \rightarrow 0$. Ассоциированным с ОДУ (3.5) является квадратное матричное уравнение

$$\Lambda^2 - A\Lambda - B = 0. \tag{3.6}$$

Нетрудно убедиться, что справедлива

Теорема 8. Пусть существуют квадратные матрицы Λ_1 и Λ_2 , $\Lambda_j \in \mathbb{L}(\mathbb{K}^N)$, $j = 1, 2$, удовлетворяющие уравнению (3.6) и условиям

$$\det(\Lambda_1 - \Lambda_2) \neq 0, \tag{3.7}$$

$$\Lambda_1 \Lambda_2 = \Lambda_2 \Lambda_1. \tag{3.8}$$

Тогда справедливы матричные формулы Виета:

$$A = \Lambda_1 + \Lambda_2, \quad B = -\Lambda_1 \Lambda_2. \tag{3.9}$$

Обратно, пусть матрицы A и B определены по формулам (3.9), где $\Lambda_j \in \mathbb{L}(\mathbb{K}^N)$, $j = 1, 2$, и справедливо равенство (3.8).

Тогда Λ_1 и Λ_2 являются решениями уравнения (3.6).

Пусть Λ_1 и Λ_2 такие решения уравнения (3.6), что справедливы равенства (3.7), (3.8). В фазовом пространстве системы (3.5) зададим соотношение

$$y' = \Lambda_1 y + \rho(y), \quad y \in \Omega_y(c), \quad (3.10)$$

где $\rho \in C_N(\Omega_y(c))$, $0 < c \leq a_0$. Тогда для того, чтобы соотношение (3.10) задавало гладкое поле для решений системы (3.5), содержащее стационарное решение $\{y, y'\} = \{0, 0\}$, необходимо и достаточно, чтобы $\rho(y)$ было решением задачи типа Ляпунова:

$$\begin{aligned} \partial \rho / \partial y (\Lambda_1 y + \rho) &= \Lambda_2 \rho + F(y, \Lambda_1 y + \rho), \\ y &\in \Omega_y(c), \end{aligned} \quad (3.11)$$

$$\rho(0) = 0. \quad (3.12)$$

Достаточные условия существования и единственности решения задачи (3.11), (3.12) получаем из теоремы 5. Таким образом вопрос сводится к существованию и построению корней квадратного матричного уравнения (3.6), удовлетворяющих соотношениям 1 и равенствам (3.8), (3.9).

Следуя [1], где, в частности, исправлены неточности, допущенные в [9], зафиксируем нужные нам свойства корней квадратного матричного уравнения (3.6) в переработанном и дополненном виде. Приведенные ниже утверждения достаточно просты, однако большая часть из них нами в литературе не обнаружена.

Известно (см. [21, §17], [9]), как получить множество решений уравнения (3.6), выделяя N -мерные инвариантные подпространства сопутствующего линейного оператора $L : \mathbb{K}^{2N} \rightarrow \mathbb{K}^{2N}$, где $L = \begin{pmatrix} 0 & E_N \\ B & A \end{pmatrix}$. В самом деле, уравнение (3.6) эквивалентно соотношению $L \begin{pmatrix} E_N \\ \Lambda \end{pmatrix} = \begin{pmatrix} E_N \\ \Lambda \end{pmatrix} \Lambda$. Возьмем какое-либо N -мерное подпространство (обозначим его $\mathbb{K}_{(L)}^N$), инвариантное относительно L . Возьмем какой-либо базис в $\mathbb{K}_{(L)}^N$ и из этих столбцов составим $2N \times N$ -матрицу $S = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}$, где $M_j \in \mathbb{L}(\mathbb{K}^N)$, $j = 1, 2$. Тогда справедливо $LS = SM$, $M_2 = M_1 M$. Если $\det M_1 \neq 0$, и только в этом случае, заменим S на \tilde{S} (также дающую базис): $\tilde{S} = SM_1^{-1} = \begin{pmatrix} E_N \\ \tilde{M} \end{pmatrix}$, $\tilde{M} = M_2 M_1^{-1}$. Тогда окончательно получаем $L \begin{pmatrix} E_N \\ \tilde{M} \end{pmatrix} = \begin{pmatrix} E_N \\ \tilde{M} \end{pmatrix} \tilde{M}$, $\Lambda = \tilde{M} = M_1 M M_1^{-1}$ (тем самым максимальное возможное число различных решений уравнения (3.6) равно C_{2N}^N ; см. [21, с. 138]).

Из этого способа сразу вытекает справедливость следующих утверждений (см., например, [9]).

Утверждение 3. Пусть Λ_1 и Λ_2 — два решения уравнения (3.6), не имеющие общих СЗ. Тогда для них выполняется требование (3.7).

Утверждение 4. Пусть Λ_1 и Λ_2 — два решения уравнения (3.6), удовлетворяющие соотношениям 1. Тогда такая пара матриц единственна.

Так как нас интересуют коммутирующие решения уравнения (3.6), то обратимся к известным свойствам коммутирующих матриц. Пусть C и D — коммутирующие матрицы порядка N . Пусть X_λ — собственное подпространство матрицы C , отвечающее СЗ λ , т.е. X_λ — множество решений уравнения $(C - \lambda E_N)x = 0$, $x \in \mathbb{K}^N$. Тогда, очевидно, X_λ инвариантно относительно D . Отсюда следует тот известный факт, что перестановочные матрицы всегда имеют хотя бы один общий СВ (см. [20, с. 221]), и, в частности, если C и D простой структуры, то существует общий базис в \mathbb{K}^N , в котором обе эти матрицы имеют диагональный вид (см. [20, с. 186]). В общем случае по индукции получаем

Утверждение 5. Пусть C и D — коммутирующие матрицы порядка N . Тогда для них существует общий ортонормированный базис в \mathbb{K}^N , в котором обе эти матрицы имеют верхний треугольный вид.

Обратимся к построению коммутирующих решений уравнения (3.6). Введем обозначения

$$D = A^2 + 4B, \quad R = \sqrt{D}, \tag{3.13}$$

$$\Lambda_1 = (A - R) / 2, \quad \Lambda_2 = (A + R) / 2. \tag{3.14}$$

Здесь под \sqrt{D} понимается пока любая матрица R такая, что $R^2 = D$, если таковая существует. Если матрица D непростой структуры, то существование R , во всяком случае, обеспечивается выполнением требования

$$\det D = \det(A^2 + 4B) \neq 0. \tag{3.15}$$

Непосредственной проверкой убеждаемся, что справедлива

Теорема 9. Матрицы Λ_1 и Λ_2 , определенные по формулам (3.14), тогда, и только тогда, являются решениями уравнения (3.6), когда A коммутирует с R . При этом справедливы равенства (3.7), (3.8).

Очевидным является

Утверждение 6. Пусть матрицы D и R определены по формулам (3.13) и пусть A коммутирует с R . Тогда A и B коммутируют.

Обратное утверждение (в частности, ошибочно сформулированное в [9]), вообще говоря, неверно, в чем легко убедиться на примерах. Во всяком случае, справедливо

Утверждение 7. Пусть A и B коммутируют и пусть определены матрицы D и R по формулам (3.13), причем R и $-R$ не имеют общих СЗ. Тогда A коммутирует с R .

Доказательство. Пусть $K = AR - RA$. Из равенства $AB = BA$ имеем, в частности, $AR^2 = R^2A$, что равносильно $KR + RK = 0$. Из последнего равенства получаем $K = 0$, так как линейное преобразование $L(K) = KR + RK$ в пространстве $N \times N$ -матриц является невырожденным.

Утверждение доказано.

Пусть далее в (3.6) будет

$$AB = BA. \tag{3.16}$$

На основании утверждения 5, возьмем матрицу T , приводящую одновременно A и B к верхнему треугольному виду, и занумеруем в порядке следования получившиеся диагональные элементы: $\lambda_1(A), \dots, \lambda_N(A)$, $\lambda_1(B), \dots, \lambda_N(B)$. Полагая в (3.14) $\tilde{\Lambda} = T^{-1}\Lambda T$, получим для $\tilde{\Lambda}$ снова уравнение вида (3.14) с верхними треугольными матрицами $\tilde{A} = T^{-1}AT$, $\tilde{B} = T^{-1}BT$. Полагая, что в уравнении (3.14) матрицы A и B уже имеют верхний треугольный вид, определим верхние треугольные матрицы Λ_j по формулам (3.13), (3.14) и занумеруем их диагональные элементы в порядке следования: $\lambda_1(\Lambda_j), \dots, \lambda_N(\Lambda_j)$, $j = 1, 2$. Введем обозначения:

$$P_j^{(1)} = [\operatorname{Re}\lambda_j(A)]^2 \operatorname{Re}\lambda_j(B) + [\operatorname{Im}\lambda_j(B)]^2 + \operatorname{Re}\lambda_j(A) \operatorname{Im}\lambda_j(A) \operatorname{Im}\lambda_j(B), \quad j = 1, \dots, N; \tag{3.17}$$

$$P_j^{(2)} = \operatorname{Re}\lambda_j(A) \operatorname{Im}\lambda_j(A) + 2\operatorname{Im}\lambda_j(B), \quad j = 1, \dots, N; \tag{3.18}$$

$$P_j^{(3)} = [\operatorname{Re}\lambda_j(A)]^2 - [\operatorname{Im}\lambda_j(A)]^2 + 4\operatorname{Re}\lambda_j(B), \quad j = 1, \dots, N; \tag{3.19}$$

$$\mu_j = -\operatorname{Re}\lambda_j(A) / 2, \quad j = 1, \dots, N. \tag{3.20}$$

Очевидно, что при вычислении СЗ треугольных матриц $\Lambda_j, j = 1, 2$, через СЗ треугольных матриц A и B можем обращаться с (3.13), (3.14) как со скалярными формулами.

Предварительно напомним скалярные формулы. Пусть a и b — вещественные числа, $a^2 + b^2 \neq 0$; положим $x + iy = \sqrt{a + ib}$, где корень берется с положительной вещественной частью; тогда $x = \sqrt{(a + \sqrt{a^2 + b^2}) / 2}$, $y = \operatorname{sign}(b)\sqrt{(\sqrt{a^2 + b^2} - a) / 2}$. Эти формулы однозначно определяют величины x и y для почти всех случаев, кроме $b = 0, a < 0$, в этом случае берем $x = 0, y = \sqrt{-a}$ или $y = -\sqrt{-a}$. Исключая из рассмотрения тривиальный случай, когда A и B простой структуры, из теоремы 9 и утверждений 6, 7 получаем следующее утверждение.

Теорема 10. Пусть в (3.6) для матриц A, B выполнены условия (3.15), (3.16), причем A или B непростои структуры, и пусть осуществлено приведенное выше преобразование. Пусть матрица R в (3.13) определена по следующим правилам:

1) если матрица D не имеет $C3$, лежащих на отрицательной вещественной полуоси комплексной плоскости λ , то под R понимаем такой корень из D , что

$$\operatorname{Re}\lambda_j(R) > 0, \quad j = 1, \dots, N \quad (3.21)$$

(этим требованием R определяется однозначно);

2) если матрица D имеет отрицательные вещественные $C3$, то под R понимаем какой-либо корень из D такой, что

$$\operatorname{Re}\lambda_j(R) \geq 0, \quad j = 1, \dots, N, \quad (3.22)$$

и справедливо (3.16) (этим требованием R определяется однозначно, если для всех тех j , для которых $\lambda_j(D) < 0$, под обозначением $\sqrt{\lambda_j(D)}$ понимаем чисто мнимое число с положительной мнимой частью, т.е. берем $\operatorname{Re}\lambda_j(R) = 0$, $\operatorname{Im}\lambda_j(R) > 0$, в противном случае R определяется неоднозначно).

Тогда формулы (3.14) определяют решения уравнения (3.6), удовлетворяющие соотношениям (3.8), (3.9).

Учитывая результаты [1], получаем

Теорема 11. Пусть выполняются условия теоремы 10.

Тогда для величин (3.14) справедливы следующие утверждения:

1) если для фиксированного $j, 1 \leq j \leq n$, в (3.17) будет $P_j^{(1)} > 0$, то справедливы неравенства $\operatorname{Re}\lambda_j(\Lambda_1) < 0, \operatorname{Re}\lambda_j(\Lambda_2) > 0$, а если будет $P_j^{(1)} = 0$, то при $\operatorname{Re}\lambda_j(A) < 0$ справедливо $\operatorname{Re}\lambda_j(\Lambda_1) < 0, \operatorname{Re}\lambda_j(\Lambda_2) = 0$, а при $\operatorname{Re}\lambda_j(A) > 0$ справедливо $\operatorname{Re}\lambda_j(\Lambda_1) = 0, \operatorname{Re}\lambda_j(\Lambda_2) > 0$;

2) если для фиксированного $j, 1 \leq j \leq n$, в (3.18), (3.19) будет $P_j^{(3)} > 0$ или $P_j^{(2)} \neq 0$, то справедливы неравенства $\operatorname{Re}\lambda_j(\Lambda_1) < -\mu_j, \operatorname{Re}\lambda_j(\Lambda_2) > -\mu_j$, а если будет $P_j^{(3)} < 0$ и $P_j^{(2)} = 0$, то $\operatorname{Re}\lambda_j(\Lambda_1) = \operatorname{Re}\lambda_j(\Lambda_2) = -\mu_j$, где $\mu_j = -\operatorname{Re}\lambda_j(A) / 2$.

Замечание 9. В заключение этого раздела отметим, что для одного автономного нелинейного ОДУ второго порядка вида (3.5), заданного на полуоси и с малым параметром при старшей производной, в [22], [23] получены, в частности, оценки погрешности переноса предельного условия $\{y, y'\} \rightarrow 0, \tau \rightarrow \infty$, из бесконечности в достаточно удаленную конечную точку $\tau = \tau_\infty$ с помощью приближенно заданного нелинейного соотношения (3.10). Эти оценки и проведенные в [22], [23] численные эксперименты подтверждают эффективность такого подхода по сравнению с заданием грубых условий, например вида $y(\tau_\infty) = 0$ или $y'(\tau_\infty) = 0$. (Задачи, рассмотренные в [22], [23], являются модельными в теории горения.)

4. О ПОВЕДЕНИИ РЕШЕНИЙ ВНУТРИ И ВНЕ УМР

4.1. О параметрических экспоненциальных рядах Ляпунова, порождающих в фазовом пространстве аналитические УМР

Вернемся к системе (1.8) на неотрицательной вещественной полуоси

$$dx/d\tau = \Lambda_1 x + F_1(x, y), \quad dy/d\tau = \Lambda_2 y + F_2(x, y), \quad 0 \leq \tau < \infty, \quad (4.1)$$

с условием на левом конце

$$x(0) = a, \quad a \in \mathbb{K}^n, \quad (4.2)$$

и с предельными условиями на бесконечности

$$\lim_{\tau \rightarrow \infty} [\exp(\mu\tau)x(\tau)] = 0, \quad \lim_{\tau \rightarrow \infty} [\exp(\mu\tau)y(\tau)] = 0, \quad (4.3)$$

где $x \in \mathbb{K}^n$, $y \in \mathbb{K}^m$, $m + n = N$, $a = (a_1, \dots, a_n)$ — вектор параметров, $\Lambda_{1,2}$ — постоянные матрицы, $\Lambda_1 \in \mathbb{L}(\mathbb{K}^n)$, $\Lambda_2 \in \mathbb{L}(\mathbb{K}^m)$, $F_1 : \bar{\Omega}_{x,y}(a_0) \rightarrow \mathbb{K}^n$, $F_2 : \bar{\Omega}_{x,y}(a_0) \rightarrow \mathbb{K}^m$, и выполняются условия теоремы 4 или теоремы 5. Рассмотрим задачу (4.1), (4.3) как сингулярную ЗК. Тогда, на основании результатов подразделов 2.1, 2.2 эта задача обладает n -параметрическим семейством решений, значения которых порождают в окрестности точки $x = 0, y = 0$ фазового пространства n -мерное гладкое УМР $M_+^{(n)}$, инвариантное относительно τ и задаваемое соотношением (2.2), где $\rho(x)$ есть решение задачи (2.5), (2.6). При этом при каждом фиксированном a , достаточно малом по норме, решение КрЗ (4.1)–(4.3) существует, единственно и принадлежит УМР для всех $\tau \geq 0$.

Пусть теперь $F_1 \in H_{n,a_0}$, $F_2 \in H_{m,a_0}$, а в остальном выполняются приведенные выше предположения. Тогда для аналитического УМР $M_+^{(n)}$ имеем

$$M_+^{(n)} : y = \rho(x), \quad \rho(x) = [x]_v, \quad v \geq 2, \quad x \in \Omega_x(c), \tag{4.4}$$

где $\rho(x)$ — голоморфное в нуле решение сингулярной задачи Ляпунова (2.5), (2.6). Из результатов Ляпунова [5, §23] следует, что решения системы (4.1), значения которых порождают УМР (4.4), представимы сходящимися n -параметрическими экспоненциальными рядами с коэффициентами, имеющими рост не выше степенного. При выполнении указанных выше предположений эти ряды могут быть построены с помощью простой рекурсивной процедуры, дающей представление об устройстве этих рядов. В самом деле, представим решения задачи (4.1)–(4.3) (т.е. решения, принадлежащие УМР (4.4) для всех $\tau \geq 0$) в виде разложений по степеням параметров a_1, \dots, a_n :

$$x(\tau, a) = \sum_{p=1}^{\infty} x_{(p)}(\tau, a), \quad y(\tau, a) = \sum_{p=v}^{\infty} y_{(p)}(\tau, a), \quad \tau \geq 0, \tag{4.5}$$

где $x_{(p)}(\tau, a)$ ($y_{(p)}(\tau, a)$) — однородная n -векторная (m -векторная) форма степени p величин a_1, \dots, a_n ,

$$x_{(p)}(\tau, a) = \sum_{|l|=p} X_l(\tau) a^l, \tag{4.6}$$

$$X_l(\cdot) \in \mathbb{K}^n, \quad |l| = p, \quad p = 1, 2, \dots,$$

$$y_{(p)}(\tau, a) = \sum_{|l|=p} Y_l(\tau) a^l, \tag{4.7}$$

$$Y_l(\cdot) \in \mathbb{K}^m, \quad |l| = p, \quad p = v, v + 1, \dots,$$

причем из (4.2), (4.3) следует выполнение условий

$$x_{(1)}(0, a) = a, \quad x_{(p)}(0, a) = 0, \quad p = 2, 3, \dots, \tag{4.8}$$

$$\lim_{\tau \rightarrow \infty} [\exp(\mu\tau) x_{(p)}(\tau, a)] = 0, \quad p = 1, 2, \dots, \tag{4.9}$$

$$\lim_{\tau \rightarrow \infty} [\exp(\mu\tau) y_{(p)}(\tau, a)] = 0, \quad p = v, v + 1, \dots \tag{4.10}$$

Далее представим

$$F_j(\sum_{p=1}^{\infty} x_{(p)}(\tau, a), \sum_{p=v}^{\infty} y_{(p)}(\tau, a)), \quad j = 1, 2,$$

в виде разложений по формам:

$$\begin{aligned}
 & F_j \left(\sum_{p=1}^{\infty} x_{(p)}(\tau, a), \sum_{p=v}^{\infty} y_{(p)}(\tau, a) \right) = \\
 = & \sum_{\substack{|q|+|s| \geq (2-j)(2-\delta_{|q|,0}) + (j-1)(v\delta_{0,|s|} + 2(1-\delta_{0,|s|})) \\ |q|, |s|=0}}^{\infty} F_{qs}^{(j)} \left(\sum_{p=1}^{\infty} x_{(p)}(\tau, a) \right)^q \left(\sum_{p=v}^{\infty} y_{(p)}(\tau, a) \right)^s = \\
 & = \sum_{p=2(2-j)+v(j-1)}^{\infty} F_{(p)}^{(j)}(x_{(1)}, \dots, x_{(p-1)}; y_{(v)}, \dots, y_{(p-1)}) + \\
 & \quad + (2-j) \partial F_1 \partial y(0,0) \sum_{p=v}^{\infty} y_{(p)},
 \end{aligned} \tag{4.11}$$

где, в частности, справедливо

$$F_{(2)}^{(1)}(x_{(1)}) = \sum_{|q|=2} F_{q0}^{(1)} x_{(1)}^q, \quad F_{(v)}^{(2)}(x_{(1)}) = \sum_{|q|=v} F_{q0}^{(2)} x_{(1)}^q, \tag{4.12}$$

а коэффициенты разложений $F_{(p)}^{(j)}$, $j = 1, 2$, по a^l , $|l| = p$, образованы из коэффициентов $F_{qs}^{(j)}$ и коэффициентов разложений предыдущих форм $x_{(1)}, \dots, x_{(p-1)}$, $y_{(v)}, \dots, y_{(p-1)}$ с помощью только действий типа сложения и умножения.

Подставляя разложения по формам (4.5), (4.11) в (4.1) и собирая формы одинаковой степени, получим рекуррентную последовательность ОДУ для определения этих форм:

$$x'_{(1)} = \Lambda_1 x_{(1)}, \quad y'_{(v)} = \Lambda_2 y_{(v)} + F_{(v)}^{(2)}(x_{(1)}), \tag{4.13}$$

$$\begin{aligned}
 x'_{(p)} &= \Lambda_1 x_{(p)} + F_{(p)}^{(1)}(x_{(1)}, \dots, x_{(p-1)}) + \\
 & \quad + \delta_{p,v} \partial F_1 \partial y(0,0) y_{(v)}, \quad 2 \leq p \leq v,
 \end{aligned} \tag{4.14}$$

$$x'_{(p)} = \Lambda_1 x_{(p)} + F_{(p)}^{(1)}(x_{(1)}, \dots, x_{(p-1)}; y_{(v)}, \dots, y_{(p-1)}) + \partial F_1 \partial y(0,0) y_{(p)}, \quad p \geq v+1, \tag{4.15}$$

$$y'_{(p)} = \Lambda_2 y_{(p)} + F_{(p)}^{(2)}((x_{(1)}, \dots, x_{(p-1)}; y_{(v)}), \dots, y_{(p-1)}), \quad p \geq v+1. \tag{4.16}$$

Отсюда, учитывая требования (4.8)–(4.10), окончательно для всех $\tau \geq 0$ получаем

$$x_{(1)}(\tau, a) = a \exp(\Lambda_1 \tau), \tag{4.17}$$

$$y_{(v)} = - \int_{\tau}^{\infty} \exp(\Lambda_2(\tau - s)) F_{(v)}^{(2)}(x_{(1)}(s, a)) ds,$$

$$\begin{aligned}
 x_{(p)}(\tau, a) &= \int_0^{\tau} \exp(\Lambda_1(\tau - s)) \times \left[F_{(p)}^{(1)}(x_{(1)}(s, a), \dots, x_{(p-1)}(s, a)) + \right. \\
 & \quad \left. + \delta_{p,v} \partial F_1 \partial y(0,0) y_{(v)}(s, a) \right] ds, \\
 & \quad 2 \leq p \leq v,
 \end{aligned} \tag{4.18}$$

$$\begin{aligned}
 x_{(p)}(\tau, a) &= \int_0^{\tau} \exp(\Lambda_1(\tau - s)) \times \left[F_{(p)}^{(1)} \left(x_{(1)}(s, a), \dots, x_{(p-1)}(s, a); \right. \right. \\
 & \quad \left. \left. y_{(v)}(s, a), \dots, y_{(p-1)}(s, a) \right) + \right. \\
 & \quad \left. + \partial F_1 \partial y(0,0) y_{(p)}(s, a) \right] ds, \\
 & \quad p \geq v+1,
 \end{aligned} \tag{4.19}$$

$$y_{(p)}(\tau, a) = -\int_{\tau}^{\infty} \exp(\Lambda_2(\tau - s)) F_{(p)}^{(2)} \times (x_{(1)}(s, a), \dots, x_{(p-1)}(s, a), y_{(v)}(s, a), \dots, y_{(p-1)}(s, a)) ds, \quad (4.20)$$

$$p \geq v + 1.$$

Замечание 10. При сделанных предположениях задача (4.1)–(4.3) эквивалентна ЗК на гладком многообразии (4.4):

$$\frac{dx}{d\tau} = \Lambda_1 x + F_1(x, \rho(x)), \quad 0 \leq \tau < \infty, \quad x(0) = a,$$

$$y(\tau) = \rho(x(\tau)), \quad \tau \geq 0.$$

Из сравнения формул (4.17)–(4.20) и (2.17), в частности, следует тот общий для теории УМР факт, что строение всего УМР является, как правило, более простым, чем поведение отдельных решений, его порождающих.

О построении УМР и рядов Ляпунова в конкретных прикладных задачах см. в статьях, указанных здесь в заключении, и цитированных там работах.

4.2. О поведении решений вне УМР

Пусть теперь $\{x(\tau); y(\tau)\}$ — какое-либо решение системы (4.1), достаточно малое в момент $\tau = 0$, но не принадлежащее УМР (4.4). Введем функцию уклонения решения от $M_+^{(n)}$:

$$\eta(\tau) = y(\tau) - \rho(x(\tau)), \quad 0 \leq \tau \leq \tau_0.$$

Для $\eta(\tau)$ получаем

$$\eta' = \Lambda_2 \eta + W(x(\tau), \eta), \quad 0 \leq \tau \leq \tau_0, \quad \eta(0) = \eta_0 \neq 0, \quad (4.21)$$

где $|\eta_0|$ — мало,

$$W(x, \eta) = F_2(x, \eta + \rho(x)) - F_2(x, \rho(x)) - \partial \rho / \partial x(x) [F_1(x, \eta + \rho(x)) - F_1(x, \rho(x))],$$

так что $W(x, 0) \equiv 0, \partial W / \partial \eta(0, 0) = 0$, т.е. $\forall \epsilon > 0 \exists \delta_\epsilon > 0$ такое, что при $0 \leq \tau \leq \tau_0$ справедливо $|W(x, \eta)| \leq \epsilon |\eta|$, если $\max\{\sup_{0 \leq \tau \leq \tau_0} |x(\tau)|; \sup_{0 \leq \tau \leq \tau_0} |y(\tau)|\} \leq \delta_\epsilon$.

Тогда при выполнении неравенства (1.9) с $\mu = 0$ стандартным образом получаем из (4.21) известную оценку удаления от $M_+^{(n)}$ решения системы (4.1), начинающегося вне этого УМР, но остающегося достаточно малым при $0 \leq \tau \leq \tau_0$:

$$|y(\tau) - \rho(x(\tau))| \geq C \exp(b\tau), \quad 0 \leq \tau \leq \tau_0, \quad (4.22)$$

где $0 < C = \text{const}$, $0 < b = \text{const}$ (не будем уточнять эти постоянные, см. [11, гл. XIII]). Эта оценка показывает, что расстояние решения $\{x(\tau); y(\tau)\}$, измеренное по нормали к поверхности $y = 0$, есть экспоненциально возрастающая функция τ . Если в (1.9) справедливо $\mu \neq 0$, то соответствующую оценку получим для величины $\exp(\mu\tau)\eta(\tau)$, на чем останавливаться не будем.

ЗАКЛЮЧЕНИЕ

В заключение отметим, что результаты изложенной теории применялись в ряде работ для корректной постановки, анализа и численного решения сингулярных нелинейных краевых задач, возникающих в различных прикладных областях. Для примера, наряду с [3], укажем статьи [24]–[26], в которых, в частности, устранены допущенные в литературе неточности при постановке, аппроксимации и численном решении некоторых сингулярных краевых задач математической физики (см. подробнее замечание 1).

СПИСОК ЛИТЕРАТУРЫ

1. *Конюхова Н. Б.* О стационарной задаче Ляпунова для системы квазилинейных уравнений с частными производными первого порядка // Дифференц. ур-ния. 1994. Т. 30. № 8. С. 1384–1395.

2. *Конюхова Н. Б.* Об устойчивых многообразиях Ляпунова для автономных систем нелинейных обыкновенных дифференциальных уравнений // *Ж. вычисл. матем. и матем. физ.* 1994. Т. 34. № 10. С. 1358–1379.
3. *Конюхова Н. Б.* Гладкие многообразия Ляпунова и сингулярные краевые задачи // *Сообщ. по прикл. матем. ВЦ РАН. М.: ВЦ РАН, 1996.*
4. *Kuznetsov Yu. A.* Elements of Applied Bifurcation Theory. New York: Springer-Verlag, 1995.
5. *Ляпунов А. М.* Общая задача об устойчивости движения. М. — Л.: Гостехтеоретиздат, 1950.
6. *Лефшец С.* Геометрическая теория дифференциальных уравнений. М.: Изд-во иностр. лит., 1961.
7. *Гельфанд И. М., Фомин С. В.* Вариационное исчисление. М.: Физматлит, 1961.
8. *Абрамов А. А.* О граничных условиях в особой точке для систем линейных обыкновенных дифференциальных уравнений // *Ж. вычисл. матем. и матем. физ.* 1971. Т. 11. № 1. С. 275–278.
9. *Абрамов А. А., Конюкхлова Н. В.* Transfer of admissible boundary conditions from a singular point of linear ordinary differential equations // *Sov. J. Numer. Anal. Math. Modelling.* 1986. 4. N 4. V. 1. P. 245–265 (VNU Science Press., Utrecht, The Netherlands).
10. *Абрамов А. А., Конюхова Н. Б.* Допустимые граничные условия на бесконечности или в особой точке для систем линейных обыкновенных дифференциальных уравнений // *Numer. Anal. Math. Modelling.* 1990. V. 24. P. 181–198 (Banach Center Publ., Warsaw, PWN-Polish Scient. Publ.).
11. *Коддингтон Э. А., Левинсон Н.* Теория обыкновенных дифференциальных уравнений. М.: Изд-во иностр. лит., 1958.
12. *Демидович Б. П.* Лекции по математической теории устойчивости. М.: Наука, 1957.
13. *Далецкий Ю. А., Крейн М. Г.* Устойчивость решений дифференциальных уравнений в банаховом пространстве. М.: Наука, 1970.
14. *Конюхова Н. Б.* О существовании и единственности решений сингулярных задач Коши для систем нелинейных функционально-дифференциальных уравнений // *Докл. АН СССР.* 1987. Т. 295. № 4. С. 798–801.
15. *Конюхова Н. Б.* О существовании устойчивых начальных многообразий для систем нелинейных функционально-дифференциальных уравнений // *Докл. АН СССР.* 1989. Т. 306. № 3. С. 535–540.
16. *Конюхова Н. Б.* Об устойчивых начальных многообразиях для систем нелинейных функционально-дифференциальных уравнений // *В сб.: Аналитич. и числ. методы решения задач матем. физ. М.: ВЦ АН СССР, 1989. С. 136–154.*
17. *Конюхова Н. Б.* Сингулярные задачи Коши для некоторых систем нелинейных функционально-дифференциальных уравнений // *Дифференц. уравнения.* 1995. Т. 31. № 8. С. 1340–1347.
18. *Конюхова Н. Б.* Сингулярные задачи Коши для систем обыкновенных дифференциальных уравнений // *Ж. вычисл. матем. и матем. физ.* 1983. Т. 23. № 3. С. 629–645.
19. *Guan P., Li Y. Y.* $C^{1,1}$ estimates for solutions of a problem of Alexandrov // *Commun. Pure and Appl. Math.* 1997. V. 50. P. 789–811.
20. *Гантмахер Ф. Р.* Теория матриц. М.: Гостехтеориздат, 1953.
21. *Икрамов Х. Д.* Численное решение матричных уравнений. М.: Наука, 1984.
22. *Задорин А. И.* Численное решение уравнения с малым параметром на бесконечном интервале // *Ж. вычисл. матем. и матем. физ.* 1998. Т. 38. № 10. С. 1671–1682.
23. *Задорин А. И.* Перенос краевого условия из бесконечности при численном решении уравнений второго порядка с малым параметром // *Сибирский ж. вычисл. матем.* 1999. Т. 2. № 1. С. 21–35.
24. *Конюхова Н. Б., Курочкин С. В.* Сингулярные нелинейные задачи для автомодельных решений уравнений пограничного слоя с нулевым градиентом давления: анализ и численное решение // *Ж. вычисл. матем. и матем. физ.* 2021. Т. 61. № 10. С. 1619–1645.
25. *Konyukhova, N.B. and Sukov, A. I.* Smooth Lyapunov manifolds and correct mathematical simulation of nonlinear singular problems in mathematical physics // *Mathematical Modeling. Problems, Methods, Applications.* — New York–Boston–Dordrecht–London–Moscow: Kluwer Academic/ Plenum Publishers, 2001. P. 205–217.
26. *Konyukhova, N.B. and Sukov, A. I.* On correct statement of singular BVPs for autonomous systems of nonlinear ODEs with the applications to hydrodynamics // *Proc. Inter. Seminar “Day on Diffraction — 2003” (St. Petersburg, Russia, June 24–27, 2003) / Ed. by I. V. Andronov.* — St. Petersburg: Faculty of Physics, SPbU, 2003. P. 99–109 (IEEE Xplore, Digital Library, 2003; <https://doi.org/10.1109/DD.2003.23818>).

УРАВНЕНИЯ В ЧАСТНЫХ ПРОИЗВОДНЫХ

УДК 519.63

ЯВНЫЕ ЧИСЛЕННО РЕАЛИЗУЕМЫЕ ФОРМУЛЫ ДЛЯ ОПЕРАТОРОВ ПУАНКАРЕ–СТЕКЛОВА

© 2024 г. А. С. Демидов^{1,*}, А. С. Самохин²

¹119991 Москва, Ленинские горы, 1, МГУ им. М. В. Ломоносова, Россия;

²117997 Москва, ул. Профсоюзная, 65, Институт проблем управления им. В. А. Трапезникова Российской академии наук, Россия

*e-mail: demidov.alexandre@gmail.com

Поступила в редакцию 17.04.2023 г.

Переработанный вариант 09.09.2023 г.

Принята к публикации 20.10.2023 г.

Представлены явные численно реализуемые формулы для операторов Пуанкаре–Стеклова, таких как операторы Дирихле–Неймана, Дирихле–Робена, Робена1–Робена2, Гринберга–Майергойза, относящихся к двумерному уравнению Лапласа. Эти формулы базируются на лемме об однолистном изометрическом отображении замкнутой аналитической кривой на окружность. Численные результаты для областей с весьма сложной геометрией получены для нескольких тестовых гармонических функций для операторов Дирихле–Неймана и Дирихле–Робена. Библ. 9. Фиг. 9.

Ключевые слова: операторы Пуанкаре–Стеклова, однолистное изометрическое отображение аналитической кривой, явные численно реализуемые формулы.

DOI: 10.31857/S0044466924020064, EDN: YJZOIN

1. ВВЕДЕНИЕ

Идея альтернирующего метода Шварца [1] получила развитие в так называемых методах декомпозиции “сложной” области $\Omega = \cup_{k \geq 1}^m \omega_k$ на “простые” подобласти при построении решения $u : \Omega \rightarrow \mathbb{R}$ краевой задачи в исходной сложной области. Это позволяет (см., например, [2]) эффективно использовать распараллеливание численных расчетов. А именно, искомое решение $u : \Omega \rightarrow \mathbb{R}$ в сложной области может быть найдено в результате “склейки” параллельно построенных решений $u_k : \omega_k \rightarrow \mathbb{R}$, если для смежных областей ω_j и ω_k известно соответствие между граничными условиями для u_j и u_k . В некоторых задачах (см., например, [3], [4]) эти условия имеют вид

$$(1 - \lambda)u_k - (1 + \lambda)u_j = 2f, \quad |\lambda| \leq 1, \quad \frac{\partial u_k}{\partial \nu} = \frac{\partial u_j}{\partial \nu}, \tag{1}$$

где ν — нормаль к границе смежных областей ω_k и ω_j . Но особую значимость имеют граничные операторы Робена¹

$$\mathcal{R}_k : u|_{\omega_k} \mapsto \left(\alpha_k u_k + \beta_k \frac{\partial u_k}{\partial \nu} \right) |_{\partial \omega_k}, \quad \alpha_k + \beta_k \neq 0, \tag{2}$$

в которых ν — внешняя нормаль к границе соответствующей «простой» области ω_k .

Впрочем, и для одной конкретной области $\Omega \subset \mathbb{R}^n$ большой интерес (см., например, [5], [6]) представляет соответствие между двумя операторами Робена, в частности, между операторами Дирихле $\mathcal{D} : u|_{\Omega} \mapsto u|_{\partial \Omega}$ и Неймана $\mathcal{N} : u|_{\Omega} \mapsto \frac{\partial u}{\partial \nu} |_{\partial \Omega}$, чьи альтернативные композиции $\mathcal{DN} : u|_{\partial \Omega} \mapsto \frac{\partial u}{\partial \nu} |_{\partial \Omega}$

¹Граничные операторы вида $u + b \frac{\partial u}{\partial \nu}$ впервые появились у Исаака Ньютона (1642–1727) и Георгия Рихмана (1711–1753) в качестве

эмпирической зависимости процесса теплообмена. Но недаром Французская академия наук дважды удостоивала Виктора Робена (1855–1897) премией Франсуа (1893 и 1897), а также премией Ponsel (1895). Ведь Робен, рассматривая проблему распределения плотности поверхностного тока в проводнике, свел ее к интегральному уравнению (задолго до появления теории интегральных уравнений) и в случае выпуклого проводника нашел решение этого уравнения методом последовательных приближений.

и $\mathcal{ND} : \frac{\partial u}{\partial \nu} |_{\partial \Omega} \mapsto u |_{\partial \Omega}$ называются, соответственно, операторами Дирихле–Неймана и Неймана–Дирихле. Соответствия граничных операторов типа (1) или композиции операторов Робена стали называть, следуя [7], операторами Пуанкаре–Стеклова, поскольку в работах [8] и [9] такого рода операторы играли важную роль при построении решений весьма общих краевых задач для уравнений математической физики 19 века.

В настоящей работе строятся и численно реализуются формулы для операторов Пуанкаре–Стеклова, относящихся к решению уравнения Лапласа в односвязной области $\Omega \subset \mathbb{R}^2$ с C^1 -границей Γ длины 2π . Если область Ω_ε отличается от Ω только тем, что ее граница $\Gamma_\varepsilon = \partial \Omega_\varepsilon$ аналитична и в хаусдорфовой C^1 -метрике отличается от Γ на величину порядка $\varepsilon > 0$, то оператор Пуанкаре–Стеклова для Ω_ε и для Ω будет отличаться в метрике C^1 на тот же порядок. Поэтому в дальнейшем будем считать, что кривая Γ аналитична. Это позволяет воспользоваться приведенной в разд. 2 леммой, которая играет ключевую роль при построении формул для операторов Пуанкаре–Стеклова. Дело в том, что в этой лемме явно строится однолистное отображение

$$z : V_{\mathbb{T}} \ni \zeta = \rho e^{i\theta} \mapsto z(\zeta) \in V_\Gamma, \left| \frac{dz(\zeta)}{d\zeta} \right| \stackrel{\text{def}}{=} 1 \text{ для } \zeta \in \mathbb{T} \tag{3}$$

окрестности $V_{\mathbb{T}}$ единичной окружности $\mathbb{T} \stackrel{\text{def}}{=} \mathbb{R} / 2\pi$ на окрестность V_Γ кривой Γ , которое изометрично преобразует \mathbb{T} в Γ (\mathbb{T} и Γ изометричны, а их окрестности связаны однолистным отображением). Поэтому при таком отображении единичная нормаль к окружности \mathbb{T} переходит в *единичную* (!) нормаль к Γ , в результате чего нормальная производная $\frac{\partial u}{\partial \nu} |_\Gamma$ совпадает с нормальной производной $\frac{\partial U(\rho e^{i\theta})}{\partial \rho} |_{\rho=1}$ функции $U(\zeta) = u(z(\zeta))$.

2. ИЗОМЕТРИЧНОЕ ОДНОЛИСТНОЕ ОТОБРАЖЕНИЕ АНАЛИТИЧЕСКОЙ КРИВОЙ НА ОКРУЖНОСТЬ

Пусть $\Omega \subset \mathbb{R}^2$ — односвязная область с аналитической границей Γ длины 2π , а $s \in \mathbb{T} = \mathbb{R} / 2\pi$ — натуральный параметр на Γ , т.е. s — длина дуги ${}_0P_s$ кривой Γ , отсчитываемой в положительном направлении от некоторой фиксированной точки P_0 с декартовыми координатами $(x(0), y(0))$ до точки P_s с координатами $(x(s), y(s))$. Без ограничения общности можно считать, что $(x(0), y(0)) = (1, 0)$. Тем самым, применив простейшие геометрические построения, получаем

$$\begin{aligned} \dot{x}(s) &= -\sin N(s), \quad x(0) = 1; \\ \dot{y}(s) &= \cos N(s), \quad y(0) = 0, \end{aligned} \tag{4}$$

где $N(s)$ — радианная мера угла между осью x и внешней нормалью ν к Γ в точке $P_s \in \Gamma$. Ввиду (4), замкнутость кривой Γ влечет условие

$$\int_0^{2\pi} \sin N(s) ds = \int_0^{2\pi} \cos N(s) ds = 0 \Leftrightarrow \int_0^{2\pi} e^{iN(\theta)} d\theta = \frac{1}{i} \int_0^{2\pi} e^{i[N(\theta)-\theta]} d\theta = 0. \tag{5}$$

Поскольку длина кривой Γ задана, эта кривая полностью определяется заданием угловой функции $s \mapsto N(s)$ и, соответственно, 2π -периодической функцией

$$Q : \mathbb{T} = \mathbb{R} / 2\pi \ni s \mapsto Q(s) \stackrel{\text{def}}{=} N(s) - s. \tag{6}$$

Обозначим через Q^σ комплексно-значное аналитическое продолжение вещественной аналитической функции Q в такую окрестность V_Γ кривой $\Gamma \subset \mathbb{R}^2 \simeq \mathbb{C} \ni z = x + iy$, которая содержит все точки, отстоящие от Γ на расстоянии $\sigma > 0$. Как известно (см., например, [10]), коэффициенты L_k и M_k в фурье-представлении

$$Q(s) = \sum_{k \geq 0} (L_k \cos ks + M_k \sin ks), \tag{7}$$

$$|L_k| \leq C e^{-k\sigma_1}, \quad |M_k| \leq C e^{-k\sigma_1},$$

с константой $\sigma_1 = \sigma_1(\sigma, \Gamma) > 0$ и константой C , мажорирующей модуль функции Q^σ .

Лемма 1². *Справедливы следующие три утверждения.*

1. Существует такая окрестность $V_{\mathbb{T}} \ni \zeta = \rho e^{i\theta}$ единичной окружности $\mathbb{T} = \{\rho = 1, \theta \in \mathbb{R} / 2\pi\}$, в которой функции A и B , заданные формулами

$$A(\rho, \theta) = \sum_{k \geq 1} \frac{\rho^k - \rho^{-k}}{2} \{M_k \cos k\theta - L_k \sin k\theta\},$$

$$B(\rho, \theta) = \sum_{k \geq 0} \frac{\rho^k + \rho^{-k}}{2} \{L_k \cos k\theta + M_k \sin k\theta\},$$

гармонически сопряжены, $A(1, \theta) = 0$, а $B(1, \theta) \stackrel{(6),(6)(7)}{=} Q(\theta)$.

2. Определено отображение

$$z : V_{\mathbb{T}} \ni \zeta = \rho e^{i\theta} \mapsto z(\zeta) = x(\rho, \theta) + iy(\rho, \theta) \in V_{\Gamma}, \tag{8}$$

заданное интегралом

$$z(\zeta) = 1 + \int_1^{\zeta} e^{A+iB} d\zeta, \quad \zeta \in V_{\mathbb{T}} \Leftrightarrow A + iB = \ln \frac{dz}{d\zeta}, \tag{9}$$

который не зависит, в силу (5), от пути интегрирования.

3. Отображение (8) является однолиственным и оно изометрично переводит окружность \mathbb{T} на кривую Γ , т.е. справедливо упомянутое в (3) тождество

$$\left| \frac{dz(\zeta)}{d\zeta} \right| \equiv 1 \text{ для } \zeta \in \mathbb{T}.$$

3. ПОСТРОЕНИЕ ОПЕРАТОРОВ ДИРИХЛЕ–РОБЕНА И РОБЕНА1–РОБЕНА2

Пусть u — решение задачи Дирихле для уравнения Лапласа, т.е. $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$ в Ω , $u(P_s) = F(s)$, где функция $F \in C^1(\mathbb{T})$ задана рядом Фурье

$$F : \mathbb{T} \ni s \mapsto \sum_{k \geq 0} (a_k \cos ks + b_k \sin ks). \tag{10}$$

Возьмем кривую $\gamma_r = z(C_r) \subset V_{\Gamma}$, где $C_r = \{\zeta \mid |\zeta| = r < 1\}$. Обозначим через g след на $\gamma = \gamma_r$ решения u исходной задачи Дирихле. Положим $f : \mathbb{T} \ni \theta \mapsto f(\theta) = g(z) \Big|_{z=z(re^{i\theta}) \in \gamma}$.

Ряд Фурье этой периодической функции имеет вид

$$f : \mathbb{T} \ni \theta \mapsto \sum_{k \geq 0} (c_k \cos k\theta + d_k \sin k\theta). \tag{11}$$

Учитывая (10) и (11), положим $p_0 = c_0$, а для $k \geq 1$ имеем

$$p_k = a_k - \frac{c_k r^k - a_k r^{2k}}{1 - r^{2k}}, \quad q_k = b_k - \frac{d_k r^k - b_k r^{2k}}{1 - r^{2k}}, \quad \varphi_k = \frac{c_k - a_k r^k}{1 - r^{2k}}, \quad \psi_k = \frac{d_k - b_k r^k}{1 - r^{2k}}.$$

Тогда гармоническая в кольце $V_{\mathbb{T}} = \{r < \rho < 1; \theta \in \mathbb{T}\}$ функция³

² Короткое доказательство, предпочтительнее изначальных [11] и [12], приведено в сноске 4 статьи [13].

³ Функция U , заданная рядом (12), зависит от параметра r , как и коэффициенты c_k и d_k в формуле (11). Однако эта зависимость может сказаться лишь на скорости сходимости ряда в формуле (12), но никоим образом не на представленных ниже результатах и их доказательствах.

$$\zeta = \rho e^{i\theta} \mapsto U(\rho, \theta) = p_0 + \sum_{k \geq 1} \left\{ \begin{aligned} &\rho^k [p_k \cos k\theta + q_k \sin k\theta] + \\ &+ \left(\frac{r}{\rho}\right)^k [\varphi_k \cos k\theta + \psi_k \sin k\theta] \end{aligned} \right\} \quad (12)$$

удовлетворяет таким граничным условиям: $U(1, \theta) = F(\theta)$, $U(r, \theta) = f(\theta)$.

Теорема 1. Для оператора Дирихле–Неймана $\mathcal{DN}(u|_\Gamma) \stackrel{\text{def}}{=} \frac{\partial u}{\partial \nu}|_\Gamma$ имеем

$$\mathcal{DN}(u|_{s \in \Gamma}) \stackrel{(3)}{=} \frac{\partial U(\rho, s)}{\partial \rho} \Big|_{\rho=1}, \quad (13)$$

а для оператора Дирихле–Робена справедлива формула

$$\mathcal{DR}(u|_{s \in \Gamma}) = (\alpha u + \beta \frac{\partial u}{\partial \nu}) \Big|_{P_s \in \Gamma} \stackrel{(12),(13)}{=} \alpha U(1, s) + \beta \frac{\partial U(\rho, s)}{\partial \rho} \Big|_{\rho=1}. \quad (14)$$

Остается представить численно реализуемые формулы для $u|_\gamma$, где u — решение исходной задачи Дирихле. Известно (см. [14] и [15]), что

$$u(x, y) = \int_0^{2\pi} \mu(t) K(x(t) - x, y(t) - y) dt, \quad \text{где } K(\xi(t), \eta(t)) = \frac{\eta'(t)\xi(t) - \eta(t)\xi'(t)}{\xi^2(t) + \eta^2(t)}, \quad (15)$$

т.е. $K(\xi(t), \eta(t)) = \frac{d}{dt} \text{arctg} \frac{\eta(t)}{\xi(t)}$, а μ — решение интегрального уравнения

$$\mu(\tau) + \frac{1}{\pi} \int_0^{2\pi} \mu(t) \frac{d}{dt} \text{arctg} \frac{y(t) - y(\tau)}{x(t) - x(\tau)} dt = \frac{1}{\pi} F(\tau). \quad (16)$$

Аппроксимируя интеграл в (16) суммой (подробности см. в [16]) и полагая последовательно $\tau = t_1, \dots, t_n$, получим систему линейных алгебраических уравнений, из которой находится приближение к плотности μ и, соответственно, приближение исходной задачи Дирихле.

Замечание 1. Ядро K интегрального оператора (15) имеет особенность $0/0$ там, где $\xi = x(t) - x$ и $\eta = y(t) - y$ обращаются в ноль. При численном счете эта особенность устраняется следующим образом. Поскольку

$$\begin{aligned} x(t) &\approx x(\tau) + x'(\tau)(t - \tau) + \frac{1}{2} x''(\tau)(t - \tau)^2, \\ y(t) &\approx y(\tau) + y'(\tau)(t - \tau) + \frac{1}{2} y''(\tau)(t - \tau)^2, \end{aligned}$$

а

$$x'(t) \approx x'(\tau) + x''(\tau)(t - \tau), \quad y'(t) \approx y'(\tau) + y''(\tau)(t - \tau),$$

то

$$\begin{aligned} K &= \frac{y'(t)[x(t) - x(\tau)] - x'(t)[y(t) - y(\tau)]}{[x(t) - x(\tau)]^2 + [y(t) - y(\tau)]^2} \approx \\ &\approx \frac{[y'(\tau) + y''(\tau)(t - \tau)][x'(\tau)(t - \tau) + \frac{1}{2} x''(\tau)(t - \tau)^2] - [x'(\tau) + x''(\tau)(t - \tau)][y'(\tau)(t - \tau) + \frac{1}{2} y''(\tau)(t - \tau)^2]}{[x'(\tau)(t - \tau)]^2 + [y'(\tau)(t - \tau)]^2} = \\ &= \frac{y''(\tau)x'(\tau) - x''(\tau)y'(\tau)}{[x'(\tau)]^2 + [y'(\tau)]^2} - \frac{1}{2} \frac{x'(\tau)y''(\tau) - y'(\tau)x''(\tau)}{[x'(\tau)]^2 + [y'(\tau)]^2} = \frac{1}{2} \frac{x'(\tau)y''(\tau) - y'(\tau)x''(\tau)}{[x'(\tau)]^2 + [y'(\tau)]^2}. \end{aligned}$$

Поэтому, полагая $x_k = k\tau$ и $y_k = k\tau$ для $k = -1, 0, 1$, получаем вблизи особой точки $t - \tau = 0$ такую аппроксимацию

$$K \approx \frac{1}{h} \frac{(y_1 - 2y_0 + y_{-1})(x_1 - x_{-1}) - (x_1 - 2x_0 + x_{-1})(y_1 - y_{-1})}{(x_1 - x_{-1})^2 + (y_1 - y_{-1})^2}. \quad (17)$$

Замечание 2. Численная реализация оператора Робена1–Робена2, т.е. оператора

$$\left(\alpha_1 u + \beta_1 \frac{\partial u}{\partial \nu} \right) \Big|_{\Gamma} \mapsto \left(\alpha_2 u + \beta_2 \frac{\partial u}{\partial \nu} \right) \Big|_{\Gamma}, \quad |\alpha_k| + |\beta_k| \neq 0,$$

сводится к рассмотренной выше задаче Дирихле–Робена2 с использованием построения (см., например, [17]) решения такой γ -задачи Робена1–Дирихле $(\alpha_1 u + \beta_1 \frac{\partial u}{\partial \nu}) \Big|_{\Gamma} \mapsto u \Big|_{\gamma}$.

Замечание 3. Еще один частный случай оператора Пуанкаре–Стеклова — это оператор Гринберга–Майергойза \mathcal{G}_M , который задается при $-1 \leq \lambda < 1$ условиями (1), т.е. так:

$$\begin{aligned} (1 - \lambda)u(P_s^-) - (1 + \lambda)u(P_s^+) &= 2F(P_s), \\ \frac{\partial u}{\partial \nu}(P_s^-) &= \frac{\partial u}{\partial \nu}(P_s^+). \end{aligned} \quad (18)$$

Здесь значение в точке P_s^+ (соответственно в P_s^-) означает предельное значение в точке $P_s \in \Gamma$ вдоль нормали ν с внешней (соответственно внутренней) стороны Γ . Явные формулы для следов $u \Big|_{\gamma}$ и $u \Big|_{\Gamma}$ решения задачи (18) обозначены выше в случае задачи Дирихле, т.е. для $\lambda = -1$, а при $|\lambda| < 1$ такого же типа формулы через потенциалы с плотностью, удовлетворяющей интегральному уравнению II рода, даны в [4]. Это позволяет получить (аналогично теореме 1) значение $\mathcal{G}_M(F)$.

Замечание 4. Рассмотренный выше вопрос о построении явных численно реализуемых формул для операторов Пуанкаре–Стеклова, относящихся к уравнению Лапласа в односвязной области Ω , позволяет, ввиду [13], получить явные численно реализуемые формулы для операторов Пуанкаре–Стеклова в случае уравнений

$$\operatorname{div}(\alpha(w)\nabla w), \quad \operatorname{div}(\beta\nabla w) = 0,$$

в том числе для некоторых многомерных областей.

4. ЧИСЛЕННАЯ РЕАЛИЗАЦИЯ И ЕЕ РЕЗУЛЬТАТЫ

Представленные далее численные результаты относятся к операторам Дирихле–Неймана и Дирихле–Робена для уравнения Лапласа в плоской односвязной области Ω с весьма сложной C^1 -границей Γ . При численном моделировании брались различные тестовые гармонические функции u . Приводимые ниже результаты получены для $u(x, y) = x + y$, $u(x, y) = \sin xy$ и $u(x, y) = x^2 - y^2$.

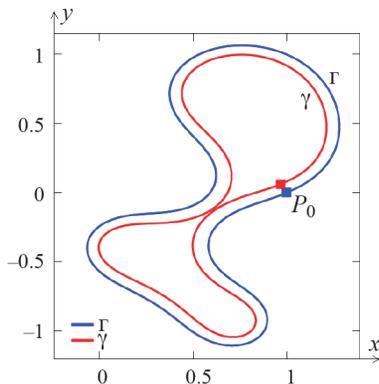
Были рассмотрены два способа задания границы Γ .

Способ 1. Этот способ предполагает задание функции (6) конечным рядом Фурье (7) с заранее фиксированным не слишком большим числом членов ряда. В этом случае возможно подобрать такие коэффициенты ряда, при которых выполнено с большой точностью условие замкнутости кривой (5). Однако таким способом затруднительно получить хорошее приближение к априори заданному, возможно весьма сложному виду кривой Γ .

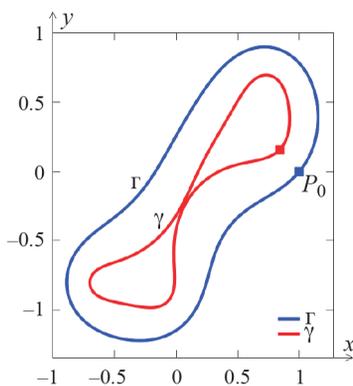
Ниже рассмотрено 2 примера для гармонической функции $x + y$. В первом примере в качестве такого ряда Фурье была взята следующая сумма. Кривая задана рядом (7): $L_1 = 0, M_1 = 0, L_2 = -0.663, M_2 = -0.793, L_3 = 0.4, M_3 = 0.6, L_4 = -0.6, M_4 = 0.2, L_5 = 0.2, M_5 = 0.2$, $r = 0.93$. Для представления ответа берется 30 членов ряда Фурье. Количество точек разбиения отрезка $[0, 2\pi]$ при дискретизации $N = 700$. Соответствующие кривые Γ и γ изображены на фиг. 1.

Во втором примере кривая задавалась рядом: $L_1 = 0, M_1 = 0, L_2 = -0.75, M_2 = -0.11, L_3 = 0, M_3 = 0, L_4 = -0.1, M_4 = 0.1, L_5 = 0.1, M_5 = 0.1$, $r = 0.75$. Для представления ответа берется 25 членов ряда Фурье. Количество точек разбиения отрезка $[0, 2\pi]$ при дискретизации $N = 2000$. Для этого случая кривые Γ и γ представлены на фиг. 2.

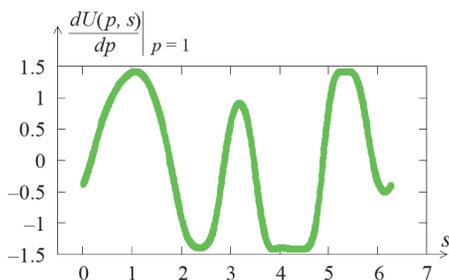
На фиг. 3, 4 изображен ответ — полученные численно производные по нормали на Γ для примеров 1 и 2. Абсолютная ошибка в норме $C[0, 2\pi]$ составляет $8 \cdot 10^{-4}$, в приведенном на фиг. масштабе численно полученный результат совпадает с известным аналитическим решением.



Фиг. 1. Пример 1 двух кривых Γ и γ в случае конечного ряда Фурье. На кривой γ точкой отмечено значение $z(\zeta)$ при $\zeta = r$, полученное интегрированием (9) от P_0 до r .



Фиг. 2. Пример 2 двух кривых Γ и γ в случае конечного ряда Фурье. На кривой γ точкой отмечено значение $z(\zeta)$ при $\zeta = r$, полученное интегрированием (9) от P_0 до r .



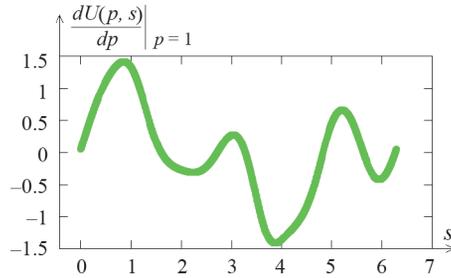
Фиг. 3. Ответ — производная по нормали ν на границе Γ из примера 1.

Более гибким в отношении задания того или иного вида кривой оказался

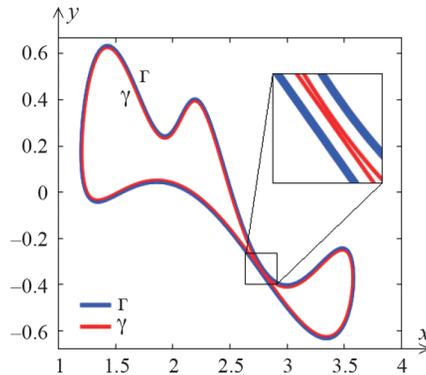
Способ 2. Этот способ основывается на параметрическом задании декартовых координат точки кривой Γ в виде $(x(t), y(t)) = \Lambda(a(t), b(t))$. Эти координаты лишь множителем

$$\Lambda = \frac{2\pi}{\int_0^{2\pi} \sqrt{\left(\frac{da}{dt}\right)^2 + \left(\frac{db}{dt}\right)^2} dt}$$

отличаются от параметризации



Фиг. 4. Ответ — производная по нормали ν на границе Γ из примера 2.



Фиг. 5. Пример 3 двух кривых Γ и γ для $r = 0.99$ в случае аналитически заданной области. Количество точек разбиения отрезка $[0, 2\pi]$ при дискретизации $N = 8000$. В разложении Γ в ряд Фурье используются первые 150 членов.

$$[0, 2\pi] \ni t \mapsto (a(t), b(t)).$$

Ниже в качестве 3-го примера взята параметризация, в которой

$$\begin{aligned} a(t) &= \cos t - 0.6 \cos 2t - 3 \sin t + 0.2 \sin 2t, \\ b(t) &= \sin t + 0.4 \sin 4t. \end{aligned} \tag{19}$$

Соответствующая кривая Γ , представленная на фиг. 5, специально была выбрана так, чтобы тестирование численной реализации оператора Дирихле–Неймана происходило в достаточно сложной области Ω , имеющей узкий перешеек или аппендикс.

Хотя второй способ достаточно гибок при задании кривой Γ , однако он связан с вычислительной сложностью представления зависимости от натурального параметра s функции $Q : s \mapsto Q(s) = N(s) - s$. Дело в том, что для этого предварительно требуется построение отображения $s \mapsto t(s)$. Оно возможно в силу формул

$$x(s(t)) = \Lambda a(t), \quad y(s(t)) = \Lambda b(t) \quad \text{и} \quad x'(s)|_{s=s(t)} \dot{s}(t) = \Lambda \dot{a}(t), \quad y'(s)|_{s=s(t)} \dot{s}(t) = \Lambda \dot{b}(t). \tag{20}$$

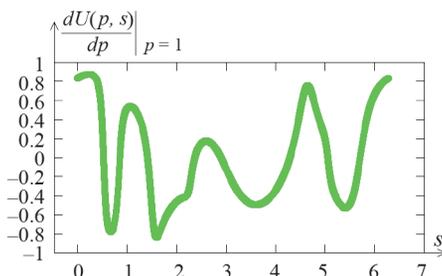
Здесь $x'(s) = -\sin N(s)$, $y'(s) = \cos N(s)$. Тем самым, $\text{tg} N(s)|_{s=s(t)} = \sigma(t)$, где $\sigma(t) = -\frac{\dot{a}(t)}{\dot{b}(t)}$. Значит,

$$Q(s) = -\text{arctg} \frac{\dot{a}(t)}{\dot{b}(t)} \Big|_{t=t(s)} - s. \tag{21}$$

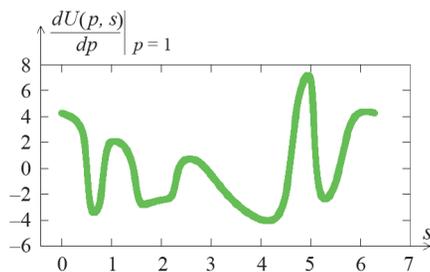
Но зависимость t от s удается получить только численно, задав зависимость $s_k \mapsto t(s_k)$, представленную при $s_k = s(t_k) = k\tau$, $\tau = 2\pi / K$ ($K \gg 1$) такими соотношениями

$$\begin{aligned}
 s_0 &= t_0 = 0, \\
 s_{k+1} - s_k &= \Lambda \int_{t_k}^{t_{k+1}} \sqrt{\left(\frac{da}{dt}\right)^2 + \left(\frac{db}{dt}\right)^2} dt = \tau, \\
 k &= 0, 1, 2, \dots, K - 1.
 \end{aligned}
 \tag{22}$$

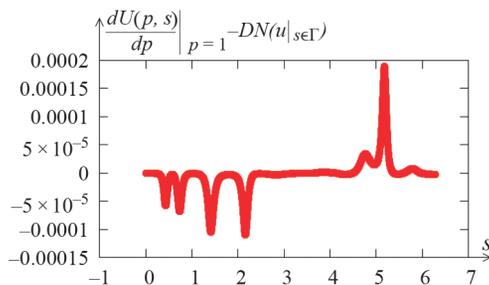
Заменяя в формуле (21) значение $t = t(s)$ на $t = t_k(s_k)$ ($k = 0 \div K$) получаем вместо функции Q ее численную аппроксимацию \tilde{Q} , разложение которой в ряд Фурье сходится медленно. Ускорить сходимость можно сгладив численную аппроксимацию \tilde{s} функции s , например, с помощью имитации регуляризации по Стеклову (свертки с δ -образной функцией). А именно, сначала зависимость $\pi(s_k) \stackrel{\text{def}}{=} t(s_k) - s_k$ для $k = 0, 1, 2, \dots, K$ продолжается периодически с периодом K по формуле: $\pi(s_k) = \pi(s_{K+k})$ для $k \in \mathbb{Z}$. Затем в качестве нового значения $\pi(s_k) = \pi_0(s_k)$ берется $\pi_1(s_k) = \frac{1}{2}(\pi_0(s_{k+1}) + \pi_0(s_{k-1}))$. Эта процедура



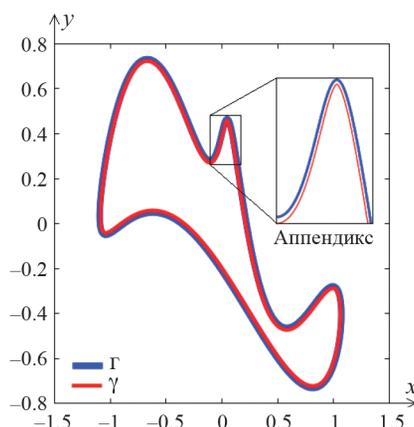
Фиг. 6. Ответ — производная по нормали ν на границе Γ для примера 3 в случае функции $x^2 - y^2$.



Фиг. 7. Ответ — производная по нормали ν на границе Γ для примера 3 в случае функции $\cosh(x) \cdot \sin(y)$.



Фиг. 8. График погрешности — разности производной по нормали на границе для примера 3 в случае функции $\cosh(x) \cdot \sin(y)$, найденной численно $\frac{\partial U(p, s)}{\partial p} \Big|_{p=1}$ и точного аналитического решения $DN(u|_{s \in \Gamma})$.



Фиг. 9. Пример 4 кривых Γ и γ с особенностью в виде аппендикса.

сглаживания (которую можно повторить несколько раз) позволяет существенно ускорить сходимость ряда Фурье для аппроксимации функции Q .

Далее по найденному фурье-разложению $\tilde{Q}(s) = \sum_{k \geq 0} (\tilde{L}_k \cos ks + \tilde{M}_k \sin ks)$ для аппроксимации функции Q (ср. с (6)) строится (в соответствии с леммой 1) отображение (9).

Ниже для второго способа задания кривой Γ представлены численные результаты для оператора Дирихле–Неймана. Результаты для Дирихле–Робена очевидным образом вытекают из результатов для Дирихле–Неймана по (13), (14).

На фиг. 6 и 7 представлены графики решения — найденные численно производные по нормали к Γ для гармонических функций $x^2 - y^2$ и $\cosh(x) \cdot \sin(y)$ соответственно.

На фиг. 8 приведен график погрешности найденного решения для функции $\cosh(x) \cdot \sin(y)$. Норма ошибки в $C[0, 2\pi]$ составила $2 \cdot 10^{-4}$.

Если в формуле (19) коэффициент в $a(t)$ при $\sin(x)$ взять равным -2.3 , то получится кривая, изображенная на фиг. 9.

Она не имеет узкого перешейка как кривая в примере 3 выше, но зато у нее есть достаточно крутой горб-аппендикс. Для данной кривой также было проведено численное моделирование, описанная выше методика работает, получились схожие результаты. Количество точек разбиения отрезка $[0, 2\pi]$ при дискретизации бралось равным $N = 8000$ для $r = 0.992$. В разложении Γ в ряд Фурье использовалось первые 550 членов. Для гармонических функций $x^2 - y^2$ и $\cosh(x) \cdot \sin(y)$ относительная погрешность в C составила 0.001 и достигалась в малой окрестности аппендикса.

Отметим также, что в случае более простых границ областей Γ результат получается намного точнее. Например, для симметричной кривой, рассмотренной в [16] абсолютная погрешность, полученная в результате численных расчетов описанным выше методом получается порядка 10^{-6} .

5. ЗАКЛЮЧЕНИЕ

Опираясь на лемму, дающую явную численно реализуемую формулу для изометричного однолистного отображения аналитической границы односвязной двумерной области на единичную окружность, представлены численные алгоритмы для операторов Пуанкаре–Стеклова, относящихся к уравнению Лапласа в двумерной односвязной области с C^1 -границей. Тестовые численные результаты для операторов Дирихле–Неймана и Дирихле–Робена показали точность приближения с относительной погрешностью порядка 0.1 процента даже для сложных областей, имеющих узкий перешеек или аппендикс.

СПИСОК ЛИТЕРАТУРЫ

1. Schwarz H. A. Uber einige Abbildungsaufgaben // Ges. Math. 1869. Abh. II. P. 65.
2. Агошков В. И. Новая методика формулировки алгоритмов разделения области // Ж. вычисл. матем. и матем. физ. 2020. Т. 60. № 3. С. 351.

3. *Гринберг Г. А.* Избранные вопросы математической теории электрических и магнитных явлений. М.: Изд. АН СССР. 1948.
4. *Маергойз И. Д.* О численном решении краевых задач теории потенциала методом интегральных уравнений // Сиб. матем. журн. 1971. Т. 12. № 6. С. 1318.
5. *Khoromskij B. N., Wittum G.* Elliptic Poincaré–Steklov Operators. Springer. Berlin. Heidelberg. Lecture Notes Comput. Science and Engng. 2004. V. 36. P. 63–81.
6. *Новиков Р. Г., Тайманов И. А.* Преобразования Дарбу–Мутара и операторы Пуанкаре–Стеклова // Труды МИАН. 2018. Т. 302. P. 334.
7. *Лебедев В. И., Агошков В. И.* Операторы Пуанкаре–Стеклова и их приложения в анализе. М.: Отдел вычислительной математики АН СССР. 1983. 184 с.
8. *Poincaré H.* La méthode de Neumann et le problème de Dirichlet // Acta Math. V. XX. 59.
9. *Stekloff W.* Les méthodes générales pour résoudre les problèmes fondamentaux de la physique mathématique // Ann. fac. sci. Toulouse. Sér. 1900. V. 2. N 2. P. 207.
10. *Арнольд В. И.* Геометрические методы в теории обыкновенных дифференциальных уравнений. М.: МЦНМО. 2012.
11. *Демидов А. С.* Полная асимптотика решения задачи Дирихле для 2-мерного уравнения Лапласа с быстро осциллирующими граничными данными // Докл. АН. 1996. Т. 346. № 6. С. 732.
12. *Демидов А. С.* Функционально-геометрический метод решения задач со свободной границей для гармонических функций // Успехи матем. наук. 2010. Т. 65. № 1. С. 3.
13. *Демидов А. С.* О численно реализуемых явных формулах для решений двумерных и трехмерных уравнений $\operatorname{div}(\alpha(w)\operatorname{div}w) = 0, \operatorname{div}(\beta\nabla w) = 0$ с данными Коши на аналитической границе // Функц. анализ и его прил. 2021. Т. 55. № 1. С. 65.
14. *Крылов Н. М., Боголюбов Н. Н.* Приближенное решение задачи Дирихле // Докл. АН СССР. 1929. № 12. 283.
15. *Канторович Л. В., Крылов В. И.* Приближенные методы высшего анализа. М.: Физмат. гиз., 1962.
16. *Власов В. К., Бакушинский А. Б.* Метод потенциалов и численное решение задачи Дирихле для уравнения Лапласа // Ж. вычисл. матем. и матем. физ. 1963. Т. 3. № 3. С. 574.
17. *Zhou Y., Cai W.* Numerical Solution of the Robin Problem of Laplace Equations with a Feynman–Kac Formula and Reflecting Brownian Motions // J. Scientific Comput. 2016 T. 69. № 1. 107. <https://doi.org/10.1007/s10915-016-0184-y>

УДК 519.63

МЕТОД ЧИСЛЕННОГО РЕШЕНИЯ НЕСТАЦИОНАРНОГО УРАВНЕНИЯ ШРЁДИНГЕРА ДЕСЯТОГО ПОРЯДКА ТОЧНОСТИ¹⁾

© 2024 г. М. А. Захаров^{1,*}¹141980 Дубна, М.о., ул. Жолио-Кюри, 6, Объединенный институт ядерных исследований, Россия

*e-mail: zakharovmax@jinr.ru

Поступила в редакцию 06.07.2023 г.

Переработанный вариант 01.09.2023 г.

Принята к публикации 19.10.2023 г.

Представлен метод численного решения нестационарного уравнения Шрёдингера десятого порядка точности, основанный на аппроксимации оператора эволюции формулой произведения. Обсуждается проблема уменьшения числа операторных экспонент в итоговой формуле за счет оптимизации их последовательности. На основе идеи, предложенной Йошида, построены два алгоритма десятого порядка точности для аппроксимации оператора эволюции. Численные тесты продемонстрировали устойчивость этих алгоритмов и их порядок точности. Метод, использованный в статье, позволил значительно уменьшить количество экспоненциальных множителей в схеме по сравнению с известной формулой Ли–Троттера–Сузуки. Библ. 25. Фиг. 2. Табл. 2.

Ключевые слова: квантовая механика, нестационарное уравнение Шрёдингера, численные методы, аппроксимация высокого порядка точности.

DOI: 10.31857/S0044466924020079, EDN: YJYRQP

1. ВВЕДЕНИЕ

Исследование методов численного решения нестационарного уравнения Шрёдингера имеет продолжительную историю и множество предложенных вычислительных схем. Одним из наиболее часто используемых методов, помимо широкого спектра разностных схем (см. [1]–[4]), является аппроксимация оператора эволюции. Достаточно давно известно (см. [5], [6]), что решение дифференциального уравнения

$$i \frac{\partial \Psi}{\partial t} = A(t) \Psi \quad (1)$$

можно представить в экспоненциальном виде

$$\Psi(t) = \exp \left\{ -i \int_{t_0}^t A(\tau) d\tau \right\} \Psi(t_0). \quad (2)$$

Экспонента в выражении (2) обычно представляет собой довольно сложный дифференциальный оператор (оператор эволюции), аналитический расчет действия которого на волновую функцию является нетривиальной задачей. В связи с этим значительный интерес представляет построение приближенного численного выражения для действия такого оператора.

Существует несколько путей аппроксимации операторной экспоненты в (2). Один из двух наиболее часто встречающихся основан на представлении экспоненты в виде разложения Магнуса (см. [5], [7], [8]). Второй путь связан с представлением экспоненты в виде произведения более простых экспонент. В статье речь пойдет именно о втором направлении, оно носит название формулы произведения Ли–Троттера–Сузуки (см. [9], [10]). За последние 30 лет вышло большое количество работ, посвященных изучению области применимости этого метода, оптимизации, улучшению точности, поискам альтернативных подходов (см. [11]–[15]).

Основными критериями при оценке вычислительной эффективности являются устойчивость и порядок точности метода. Однако сегодня появляются задачи (см., например, [16], [17]), которые могут требовать не только высокой точности по временному шагу, но и высокого пространственного разрешения

¹⁾ Работа выполнена при финансовой поддержке РФФИ (грант № 20-11-20257).

при достаточно большой исследуемой области пространства. Такими требованиями, например, обладает любая задача, в которой волновой пакет с узким энергетическим спектром взаимодействует с тонкой потенциальной структурой, движущейся достаточно быстро.

В формуле произведения Ли–Троттера–Сузуки основной объем вычислений затрачивается на дискретное преобразование Фурье. Сегодня существуют хорошо оптимизированные алгоритмы и библиотеки для расчета дискретного преобразования Фурье (например, библиотека FFTW3, см. [18]), в том числе для распределенных вычислений (технология CUDA Nvidia¹). Тем не менее, рост необходимого для расчета времени происходит нелинейно с ростом числа точек в массиве волновой функции, и использование больших сеток по координате и времени сопряжено с большим временем расчета.

В связи с этим становятся актуальными численные схемы решения нестационарного уравнения Шрёдингера высокого порядка точности, в которых используется наименьшее количество экспонент-множителей. Настоящая статья посвящена построению схемы 10-го порядка точности с наименее возможным количеством множителей в формуле произведения.

Статья состоит из трех частей. В разд. 2 приведен общий обзор и краткий вывод формулы Ли–Троттера–Сузуки. Раздел 3 посвящен подробному построению схем аппроксимации 10-го порядка точности оператора эволюции с использованием аппроксимаций более низкого порядка, а именно, аппроксимаций 2-го и 4-го порядков точности. Такие численные схемы требуют гораздо меньшего числа операторных экспонент, чем в формуле Ли–Троттера–Сузуки. Раздел 4 включает численные тесты и обсуждение построенных схем.

2. ОПЕРАТОР ЭВОЛЮЦИИ

Решение нестационарного уравнения Шрёдингера

$$i\hbar \frac{\partial \Psi(x,t)}{\partial t} = H(x,t) \Psi(x,t) \quad (3)$$

можно представить в виде результата действия унитарного оператора эволюции на начальное волновое состояние:

$$\Psi(x,t) = \hat{S}(x,t,t_0) \Psi(x,t_0). \quad (4)$$

Далее будет использоваться безразмерная система единиц, где $\hbar = 1$. В случае, если гамильтониан в (3) не зависит явно от времени, оператор эволюции запишется в виде

$$\hat{S}(x,t,t_0) = \exp\{-i(t-t_0)H(x)\}. \quad (5)$$

В случае гамильтониана, зависящего от времени, оператор \hat{S} может быть выражен через упорядоченную операторную экспоненту:

$$\hat{S}(x,t,t_0) = T \left\{ \exp \left[-i \int_{t_0}^t H(x,t') dt' \right] \right\} \quad (6)$$

или в более интуитивно понятной записи

$$\begin{aligned} \hat{S}(x,t,t_0) &= \lim_{n \rightarrow \infty} T \left\{ e^{-i \frac{\Delta t}{n} \sum_{j=1}^n H\left(t_0 + j \frac{\Delta t}{n}\right)} \right\} = \\ &= \lim_{n \rightarrow \infty} e^{-i \frac{\Delta t}{n} H(t_0 + \Delta t)} e^{-i \frac{\Delta t}{n} H\left(t_0 + \frac{n-1}{n} \Delta t\right)} \dots e^{-i \frac{\Delta t}{n} H\left(t_0 + \frac{2}{n} \Delta t\right)} e^{-i \frac{\Delta t}{n} H\left(t_0 + \frac{1}{n} \Delta t\right)}. \end{aligned} \quad (6.1)$$

Здесь $\Delta t = t - t_0$.

Рассмотрим оператор сдвига по времени, действующий *влево* (см. [19]):

¹ <https://docs.nvidia.com/cuda/cufft>

$$\tilde{T}^t(\tau) = \exp(\tau \tilde{D}_t) = \exp\left\{\tau \frac{\tilde{\partial}}{\partial t}\right\}, \quad (7)$$

который определяется свойством

$$A(t) \tilde{T}^t(\tau) B(t) = \tilde{T}^t(\tau) A(t + \tau) B(t). \quad (7.1)$$

Используя формулу Троттера (см. [19], [20]) можно показать, что

$$\begin{aligned} \hat{S}(x, t, t_0) &= \lim_{n \rightarrow \infty} T \left\{ \exp \left[-i \frac{\Delta t}{n} \sum_{j=1}^n H \left(x, t_0 + j \frac{\Delta t}{n} \right) \right] \right\} = \\ &= \exp \left\{ -i \Delta t [H(x, t) + \tilde{D}_t] \right\}. \end{aligned} \quad (8)$$

Доказательство утверждения (8) следующее:

$$\begin{aligned} e^{-i \Delta t [H(x, t) + \tilde{D}_t]} &= \lim_{n \rightarrow \infty} \left(e^{-i \frac{\Delta t}{n} H(x, t)} e^{-i \frac{\Delta t}{n} \tilde{D}_t} \right)^n = \\ &= \lim_{n \rightarrow \infty} \overbrace{e^{-i \frac{\Delta t}{n} H(x, t)} e^{-i \frac{\Delta t}{n} \tilde{D}_t} \dots e^{-i \frac{\Delta t}{n} H(x, t)} e^{-i \frac{\Delta t}{n} \tilde{D}_t}}^n = \\ &= \lim_{n \rightarrow \infty} e^{-i \frac{\Delta t}{n} H(x, t + \Delta)} e^{-i \frac{\Delta t}{n} H(x, t + \frac{n-1}{n} \Delta t)} \dots e^{-i \frac{\Delta t}{n} H(x, t + 2 \frac{\Delta t}{n})} e^{-i \frac{\Delta t}{n} H(x, t + \frac{\Delta t}{n})} = \hat{S}(x, t, t_0) \end{aligned} \quad (8.1)$$

Гамильтониан можно представить в виде суммы кинетической и потенциальной частей, где от времени явно зависит лишь потенциальная часть:

$$H(x, t) = \hat{T}(x) + \hat{U}(x, t). \quad (9)$$

Поскольку кинетическая часть и оператор сдвига по времени (7) коммутируют, $[\hat{T}(x), \tilde{D}_t] = 0$, эти операторы можно объединить в один:

$$\exp\{\tilde{T}(x, t)\} = \exp\{\hat{T}(x) + \tilde{D}_t\} = \exp\{\hat{T}(x)\} \exp\{\tilde{D}_t\} = \exp\{\tilde{D}_t\} \exp\{\hat{T}(x)\}. \quad (10)$$

Тогда оператор эволюции S в итоге будет представлен в виде

$$\hat{S}(x, t, t_0) = \exp\left\{-i \Delta t [\tilde{T}(x, t) + \hat{U}(x, t)]\right\}. \quad (11)$$

Таким образом, проблема аппроксимации оператора эволюции (4) сводится к проблеме аппроксимации оператора вида $\exp\{\Delta t(A + B)\}$.

2.1. Аппроксимация оператора эволюции. Формула произведения

Рассмотрим аппроксимацию экспоненциального оператора формулой произведения (в матричном виде известной со времен Ли, однако, к численному решению нестационарного уравнения Шрёдингера впервые примененная, по-видимому, в работе [21]):

$$\exp\{\tau(A + B)\} = \prod_{j=1}^k \exp(a_j \tau A) \exp(b_j \tau B) + o(\tau^n). \quad (12)$$

Здесь k — необходимое количество пар экспоненциальных операторов A, B для аппроксимации оператора $(A + B)$ с точностью $o(\tau^n)$, a_i, b_i — вещественные коэффициенты, которые предстоит определить. Кратко изложим аналитический метод определения коэффициентов a_i, b_i , следуя работам [9], [10].

Прежде всего приведем необходимую для построения решения формулу Бэйкера–Кэмбелла–Хаслдорфа (БКХ):

$$\exp\{X\}\exp\{Y\} = \exp\left\{ \begin{array}{l} X + Y + \frac{1}{2}[X, Y] + \\ + \frac{1}{12}([X, [X, Y]] + [Y, [Y, X]]) + \dots \end{array} \right\}. \quad (13)$$

Используя формулу БКХ дважды, можно получить аналогичное выражение для симметричного оператора:

$$\exp\{X\}\exp\{Y\}\exp\{X\} = \exp\left\{2X + Y - \frac{1}{6}([X, [X, Y]] - [Y, [Y, X]]) + \dots\right\}. \quad (14)$$

Из формул (13), (14) можно получить выражения для формулы (12) 2-го и 3-го порядков соответственно, отбрасывая все слагаемые с коммутаторами.

Аппроксимация формулы (12) второго порядка

$$\exp\{\tau(A + B)\} = \exp\{\tau A\}\exp\{\tau B\} + o(\tau^2). \quad (15)$$

Аппроксимация формулы (12) третьего порядка

$$\exp\{\tau(A + B)\} = \exp\left\{\frac{1}{2}\tau A\right\}\exp\{\tau B\}\exp\left\{\frac{1}{2}\tau A\right\} + o(\tau^3). \quad (16)$$

При определении точности выражений (15), (16) учитывалось следующее тождество:

$$\exp\{\tau(A + B) + o(\tau^n)\} = \exp\{\tau(A + B)\} + o(\tau^n). \quad (17)$$

Для определения выражения формулы (12) 4-го порядка обозначим выражение (16) следующим образом:

$$S_2(\tau) \equiv \exp\left\{\frac{\tau}{2}A\right\}\exp\{\tau B\}\exp\left\{\frac{\tau}{2}A\right\} = \exp\{\tau\Theta_1 + \tau^3\Theta_3 + \tau^5\Theta_5 + \dots\} \equiv \exp\{\Omega(\tau)\}. \quad (18)$$

Здесь под функциями Θ_n обозначены коммутаторные выражения из (14):

$$\begin{aligned} \Theta_1 &= A + B; \\ \Theta_3 &= \frac{1}{12}[B, [B, A]] - \frac{1}{24}[A, [A, B]]; \\ \Theta_5 &= \frac{7}{5760}[A, [A, [A, [A, B]]]] + \dots; \dots \end{aligned}$$

Рассмотрим следующую операторную конструкцию:

$$S_4 \equiv S_2(\alpha_1\tau)S_2(\alpha_0\tau)S_2(\alpha_1\tau) = \exp\{\Omega(\alpha_1\tau)\}\exp\{\Omega(\alpha_0\tau)\}\exp\{\Omega(\alpha_1\tau)\}, \quad (19)$$

где α_i — вещественные коэффициенты. Применим к формуле (19) формулу (14), оставляя в степени экспоненты только слагаемые меньше 5-го порядка:

$$S_4 = \exp\left\{ \begin{array}{l} \tau(\alpha_0 + 2\alpha_1)\Theta_1 + \\ + \tau^3(\alpha_0^3 + 2\alpha_1^3)\Theta_3 + o(\tau^5) \end{array} \right\}. \quad (20)$$

Выражение (20) будет аппроксимацией оператора (12) 4-го порядка в том случае, когда коэффициенты α_i удовлетворяют следующим условиям:

$$\begin{cases} \alpha_0 + 2\alpha_1 = 1 \\ \alpha_0^3 + 2\alpha_1^3 = 0 \end{cases} \rightarrow \alpha_0 = -\frac{2^{1/3}}{2 - 2^{1/3}}; \alpha_1 = \frac{1}{2 - 2^{1/3}}. \quad (21)$$

Подставив результаты (21) в (19), с учетом (18) получим в итоге

$$S_4 = e^{\frac{1}{2}\alpha_1\tau A} e^{\alpha_1\tau B} e^{\frac{1}{2}(\alpha_1+\alpha_0)\tau A} e^{\alpha_0\tau B} e^{\frac{1}{2}(\alpha_1+\alpha_0)\tau A} e^{\alpha_1\tau B} e^{\frac{1}{2}\alpha_1\tau A} = \prod_{i=1}^4 e^{a_{4,i}\tau A} e^{b_{4,i}\tau B} + o(\tau^5). \quad (22)$$

Аналогичным образом строятся схемы порядков 6, 8, 10 соответственно:

$$S_6(\tau) \equiv S_4(\beta_1\tau)S_4(\beta_0\tau)S_4(\beta_1\tau) = \prod_{i=1}^{10} e^{a_i\tau A} e^{b_i\tau B} + o(\tau^7), \quad (23)$$

$$S_8(\tau) \equiv S_6(\gamma_1\tau)S_6(\gamma_0\tau)S_6(\gamma_1\tau) = \prod_{i=1}^{28} e^{a_i\tau A} e^{b_i\tau B} + o(\tau^9), \quad (24)$$

$$S_{10}(\tau) \equiv S_8(\delta_1\tau)S_8(\delta_0\tau)S_8(\delta_1\tau) = \prod_{i=1}^{82} e^{a_i\tau A} e^{b_i\tau B} + o(\tau^{11}). \quad (25)$$

Важным преимуществом формул (22)–(25) является тот факт, что для любого порядка точности коэффициенты a_i, b_i можно найти аналитически. Вместе с тем за подобную простоту и удобство приходится платить большим количеством используемых экспоненциальных операторов. В схеме 8-го порядка количество конструкций S_2 равно 28, в схеме 10-го порядка — уже 82.

3. СХЕМЫ АППРОКСИМАЦИИ ОПЕРАТОРА ЭВОЛЮЦИИ С МЕНЬШИМ КОЛИЧЕСТВОМ ЭКСПОНЕНТ

Количество используемых конструкций S_2 можно значительно сократить, если воспользоваться следующей симметричной операторной конструкцией:

$$S^{(m)}(\tau) = S_k(n_m\tau)S_k(n_{m-1}\tau)\dots S_k(n_1\tau)S_k(n_0\tau)S_k(n_1\tau)\dots S_k(n_{m-1}\tau)S_k(n_m\tau). \quad (26)$$

Метод аппроксимации оператора эволюции формулой (26) был предложен в работе [9], где представлены результаты для схем 6-го и 8-го порядков для конструкций S_2 . В формуле (26) возможно также использование конструкций S_4, S_6, S_8 и т.д. Данный раздел посвящен расчету коэффициентов n_i для схем 10-го порядка по формуле (26) для конструкций S_2 и S_4 .

Идея метода состоит в применении формулы БКХ к выражению (26), приведению его к виду одной экспоненты и определению условий, при которых коэффициенты при всех выражениях с коммутаторами становятся равными нулю.

3.1. Схема аппроксимации 10-го порядка с использованием конструкций S_2

Рассмотрим вначале решение с операторами S_2 , поскольку оно является наиболее общим для данного подхода. Для этого представим операторы $S_2(n_i\tau)$ в виде экспонент, аналогично формуле (18):

$$S_2(n_0\tau) = \exp\{\Omega_0\}; S_2(n_1\tau) = \exp\{\Omega_1\}; \dots; S_2(n_m\tau) = \exp\{\Omega_m\}. \quad (27)$$

Строим симметричную операторную конструкцию, последовательно «приставляя» слева и справа очередной оператор $S_2(n_i\tau)$. Полученную конструкцию с помощью формулы БКХ вновь представляем в виде одной экспоненты:

$$S_2(n_1\tau)S_2(n_0\tau)S_2(n_1\tau) = \exp\{\Omega_1\}\exp\{\Omega_0\}\exp\{\Omega_1\} = \exp\{\Omega^{(1)}\}, \quad (28.1)$$

$$S_2(n_2\tau)S_2(n_1\tau)S_2(n_0\tau)S_2(n_1\tau)S_2(n_2\tau) = \exp\{\Omega_2\}\exp\{\Omega^{(1)}\}\exp\{\Omega_2\} = \exp\{\Omega^{(2)}\}, \quad (28.2)$$

, ...,

$$S_2(n_m\tau)\dots S_2(n_0\tau)\dots S_2(n_m\tau) = \exp\{\Omega_m\}\exp\{\Omega^{(m-1)}\}\exp\{\Omega_m\} = \exp\{\Omega^{(m)}\}. \quad (28.3)$$

Итоговую экспоненту будем искать в виде

$$\Omega^{(m)} = \tau F_1^{(m)}(\Theta_1, n_0, n_1, \dots, n_m) + \tau^3 F_3^{(m)}(\Theta_3, n_0, n_1, \dots, n_m) + \tau^5 F_5^{(m)}(\Theta_1, \Theta_3, \Theta_5, n_0, n_1, \dots, n_m) + \tau^7 F_7^{(m)}(\Theta_1, \Theta_3, \Theta_5, \Theta_7, n_0, n_1, \dots, n_m) + \tau^9 F_9^{(m)}(\Theta_1, \Theta_3, \Theta_5, \Theta_7, \Theta_9, n_0, n_1, \dots, n_m), \tag{29}$$

где Θ_i — выражения, содержащие коммутаторы, определяемые из формулы БКХ.

Таким образом, задача сводится к определению всех функций $F_i^{(m)}$, которые будем искать рекуррентным образом. Чтобы рассчитать схему 10-го порядка, необходимо знать слагаемые формулы БКХ до 7-го порядка. В работе будут использоваться результаты расчета слагаемых формулы БКХ в базе Холла–Веннота (см. [23]). Все слагаемые формулы БКХ вплоть до 7-го порядка в базе Холла–Веннота приведены в приложении А.

Введем следующие обозначения для коммутаторных конструкций:

$$\begin{aligned} \Xi_a &= a, \\ \Xi_{abc} &= [[a, b], c], \\ \Xi_{abcde} &= [[[[a, b], c], d], e], \\ \Xi_{abc, de} &= [[[a, b], c], [d, e]], \\ \Xi_{abs, de, fg} &= [[[a, b], c], [d, e], [f, g]], \\ \Xi_{abcd, efg} &= [[[[a, b], c], d], [e, f], g]. \end{aligned} \tag{30}$$

В итоге формула БКХ до 7-го порядка точности запишется в виде

$$\begin{aligned} Z(X, Y)_{e^x e^y e^x} &= 2\Xi_X + \Xi_Y + N_{YXX} \Xi_{YXX} + N_{YXY} \Xi_{YXY} + N_{YXXXX} \Xi_{YXXXX} + \\ &+ N_{YXXXY} \Xi_{YXXXY} + N_{YXXYY} \Xi_{YXXYY} + N_{YXYYY} \Xi_{YXYYY} + N_{YXX, YX} \Xi_{YXX, YX} + \\ &+ N_{YXY, YX} \Xi_{YXY, YX} + N_{YXXXXXX} \Xi_{YXXXXXX} + N_{YXXXXXY} \Xi_{YXXXXXY} + \\ &+ N_{YXXXXYY} \Xi_{YXXXXYY} + N_{YXXYYYY} \Xi_{YXXYYYY} + N_{YXXYY, YX} \Xi_{YXXYY, YX} + \\ &+ N_{YXYYY, YX} \Xi_{YXYYY, YX} + N_{YXXXX, YX} \Xi_{YXXXX, YX} + N_{YXXY, YX} \Xi_{YXXY, YX} + \\ &+ N_{YXXY, YX} \Xi_{YXXY, YX} + N_{YXY, YX, YX} \Xi_{YXY, YX, YX} + N_{YXXX, YXX} \Xi_{YXXX, YXX} + N_{YXXY, YXX} \Xi_{YXXY, YXX} + \\ &+ N_{YXY, YX, YX} \Xi_{YXY, YX, YX} + N_{YXXX, YXY} \Xi_{YXXX, YXY} + N_{YXXY, YXY} \Xi_{YXXY, YXY} + \\ &+ N_{YXY, YXY} \Xi_{YXY, YXY}. \end{aligned} \tag{31}$$

Символами N_i обозначены коэффициенты при коммутаторах. Количественное значение коэффициентов N_i также приведено в приложении А.

Чтобы определить рекуррентные формулы для функций $F_i^{(m)}$, достаточно посчитать первые два шага $\Omega^{(1)}$ и $\Omega^{(2)}$. Подставляя выражения $\Omega_0(\Theta_1, \Theta_3, \Theta_5, \Theta_7)$ и $\Omega_1(\Theta_1, \Theta_3, \Theta_5, \Theta_7)$ в формулу (31), можно получить выражение для $\Omega^{(1)}$:

$$\begin{aligned} \Omega^{(1)} &= \tau A_{1,1} \Xi_{\Theta_1} + \tau^3 A_{3,1} \Xi_{\Theta_3} + \tau^5 (A_{5,1} \Xi_{\Theta_5} + B_{5,1} \Xi_{\Theta_3, \Theta_1}) + \\ &+ \tau^7 (A_{7,1} \Xi_{\Theta_7} + B_{7,1} \Xi_{\Theta_5, \Theta_1} + B_{7,1}^3 \Xi_{\Theta_3, \Theta_1, \Theta_3} + C_{7,1} \Xi_{\Theta_3, \Theta_1, \Theta_1, \Theta_1}) + \\ &+ \tau^9 (A_{9,1} \Xi_{\Theta_9} + B_{9,1} \Xi_{\Theta_7, \Theta_1} + B_{9,1}^1 \Xi_{\Theta_5, \Theta_3, \Theta_1} + B_{9,1}^3 \Xi_{\Theta_3, \Theta_1, \Theta_3} + B_{9,1}^5 \Xi_{\Theta_3, \Theta_1, \Theta_3} + C_{9,1} \Xi_{\Theta_5, \Theta_1, \Theta_1, \Theta_1} + \\ &+ C_{9,1}^{311} \Xi_{\Theta_3, \Theta_1, \Theta_3, \Theta_1} + C_{9,1}^{131} \Xi_{\Theta_3, \Theta_1, \Theta_3, \Theta_1} + C_{9,1}^{113} \Xi_{\Theta_3, \Theta_1, \Theta_1, \Theta_3} + E_{9,1} \Xi_{\Theta_3, \Theta_1, \Theta_3, \Theta_1} + D_{9,1} \Xi_{\Theta_3, \Theta_1, \Theta_1, \Theta_1, \Theta_1}). \end{aligned} \tag{32}$$

Здесь функции A, B, C, D, E содержат неизвестные n_0, n_1 . Подставляя так же выражения $\Omega^{(1)}$ и Ω_2 в формулу (31), можно получить выражения для $\Omega^{(2)}$ в форме, аналогичной (32), и определить, какие новые слагаемые по сравнению с $\Omega^{(1)}$ появляются. Используя результаты первых двух итераций, можно

записать рекуррентные формулы для $F_i^{(m)}$, поскольку следующие итерации уже не принесут никаких новых слагаемых.

Окончательная формула для $\Omega^{(m)}$ может быть представлена в виде

$$\begin{aligned} \Omega^{(m)} = & \tau A_{1,m} \Xi_{\Theta_1} + \tau^3 A_{3,m} \Xi_{\Theta_3} + \tau^5 (A_{5,m} \Xi_{\Theta_5} + B_{5,m} \Xi_{\Theta_3 \Theta_1 \Theta_1}) + \\ & + \tau^7 (A_{7,m} \Xi_{\Theta_7} + B_{7,m} \Xi_{\Theta_5 \Theta_1 \Theta_1} + B_{7,m}^3 \Xi_{\Theta_3 \Theta_1 \Theta_3} + C_{7,m} \Xi_{\Theta_3 \Theta_1 \Theta_1 \Theta_1}) + \\ & + \tau^9 (A_{9,m} \Xi_{\Theta_9} + B_{9,m} \Xi_{\Theta_7 \Theta_1 \Theta_1} + B_{9,m}^1 \Xi_{\Theta_5 \Theta_3 \Theta_1} + B_{9,m}^3 \Xi_{\Theta_3 \Theta_1 \Theta_3} + B_{9,m}^5 \Xi_{\Theta_3 \Theta_1 \Theta_5} + C_{9,m} \Xi_{\Theta_5 \Theta_1 \Theta_1 \Theta_1} + \\ & + C_{9,m}^{311} \Xi_{\Theta_3 \Theta_1 \Theta_3 \Theta_1 \Theta_1} + C_{9,m}^{131} \Xi_{\Theta_3 \Theta_1 \Theta_1 \Theta_3 \Theta_1} + C_{9,m}^{113} \Xi_{\Theta_3 \Theta_1 \Theta_1 \Theta_1 \Theta_3} + E_{9,m} \Xi_{\Theta_3 \Theta_1 \Theta_1 \Theta_3 \Theta_1} + D_{9,m} \Xi_{\Theta_3 \Theta_1 \Theta_1 \Theta_1 \Theta_1 \Theta_1}). \end{aligned} \tag{33}$$

Здесь функции A, B, C, D, E содержат неизвестные n_0, \dots, n_{18} . Полные математические выкладки и точный вид коэффициентов функции (33) приведены в приложении В.

Условием аппроксимации оператора эволюции (12) является равенство нулю всех функций A, B, C, D, E (кроме $A_{1,m} = 1$):

$$\begin{aligned} A_{1,m} - 1 = A_{3,m} = A_{5,m} = A_{7,m} = A_{9,m} = 0, \\ B_{5,m} = B_{7,m} = B_{7,m}^3 = B_{9,m} = B_{9,m}^1 = B_{9,m}^3 = B_{9,m}^5 = 0, \\ C_{7,m} = C_{9,m} = C_{9,m}^{311} = C_{9,m}^{131} = C_{9,m}^{113} = E_{9,m} = D_{9,m} = 0. \end{aligned} \tag{34}$$

Таблица 1. Некоторые результаты численного решения системы (34)

№	1	2	3
Значение функции невязки	0.00062	0.00061	0.00075
n_1	0.25596748726158713	0.15258923803377555	-0.13729497885984
n_2	-0.007803892018322212	0.008013207919534027	0.00165029350899004
n_3	0.06078650054710946	0.04034373194522631	0.068057054219662
n_4	-0.2106213291618254	-0.1619185810304815	0.0706939494581959
n_5	0.013374554087768997	0.11339287384911442	0.0419912057755932
n_6	0.040828812743666815	-0.06008855109018835	0.091334638627057
n_7	0.11754081338889102	0.0927048339212293	-0.000255226291493715
n_8	0.06355483224567013	0.12934029854283247	-0.1404231284529
n_9	0.06070957480623204	0.057849977072719553	0.06448480160075
n_{10}	-0.18375967470208748	-0.17035112690414977	0.117633619535249
n_{11}	0.09163708826251835	0.10813778785242215	-0.162841812844147
n_{12}	0.028120797452049817	0.06970196439055515	0.062471615933172
n_{13}	0.09425234692400111	0.01830413288019983	-0.199629079192461
n_{14}	-0.03201600580713401	0.0849105484960073	0.152547998021664
n_{15}	-0.044068114910998546	-0.003814424982338	0.203136087501066
n_{16}	0.12566396849270714	0.087081938943613	0.059142843468599
n_{17}	-0.0257093171535282	-0.0522776167118995	0.064429738489327
n_{18}	0.08793327900898343	0.0154549048217179	-0.015307036307547
n_{19}	0.09159202216350396	0.046919487412522	0.089529926379149

Таким образом, мы получаем систему из 19 нелинейных алгебраических уравнений, которую необходимо решить численно. Система (34) чрезвычайно громоздка и хаотична, с большим количеством локальных минимумов, поэтому большинство известных методов решения систем нелинейных алгебраических уравнений (см. [24]) здесь не сходятся.

Для нахождения приближенных решений системы в Mathematica был реализован алгоритм решения системы методом минимизации невязки. Для этого методом Монте-Карло разыгрывалось приближенное решение. Затем каждый неизвестный параметр в решении варьировался, пока невязка не окажется минимальной. В качестве значения невязки принималась сумма значений функций A, B, C, D, E (34) для данного набора параметров n_i . Некоторые найденные решения системы (34) приведены в табл. 1.

Итоговая формула для аппроксимации оператора эволюции (12) будет выглядеть следующим образом:

$$S_{S_2}(\tau) = S_2(n_{18}\tau) \dots S_2(n_0\tau) \dots S_k(n_{18}\tau) = \prod_{i=1}^{37} e^{\bar{a}_{10,i}\tau A} e^{\bar{b}_{10,i}\tau B} + o(\tau^{11}). \tag{35}$$

3.2. Схема аппроксимации 10-го порядка с использованием конструкций S_4

Гораздо более простое решение представляет построение формулы (26) через конструкции S_4 , оператор которого можно представить в виде

$$S_4(\tau) \equiv \exp\{\tau\Theta_1 + \tau^5\tilde{\Theta}_5 + \tau^7\tilde{\Theta}_7 + \tau^9\tilde{\Theta}_9 + \dots\}. \tag{36}$$

Здесь $\tilde{\Theta}_i$ по аналогии с Θ_i представляют собой некоторые коммутаторные выражения. Интересно отметить, что $\tilde{\Theta}_i$ отличаются от Θ_i , но их точный вид в данной работе значения не имеет.

По аналогии с предыдущим пунктом строится функция

$$\begin{aligned} \Omega^{(m)} = & \tau F_1^{(m)}(\Theta_1, n_0, n_1, \dots, n_m) + \tau^5 F_5^{(m)}(\tilde{\Theta}_5, n_0, n_1, \dots, n_m) + \\ & + \tau^7 F_7^{(m)}(\Theta_1, \tilde{\Theta}_5, \tilde{\Theta}_7, n_0, n_1, \dots, n_m) + \tau^9 F_9^{(m)}(\Theta_1, \tilde{\Theta}_5, \tilde{\Theta}_7, \tilde{\Theta}_9, n_0, n_1, \dots, n_m) \end{aligned} \tag{37}$$

и определяется рекурсивными формулами. Для этого необходимо найти функции $\Omega^{(1)}$ и $\Omega^{(2)}$, подобные (28), но записанные для S_4 :

$$S_4(n_1\tau)S_4(n_0\tau)S_4(n_1\tau) = \exp\{\Omega_1\}\exp\{\Omega_0\}\exp\{\Omega_1\} = \exp\{\Omega^{(1)}\}, \tag{38.1}$$

$$S_4(n_2\tau)S_4(n_1\tau)S_4(n_0\tau)S_4(n_1\tau)S_4(n_2\tau) = \exp\{\Omega_2\}\exp\{\Omega^{(1)}\}\exp\{\Omega_2\} = \exp\{\Omega^{(2)}\}, \tag{38.2}$$

, ...,

$$S_4(n_m\tau) \dots S_4(n_0\tau) \dots S_4(n_m\tau) = \exp\{\Omega_m\}\exp\{\Omega^{(m-1)}\}\exp\{\Omega_m\} = \exp\{\Omega^{(m)}\}. \tag{38.3}$$

Проделав все математические операции, как и в предыдущем пункте, можно получить рекуррентную формулу для выражения $\Omega^{(m)}$:

$$\begin{aligned} \Omega^{(m)} = & \tau A_{1,m} \Xi_{\Theta_1} + \tau^5 A_{5,m} \Xi_{\tilde{\Theta}_5} + \tau^7 (A_{7,m} \Xi_{\tilde{\Theta}_7} + B_{7,m} \Xi_{\tilde{\Theta}_5\Theta_1}) + \\ & + \tau^9 (A_{9,m} \Xi_{\tilde{\Theta}_9} + B_{9,m} \Xi_{\tilde{\Theta}_7\Theta_1} + C_{9,m} \Xi_{\tilde{\Theta}_5\Theta_1\Theta_1}) \end{aligned} \tag{39}$$

Здесь функции A, B, C содержат неизвестные n_0, \dots, n_6 . Полные математические выкладки и точный вид функций приведены в приложении С.

Условием аппроксимации оператора эволюции в данном случае является система семи нелинейных алгебраических уравнений:

$$\begin{aligned} A_{1,m} - 1 = A_{5,m} = A_{7,m} = A_{9,m} = 0, \\ B_{7,m} = B_{9,m} = C_{9,m} = 0. \end{aligned} \tag{40}$$

Таблица 2. Шесть решений системы (40) с наименьшими по величине значениями n_i

Решение 1	Решение 2	Решение 3
$n_0 \rightarrow -0.944263162369538$	$n_0 \rightarrow 0.9138389914113408$	$n_0 \rightarrow 0.8799651058319341$
$n_1 \rightarrow -0.001699669869606$	$n_1 \rightarrow -0.5761621584448674$	$n_1 \rightarrow -0.6895760778774856$
$n_2 \rightarrow 0.901856109289368$	$n_2 \rightarrow 0.2847691807605744$	$n_2 \rightarrow 0.5022004328862611$
$n_3 \rightarrow -0.776519231974205$	$n_3 \rightarrow 0.2097456771750264$	$n_3 \rightarrow 0.2422465593742511$
$n_4 \rightarrow 0.042244707029581$	$n_4 \rightarrow -0.8748957631571992$	$n_4 \rightarrow -0.8650419726013663$
$n_5 \rightarrow 0.570335703509034$	$n_5 \rightarrow 0.2391267522973746$	$n_5 \rightarrow 0.0625099903030752$
$n_6 \rightarrow 0.235913963200597$	$n_6 \rightarrow 0.7604968156634205$	$n_6 \rightarrow 0.8076785149992974$
Решение 4	Решение 5	Решение 6
$n_0 \rightarrow -0.96028501069323$	$n_0 \rightarrow -1.046841102301594$	$n_0 \rightarrow 0.8796664129952768$
$n_1 \rightarrow 0.43745939094565$	$n_1 \rightarrow -1.054334166569394$	$n_1 \rightarrow -0.7725703942483008$
$n_2 \rightarrow 0.92116736423630$	$n_2 \rightarrow 1.051838045363935$	$n_2 \rightarrow 0.5660708443785679$
$n_3 \rightarrow 0.32276673964696$	$n_3 \rightarrow 0.250073053930700$	$n_3 \rightarrow 0.2599557921889106$
$n_4 \rightarrow 0.65920326505061$	$n_4 \rightarrow 1.004899649890149$	$n_4 \rightarrow -0.9423452393494703$
$n_5 \rightarrow -0.55059025299523$	$n_5 \rightarrow -0.865766610376252$	$n_5 \rightarrow 0.0190648617779735$
$n_6 \rightarrow -0.80986400153768$	$n_6 \rightarrow 0.636710578911658$	$n_6 \rightarrow 0.9299909287546806$

Система (40) оказывается намного проще системы (34) и может быть решена с помощью встроенных средств Mathematica. Список некоторых полученных решений приведен в табл. 2.

Используя значения n_i , полученные при решении системы (40), можно получить численную аппроксимацию оператора эволюции (12):

$$S_{S_4}(\tau) = S_4(n_6\tau) \dots S_4(n_0\tau) \dots S_4(n_6\tau) = \prod_{i=1}^{39} e^{\tilde{a}_{10,i}\tau A} e^{\tilde{b}_{10,i}\tau B} + o(\tau^{11}). \tag{41}$$

4. ОБСУЖДЕНИЕ

При сравнении становится ясно, что наименьшее количество экспонент потребуется при использовании в формуле (26) конструкций S_2 , и что при увеличении k формула (26) становится все менее выгодной по сравнению с формулами (22)–(25). С другой стороны, построение формулы (26) и расчет коэффициентов n_i для конструкций S_2 значительно более сложные, чем расчет аналогичных коэффициентов для конструкций S_4 , поскольку требует решения гораздо более сложной системы нелинейных алгебраических уравнений. Кроме того, количество экспонент в формуле (26) с S_4 лишь ненамного больше варианта формулы (26) с S_2 .

Были проведены тестовые расчеты, в которых сравнивались порядки точности схем, основанных на формулах (25), (35), (41). В качестве тестовой задачи был взят одномерный осциллятор с частотой, зависящей от времени, по аналогии с работой [8]. Уравнение Шрёдингера такой системы выглядит следующим образом:

$$i \frac{\partial \Psi(x,t)}{\partial t} = \left(-\frac{1}{2} \frac{\partial^2}{\partial x^2} + \frac{\omega^2(t)x^2}{2} \right) \Psi(x,t), \tag{42}$$

где $\omega^2(t) = 4 - 3\exp(-t)$ (см. [25]).

В качестве начального состояния берется волновой пакет

$$\Psi(x,0) = \frac{1}{\sqrt[4]{A}} \exp\left[-\frac{1}{2}(x - \sqrt{2})^2\right]. \quad (43)$$

Точное решение задачи известно и может быть записано как

$$\Psi(x,t) = \frac{1}{\sqrt[4]{\pi}} \exp\left[-X(t)x^2 + 2Y(t)x - Z(t)\right]. \quad (44)$$

Здесь функции $X(t)$, $Y(t)$, $Z(t)$ удовлетворяют следующей задаче Коши:

$$\begin{aligned} i \frac{d}{dt} X(t) &= 2X^2(t) - \frac{\omega^2(t)}{2}, & X(0) &= \frac{1}{2}, \\ i \frac{d}{dt} Y(t) &= 2X(t)Y(t), & Y(0) &= \frac{\sqrt{2}}{2}, \\ i \frac{d}{dt} Z(t) &= -X(t) + 2Y^2(t), & Z(0) &= 1. \end{aligned} \quad (45)$$

Рассматриваемая область пространства определена как $[x_{\min} = -10, x_{\max} = 10]$. Количество узлов в сетке было взято 1000. Шаг по времени $\tau = 0.1$.

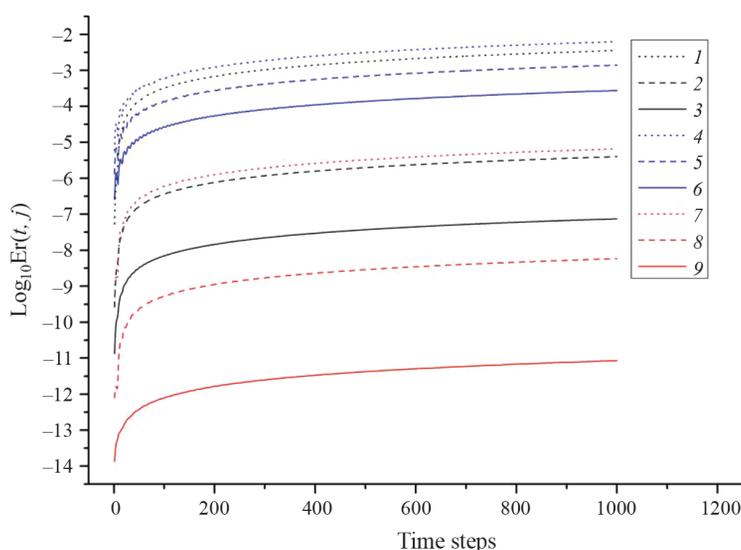
Для анализа сходимости представленных схем на последовательности трех сгущающихся по времени временных сетках вычислялась функция погрешности

$$\text{Er}^2(t,j) = \int_{x_{\min}}^{x_{\max}} |\Psi(x,t) - \Psi_{\tau_j}(x,t)|^2 dx. \quad (46)$$

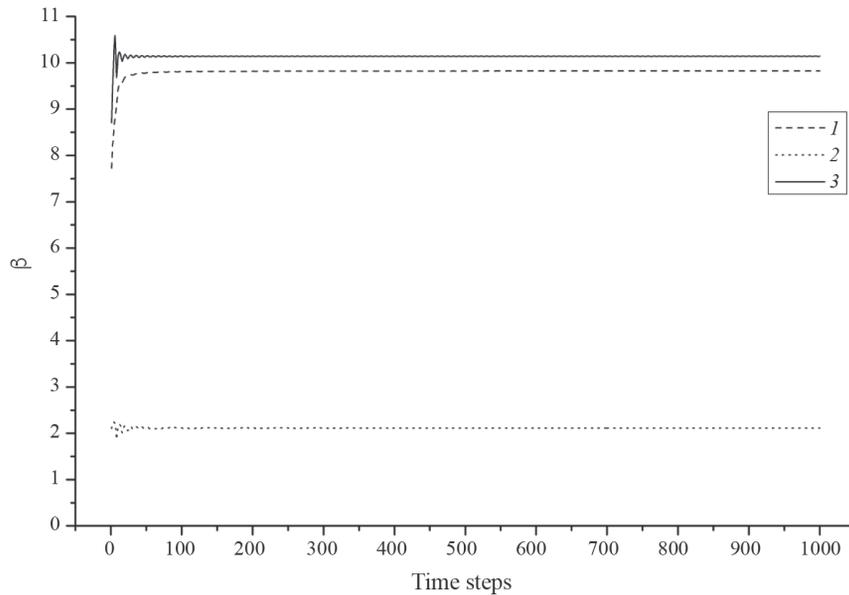
Здесь индекс $j = 1, 2, 3$ нумерует решения, полученные для шагов $\tau, \tau/2, \tau/4$ по временной переменной. В качестве функции $\Psi(x, t)$ берется решение (44), (45).

Затем рассчитывался коэффициент Рунге

$$\beta(t) = \log_2 \left| \frac{\text{Er}(t,1) - \text{Er}(t,2)}{\text{Er}(t,2) - \text{Er}(t,3)} \right|. \quad (47)$$



Фиг. 1. Результаты тестовых расчетов в виде функций $\text{Er}(t,j)$, $j = 1, 2, 3$. Цифрами обозначены формулы, по которым проводился численный расчет: 1, 2, 3 — (25), 4, 5, 6 — (35), 7, 8, 9 — (41) с шагом по времени $\tau, \tau/2, \tau/4$ соответственно.



Фиг. 2. Поведение функции $\beta(t)$ для схем расчета по формулам: 1 — (25), 2 — (35), 3 — (41) с течением времени.

На фиг. 1 приведены кривые $\log_{10}Er(t, j)$, $j = 1, 2, 3$, для схем 10-го порядка точности (25), (35), (41). Как следует из графиков, среди трех представленных схем наилучший результат показывает схема (41). Несколько хуже результаты для схемы (25). Наихудший результат выдает формула (35).

На фиг. 2 представлены графики функции $\beta(t)$ для (25), (35), (41) в зависимости от времени. По результатам тестовых расчетов численная схема (35) не оправдала ожиданий. Среднее значение коэффициента Рунге равно 2. Наиболее вероятной причиной подобного поведения является относительно низкая точность численного решения нелинейной системы алгебраических уравнений (34). Несмотря на то что в формуле (41) содержится на 4 экспоненциальных множителя больше, чем в формуле (35), в силу ее точности и более простой практической реализации предпочтительнее использовать именно формулу (41) для численных расчетов.

5. ЗАКЛЮЧЕНИЕ

Представлена схема аппроксимации оператора эволюции (12) 10-го порядка точности с гораздо меньшим числом показателей, чем в формуле Ли–Троттера–Сузуки.

Рассмотрены две численные схемы с использованием операторных конструкций S_2 и S_4 . Для обоих случаев определены системы нелинейных алгебраических уравнений для вычисления коэффициентов при операторных показателях. Решения систем для случая S_2 находились методом Монте-Карло с последующим варьированием коэффициентов для минимизации функции невязки, для случая S_4 система решалась в программном продукте Mathematica.

Были проведены тестовые расчеты для сравнения эффективности трех численных схем 10-го порядка: (25), (35) и (41). Результаты показывают хорошую относительную точность схемы (41). Крайне слабая эффективность схемы (35), по-видимому, объясняется недостаточной точностью найденных коэффициентов при экспонентах.

Автор благодарит д.ф.-м.н. Владимира Степановича Мележика за ценные замечания и полезные обсуждения. Автор также благодарит команду гетерогенной платформы кластера «HybriLIT» Лаборатории информационных технологий им. М. Г. Мещерякова Объединенного института ядерных исследований за предоставленные расчетные ресурсы.

ПРИЛОЖЕНИЕ А

Список необходимых слагаемых формулы Бэйкера–Кэмпбелла–Хаусдорфа

$$\exp\left(\frac{1}{2}X\right)\exp(Y)\exp\left(\frac{1}{2}X\right) = \exp\{Z(X, Y)\}:$$

Список необходимых слагаемых формулы Бэйкера-Кэмпбелла-Хаусдорфа $\exp\left(\frac{1}{2}X\right)\exp(Y)\exp\left(\frac{1}{2}X\right) = \exp\{Z(X, Y)\}$:

$$E_i = [E_i', E_i'']$$

$E_1 = N_X X$	$N_X = 2$	$E_{16} = [E_9, E_2] = N_{YXXXXY} \Xi_{YXXXXY}$	$N_{YXXXXY} = 0$
$E_2 = N_Y Y$	$N_Y = 1$	$E_{17} = [E_{10}, E_2] = N_{YXXXY} \Xi_{YXXXY}$	$N_{YXXXY} = 0$
$E_3 = [E_2, E_1] = N_{YX} \Xi_{YX}$	$N_{YX} = 0$	$E_{18} = [E_{11}, E_2] = N_{YXXYY} \Xi_{YXXYY}$	$N_{YXXYY} = 0$
$E_4 = [E_3, E_1] = N_{YXX} \Xi_{YXX}$	$N_{YXX} = -\frac{1}{6}$	$E_{19} = [E_{12}, E_2] = N_{YXYYY} \Xi_{YXYYY}$	$N_{YXYYY} = 0$
$E_5 = [E_3, E_2] = N_{YXY} \Xi_{YXY}$	$N_{YXY} = -\frac{1}{6}$	$E_{20} = [E_6, E_3] = N_{YXXX, YX} \Xi_{YXXX, YX}$	$N_{YXXX, YX} = 0$
$E_6 = [E_4, E_1] = N_{YXXX} \Xi_{YXXX}$	$N_{YXXX} = 0$	$E_{21} = [E_7, E_3] = N_{YXXY, YX} \Xi_{YXXY, YX}$	$N_{YXXY, YX} = 0$
$E_7 = [E_4, E_2] = N_{YXXY} \Xi_{YXXY}$	$N_{YXXY} = 0$	$E_{22} = [E_8, E_3] = N_{YXY, YX} \Xi_{YXY, YX}$	$N_{YXY, YX} = 0$
$E_8 = [E_5, E_2] = N_{YXY} \Xi_{YXY}$	$N_{YXY} = 0$	$E_{23} = [E_5, E_4] = N_{YXY, YXX} \Xi_{YXY, YXX}$	$N_{YXY, YXX} = 0$
$E_9 = [E_6, E_1] = N_{YXXXX} \Xi_{YXXXX}$	$N_{YXXXX} = \frac{7}{360}$	$E_{24} = [E_{15}, E_1] = N_{YXXXXXXXX} \Xi_{YXXXXXXXX}$	$N_{YXXXXXXXX} = -\frac{31}{15120}$
$E_{10} = [E_6, E_2] = N_{YXXXY} \Xi_{YXXXY}$	$N_{YXXXY} = \frac{7}{180}$	$E_{25} = [E_{15}, E_2] = N_{YXXXXXY} \Xi_{YXXXXXY}$	$N_{YXXXXXY} = -\frac{31}{5040}$
$E_{11} = [E_7, E_2] = N_{YXXY} \Xi_{YXXY}$	$N_{YXXY} = \frac{1}{45}$	$E_{26} = [E_{16}, E_2] = N_{YXXXY} \Xi_{YXXXY}$	$N_{YXXXY} = -\frac{13}{1890}$
$E_{12} = [E_8, E_2] = N_{YXY} \Xi_{YXY}$	$N_{YXY} = \frac{1}{360}$	$E_{27} = [E_{17}, E_2] = N_{YXXYY} \Xi_{YXXYY}$	$N_{YXXYY} = -\frac{53}{15120}$
$E_{13} = [E_4, E_3] = N_{YXX, YX} \Xi_{YXX, YX}$	$N_{YXX, YX} = \frac{1}{60}$	$E_{28} = [E_{18}, E_2] = N_{YXXYYY} \Xi_{YXXYYY}$	$N_{YXXYYY} = -\frac{1}{1260}$
$E_{14} = [E_5, E_3] = N_{YXY, YX} \Xi_{YXY, YX}$	$N_{YXY, YX} = -\frac{1}{90}$	$E_{29} = [E_{19}, E_2] = N_{YXYYY} \Xi_{YXYYY}$	$N_{YXYYY} = -\frac{1}{15120}$
$E_{15} = [E_9, E_1] = N_{YXXXX} \Xi_{YXXXX}$	$N_{YXXXX} = 0$		

Здесь

$$\Xi_{ab} = [a, b]$$

$$\Xi_{abc} = [[a, b], c]$$

$$\Xi_{abcde} = [[[[a, b], c], d], e]$$

$$\Xi_{abc, de} = [[a, b], c], [d, e]$$

$$\Xi_{abs, defg} = [[a, b], c], [d, e], [f, g]$$

$$\Xi_{abcd, efg} = [[[[a, b], c], d], [e, f], g]$$

ПРИЛОЖЕНИЕ В

Подставим $\Omega_0(\Theta_1, \Theta_3, \Theta_5, \Theta_7)$ и $\Omega_1(\Theta_1, \Theta_3, \Theta_5, \Theta_7)$ в формулу (31).

Введем следующие обозначения:

$$A_{1,0} = n_0; A_{3,0} = n_0^3; A_{5,0} = n_0^5; A_{7,0} = n_0^7; A_{9,0} = n_0^9. \quad (B.1)$$

Первые два слагаемых, $2\Xi_X + \Xi_Y$, можно записать как

$$\begin{aligned} \Omega_0 + 2\Omega_1 &= \tau A_{1,0}\Theta_1 + \tau^3 A_{3,0}\Theta_3 + \\ &+ \tau^5 A_{5,0}\Theta_5 + \tau^7 A_{7,0}\Theta_7 + \tau^9 A_{9,0}\Theta_9 + \\ &+ 2\tau n_1\Theta_1 + 2\tau^3 n_1^3\Theta_3 + 2\tau^5 n_1^5\Theta_5 + 2\tau^7 n_1^7\Theta_7 + 2\tau^9 n_1^9\Theta_9. \end{aligned} \quad (B.2.1)$$

Остальные слагаемые:

$$\begin{aligned} N_{YXX}\Xi_{\Omega_0\Omega_1\Omega_1} &= \tau^5 N_{YXX} \left\{ A_{1,0}n_1^4\Xi_{\Theta_1\Theta_3\Theta_1} + A_{3,0}n_1^2\Xi_{\Theta_3\Theta_1\Theta_1} \right\} + \\ &+ \tau^7 N_{YXX} \left\{ A_{1,0}n_1^6\Xi_{\Theta_1\Theta_3\Theta_3} + A_{3,0}n_1^4\Xi_{\Theta_3\Theta_1\Theta_3} + A_{1,0}n_1^6\Xi_{\Theta_1\Theta_3\Theta_1} + A_{5,0}n_1^2\Xi_{\Theta_5\Theta_1\Theta_1} \right\} + \\ &+ \tau^9 N_{YXX} \left\{ A_{1,0}n_1^8\Xi_{\Theta_1\Theta_3\Theta_5} + A_{3,0}n_1^6\Xi_{\Theta_3\Theta_1\Theta_5} + A_{1,0}n_1^8\Xi_{\Theta_1\Theta_3\Theta_3} + A_{5,0}n_1^4\Xi_{\Theta_5\Theta_1\Theta_3} + \right. \\ &\left. + A_{3,0}n_1^6\Xi_{\Theta_3\Theta_3\Theta_1} + A_{5,0}n_1^4\Xi_{\Theta_3\Theta_3\Theta_1} + A_{1,0}n_1^8\Xi_{\Theta_1\Theta_7\Theta_1} + A_{7,0}n_1^2\Xi_{\Theta_7\Theta_1\Theta_1} \right\}, \end{aligned} \quad (B.2.2)$$

$$\begin{aligned} N_{YXY}\Xi_{\Omega_0\Omega_1\Omega_0} &= \tau^5 N_{YXY} \left\{ (A_{1,0})^2 n_1^3\Xi_{\Theta_1\Theta_3\Theta_1} + A_{1,0}A_{3,0}n_1\Xi_{\Theta_3\Theta_1\Theta_1} \right\} + \\ &+ \tau^7 N_{YXY} \left\{ A_{1,0}A_{3,0}n_1^3\Xi_{\Theta_1\Theta_3\Theta_3} + (A_{3,0})^2 n_1\Xi_{\Theta_3\Theta_1\Theta_3} + (A_{1,0})^2 n_1^5\Xi_{\Theta_1\Theta_5\Theta_1} + A_{1,0}A_{5,0}n_1\Xi_{\Theta_5\Theta_1\Theta_1} \right\} + \\ &+ \tau^9 N_{YXY} \left\{ A_{1,0}A_{5,0}n_1^3\Xi_{\Theta_1\Theta_3\Theta_5} + A_{3,0}A_{5,0}n_1\Xi_{\Theta_3\Theta_1\Theta_5} + A_{1,0}A_{3,0}n_1^5\Xi_{\Theta_1\Theta_3\Theta_3} + A_{3,0}A_{5,0}n_1\Xi_{\Theta_5\Theta_1\Theta_3} + \right. \\ &\left. + A_{1,0}A_{3,0}n_1^5\Xi_{\Theta_3\Theta_3\Theta_1} + A_{1,0}A_{5,0}n_1^3\Xi_{\Theta_3\Theta_3\Theta_1} + (A_{1,0})^2 n_1^7\Xi_{\Theta_1\Theta_7\Theta_1} + A_{1,0}A_{7,0}n_1\Xi_{\Theta_7\Theta_1\Theta_1} \right\}, \end{aligned} \quad (B.2.3)$$

$$\begin{aligned} N_{YXXXX}\Xi_{\Omega_0\Omega_1\Omega_1\Omega_1} &= \tau^7 N_{YXXXX} \left\{ A_{1,0}n_1^6\Xi_{\Theta_1\Theta_3\Theta_1\Theta_1} + A_{3,0}n_1^4\Xi_{\Theta_3\Theta_1\Theta_1\Theta_1} \right\} + \\ &+ \tau^9 N_{YXXXX} \left\{ A_{1,0}n_1^8\Xi_{\Theta_1\Theta_3\Theta_3\Theta_1} + A_{3,0}n_1^6\Xi_{\Theta_3\Theta_1\Theta_3\Theta_1} + A_{1,0}n_1^8\Xi_{\Theta_1\Theta_3\Theta_1\Theta_3} + \right. \\ &\left. + A_{3,0}n_1^6\Xi_{\Theta_3\Theta_1\Theta_1\Theta_3} + A_{1,0}n_1^8\Xi_{\Theta_1\Theta_3\Theta_1\Theta_3} + A_{3,0}n_1^6\Xi_{\Theta_3\Theta_1\Theta_1\Theta_3} + \right. \\ &\left. + A_{1,0}n_1^8\Xi_{\Theta_1\Theta_3\Theta_1\Theta_1} + A_{5,0}n_1^4\Xi_{\Theta_5\Theta_1\Theta_1\Theta_1} \right\}, \end{aligned} \quad (B.2.4)$$

$$\begin{aligned} N_{YXXXY}\Xi_{\Omega_0\Omega_1\Omega_1\Omega_0} &= \tau^7 N_{YXXXY} \left\{ (A_{1,0})^2 n_1^5\Xi_{\Theta_1\Theta_3\Theta_1\Theta_1} + A_{1,0}A_{3,0}n_1^3\Xi_{\Theta_3\Theta_1\Theta_1\Theta_1} \right\} + \\ &+ \tau^9 N_{YXXXY} \left\{ (A_{1,0})^2 n_1^7\Xi_{\Theta_1\Theta_3\Theta_3\Theta_1} + A_{1,0}A_{3,0}n_1^5\Xi_{\Theta_3\Theta_1\Theta_3\Theta_1} + (A_{1,0})^2 n_1^7\Xi_{\Theta_1\Theta_3\Theta_1\Theta_3} + \right. \\ &\left. + A_{1,0}A_{3,0}n_1^5\Xi_{\Theta_3\Theta_1\Theta_1\Theta_3} + A_{1,0}A_{3,0}n_1^5\Xi_{\Theta_3\Theta_1\Theta_1\Theta_3} + (A_{3,0})^2 n_1^3\Xi_{\Theta_3\Theta_1\Theta_1\Theta_3} + \right. \\ &\left. + (A_{1,0})^2 n_1^7\Xi_{\Theta_1\Theta_5\Theta_1\Theta_1} + A_{1,0}A_{5,0}n_1^3\Xi_{\Theta_5\Theta_1\Theta_1\Theta_1} \right\}, \end{aligned} \quad (B.2.5)$$

$$\begin{aligned} N_{YXXYY}\Xi_{\Omega_0\Omega_1\Omega_1\Omega_0} &= \tau^7 N_{YXXYY} \left\{ (A_{1,0})^3 n_1^4\Xi_{\Theta_1\Theta_3\Theta_1\Theta_1} + A_{3,0}(A_{1,0})^2 n_1^2\Xi_{\Theta_3\Theta_1\Theta_1\Theta_1} \right\} + \\ &+ \tau^9 N_{YXXYY} \left\{ (A_{1,0})^3 n_1^6\Xi_{\Theta_1\Theta_3\Theta_3\Theta_1} + A_{3,0}(A_{1,0})^2 n_1^4\Xi_{\Theta_3\Theta_1\Theta_3\Theta_1} + A_{3,0}(A_{1,0})^2 n_1^4\Xi_{\Theta_1\Theta_3\Theta_1\Theta_3} + \right. \end{aligned} \quad (B.2.6)$$

$$\begin{aligned}
 & +A_{1,0} (A_{3,0})^2 n_1^2 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_3\Theta_1} + A_{3,0} (A_{1,0})^2 n_1^4 \Xi_{\Theta_1\Theta_3\Theta_1\Theta_1\Theta_3} + A_{1,0} (A_{3,0})^2 n_1^2 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_3} + \\
 & \quad + (A_{1,0})^3 n_1^6 \Xi_{\Theta_1\Theta_3\Theta_1\Theta_1\Theta_1} + (A_{1,0})^2 A_{5,0} n_1^2 \Xi_{\Theta_5\Theta_1\Theta_1\Theta_1\Theta_1} \Big\}, \\
 N_{YXYYY} \Xi_{\Omega_0\Omega_1\Omega_0\Omega_0\Omega_0} & = \tau^7 N_{YXYYY} \left\{ (A_{1,0})^4 n_1^3 \Xi_{\Theta_1\Theta_3\Theta_1\Theta_1\Theta_1} + (A_{1,0})^3 A_{3,0} n_1 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1} \right\} + \\
 + \tau^9 N_{YXYYY} & \left\{ (A_{1,0})^3 A_{3,0} n_1^3 \Xi_{\Theta_1\Theta_3\Theta_3\Theta_1\Theta_1} + (A_{1,0})^2 (A_{3,0})^2 n_1 \Xi_{\Theta_3\Theta_1\Theta_3\Theta_1\Theta_1} + (A_{1,0})^3 A_{3,0} n_1^3 \Xi_{\Theta_1\Theta_3\Theta_1\Theta_3\Theta_1} + \right. \\
 & \quad + (A_{1,0})^2 (A_{3,0})^2 n_1 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_3\Theta_1} + (A_{1,0})^3 A_{3,0} n_1^3 \Xi_{\Theta_1\Theta_3\Theta_1\Theta_1\Theta_3} + (A_{1,0})^2 (A_{3,0})^2 n_1 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_3} + \\
 & \quad \left. + (A_{1,0})^4 n_1^5 \Xi_{\Theta_1\Theta_3\Theta_1\Theta_1\Theta_1} + (A_{1,0})^3 A_{5,0} n_1 \Xi_{\Theta_5\Theta_1\Theta_1\Theta_1\Theta_1} \right\}, \tag{B.2.7}
 \end{aligned}$$

$$\begin{aligned}
 N_{YXX,YX} \Xi_{\Omega_0\Omega_1\Omega_1,\Omega_0\Omega_1} & = N_{YXX,YX} \tau^9 \left\{ (A_{1,0})^2 n_1^7 \Xi_{\Theta_1\Theta_3\Theta_1,\Theta_1\Theta_3} + A_{1,0} A_{3,0} n_1^5 \Xi_{\Theta_3\Theta_1\Theta_1,\Theta_1\Theta_3} \right. \\
 & \quad \left. + A_{1,0} A_{3,0} n_1^5 \Xi_{\Theta_1\Theta_3\Theta_1,\Theta_3\Theta_1} + (A_{3,0})^2 n_1^3 \Xi_{\Theta_3\Theta_1\Theta_1,\Theta_3\Theta_1} \right\}, \tag{B.2.8}
 \end{aligned}$$

$$\begin{aligned}
 N_{YXY,YX} \Xi_{\Omega_0\Omega_1\Omega_0,\Omega_0\Omega_1} & = N_{YXY,YX} \tau^9 \left\{ (A_{1,0})^3 n_1^6 \Xi_{\Theta_1\Theta_3\Theta_1,\Theta_1\Theta_3} + (A_{1,0})^2 A_{3,0} n_1^4 \Xi_{\Theta_3\Theta_1\Theta_1,\Theta_1\Theta_3} \right. \\
 & \quad \left. + (A_{1,0})^2 A_{3,0} n_1^4 \Xi_{\Theta_1\Theta_3\Theta_1,\Theta_3\Theta_1} + A_{1,0} (A_{3,0})^2 n_1^2 \Xi_{\Theta_3\Theta_1\Theta_1,\Theta_3\Theta_1} \right\}, \tag{B.2.9}
 \end{aligned}$$

$$N_{YXXXXXX} \Xi_{\Omega_0\Omega_1\Omega_1\Omega_1\Omega_1\Omega_1\Omega_1} = \tau^9 N_{YXXXXXX} \left\{ A_{1,0} n_1^8 \Xi_{\Theta_1\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1\Theta_1} + A_{3,0} n_1^6 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1\Theta_1} \right\}, \tag{B.2.10}$$

$$N_{YXXXXXY} \Xi_{\Omega_0\Omega_1\Omega_1\Omega_1\Omega_1\Omega_1\Omega_0} = \tau^9 N_{YXXXXXY} \left\{ (A_{1,0})^2 n_1^7 \Xi_{\Theta_1\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1\Theta_1} + A_{3,0} A_{1,0} n_1^5 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1\Theta_1} \right\}, \tag{B.2.11}$$

$$N_{YXXXXYY} \Xi_{\Omega_0\Omega_1\Omega_1\Omega_1\Omega_1\Omega_0\Omega_0} = \tau^9 N_{YXXXXYY} \left\{ (A_{1,0})^3 n_1^6 \Xi_{\Theta_1\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1\Theta_1} + A_{3,0} (A_{1,0})^2 n_1^4 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1\Theta_1} \right\}, \tag{B.2.12}$$

$$N_{YXXXXYY} \Xi_{\Omega_0\Omega_1\Omega_1\Omega_1\Omega_0\Omega_0\Omega_0} = \tau^9 N_{YXXXXYY} \left\{ (A_{1,0})^4 n_1^5 \Xi_{\Theta_1\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1\Theta_1} + A_{3,0} (A_{1,0})^3 n_1^3 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1\Theta_1} \right\}, \tag{B.2.13}$$

$$N_{YXXXXYY} \Xi_{\Omega_0\Omega_1\Omega_1\Omega_0\Omega_0\Omega_0\Omega_0} = \tau^9 N_{YXXXXYY} \left\{ (A_{1,0})^5 n_1^4 \Xi_{\Theta_1\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1\Theta_1} + A_{3,0} (A_{1,0})^4 n_1^2 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1\Theta_1} \right\}, \tag{B.2.14}$$

$$N_{YXXXXYY} \Xi_{\Omega_0\Omega_1\Omega_0\Omega_0\Omega_0\Omega_0\Omega_0} = \tau^9 N_{YXXXXYY} \left\{ (A_{1,0})^6 n_1^3 \Xi_{\Theta_1\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1\Theta_1} + A_{3,0} (A_{1,0})^5 n_1 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1\Theta_1} \right\}. \tag{B.2.15}$$

В итоге выражение $\Omega^{(1)}$ можно представить в виде

$$\begin{aligned}
 \Omega^{(1)} & = \tau A_{1,1} \Xi_{\Theta_1} + \tau^3 A_{3,1} \Xi_{\Theta_3} + \tau^5 (A_{5,1} \Xi_{\Theta_5} + B_{5,1} \Xi_{\Theta_3\Theta_1\Theta_1}) \\
 & + \tau^7 (A_{7,1} \Xi_{\Theta_7} + B_{7,1} \Xi_{\Theta_5\Theta_1\Theta_1} + B_{7,1}^3 \Xi_{\Theta_3\Theta_1\Theta_3} + C_{7,1} \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1}) \\
 & + \tau^9 (A_{9,1} \Xi_{\Theta_9} + B_{9,1} \Xi_{\Theta_7\Theta_1\Theta_1} + B_{9,1}^1 \Xi_{\Theta_5\Theta_3\Theta_1} + B_{9,1}^3 \Xi_{\Theta_5\Theta_3\Theta_3} + B_{9,1}^5 \Xi_{\Theta_3\Theta_1\Theta_5} + C_{9,1} \Xi_{\Theta_5\Theta_3\Theta_1\Theta_1\Theta_1} \\
 & + C_{9,1}^{311} \Xi_{\Theta_3\Theta_1\Theta_3\Theta_1\Theta_1} + C_{9,1}^{131} \Xi_{\Theta_3\Theta_1\Theta_1\Theta_3\Theta_1} + C_{9,1}^{113} \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_3} + E_{9,1} \Xi_{\Theta_3\Theta_1\Theta_1,\Theta_3\Theta_1} + D_{9,1} \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1\Theta_1}). \tag{B.3}
 \end{aligned}$$

Здесь

$$A_{1,1} \equiv A_{1,0} + 2n_1, \tag{B.4.1}$$

$$A_{3,1} \equiv A_{3,0} + 2n_1^3, \tag{B.4.2}$$

$$A_{5,1} \equiv A_{5,0} + 2n_1^5, \quad (\text{B.4.3})$$

$$A_{7,1} \equiv A_{7,0} + 2n_1^7, \quad (\text{B.4.4})$$

$$A_{9,1} \equiv A_{9,0} + 2n_1^9, \quad (\text{B.4.5})$$

$$B_{5,1} = N_{YXX} \{A_{3,0}n_1^2 - A_{1,0}n_1^4\} + N_{YXY} \{A_{1,0}A_{3,0}n_1 - (A_{1,0})^2 n_1^3\}, \quad (\text{B.4.6})$$

$$B_{7,1} = N_{YXX} \{A_{5,0}n_1^2 - A_{1,0}n_1^6\} + N_{YXY} \{A_{1,0}A_{5,0}n_1 - (A_{1,0})^2 n_1^5\}, \quad (\text{B.4.7})$$

$$B_{7,1}^3 = N_{YXX} \{A_{3,0}n_1^4 - A_{1,0}n_1^6\} + N_{YXY} \{(A_{3,0})^2 n_1 - A_{1,0}A_{3,0}n_1^3\}, \quad (\text{B.4.8})$$

$$B_{9,1} = N_{YXX} \{A_{7,0}n_1^2 - A_{1,0}n_1^8\} + N_{YXY} \{A_{1,0}A_{7,0}n_1 - (A_{1,0})^2 n_1^7\}, \quad (\text{B.4.9})$$

$$B_{9,1}^1 = N_{YXX} \{A_{5,0}n_1^4 - A_{3,0}n_1^6\} + N_{YXY} \{A_{1,0}A_{5,0}n_1^3 - A_{1,0}A_{3,0}n_1^5\}, \quad (\text{B.4.10})$$

$$B_{9,1}^3 = N_{YXX} \{A_{5,0}n_1^4 - A_{1,0}n_1^8\} + N_{YXY} \{A_{3,0}A_{5,0}n_1 - A_{1,0}A_{3,0}n_1^5\}, \quad (\text{B.4.11})$$

$$B_{9,1}^5 = N_{YXX} \{A_{3,0}n_1^6 - A_{1,0}n_1^8\} + N_{YXY} \{A_{3,0}A_{5,0}n_1 - A_{1,0}A_{5,0}n_1^3\}, \quad (\text{B.4.12})$$

$$C_{7,1} = N_{YXXXX} \{A_{3,0}n_1^4 - A_{1,0}n_1^6\} + N_{YXXXY} \{A_{1,0}A_{3,0}n_1^3 - (A_{1,0})^2 n_1^5\} \\ + N_{YXXYY} \{A_{3,0}(A_{1,0})^2 n_1^2 - (A_{1,0})^3 n_1^4\} + N_{YXYYY} \{(A_{1,0})^3 A_{3,0}n_1 - (A_{1,0})^4 n_1^3\}, \quad (\text{B.4.13})$$

$$C_{9,1} = N_{YXXXX} \{A_{5,0}n_1^4 - A_{1,0}n_1^8\} + N_{YXXXY} \{A_{1,0}A_{5,0}n_1^3 - (A_{1,0})^2 n_1^7\} \\ + N_{YXXYY} \{(A_{1,0})^2 A_{5,0}n_1^2 - (A_{1,0})^3 n_1^6\} + N_{YXYYY} \{(A_{1,0})^3 A_{5,0}n_1 - (A_{1,0})^4 n_1^5\}, \quad (\text{B.4.14})$$

$$C_{9,1}^{311} = N_{YXXXX} \{A_{3,0}n_1^6 - A_{1,0}n_1^8\} + N_{YXXXY} \{A_{1,0}A_{3,0}n_1^5 - (A_{1,0})^2 n_1^7\} \\ + N_{YXXYY} \{A_{3,0}(A_{1,0})^2 n_1^4 - (A_{1,0})^3 n_1^6\} + N_{YXYYY} \{(A_{1,0})^2 (A_{3,0})^2 n_1 - (A_{1,0})^3 A_{3,0}n_1^3\}, \quad (\text{B.4.15})$$

$$C_{9,1}^{131} = N_{YXXXX} \{A_{3,0}n_1^6 - A_{1,0}n_1^8\} + N_{YXXXY} \{A_{1,0}A_{3,0}n_1^5 - (A_{1,0})^2 n_1^7\} \\ + N_{YXXYY} \{A_{1,0}(A_{3,0})^2 n_1^2 - A_{3,0}(A_{1,0})^2 n_1^4\} + N_{YXYYY} \{(A_{1,0})^2 (A_{3,0})^2 n_1 - (A_{1,0})^3 A_{3,0}n_1^3\}, \quad (\text{B.4.16})$$

$$C_{9,1}^{113} = N_{YXXXX} \{A_{3,0}n_1^6 - A_{1,0}n_1^8\} + N_{YXXXY} \{(A_{3,0})^2 n_1^3 - A_{1,0}A_{3,0}n_1^5\} \\ + N_{YXXYY} \{A_{1,0}(A_{3,0})^2 n_1^2 - A_{3,0}(A_{1,0})^2 n_1^4\} + N_{YXYYY} \{(A_{1,0})^2 (A_{3,0})^2 n_1 - (A_{1,0})^3 A_{3,0}n_1^3\}, \quad (\text{B.4.17})$$

$$E_{9,1} = N_{YXX,YX} \left\{ (A_{3,0})^2 n_1^3 + (A_{1,0})^2 n_1^7 - 2A_{1,0}A_{3,0}n_1^5 \right\} + N_{YXY,YX} \left\{ A_{1,0} (A_{3,0})^2 n_1^2 + (A_{1,0})^3 n_1^6 - 2(A_{1,0})^2 A_{3,0}n_1^4 \right\}, \tag{B.4.18}$$

$$D_{9,1} = N_{YXXXXXX} \left\{ A_{3,0}n_1^6 - A_{1,0}n_1^8 \right\} + N_{YXXXXXY} \left\{ A_{3,0}A_{1,0}n_1^5 - (A_{1,0})^2 n_1^7 \right\} + N_{YXXXXYY} \left\{ A_{3,0} (A_{1,0})^2 n_1^4 - (A_{1,0})^3 n_1^6 \right\} + N_{YXXYYYY} \left\{ A_{3,0} (A_{1,0})^3 n_1^3 - (A_{1,0})^4 n_1^5 \right\} + N_{YXXYYYY} \left\{ A_{3,0} (A_{1,0})^4 n_1^2 - (A_{1,0})^5 n_1^4 \right\} + N_{YXYYYYY} \left\{ A_{3,0} (A_{1,0})^5 n_1 - (A_{1,0})^6 n_1^3 \right\}. \tag{B.4.19}$$

Подставим теперь $\Omega^{(1)}(\Theta_1, \Theta_3, \Theta_5, \Theta_7)$ и $\Omega_2(\Theta_1, \Theta_3, \Theta_5, \Theta_7)$ в формулу (31).

Первые два слагаемых, $2\Xi_X + \Xi_Y$, можно записать в виде

$$\Omega^{(1)} + 2\Omega_2 = \tau A_{1,1}\Theta_1 + \tau^3 A_{3,1}\Theta_3 + \tau^5 A_{5,1}\Theta_5 + \tau^7 A_{7,1}\Theta_7 + \tau^9 A_{9,1}\Theta_9 + 2\tau n_2\Theta_1 + 2\tau^3 n_2^3\Theta_3 + 2\tau^5 n_2^5\Theta_5 + 2\tau^7 n_2^7\Theta_7 + 2\tau^9 n_2^9\Theta_9. \tag{B.5.1}$$

Обозначим $G(\Theta_1, \Theta_3, \Theta_5, \Theta_7)$ как вклад, схожий по форме для предыдущего шага. Тогда выражения с новыми слагаемыми запишутся как

$$N_{YXX}\Xi_{\Omega^{(1)}\Omega_2\Omega_2} = G(\Theta_1, \Theta_3, \Theta_5, \Theta_7) + \tau^7 N_{YXX} B_{5,1} n_2^2 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1} + \tau^9 N_{YXX} B_{5,1} n_2^4 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_3} + \tau^9 N_{YXX} B_{5,1} n_2^4 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_3} + \tau^9 N_{YXX} B_{7,1} n_2^2 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1} + \tau^9 N_{YXX} B_{7,1} n_2^2 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1} + \tau^9 N_{YXX} C_{7,1} n_2^2 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1}, \tag{B.5.2}$$

$$N_{YXY}\Xi_{\Omega^{(1)}\Omega_2\Omega^{(1)}} = G(\Theta_1, \Theta_3, \Theta_5, \Theta_7) + \tau^7 N_{YXY} B_{5,1} A_{1,1} n_2 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1} + \tau^9 N_{YXY} B_{5,1} A_{1,1} n_2^3 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_3} + \tau^9 N_{YXY} B_{5,1} A_{3,1} n_2 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_3} + \tau^9 N_{YXY} B_{7,1} A_{1,1} n_2 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1} + \tau^9 N_{YXY} B_{7,1} A_{1,1} n_2 \Xi_{\Theta_3\Theta_1\Theta_3\Theta_1} + \tau^9 N_{YXY} C_{7,1} A_{1,1} n_2 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1} + N_{YXY} \tau^9 A_{1,1} n_2^3 B_{5,1} \Xi_{\Theta_1\Theta_3\Theta_3\Theta_1} + N_{YXY} \tau^9 A_{3,1} n_2 B_{5,1} \Xi_{\Theta_3\Theta_1\Theta_3\Theta_1}, \tag{B.5.3}$$

$$N_{YXXXX}\Xi_{\Omega^{(1)}\Omega_1\Omega_1\Omega_1\Omega_1} = G(\Theta_1, \Theta_3, \Theta_5) + \tau^9 N_{YXXXX} B_{5,1} n_2^4 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1}, \tag{B.5.4}$$

$$N_{YXXXXY}\Xi_{\Omega^{(1)}\Omega_1\Omega_1\Omega_1\Omega^{(1)}} = G(\Theta_1, \Theta_3, \Theta_5) + \tau^9 N_{YXXXXY} B_{5,1} A_{1,1} n_2^3 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1}, \tag{B.5.5}$$

$$N_{YXXYY}\Xi_{\Omega^{(1)}\Omega_1\Omega_1\Omega^{(1)}\Omega^{(1)}} = G(\Theta_1, \Theta_3, \Theta_5) + \tau^9 N_{YXXYY} B_{5,1} (A_{1,1})^2 n_2^2 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1}, \tag{B.5.6}$$

$$N_{YXYYY}\Xi_{\Omega^{(1)}\Omega_1\Omega^{(1)}\Omega^{(1)}\Omega^{(1)}} = G(\Theta_1, \Theta_3, \Theta_5) + \tau^9 N_{YXYYY} B_{5,1} (A_{1,1})^3 n_2 \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1}. \tag{B.5.7}$$

Подстановка функций более высокого порядка $\Omega^{(m)}$ не обнаружит новых слагаемых. Поэтому окончательный результат может быть записан следующей рекуррентной формулой:

$$\begin{aligned} \Omega^{(m)} = & \tau A_{1,m} \Xi_{\Theta_1} + \tau^3 A_{3,m} \Xi_{\Theta_3} + \tau^5 (A_{5,m} \Xi_{\Theta_5} + B_{5,m} \Xi_{\Theta_3\Theta_1\Theta_1}) \\ & + \tau^7 (A_{7,m} \Xi_{\Theta_7} + B_{7,m} \Xi_{\Theta_3\Theta_1\Theta_1} + B_{7,m}^3 \Xi_{\Theta_3\Theta_1\Theta_3} + C_{7,m} \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1}) \\ & + \tau^9 (A_{9,m} \Xi_{\Theta_9} + B_{9,m} \Xi_{\Theta_7\Theta_1\Theta_1} + B_{9,m}^1 \Xi_{\Theta_5\Theta_3\Theta_1} + B_{9,m}^3 \Xi_{\Theta_5\Theta_1\Theta_3} + B_{9,m}^5 \Xi_{\Theta_5\Theta_1\Theta_5} + C_{9,m} \Xi_{\Theta_5\Theta_1\Theta_1\Theta_1}) \\ & + C_{9,m}^{311} \Xi_{\Theta_3\Theta_1\Theta_3\Theta_1\Theta_1} + C_{9,m}^{131} \Xi_{\Theta_3\Theta_1\Theta_1\Theta_3\Theta_1} + C_{9,m}^{113} \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_3} + E_{9,m} \Xi_{\Theta_3\Theta_1\Theta_1\Theta_3\Theta_1} + D_{9,m} \Xi_{\Theta_3\Theta_1\Theta_1\Theta_1\Theta_1\Theta_1}. \end{aligned} \tag{B.6}$$

Здесь

$$A_{1,m} = A_{1,m-1} + 2n_m, \quad (\text{B.7.1})$$

$$A_{3,m} = A_{3,m-1} + 2n_m^3, \quad (\text{B.7.2})$$

$$A_{5,m} = A_{5,m-1} + 2n_m^5, \quad (\text{B.7.3})$$

$$A_{7,m} = A_{7,m-1} + 2n_m^7, \quad (\text{B.7.4})$$

$$A_{9,m} = A_{9,m-1} + 2n_m^9, \quad (\text{B.7.5})$$

$$B_{5,m} = B_{5,m-1} + N_{YXX} \left\{ A_{3,m-1} n_m^2 - A_{1,m-1} n_m^4 \right\} + N_{YXY} \left\{ A_{1,m-1} A_{3,m-1} n_m - (A_{1,m-1})^2 n_m^3 \right\}, \quad (\text{B.7.6})$$

$$B_{7,m} = B_{7,m-1} + N_{YXX} \left\{ A_{5,m-1} n_m^2 - A_{1,m-1} n_m^6 \right\} + N_{YXY} \left\{ A_{1,m-1} A_{5,m-1} n_m - (A_{1,m-1})^2 n_m^5 \right\}, \quad (\text{B.7.7})$$

$$B_{7,m}^3 = B_{7,m-1}^3 + N_{YXX} \left\{ A_{3,m-1} n_m^4 - A_{1,m-1} n_m^6 \right\} + N_{YXY} \left\{ (A_{3,m-1})^2 n_m - A_{1,m-1} A_{3,m-1} n_m^3 \right\}, \quad (\text{B.7.8})$$

$$B_{9,m} = B_{9,m-1} + N_{YXX} \left\{ A_{7,m-1} n_m^2 - A_{1,m-1} n_m^8 \right\} + N_{YXY} \left\{ A_{1,m-1} A_{7,m-1} n_m - (A_{1,m-1})^2 n_m^7 \right\}, \quad (\text{B.7.9})$$

$$B_{9,m}^1 = B_{9,m-1}^1 + N_{YXX} \left\{ A_{5,m-1} n_m^4 - A_{3,m-1} n_m^6 \right\} + N_{YXY} \left\{ A_{1,m-1} A_{5,m-1} n_m^3 - A_{1,m-1} A_{3,m-1} n_m^5 \right\}, \quad (\text{B.7.10})$$

$$B_{9,m}^3 = B_{9,m-1}^3 + N_{YXX} \left\{ A_{5,m-1} n_m^4 - A_{1,m-1} n_m^8 \right\} + N_{YXY} \left\{ A_{3,m-1} A_{5,m-1} n_m - A_{1,m-1} A_{3,m-1} n_m^5 \right\}, \quad (\text{B.7.11})$$

$$B_{9,m}^5 = B_{9,m-1}^5 + N_{YXX} \left\{ A_{3,m-1} n_m^6 - A_{1,m-1} n_m^8 \right\} + N_{YXY} \left\{ A_{3,m-1} A_{5,m-1} n_m - A_{1,m-1} A_{5,m-1} n_m^3 \right\}, \quad (\text{B.7.12})$$

$$\begin{aligned} C_{7,m} = C_{7,m-1} + N_{YXXXX} \left\{ A_{3,m-1} n_m^4 - A_{1,m-1} n_m^6 \right\} + N_{YXXXY} \left\{ A_{1,m-1} A_{3,m-1} n_m^3 - (A_{1,m-1})^2 n_m^5 \right\} \\ + N_{YXXYY} \left\{ A_{3,m-1} (A_{1,m-1})^2 n_m^2 - (A_{1,m-1})^3 n_m^4 \right\} + N_{YXYYY} \left\{ (A_{1,m-1})^3 A_{3,m-1} n_m - (A_{1,m-1})^4 n_m^3 \right\} \\ + N_{YXX} B_{5,m-1} n_m^2 + N_{YXY} B_{5,m-1} A_{1,m-1} n_m, \end{aligned} \quad (\text{B.7.13})$$

$$\begin{aligned} C_{9,m} = C_{9,m-1} + N_{YXXXX} \left\{ A_{5,m-1} n_m^4 - A_{1,m-1} n_m^8 \right\} + N_{YXXXY} \left\{ A_{1,m-1} A_{5,m-1} n_m^3 - (A_{1,m-1})^2 n_m^7 \right\} \\ + N_{YXXYY} \left\{ (A_{1,m-1})^2 A_{5,m-1} n_m^2 - (A_{1,m-1})^3 n_m^6 \right\} + N_{YXYYY} \left\{ (A_{1,m-1})^3 A_{5,m-1} n_m - (A_{1,m-1})^4 n_m^5 \right\} \\ + N_{YXX} B_{7,m-1} n_m^2 + N_{YXY} B_{7,m-1} A_{1,m-1} n_m, \end{aligned} \quad (\text{B.7.14})$$

$$\begin{aligned} C_{9,m}^{311} = C_{9,m-1}^{311} + N_{YXXXX} \left\{ A_{3,m-1} n_m^6 - A_{1,m-1} n_m^8 \right\} + N_{YXXXY} \left\{ A_{1,m-1} A_{3,m-1} n_m^5 - (A_{1,m-1})^2 n_m^7 \right\} \\ + N_{YXXYY} \left\{ A_{3,m-1} (A_{1,m-1})^2 n_m^4 - (A_{1,m-1})^3 n_m^6 \right\} + N_{YXYYY} \left\{ (A_{1,m-1})^2 (A_{3,m-1})^2 n_m - (A_{1,m-1})^3 A_{3,m-1} n_m^3 \right\} \\ + N_{YXX} B_{7,m-1}^3 n_m^2 + N_{YXY} B_{7,m-1}^3 A_{1,m-1} n_m, \end{aligned} \quad (\text{B.7.15})$$

$$\begin{aligned} C_{9,m}^{131} = C_{9,m-1}^{131} + N_{YXXXX} \left\{ A_{3,m-1} n_m^6 - A_{1,m-1} n_m^8 \right\} + N_{YXXXY} \left\{ A_{1,m-1} A_{3,m-1} n_m^5 - (A_{1,m-1})^2 n_m^7 \right\} \\ + N_{YXXYY} \left\{ A_{1,m-1} (A_{3,m-1})^2 n_m^2 - A_{3,m-1} (A_{1,m-1})^2 n_m^4 \right\} \\ + N_{YXYYY} \left\{ (A_{1,m-1})^2 (A_{3,m-1})^2 n_m - (A_{1,m-1})^3 A_{3,m-1} n_m^3 \right\} + N_{YXX} B_{5,m-1} n_m^4 + N_{YXY} B_{5,m-1} A_{1,m-1} n_m^3, \end{aligned} \quad (\text{B.7.16})$$

$$C_{9,m}^{113} = C_{9,m-1}^{113} + N_{YXXXXX} \left\{ A_{3,m-1} n_m^6 - A_{1,m-1} n_m^8 \right\} + N_{YXXXY} \left\{ (A_{3,m-1})^2 n_m^3 - A_{1,m-1} A_{3,m-1} n_m^5 \right\} \\ + N_{YXXYY} \left\{ A_{1,m-1} (A_{3,m-1})^2 n_m^2 - A_{3,m-1} (A_{1,m-1})^2 n_m^4 \right\} \quad (\text{B.7.17})$$

$$+ N_{YXYYY} \left\{ (A_{1,m-1})^2 (A_{3,m-1})^2 n_m - (A_{1,m-1})^3 A_{3,m-1} n_m^3 \right\} + N_{YXX} B_{5,m-1} n_m^4 + N_{YXY} B_{5,m-1} A_{3,m-1} n_m,$$

$$E_{9,m} = E_{9,m-1} + N_{YXX,YX} \left\{ (A_{3,m-1})^2 n_m^3 + (A_{1,m-1})^2 n_m^7 - 2A_{1,m-1} A_{3,m-1} n_m^5 \right\} \\ + N_{YXY,YX} \left\{ A_{1,m-1} (A_{3,m-1})^2 n_m^2 + (A_{1,m-1})^3 n_m^6 - 2(A_{1,m-1})^2 A_{3,m-1} n_m^4 \right\} \quad (\text{B.7.18}) \\ + N_{YXY} A_{1,m-1} B_{5,m-1} n_m^3 - N_{YXY} A_{3,m-1} B_{5,m-1} n_m,$$

$$D_{9,m} = D_{9,m-1} + N_{YXXXXXX} \left\{ A_{3,m-1} n_m^6 - A_{1,m-1} n_m^8 \right\} + N_{YXXXXXY} \left\{ A_{3,m-1} A_{1,m-1} n_m^5 - (A_{1,m-1})^2 n_m^7 \right\} \\ + N_{YXXXXYY} \left\{ A_{3,m-1} (A_{1,m-1})^2 n_m^4 - (A_{1,m-1})^3 n_m^6 \right\} + N_{YXXYYYY} \left\{ A_{3,m-1} (A_{1,m-1})^3 n_m^3 - (A_{1,m-1})^4 n_m^5 \right\} \\ + N_{YXXYYYY} \left\{ A_{3,m-1} (A_{1,m-1})^4 n_m^2 - (A_{1,m-1})^5 n_m^4 \right\} + N_{YXYYYYY} \left\{ A_{3,m-1} (A_{1,m-1})^5 n_m - (A_{1,m-1})^6 n_m^3 \right\} \quad (\text{B.7.19}) \\ + N_{YXX} C_{7,m-1} n_m^2 + N_{YXY} C_{7,m-1} A_{1,m-1} n_m + N_{YXXXX} B_{5,m-1} n_m^4 + N_{YXXY} B_{5,m-1} A_{1,m-1} n_m^3 \\ + N_{YXXYY} B_{5,m-1} (A_{1,m-1})^2 n_m^2 + N_{YXYYY} B_{5,m-1} (A_{1,m-1})^3 n_m,$$

ПРИЛОЖЕНИЕ С

Подставим $\Omega_0(\Theta_1, \tilde{\Theta}_5, \tilde{\Theta}_7, \tilde{\Theta}_9)$ и $\Omega_1(\Theta_1, \tilde{\Theta}_5, \tilde{\Theta}_7, \tilde{\Theta}_9)$ в формулу (31).

Введем следующие обозначения:

$$A_{1,0} = n_0; \quad A_{3,0} = n_0^3; \quad A_{5,0} = n_0^5; \quad A_{7,0} = n_0^7; \quad A_{9,0} = n_0^9. \quad (\text{C.1})$$

Первые два слагаемых, $2\Xi_X + \Xi_Y$, будут записаны как

$$\Omega_0 + 2\Omega_1 = \tau A_{1,0} \Theta_1 + \tau^3 A_{3,0} \tilde{\Theta}_3 + \tau^5 A_{5,0} \tilde{\Theta}_5 + \tau^7 A_{7,0} \tilde{\Theta}_7 + \tau^9 A_{9,0} \tilde{\Theta}_9 + \\ + 2\tau n_1 \Theta_1 + 2\tau^3 n_1^3 \tilde{\Theta}_3 + 2\tau^5 n_1^5 \tilde{\Theta}_5 + 2\tau^7 n_1^7 \tilde{\Theta}_7 + 2\tau^9 n_1^9 \tilde{\Theta}_9. \quad (\text{C.2})$$

Остальные слагаемые:

$$N_{YXX} \Xi_{\Omega_0 \Omega_1 \Omega_1} = N_{YXX} \left[\left[\tau A_{1,0} \Theta_1 + \tau^5 A_{5,0} \tilde{\Theta}_5 + \tau^7 A_{7,0} \tilde{\Theta}_7, \tau n_1 \Theta_1 + \tau^5 n_1^5 \tilde{\Theta}_5 + \tau^7 n_1^7 \tilde{\Theta}_7 \right], \tau n_1 \Theta_1 \right] = \\ = \tau^7 N_{YXX} \left\{ A_{1,0} n_1^6 \Xi_{\Theta_1 \Theta_5 \Theta_1} + A_{5,0} n_1^2 \Xi_{\Theta_5 \Theta_1 \Theta_1} \right\} + \tau^9 N_{YXX} \left(A_{1,0} n_1^8 \Xi_{\Theta_1 \Theta_7 \Theta_1} + A_{7,0} n_1^2 \Xi_{\Theta_7 \Theta_1 \Theta_1} \right), \quad (\text{C.3.1})$$

$$N_{YXY} \Xi_{\Omega_0 \Omega_1 \Omega_0} = N_{YXY} \left[\left[\tau A_{1,0} \Theta_1 + \tau^5 A_{5,0} \tilde{\Theta}_5 + \tau^7 A_{7,0} \tilde{\Theta}_7, \tau n_1 \Theta_1 + \tau^5 n_1^5 \tilde{\Theta}_5 + \tau^7 n_1^7 \tilde{\Theta}_7 \right], \tau A_{1,0} \Theta_1 \right] = \\ = \tau^7 N_{YXY} \left\{ (A_{1,0})^2 n_1^5 \Xi_{\Theta_1 \tilde{\Theta}_5 \Theta_1} + A_{1,0} A_{5,0} n_1 \Xi_{\tilde{\Theta}_5 \Theta_1 \Theta_1} \right\} + \tau^9 N_{YXY} \left\{ (A_{1,0})^2 n_1^7 \Xi_{\Theta_1 \tilde{\Theta}_7 \Theta_1} + A_{1,0} A_{7,0} n_1 \Xi_{\tilde{\Theta}_7 \Theta_1 \Theta_1} \right\}, \quad (\text{C.3.2})$$

$$N_{YXXX} \Xi_{\Omega_0 \Omega_1 \Omega_1 \Omega_1} = \tau^9 N_{YXXX} \left(A_{1,0} n_1^8 \Xi_{\Theta_1 \tilde{\Theta}_5 \Theta_1 \Theta_1} + A_{5,0} n_1^4 \Xi_{\tilde{\Theta}_5 \Theta_1 \Theta_1 \Theta_1} \right), \quad (\text{C.3.3})$$

$$N_{YXXXY} \Xi_{\Omega_0 \Omega_1 \Omega_1 \Omega_1 \Omega_0} = \tau^9 N_{YXXXY} \left((A_{1,0})^2 n_1^7 \Xi_{\Theta_1 \tilde{\Theta}_5 \Theta_1 \Theta_1} + A_{1,0} A_{5,0} n_1^3 \Xi_{\tilde{\Theta}_5 \Theta_1 \Theta_1 \Theta_1} \right), \quad (\text{C.3.4})$$

$$N_{YXXYY} \Xi_{\Omega_0 \Omega_1 \Omega_1 \Omega_0 \Omega_0} = \tau^9 N_{YXXYY} \left((A_{1,0})^3 n_1^6 \Xi_{\Theta_1 \tilde{\Theta}_5 \Theta_1 \Theta_1} + (A_{1,0})^2 A_{5,0} n_1^2 \Xi_{\tilde{\Theta}_5 \Theta_1 \Theta_1 \Theta_1} \right), \quad (\text{C.3.5})$$

$$N_{YXYXY} \Xi_{\Omega_0 \Omega_1 \Omega_0 \Omega_0 \Omega_0} = \tau^9 N_{YXYXY} \left((A_{1,0})^4 n_1^5 \Xi_{\Theta_1 \tilde{\Theta}_5 \Theta_1 \Theta_1 \Theta_1} + (A_{1,0})^3 A_{5,0} n_1 \Xi_{\tilde{\Theta}_5 \Theta_1 \Theta_1 \Theta_1} \right). \quad (C.3.6)$$

Итоговое выражение для $\Omega^{(1)}$:

$$\begin{aligned} \Omega^{(1)} = & \tau A_{1,1} \Xi_{\Theta_1} + \tau^5 A_{5,1} \Xi_{\tilde{\Theta}_5} + \tau^7 \left(A_{7,1} \Xi_{\tilde{\Theta}_7} + B_{7,1} \Xi_{\tilde{\Theta}_5 \Theta_1 \Theta_1} \right) \\ & + \tau^9 \left(A_{9,1} \Xi_{\tilde{\Theta}_9} + B_{9,1} \Xi_{\tilde{\Theta}_7 \Theta_1 \Theta_1} + C_{9,1} \Xi_{\tilde{\Theta}_5 \Theta_1 \Theta_1 \Theta_1} \right). \end{aligned} \quad (C.4)$$

Здесь

$$A_{1,1} = n_0 + 2n_1 \equiv A_{1,0} + 2n_1, \quad (C.4.1)$$

$$A_{3,1} = n_0^3 + 2n_1^3 \equiv A_{3,0} + 2n_1^3, \quad (C.4.2)$$

$$A_{5,1} = n_0^5 + 2n_1^5 \equiv A_{5,0} + 2n_1^5, \quad (C.4.3)$$

$$A_{7,1} = n_0^7 + 2n_1^7 \equiv A_{7,0} + 2n_1^7, \quad (C.4.4)$$

$$A_{9,1} = n_0^9 + 2n_1^9 \equiv A_{9,0} + 2n_1^9, \quad (C.4.5)$$

$$B_{7,1} = B_{7,0} + N_{YXY} \left\{ A_{1,0} A_{5,0} n_1 - (A_{1,0})^2 n_1^5 \right\} + N_{YXX} \left\{ A_{5,0} n_1^2 - A_{1,0} n_1^6 \right\}, \quad (C.4.6)$$

$$B_{9,1} = B_{9,0} + N_{YXX} \left\{ A_{7,0} n_1^2 - A_{1,0} n_1^8 \right\} + N_{YXY} \left\{ A_{1,0} A_{7,0} n_1 - (A_{1,0})^2 n_1^7 \right\}, \quad (C.4.7)$$

$$\begin{aligned} C_{9,1} = & C_{9,0} + N_{YXXX} \left\{ A_{5,0} n_1^4 - A_{1,0} n_1^8 \right\} + N_{YXXX} \left\{ A_{1,0} A_{5,0} n_1^3 - (A_{1,0})^2 n_1^7 \right\} \\ & + N_{YXYXY} \left\{ (A_{1,0})^2 A_{5,0} n_1^2 - (A_{1,0})^3 n_1^6 \right\} + N_{YXYXY} \left\{ (A_{1,0})^3 A_{5,0} n_1 - (A_{1,0})^4 n_1^5 \right\}, \end{aligned} \quad (C.4.8)$$

Вычисление выражения $\Omega^{(2)}$ показывает, что только одна функция, C_9 , будет иметь слагаемые, отличающиеся по форме от предыдущего шага.

В итоге

$$\begin{aligned} \Omega^{(m)} = & \tau A_{1,m} \Xi_{\Theta_1} + \tau^5 A_{5,m} \Xi_{\tilde{\Theta}_5} + \tau^7 \left(A_{7,m} \Xi_{\tilde{\Theta}_7} + B_{7,m} \Xi_{\tilde{\Theta}_5 \Theta_1 \Theta_1} \right) + \\ & + \tau^9 \left(A_{9,m} \Xi_{\tilde{\Theta}_9} + B_{9,m} \Xi_{\tilde{\Theta}_7 \Theta_1 \Theta_1} + C_{9,m} \Xi_{\tilde{\Theta}_5 \Theta_1 \Theta_1 \Theta_1} \right). \end{aligned} \quad (C.5)$$

Здесь

$$A_{1,m} = A_{1,m-1} + 2n_m, \quad (C.6.1)$$

$$A_{5,m} = A_{5,m-1} + 2n_m^5, \quad (C.6.2)$$

$$A_{7,m} = A_{7,m-1} + 2n_m^7, \quad (C.6.3)$$

$$A_{9,m} = A_{9,m-1} + 2n_m^9, \quad (C.6.4)$$

$$B_{7,m} = B_{7,m-1} + N_{YXY} \left\{ A_{1,m-1} A_{5,m-1} n_m - (A_{1,m-1})^2 n_m^5 \right\} + N_{YXX} \left\{ A_{5,m-1} n_m^2 - A_{1,m-1} n_m^6 \right\}, \quad (C.6.5)$$

$$B_{9,m} = B_{9,m-1} + N_{YXX} \left\{ A_{7,m-1} n_m^2 - A_{1,m-1} n_m^8 \right\} + N_{YXY} \left\{ A_{1,m-1} A_{7,m-1} n_m - (A_{1,m-1})^2 n_m^7 \right\}, \quad (C.6.6)$$

$$\begin{aligned}
C_{9,m} = & C_{9,m-1} + N_{YXXX} \left\{ A_{5,m-1} n_m^4 - A_{1,m-1} n_m^8 \right\} + N_{YXXY} \left\{ A_{1,m-1} A_{5,m-1} n_m^3 - (A_{1,m-1})^2 n_m^7 \right\} + \\
& + N_{YXXY} \left\{ (A_{1,m-1})^2 A_{5,m-1} n_m^2 - (A_{1,m-1})^3 n_m^6 \right\} + N_{YXYY} \left\{ (A_{1,m-1})^3 A_{5,m-1} n_m - (A_{1,m-1})^4 n_m^5 \right\} + \\
& + N_{YXX} B_{7,m-1} n_m^2 + N_{YXY} B_{7,m-1} A_{1,m-1} n_m.
\end{aligned} \tag{C.6.7}$$

СПИСОК ЛИТЕРАТУРЫ

1. *Marchuk G. I.* Partial Differential Equations: II SYNSPADE-1970. New York: Academic, 1971.
2. *Samarskii A. A.* Teoriya raznostnykh skhem (The Theory of Difference Schemes). Moscow: Nauka, 1977.
3. *Strang G., Fix G.* An Analysis of the Finite Element Method. Englewood Cliffs: Prentice-Hall, 1973.
4. *Bathe K. J.* Finite Element Procedures in Engineering Analysis. New York: Englewood Cliffs, Prentice Hall, 1982.
5. *Magnus W.* On the Exponential solution of differential equations for a linear operator // Commun. Pure Appl. Math. 1954. V. 7. P. 649.
6. *Wilcox R. M.* Exponential operators and parameter differentiation in quantum physics // J. Math. Phys. 1967. V. 8. P. 962.
7. *Blanes S., Casas F., Ros J.* Improved high order integrators based on the Magnus expansion // BIT Numer. Math. 2000. V. 40. P. 434.
8. *Chuluunbaatar O., Derbov V. L., Galtbayar A., Gusev A. A., Kaschiev M. S., Vinitzky S. I., Zhanlav T.* Explicit Magnus expansions for solving the time-dependent Schrödinger equation // J. Phys. A: Math. Theor. 2008. V. 41. P. 295203.
9. *Yoshida H.* Construction of higher order symplectic integrators // Phys. Lett. A. 1990. V. 150. P. 262.
10. *Suzuki M.* Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations // Phys. Lett. A. 1990. V. 146. № 6. P. 319.
11. *Chin S. A., Chen C. R.* Gradient symplectic algorithms for solving the Schrödinger equation with time-dependent potentials // J. Chem. Phys. 2002. V. 117. P. 1409.
12. *McLachlan R. I.* On the Numerical Integration of Ordinary Differential Equations by Symmetric Composition // SIAM J. Sci. Comput. 1995. V. 16. P. 1.
13. *McLachlan R. I.* Families of High-Order Composition Methods // Numer. Alg. 2002. V. 31. P. 233.
14. *Blanes S.* High order numerical integrators for differential equations using composition and processing of low order methods // Appl. Numer. Math. 2001. V. 37. P. 289.
15. *Blanes S., Casas F., Ros J.* Symplection integration with processing: A general study // SIAM J. Sci. Comput. 1999. V. 21. P. 711.
16. *Zakharov M. A., Frank A. I., Kulin G. V., Goryunov S. V.* Interaction of Ultracold Neutrons with a Neutron Interference Filter Oscillating in Space // J. Surf. Invest.: X-Ray, Synchrotron Neutron Tech. 2020. V. 14. P. 6.
17. *Zakharov M. A., Frank A. I., Kulin G. V.* Reflection of neutrons from a resonant potential structure oscillating in space // Phys. Lett. A. 2021. V. 420. P. 127748.
18. *Frigo M., Johnson S. G.* The Design and Implementation of FFTW3 // Proc. IEEE. 2005. V. 93. P. 216.
19. *Suzuki M.* General Decomposition Theory of Ordered Exponentials // Proc. Japan Acad. B. 1993. V. 69. P. 161.
20. *Trotter H.* On the product of semi-groups of operators // Proc. Am. Math. Soc. 1959. V. 10. P. 545.
21. *Feit M. D., Jr. Fleck J. A., Steiger A.* Solution of the Schrödinger equation by a spectral method // J. Comp. Phys. 1982. V. 47. P. 412.
22. *Wiebe N., Berry D., Hoyer P., Sanders B.* Higher order decompositions of ordered operator exponentials // J. Phys. A: Math. Theor. 2010. V. 43. P. 065203.
23. *Casas F., Murua A.* An efficient algorithm for computing the Baker–Campbell–Hausdorff series and some of its applications // J. Math. Phys. 2009. V. 50. P. 033513.
24. *Bakhvalov N. S.* Numerical methods, Analysis, Algebra, Ordinary Differential Equations. MIR Publ., 1977.
25. *Puzynin I. V., Selin A. V., Vinitzky S. I.* A high-order accuracy method for numerical solving of the time-dependent Schrödinger equation // Comput. Phys. Commun. 1999. V. 123. P. 1.

УДК 519.633.6+517.956.45

РЕШЕНИЕ ДВУМЕРНОГО НЕЛИНЕЙНОГО ПАРАБОЛИЧЕСКОГО УРАВНЕНИЯ ТЕПЛОПРОВОДНОСТИ ПРИ КРАЕВОМ РЕЖИМЕ, ЗАДАННОМ НА ПОДВИЖНОМ МНОГООБРАЗИИ

© 2024 г. А. Л. Казаков^{1,*}, О. А. Нефедова^{2,**}, Л. Ф. Слевак^{2,***}

¹664033 Иркутск, ул. Лермонтова, 134, ИДСТУ СО РАН, Россия

²620049 Екатеринбург, ул. Комсомольская, 34, ИМАШ УрО РАН, Россия

*e-mail: kazakov@icc.ru;

**e-mail: nefedova@imach.uran.ru;

***e-mail: lfs@imach.uran.ru

Поступила в редакцию 07.08.2023 г.

Переработанный вариант 20.09.2023 г.

Принята к публикации 14.10.2023 г.

Статья посвящена исследованию вырождающегося параболического уравнения теплопроводности с нелинейностями общего вида при наличии источника (стока) в случае двух пространственных переменных. Рассмотрена задача об иницировании тепловой волны, распространяющейся по холодному (нулевому) фону с конечной скоростью, краевым режимом, заданным на подвижном многообразии — замкнутой линии. Для нее доказана новая теорема существования и единственности и предложен численный алгоритм построения решения, основанный на методе граничных элементов, методе коллокаций и разностной аппроксимации по времени, при этом использована специальная замена переменных типа преобразования годографа. Найдены новые точные решения рассматриваемого уравнения в случае нелинейностей степенного вида. Численный алгоритм реализован в виде программы, проведен комплексный вычислительный эксперимент. Сравнение построенных численных решений с точными (как найденными в работе, так и ранее известными) показало хорошее соответствие, установлена численная сходимость относительно шага по времени и числа точек коллокации. Библ. 28. Фиг. 3. Табл. 4.

Ключевые слова: нелинейное параболическое уравнение теплопроводности, вырождение, теорема существования и единственности, точное решение, численное решение, метод граничных элементов, метод коллокаций, радиальные базисные функции.

DOI: 10.31857/S0044466924020081, EDN: YJLXMI

ВВЕДЕНИЕ

Для процессов высокотемпературной теплопроводности [1] и фильтрации жидкостей и газов в пористых средах [2] характерно распространение возмущений по покоящемуся фону с конечной скоростью, описать которые в рамках традиционных линейных параболических моделей [3], [4] не представляется возможным. В качестве альтернативы иногда используют гиперболические уравнения [5], [6], однако они, зачастую, не обеспечивают достаточной точности моделирования. Другим возможным способом решения указанной проблемы является применение нелинейных параболических уравнений с вырождением [4], которые обладают решениями с требуемыми свойствами. Впервые последние, по-видимому, были предложены в работе [7] в связи с математическим описанием процессов, происходящих после ядерного взрыва (сами авторы указывают, что “с такими задачами мы встречаемся в теории распространения пламени и теории звезд”). Почти одновременно аналогичные решения были получены при моделировании фильтрации газов в пористых пластах [8]. В дальнейшем в научной литературе появился термин “тепловая” [9] или “фильтрационная волна” [10], под которой понимается решение, состоящее из двух частей: положительной и нулевой, состыкованных между собой вдоль некоторой поверхности (линии), именуемой фронтом волны (нулевым фронтом). Причиной появления решений с подобными нелинейными свойствами у параболических уравнений является, по-видимому, вырождение последних на фронте волны, вследствие чего у них появляются некие “локально-гиперболические” свойства. Вырождающиеся параболические уравнения являются в последние годы предметом активного изучения, однако обычно речь идет о доказательстве существования решений и изучении их свойств в различных функциональных пространствах без построения решения в явном виде [11], [12], хотя встречаются и исключения [13].

Авторы данной работы исследуют решения нелинейного уравнения теплопроводности (the porous medium equation [14]), имеющие тип тепловой волны, уже более 10 лет [15]. Дать полное описание работ, опубликованных за эти годы, не представляется возможным, тем более, что в них в основном рассматривались уравнения с одной пространственной переменной. Отметим здесь наиболее релевантные статьи: [16], в которой впервые построено аналитическое решение задачи об иницировании тепловой волны для нелинейного уравнения теплопроводности (без источника) с двумя пространственными переменными в полярных координатах; [17], в которой были построены новые классы точных решений нелинейного уравнения теплопроводности с источником.

Выше обсуждались аналитические результаты, однако численному решению нелинейных параболических уравнений также посвящено большое количество работ. Дать их обзор в рамках настоящей статьи невозможно, отметим лишь то, что особое место здесь занимает метод граничных элементов (МГЭ) и его модификации, в частности, с применением метода двойственной взаимности (МДВ) [18]. В основе МДВ лежит использование разложений по радиальным базисным функциям (РБФ) [19], [20] при решении неоднородных, нелинейных и нестационарных уравнений. Альтернативой МДВ может быть применение метода коллокаций (МК) на основе РБФ [21]. В большинстве работ при этом по РБФ раскладывается искомая функция, что затруднительно использовать в случае вырождающегося нелинейного уравнения. В работах авторов РБФ применяются для аппроксимации неоднородностей уравнений. Численные алгоритмы решения задач для двумерного вырождающегося параболического уравнения были построены в работах [22]–[24]. Непосредственным предшественником представляемого исследования является работа [22], в которой были предложены новые граничноэлементные алгоритмы решения одно- и двумерной задач об иницировании тепловой волны краевым режимом, заданным на неподвижном многообразии. При этом рассматривалось нелинейное уравнение теплопроводности со степенной нелинейностью без источника.

В настоящей статье исследуется двумерное вырождающееся параболическое уравнение с источником при нелинейностях общего вида, для которого рассматривается задача об иницировании тепловой волны краевым режимом, заданным на подвижном многообразии. Принятые обобщения уравнения и краевого режима потребовали вывода новых дополнительных краевых условий на заданном многообразии, необходимых для корректного численного решения. Кроме этого, поскольку заданное многообразие лежит внутри области решения, на каждом шаге по времени в новом алгоритме (в отличие от построенных нами ранее) последовательно решаются две связанные задачи для областей по обе стороны от него. Также для рассмотренной задачи доказана оригинальная теорема существования и единственности кусочно-аналитического решения, и построено новое точное решение, которое используется в вычислительном эксперименте для верификации результатов расчетов.

1. ПОСТАНОВКА ЗАДАЧИ

Рассмотрим нелинейное параболическое уравнение второго порядка

$$T_t = \operatorname{div}(k(T) \nabla T) + h(T), \quad (1.1)$$

которое описывает процесс распространения тепла в случае общего вида нелинейной зависимости коэффициента теплопроводности от температуры при наличии распределенного источника (стока) тепла. Здесь T — температура, зависящая от времени t и пространственных координат; div и ∇ — операторы дивергенции и градиента по пространственным переменным; $k(T)$ — зависимость коэффициента теплопроводности от температуры, $k(0) = 0$; $h(T)$ — функция источника (стока), $h(0) = 0$.

Пусть функция $k(T)$ является монотонной, достаточно гладкой и обратная к ней функция имеет ненулевую производную. Тогда заменой $u = k(T)$ уравнение (1.1) приводится к виду

$$u_t = u \Delta u + F(u) (\nabla u)^2 + H(u). \quad (1.2)$$

Здесь Δ — оператор Лапласа;

$$F(u) = \frac{u \psi''(u)}{\psi'(u)} + 1, \quad H(u) = \frac{h(\psi(u))}{\psi'(u)},$$

где $T = \psi(u)$ — функция, обратная к $u = k(T)$. Будем считать, что $F(u) > 0$. Можно легко убедиться в том, что, если $\psi(0) = 0$ (это следует из равенства $k(0) = 0$), то $H(0) = 0$ и уравнение (1.2) имеет тривиальное решение $u \equiv 0$.

Рассмотрим уравнение (1.2) при краевом режиме

$$u|_{b(t,\mathbf{x})=0} = f(t, \mathbf{x}), \quad (1.3)$$

где \mathbf{x} — вектор декартовых координат точки, $f(0, \mathbf{x}) \equiv 0$, а уравнение $b(t, \mathbf{x}) = 0$ задает в каждый момент времени гладкую замкнутую гиперповерхность $S^{(t)}$, ограничивающую область $V^{(t)}$, содержащую начало координат. Предполагается, что если $t_1 < t_2$, то $V^{(t_1)} \subset V^{(t_2)}$.

Будем искать для (1.2), (1.3) нетривиальные решения, для которых справедливо условие

$$u|_{a(t,\mathbf{x})=0} = 0, \quad (1.4)$$

где неизвестное априори уравнение $a(t, \mathbf{x}) = 0$ задает в каждый момент времени нулевой фронт тепловой волны $C^{(t)}$ — замкнутую поверхность, ограничивающую область $W^{(t)}$, содержащую начало координат. При этом $a(0, \mathbf{x}) = b(0, \mathbf{x})$, т.е. $S^{(0)} = C^{(0)}$, и при $t > 0$ $V^{(t)} \subset W^{(t)}$. И тогда за пределами области $W^{(t)}$ можно будет заменить отрицательные значения u нулевыми (напомним, что функция $u \equiv 0$ удовлетворяет уравнению (1.2)), и, тем самым, получить искомую тепловую волну. Существование подобных решений обеспечивает теорема, которая будет сформулирована и доказана в следующем разделе.

2. ОСНОВНАЯ ТЕОРЕМА

Будем далее для простоты и наглядности предполагать, что $\mathbf{x} = (x, y)$. В полярной системе координат, если область $V^{(t)}$ является звездной относительно точки $O(0, 0)$, задача (1.2), (1.3) имеет следующий вид:

$$u_t = uu_{\rho\rho} + F(u)u_{\rho}^2 + \frac{uu_{\rho}}{\rho} + \frac{1}{\rho^2} [F(u)u_{\phi}^2 + uu_{\phi\phi}] + H(u), \quad (2.1)$$

$$u|_{\rho=g(t,\phi)} = f(t, \phi), \quad (2.2)$$

где $t \geq 0$, $-\pi \leq \phi \leq \pi$, уравнение $\rho = g(t, \phi)$ задает линию $S^{(t)}$; для удобства обозначений мы сохранили обозначение f для функции в правой части краевого условия.

Здесь и далее будем считать, что $f(t, \pi) = f(t, -\pi)$, $g(t, \pi) = g(t, -\pi)$, $f_{\phi}(t, \pi) = f_{\phi}(t, -\pi)$, $g_{\phi}(t, \pi) = g_{\phi}(t, -\pi)$ при всех $t \geq 0$; $f(0, \phi) = 0$, $f_t(0, \phi) > 0$, $g(0, \phi) > 0$ при всех $-\pi \leq \phi \leq \pi$; $F(0) > 0$, $H(0) = 0$.

Под “аналитической” далее понимается функция, совпадающая на некотором множестве со своим тейлоровским разложением.

Теорема 1. Пусть функции $f(t, \phi)$, $g(t, \phi)$ являются аналитическими в некоторой окрестности начального момента времени $t = 0$ при $-\pi \leq \phi \leq \pi$, функции $F(u)$, $H(u)$ являются аналитическими в окрестности $u = 0$. Тогда задача (2.1), (2.2) имеет единственное (с точностью до выбора направления движения тепловой волны) аналитическое решение.

Доказательство. Ранее авторами была сформулирована и подробно доказана теорема существования и единственности аналитического решения задачи (2.1), (2.2) в случае, когда $F(u) = 1/\sigma > 0$, $H(u) = 0$, что соответствует степенной зависимости коэффициента теплопроводности от температуры (наиболее распространенной в приложениях) при отсутствии источников (стоков) тепла [16]. Как показывают ранее выполненные авторами исследования подобных задач для одномерного случая (см., например, [17]), обобщение теоремы существования и единственности со случая степенной нелинейности на случай нелинейности более общего вида с добавлением источника осуществляется достаточно естественно, поэтому ограничимся кратким обоснованием сформулированного выше утверждения. Доказательство разбивается на два этапа: вначале строится формальное решение в виде кратного степенного ряда, затем доказывается его сходимость методом мажорант.

Этап 1. Прежде всего, необходимо сделать замену независимой переменной в уравнении (2.1), $r = \rho - g(t, \phi)$, в результате которой оно примет вид

$$u_t - g_t u_r = \left(1 + \frac{g_{\phi}^2}{(g+r)^2}\right) uu_{rr} + F(u) \left(1 + \frac{g_{\phi}^2}{(g+r)^2}\right) u_r^2 + \left(\frac{1}{g+r} - \frac{g_{\phi\phi}}{(g+r)^2}\right) uu_r + \frac{uu_{\phi\phi}}{(g+r)^2} - \frac{2g_{\phi}uu_r}{(g+r)^2} - \frac{2F(u)g_{\phi}u_r u_{\phi}}{(g+r)^2} + \frac{F(u)u_{\phi}^2}{(g+r)^2} + H(u). \quad (2.3)$$

Граничное условие (2.2) тогда преобразуется к виду

$$u|_{r=0} = f(t, \varphi). \tag{2.4}$$

Задачи (2.1), (2.2) и (2.3), (2.4) эквивалентны вследствие очевидной невырожденности сделанной замены.

Решение задачи (2.3), (2.4) будем строить в виде кратного степенного ряда

$$u = \sum_{m,l=0}^{\infty} u_{m,l}(\varphi) \frac{t^m r^l}{m!l!} = \sum_{m,l=0}^{\infty} u_{m,l}(\varphi) \frac{t^m [\rho - g(t, \varphi)]^l}{m!l!},$$

$$u_{m,l} = \left. \frac{\partial^{m+l} u}{\partial t^m \partial r^l} \right|_{t=0, r=0}. \tag{2.5}$$

При этом из условия (2.4) имеем, что коэффициенты $u_{0,0} = f(0,0) = 0$, $u_{m,0} = f_m = f^{(m)}(0, \varphi)$, $m = 1, 2, \dots$, известны. Для нахождения $u_{0,1}$ получаем квадратное уравнение

$$F_0 \left(1 + \frac{g_{0,1}^2}{g_{0,0}^2} \right) u_{0,1}^2 + g_{1,0} u_{0,1} - f_1 = 0, \tag{2.6}$$

где $F_0 = F(0)$, $g_{0,0} = g(0, \varphi)$, $g_{1,0} = g_t(0, \varphi)$, $g_{0,1} = g_\phi(0, \varphi)$. Уравнение (2.6) при $F_0 f_1 > 0$ имеет два действительных ненулевых корня (выполнение неравенств $f_t(0, \varphi) > 0$, $g(0, \varphi) > 0$, $F(0) > 0$ оговорено ранее). Выбор одного из них и определяет направление движения тепловой волны, упомянутое в условии теоремы. Отметим, что $u_{0,1} \neq 0$, $u_{1,0} \neq 0$.

Дальнейшее построение производится индукцией по суммарному порядку дифференцирования. Пусть известны коэффициенты $u_{m,l}(\varphi)$ при $n \geq m + l$. На $n + 1$ -м шаге решается система линейных алгебраических уравнений (СЛАУ)

$$u_{k+1,n-k} - a_{n-k} u_{k,n-k+1} - b_k u_{k-1,n-k+2} = p_{k,n-k}, \quad k = 0, 1, \dots, n, \tag{2.7}$$

где $a_m(\varphi) = g_{1,0} + (m + 2F_0) \left(1 + \frac{g_{0,1}^2}{g_{0,0}^2} \right) u_{0,1}$, $b_l(\varphi) = l \left(1 + \frac{g_{0,1}^2}{g_{0,0}^2} \right) f_1$, а значения $p_{k,n-k}$ известны в силу предположения индукции.

СЛАУ (2.7) имеет квадратную матрицу (напомним, что $u_{n+1,0} = f_{n+1}$ — известно; считаем, что $u_{-1,n+2} = 0$), которая, начиная с $n = 2$, является трехдиагональной. Хотя условие диагонального преобладания в данном случае, очевидно, не выполняется, используя ранее доказанную лемму [16], можно показать, что СЛАУ (2.7) — крамеровская (при всех $-\pi \leq \phi \leq \pi$). Таким образом, все коэффициенты ряда (2.5) определяются однозначно. Первый этап доказательства теоремы завершен.

Этап 2. При изложении доказательства сходимости построенного ряда также будем кратки. В результате серии невырожденных замен, ключевая из которых состоит в том, что в качестве новой искомой функции рассматривается $w = r + g(t, \varphi) - a(t, \varphi)$ — отклонение от фронта (пока неизвестного) тепловой волны, а u становится независимой переменной, задача (2.3), (2.4) преобразуется к эквивалентному виду

$$2[1 + F(u)]W + [4 + F(u)]uW_u + u^2W_{uu} +$$

$$+ \Phi_0(\tau, \varphi)(W|_{\tau=0}) + \Phi_1(\tau, \varphi)u(W_u|_{\tau=0}) +$$

$$+ \Phi_2(\tau, \varphi)\left(u^2W_{uu}|_{\tau=0}\right) =$$

$$= \Psi_0(\tau, \varphi) + u\Psi_1 + u^2\Psi_2 + u^3\Psi_3. \tag{2.8}$$

Здесь $\tau = u - f$, $w = uw_1(\tau, \varphi) + u^2W$; w_1 , Φ_0 , Φ_1 , Φ_2 , Ψ_0 , Ψ_1 , Ψ_2 , Ψ_3 — известные аналитические функции своих аргументов, формулы для которых не приводятся из-за громоздкости. При этом Ψ_i , $i = 1, 2, 3$, зависят от производных искомой функции W по переменной u порядка не выше $i - 1$; по $u_{0,1}(\tau, \varphi)$ однозначно определяется $w_1(\tau, \varphi)$. Условия Коши для уравнения (2.8) при $u = 0$ получаются из требования его совместности (см. [16]) и имеют вид

$$W|_{u=0} = \frac{\Psi_0(\tau, \varphi) - w_0(\varphi)\Phi_0(\tau, \varphi)}{2[1 + F(u)]}, \quad W_u|_{u=0} = \frac{\Psi_1(\tau, \varphi) - w_1(\varphi)\Phi_1(\tau, \varphi)}{4 + F(u)}, \quad (2.9)$$

где $w_0(\varphi)$, $w_1(\varphi)$ — известные функции. Условия (2.9) не нуждаются в специальном обсуждении, поскольку условие (2.2) уже было учтено в ходе выполненных замен.

Уравнение (2.8) подпадает под действие леммы 2 из работы [16], которая обеспечивает существование и единственность у него аналитического решения в некоторой окрестности координатной плоскости $u = 0$. Это (по построению уравнения (2.8)), во-первых, гарантирует сходимость ряда (2.5), во-вторых, означает существование фронта тепловой волны (это прямо следует из того, что $u_{0,0} = 0$, $u_{1,0} > 0$, $u_{0,1} \neq 0$), т.е. подвижной линии $\rho = a(t, \varphi)$ такой, что выполняется условие

$$u|_{\rho=a(t, \varphi)} = 0.$$

Доказательство теоремы завершено.

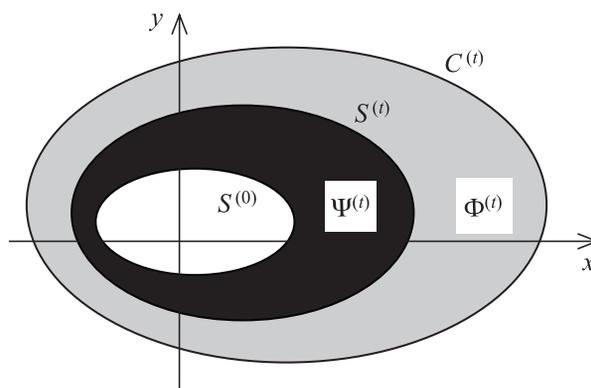
3. АЛГОРИТМ ЧИСЛЕННОГО РЕШЕНИЯ

Построение численного решения задачи (1.2), (1.3) на основе МГЭ в случае степенной нелинейности, когда $H(u) \equiv 0$ и функция $b(t, \mathbf{x})$ не зависит от времени, рассмотрено авторами в работе [22]. Также ранее были построены численные алгоритмы решения задачи (1.2), (1.4), см. [23], [24]. Решение задачи (1.2), (1.3) в двумерном случае ранее нами не рассматривалось. В доступных источниках таких решений также не обнаружено. Принципиальное отличие рассматриваемой в настоящей статье задачи, кроме произвольного вида нелинейности уравнения, заключается в том, что с течением времени две границы области определения искомой функции движутся, тогда как в рассмотренных ранее задачах одна из границ была неподвижной.

Далее в этом разделе предложен алгоритм решения задачи (1.2), (1.3). При этом используется постановка в полярной системе координат (2.1), (2.2). Пошаговое решение строится итерационно с использованием метода граничных элементов, радиальных базисных функций и метода коллокаций.

Задача состоит в определении функции u в каждый момент времени $t \in (0, T]$ в области $\Omega^{(t)}$, ограниченной линией $S^{(0)}$ и неизвестным фронтом $C^{(t)}$, который имеет уравнение $\rho = a(t, \varphi)$ (фиг. 1). Таким образом, область решения задачи нам неизвестна. Кроме этого, следует отметить, что краевой режим (2.2), как и (1.3), не является в классическом смысле краевым условием, поскольку он определяет значение функции не на границе области рассмотрения, а на некотором внутреннем многообразии. В связи с этим, будем строить решение в два этапа (последовательно с двух сторон от заданного многообразия), используя на первом из них специальную замену переменных.

Этап 1. Рассмотрим сначала задачу (2.1), (2.2) в произвольный момент времени t в области $\Phi^{(t)}$, ограниченной линиями $S^{(t)}$ и $C^{(t)}$ (см. фиг. 1). Будем предполагать при этом, что при произвольных значениях t , φ функция u в этой области монотонно зависит от ρ . Поменяем ролями в задаче (2.1), (2.2) искомую функцию u и координату ρ . Задача (2.1), (2.2) в новых обозначениях может быть записана в виде



Фиг. 1. Область $\Omega^{(t)} = \Phi^{(t)} \cup \Psi^{(t)}$

$$\begin{aligned} \rho_t \rho_u^2 = u \left(\rho_{uu} - \frac{\rho_u^2}{\rho} - \frac{\rho_u (\rho_{u\varphi} \rho_\varphi - \rho_{\varphi\varphi} \rho_u)}{\rho^2} \right) - \\ - \frac{F(u) \rho_u (\rho^2 + \rho_\varphi^2)}{\rho^2} - H(u) \rho_u^3, \quad \rho|_{u=f(t,\varphi)} = g(t,\varphi). \end{aligned} \quad (3.1)$$

Вдоль фронта $C^{(t)}$, где $u = 0$, из (3.1) следует соотношение

$$\rho_u|_{u=0} = - \frac{F(0)(\rho^2 + \rho_\varphi^2)}{\rho_t \rho^2}. \quad (3.2)$$

Замечание 1. Уравнение (3.2) справедливо при условии ограниченности множителя при u в правой части (3.1). Последнее условие, как и монотонность функции u , справедливо вблизи носителя краевого условия (локально по t, r), что следует из доказательства теоремы 1, в частности, из того, что $u_{0,1} \neq 0$. Априори обеспечить выполнение этих предположений в заданной области затруднительно. Отметим, однако, что аналитически доказать сходимость какого-либо приближенного метода решения применительно к рассматриваемому нелинейному вырождающемуся уравнению в настоящее время не представляется возможным, по крайней мере, нам подобные результаты не известны. В связи с этим, описанный ниже алгоритм является эвристическим, и вопрос о корректности численного решения, включая выдвинутые предположения, решается эмпирически. Во всех рассмотренных ниже примерах монотонность u следует, в частности, из возрастания по t функций $f(t, \varphi)$ и $g(t, \varphi)$, задающих краевое условие при $\rho = R$, и наблюдается ограниченность входящих в (3.1) производных в рассмотренных областях.

Сделаем далее замену независимой переменной $v = u + 1$, и будем рассматривать переменные v, φ как полярные координаты в плоскости декартовых координат ξ, η : $\xi = v \cos \varphi, \eta = v \sin \varphi$. В результате, выделив в (3.1) лапласиан искомой функции, получим следующую краевую задачу:

$$\Delta \rho = G(v, \rho), \quad (3.3)$$

$$\rho|_{v=1+f(t,\varphi)} = g(t, \varphi). \quad (3.4)$$

Здесь

$$\begin{aligned} G(v, \rho) = \frac{\rho_v}{v-1} \left(\rho_t \rho_v + \frac{F(v-1)(\rho^2 + \rho_\varphi^2)}{\rho^2} + H(v-1) \rho_v^2 \right) + \\ + \frac{\rho_v}{v} + \frac{\rho_{\varphi\varphi}}{v^2} + \frac{\rho_v^2}{\rho} + \frac{\rho_v (\rho_{v\varphi} \rho_\varphi - \rho_{\varphi\varphi} \rho_v)}{\rho^2}. \end{aligned} \quad (3.5)$$

Уравнение $v = 1 + f(t, \varphi)$ задает в момент времени t замкнутую линию $P^{(t)}$, ограничивающую область $U^{(t)}$, содержащую начало координат, при этом $P^{(0)}$ — окружность единичного радиуса с центром в начале координат. Решению задачи (2.1), (2.2) на этапе 1 в новых переменных соответствует задача (3.3), (3.4), состоящая в определении функции $\rho = \rho(t, v, \varphi)$ в области $t \in (0, T]$, $(v, \varphi) \in Y^{(t)}$, где $Y^{(t)}$ — известная в каждый момент область, ограниченная линиями $P^{(0)}$ и $P^{(t)}$. Отметим, что корректное задание области $Y^{(t)}$ стало возможным благодаря переходу к переменной v , поскольку при аналогичных рассуждениях для переменной u линия $P^{(0)}$ вырождается в точку, приводя к неопределенности.

Поскольку, как уже отмечено, $P^{(0)}$ — окружность, из (3.2) следует условие для потока вдоль этой границы:

$$q(\rho)|_{v=1} = \frac{\partial \rho}{\partial \mathbf{n}}|_{v=1} = \frac{F(0)(\rho^2 + \rho_\varphi^2)}{\rho_t \rho^2}, \quad (3.6)$$

где \mathbf{n} — вектор внешней нормали к границе области $Y^{(t)}$.

Итак, в каждый момент времени мы имеем краевую задачу (3.3), (3.4), (3.6). Решение будем строить по шагам по времени. На шаге $t = t_k = kh$, h — размер шага, представим его в виде

$$\rho(t_k, \nu, \Phi) = \mu(\nu, \Phi) + \lambda(\nu, \Phi),$$

где $\mu(\nu, \Phi)$ — частное решение уравнения (3.3) в момент t_k , $\lambda(\nu, \Phi)$ — решение соответствующей однородной задачи

$$\Delta\lambda = 0,$$

$$\lambda|_{\nu=1+f(t_k, \Phi)} = g(t_k, \Phi) - \mu(1+f(t_k, \Phi), \Phi),$$

$$\widehat{q}\lambda|_{\nu=1} = \frac{\partial\lambda}{\partial\mathbf{n}}|_{\nu=1} = \frac{F(0)(\rho^2 + \rho_\Phi^2)}{\rho_t \rho^2} - \frac{\partial\mu}{\partial\mathbf{n}}|_{\nu=1}.$$

Решение строится итерационно, по следующей процедуре:

$$\mu^{(0)} \equiv 0, \quad (3.7)$$

$$\Delta\lambda^{(n)} = 0, \quad (3.8)$$

$$\lambda^{(n)}|_{\nu=1+f(t_k, \Phi)} = g(t_k, \Phi) - \mu^{(n)}(1+f(t_k, \Phi), \Phi), \quad (3.9)$$

$$\widehat{q}\lambda^{(n)}|_{\nu=1} = \frac{F(0)\left(\left(\rho^{(n)}\right)^2 + \left(\rho_\Phi^{(n)}\right)^2\right)}{\rho_t^{(n)}\left(\rho^{(n)}\right)^2} - \frac{\partial\mu^{(n)}}{\partial\mathbf{n}}|_{\nu=1}, \quad (3.10)$$

$$\Delta\mu^{(n+1)} = G\left(\nu, \rho^{(n)}\right). \quad (3.11)$$

Здесь $\mu^{(n)}$, $\lambda^{(n)}$, $\rho^{(n)} = \mu^{(n)} + \lambda^{(n)}$ суть n -е итерации решений.

Задача (3.8)–(3.10) на каждой итерации решается с помощью МГЭ. Для этого граница области $Y^{(t_k)}$ разбивается на $2N$ прямолинейных граничных элементов: $e_{1k}, e_{2k}, \dots, e_{Nk}$ — на границе $P^{(0)}$, $e_{(N+1)k}, e_{(N+2)k}, \dots, e_{(2N)k}$ — на границе $P^{(t_k)}$. Равное количество элементов на границах $P^{(0)}$ и $P^{(t_k)}$ необходимо для однозначной разрешимости задачи. Будем использовать постоянную аппроксимацию искомой функции и потока на элементе. Тогда решение задачи методом граничных элементов имеет следующий вид [25]:

$$\lambda^{(n)}(\xi) = \sum_{i=1}^{2N} \left(q_{ik}^{(\lambda)(n)} \int_{e_{ik}} u^*(\xi, \mathbf{z}) dS(\mathbf{z}) - \lambda_{ik}^{(n)} \int_{e_{ik}} q^*(\xi, \mathbf{z}) dS(\mathbf{z}) \right), \quad (3.12)$$

где $\xi(\nu, \Phi)$ — внутренняя точка области $Y^{(t_k)}$, $\lambda_{ik}^{(n)}$ и $q_{ik}^{(\lambda)(n)}$ — итерации значений искомой функции и потока в момент t_k на элементе e_{ik} , $u^*(\xi, \mathbf{z})$ — фундаментальное решение для двумерной задачи теории потенциала, $q^*(\xi, \mathbf{z}) = \frac{\partial u^*(\xi, \mathbf{z})}{\partial\mathbf{n}}$. В узлах граничных элементов \mathbf{z}_{ik} , расположенных в средних точках элементов, справедливо

$$\frac{1}{2}\lambda_{ik}^{(n)} = \sum_{j=1}^{2N} \left(q_{jk}^{(\lambda)(n)} \int_{e_{jk}} u^*(\mathbf{z}_{ik}, \mathbf{z}) dS(\mathbf{z}) - \lambda_{jk}^{(n)} \int_{e_{jk}} q^*(\mathbf{z}_{ik}, \mathbf{z}) dS(\mathbf{z}) \right), \quad i = 1, 2, \dots, 2N. \quad (3.13)$$

В соотношениях (3.12) и (3.13) значения $\lambda_{ik}^{(n)}$, $i = N + 1, \dots, 2N$, заданы граничным условием (3.9). Значения $\lambda_{ik}^{(n)}$, $i = 1, \dots, N$, и $q_{ik}^{(\lambda)(n)}$, $i = 1, \dots, 2N$, составляют $3N$ неизвестных, которые должны быть определены при решении задачи. Для этого к $2N$ граничным интегральным уравнениям (3.13) добавим N уравнений, полученных с помощью разностного представления производных в уравнении (3.10)

$$\begin{aligned}
 q_{1k}^{(\lambda)(n)} &= \frac{F(0)h \left[4h_\phi^2 \left(\lambda_{1k}^{(n)} \right)^2 + \left(\lambda_{2k}^{(n)} - \lambda_{Nk}^{(n)} \right)^2 \right]}{4h_\phi^2 \left(\lambda_{1k}^{(n)} \right)^2 \left(\lambda_{1k}^{(n)} - \lambda_{1(k-1)}^{(n)} \right)}, \\
 q_{ik}^{(\lambda)(n)} &= \frac{F(0)h \left[4h_\phi^2 \left(\lambda_{ik}^{(n)} \right)^2 + \left(\lambda_{(i+1)k}^{(n)} - \lambda_{(i-1)k}^{(n)} \right)^2 \right]}{4h_\phi^2 \left(\lambda_{ik}^{(n)} \right)^2 \left(\lambda_{ik}^{(n)} - \lambda_{i(k-1)}^{(n)} \right)}, \quad i = 2, \dots, N - 1, \\
 q_{Nk}^{(\lambda)(n)} &= \frac{F(0)h \left[4h_\phi^2 \left(\lambda_{Nk}^{(n)} \right)^2 + \left(\lambda_{1k}^{(n)} - \lambda_{(N-1)k}^{(n)} \right)^2 \right]}{4h_\phi^2 \left(\lambda_{Nk}^{(n)} \right)^2 \left(\lambda_{Nk}^{(n)} - \lambda_{N(k-1)}^{(n)} \right)}.
 \end{aligned} \tag{3.14}$$

Здесь $h_\phi = 2\pi/N$ — шаг по переменной ϕ . Уравнения (3.13), (3.14) образуют систему $3N$ уравнений для $3N$ неизвестных.

Поскольку эта система нелинейная, она решается методом Ньютона, и большое значение имеет выбор начального приближения. Выбор случайных, например, нулевых, начальных приближений не позволил гарантированно решить полученную систему. Если в качестве начальных приближений (на нулевой итерации) принять значения на предыдущем шаге по времени (при $k > 1$), то наблюдается стабильная сходимость. На первом шаге ($k = 1$) начальные значения искомой функции могут быть взяты из условия (3.4) при $t = 0$, тогда как начальные значения потока необходимо найти дополнительно.

Возьмем полную производную по времени в граничном условии (3.4):

$$\left(\rho_t + \rho_v v_t + \rho_\phi \Phi_t \right) \Big|_{v=1+f(t,\phi)} = \left(\rho_t + \rho_v f_t \right) \Big|_{v=1+f(t,\phi)} = g_t.$$

Отсюда $\rho_t = g_t - \rho_v f_t$ при $v = 1 + f(t, \phi)$. Подставляя полученное в уравнение (3.1), при $t = 0$, и, соответственно, при $v = 1$, имеем

$$\rho_v^2 f_t - \rho_v g_t - \frac{F(0)(\rho^2 + \rho_\phi^2)}{\rho^2} = 0.$$

Из принятых ранее предположений (если $t_1 < t_2$, то $V^{(t_1)} \subset V^{(t_2)}$) движение тепловой волны направлено от начала координат, тогда производная ρ_v должна быть отрицательной. Отсюда

$$\rho_v(0,1,\phi) = \frac{g_t(0,\phi) - \sqrt{\left(g_t(0,\phi) \right)^4 + 4F(0)f_t(0,\phi) \left[\left(g(0,\phi) \right)^2 + \left(g_\phi(0,\phi) \right)^2 \right]}}{2f_t(0,\phi)g(0,\phi)}. \tag{3.15}$$

Поскольку $P^{(0)}$ — окружность с центром в начале координат, то в качестве начального приближения на первом шаге по времени можно принять значения потока $q^{(\lambda)}(\phi) = -\rho_v(0,1,\phi)$ на границе $P^{(0)}$ и $q^{(\lambda)}(\phi) = \rho_v(0,1,\phi)$ на границе $P^{(t_1)}$. Использование полученного соотношения (3.15) обеспечивает сходимость метода Ньютона и позволяет раскрыть неопределенность в начальный момент времени.

Найденные из решения системы значения $\lambda_{ik}^{(n)}$, $i = 1, \dots, N$, и $q_{ik}^{(\lambda)(n)}$, $i = 1, \dots, 2N$, определяют решение (3.12) задачи (3.8)–(3.10). Интегралы по граничным элементам от функций $u^*(\xi, \mathbf{z})$ и $q^*(\xi, \mathbf{z})$ в соотношениях (3.12) и (3.13) вычисляются точно с помощью аналитических формул [26, 27]. Отметим, что при найденном частном решении $\mu^{(n)}$ функция $\lambda^{(n)}$ определяется однозначно.

Для решения уравнения (3.11) на каждой итерации применим метод коллокаций [21]. Разложим правую часть уравнения по некоторой системе РБФ

$$G(v, \rho^{(n)}) = \sum_{i=1}^M \alpha_i^{(n)} f_i(\xi), \quad (3.16)$$

где $f_i(\xi) = f_i(|\xi - \xi_i|)$ — РБФ, значения которых зависят от расстояния между текущей точкой и заданными в области $Y^{(t_k)}$ точками коллокации $\xi_1, \xi_2, \dots, \xi_M$; для каждой функции f_i существует такая функция \hat{u}_i , что $f_i = \Delta \hat{u}_i$. Коэффициенты $\alpha_i^{(n)}$ определяются из решения системы линейных уравнений

$$G(v, \rho^{(n)}) \Big|_{\xi=\xi_j} = \sum_{i=1}^M \alpha_i^{(n)} f_i(\xi_j), \quad j = 1, \dots, M. \quad (3.17)$$

Тогда решение уравнения (3.11) имеет вид

$$\mu^{(n+1)}(\xi) = \sum_{i=1}^M \alpha_i^{(n)} \hat{u}_i(\xi). \quad (3.18)$$

Производная по времени в (3.17) (см. (3.5)) вычисляется разностно, с использованием решения на предыдущем шаге

$$\rho_i^{(n)}(\xi_j) = \begin{cases} \frac{\rho^{(n)}(\xi_j) - \rho(t_{k-1}, \xi_j)}{h}, & \xi_i \in Y^{(t_{k-1})}; \\ \frac{\rho^{(n)}(\xi_j)}{h_{k,j}}, & \xi_i \notin Y^{(t_{k-1})}. \end{cases}$$

Здесь $h_{k,j} = t_k - \tau_j$, τ_j — момент времени, в который точка ξ_j принадлежит границе $P^{(t_k)}$.

Итерационный процесс (3.8)–(3.18) заканчивается на n -й итерации при ее достаточной близости к $(n-1)$ -й. Тогда в качестве решения задачи (3.3), (3.4), (3.6) в момент t_k принимается непрерывная по пространственным переменным функция

$$\rho(t_k, v, \varphi) = \mu^{(n)}(v, \varphi) + \lambda^{(n)}(v, \varphi).$$

При этом уравнение $\rho = \rho(t_k, 1, \varphi)$ определит нулевой фронт тепловой волны $C^{(t_k)}$. Решение исходной задачи (2.1), (2.2) в области $\Phi^{(t_k)}$ в момент $t = t_k$, $u = u_1(t_k, \rho, \varphi)$, получается из $\rho(t_k, v, \varphi)$ без потери точности. Отметим, что применение в итерационной процедуре (3.8)–(3.18) для решения неоднородной задачи метода коллокаций приводит к повышению точности расчетов и скорости сходимости (уменьшению числа итераций) по сравнению с методом двойственной взаимности, используемым в [22]. Это является следствием того, что применение МГЭ с МДВ приводит на каждой итерации к решению системы, состоящей из линейных и нелинейных алгебраических уравнений, а при использовании МК их основная линейная часть вынесена в отдельную систему линейных уравнений, что упрощает задачу.

Этап 2. Напомним, что конечной целью является построение решения задачи (2.1), (2.2) в каждый момент времени $t \in (0, T]$ в области $\Omega^{(t)}$, ограниченной линией $S^{(t)}$ и нулевым фронтом $C^{(t)}$. При этом $\Omega^{(t)} = \Phi^{(t)} \cup \Psi^{(t)}$; область $\Psi^{(t)}$ ограничена линиями $S^{(0)}$ и $S^{(t)}$; $\Phi^{(t)}$ и $\Psi^{(t)}$ имеют общую границу $S^{(t)}$ (см. фиг. 1). Учитывая результаты Этапа 1, чтобы достроить решение в момент $t = t_k$ в области $\Omega^{(t_k)}$, нужно решить в области $\Psi^{(t_k)}$ следующую задачу:

$$\Delta u = \frac{1}{u} \left(u_t - F(u)(\nabla u)^2 - H(u) \right), \quad (3.19)$$

$$u|_{\rho=g(t_k, \varphi)} = f(t_k, \varphi), \quad (3.20)$$

$$q|_{\rho=g(t_k, \varphi)} = \frac{\partial u_1(t_k, g(t_k, \varphi), \varphi)}{\partial \mathbf{n}} = q^*(\varphi), \quad (3.21)$$

где u_1 — решение, найденное на этапе 1, \mathbf{n} — вектор внешней относительно $\Psi^{(t)}$ нормали к границе $S^{(t)}$. Аналогично этапу 1, решение (3.19)–(3.21) будем строить в виде

$$u(t_k, \rho, \varphi) = u_p(\rho, \varphi) + u_h(\rho, \varphi),$$

где u_p — частное решение уравнения (3.19) в момент t_k , u_h — решение соответствующей однородной задачи

$$\begin{aligned} \Delta u_h &= 0, \\ u_h|_{\rho=g(t_k, \varphi)} &= f(t_k, \varphi) - u_p(g(t_k, \varphi), \varphi), \\ q_h|_{\rho=g(t_k, \varphi)} &= \frac{\partial u_h}{\partial \mathbf{n}}|_{\rho=g(t_k, \varphi)} = q^*(\varphi) - \frac{\partial u_p(g(t_k, \varphi), \varphi)}{\partial \mathbf{n}}. \end{aligned}$$

Итерационная процедура выглядит следующим образом:

$$u_p^{(0)} \equiv 0, \quad (3.22)$$

$$\Delta u_h^{(n)} = 0, \quad (3.23)$$

$$u_h^{(n)}|_{\rho=g(t_k, \varphi)} = f(t_k, \varphi) - u_p^{(n)}(g(t_k, \varphi), \varphi), \quad (3.24)$$

$$q_h^{(n)}|_{\rho=g(t_k, \varphi)} = q^*(\varphi) - \frac{\partial u_p^{(n)}(g(t_k, \varphi), \varphi)}{\partial \mathbf{n}}, \quad (3.25)$$

$$\Delta u_p^{(n+1)} = \frac{1}{u^{(n)}} \left(u_t^{(n)} - F(u^{(n)}) \left(\nabla u^{(n)} \right)^2 - H(u^{(n)}) \right). \quad (3.26)$$

Здесь $u_p^{(n)}$, $u_h^{(n)}$, $u^{(n)} = u_p^{(n)} + u_h^{(n)}$ суть n -е итерации решений.

Задача (3.23)–(3.25) на каждой итерации решается методом граничных элементов. При этом наличие двух граничных условий классического вида позволяет найти недостающие граничные значения искомой функции и потока решением системы линейных граничных интегральных уравнений. Для этого требуется, как и на первом этапе, взять равное количество граничных элементов на границах $S^{(0)}$ и $S^{(t_k)}$.

Задача (3.26) на каждой итерации решается методом коллокаций, аналогично (3.16)–(3.18). При остановке итерационного процесса на n -й итерации мы имеем решение задачи (3.19)–(3.21) в области $\Psi^{(t_k)}$:

$$u_2(t_k, \rho, \varphi) = u_p^{(n)}(\rho, \varphi) + u_h^{(n)}(\rho, \varphi).$$

Тогда решение задачи (2.1), (2.2) в момент t_k в области $\Omega^{(t_k)}$ гладко “склеивается” из двух частей

$$u(t_k, \rho, \varphi) = \begin{cases} u_1(t_k, \rho, \varphi), & (\rho, \varphi) \in \Phi^{(t_k)}, \\ u_2(t_k, \rho, \varphi), & (\rho, \varphi) \in \Psi^{(t_k)}. \end{cases}$$

4. ТОЧНЫЕ РЕШЕНИЯ

Верификация разработанного алгоритма может быть проведена с помощью точных решений, нахождение которых само по себе является нетривиальной задачей.

Редукция к ОДУ. Пусть $F(u) = 1/\sigma$, $\sigma > 0$, $H(u) = \alpha u^\beta$, что соответствует степенному характеру нелинейностей в уравнении (2.1). Предположим также, что задача (2.1), (2.2) зависит только от одной пространственной координаты ρ . Тогда уравнение (2.1) примет следующий вид:

$$u_t = uu_{\rho\rho} + \frac{1}{\sigma}u_\rho^2 + \frac{uu_\rho}{\rho} + \alpha u^\beta. \quad (4.1)$$

Следуя [17], выполним редукцию уравнения (4.1) к обыкновенному дифференциальному уравнению (ОДУ), используя анзац

$$u = \phi(t)v(z), \quad z = \frac{\rho}{a(t)} - 1,$$

где $\rho = a(t)$ — фронт тепловой волны. Тогда удастся выполнить разделение переменных и получить функции $a(t)$, $\phi(t)$ в явном виде:

$$\begin{aligned} a(t) &= c_1 e^{c_2 t}, \quad \phi(t) = c_1^2 e^{2c_2 t}, \quad \beta = 1; \\ a(t) &= (c_3 + c_4 t)^\omega, \quad \phi(t) = (c_3 + c_4 t)^{2\omega-1}, \\ \omega &= \frac{\beta - 2}{2(\beta - 1)}, \quad \beta \neq 1, 2. \end{aligned} \quad (4.2)$$

Здесь c_i , $i = 1, \dots, 4$, — произвольные константы. Таким образом, фронт тепловой волны в каждом из случаев (4.2) известен (с точностью до выбора двух констант).

В свою очередь, функция $v(z)$ определяется из ОДУ

$$vv'' + \frac{(v')^2}{\sigma} + \frac{vv'}{z+1} + \gamma_1(z+1)v' + \alpha v^\beta + \gamma_2 v = 0, \quad (4.3)$$

где $\gamma_1 = c_2$, $\gamma_2 = -2c_2$ при $a(t) = c_1 e^{c_2 t}$; $\gamma_1 = c_4 \omega$, $\gamma_2 = c_4/(\beta - 1)$ при $a(t) = (c_3 + c_4 t)^\omega$. Отметим, что при $\alpha = 0$ в качестве β можно использовать любое положительное значение, за исключением $\beta = 1$.

Из условия на фронте тепловой волны

$$u|_{\rho=a(t)} = 0 \quad (4.4)$$

следуют начальные условия для уравнения (4.3)

$$v(0) = 0, \quad v'(0) = -\gamma_1 \sigma, \quad (4.5)$$

где второе условие — единственное возможное ненулевое значение производной, которое обеспечивает совместность: как легко убедиться, задача Коши (4.3), (4.5) становится несовместной при любом другом задании $v'(0) \neq 0$.

Численное интегрирование ОДУ. Известно [13], что (4.3) относится к уравнениям Лье́нара, которые, как правило, не интегрируются в квадратурах и решаются численно. При этом, чтобы построить удовлетворяющее условию (4.4) решение уравнения (4.1) на интервале времени $t \in [0, T]$, нужно найти решение задачи (4.3), (4.5) на отрезке $z \in [z^*, 0]$, где $z^* = a(0)/a(T) - 1 > -1$.

Особенность при старшей (второй) производной затрудняет применение классических разностных методов для решения (4.3), (4.5). Однако использование итерационного подхода, аналогичного предложенному в разд. 3 на каждом шаге по времени, позволяет преодолеть эти сложности. Разрешив ОДУ (4.3) относительно старшей производной, мы можем построить решение задачи Коши итерационно:

$$v(z) = r(z) + s(z), \quad r^{(0)} \equiv 0, \quad (4.6)$$

$$s^{(n)} = -\left(\gamma_1 \sigma + r^{(n)}(0)\right)z, \quad (4.7)$$

$$v^{(n)}(z) = r^{(n)}(z) + s^{(n)}(z), \quad (4.8)$$

$$\begin{aligned} \frac{d^2 r^{(n+1)}}{dz^2} = & -\frac{1}{v^{(n)}} \left(\frac{1}{\sigma} \left[\frac{dv^{(n)}}{dz} \right]^2 + \gamma_1(z+1) \frac{dv^{(n)}}{dz} \right) - \\ & - \frac{1}{z+1} \frac{dv^{(n)}}{dz} - \alpha v^{\beta-1} - \gamma_2. \end{aligned} \quad (4.9)$$

Здесь $r(z)$ — частное решение уравнения (4.3), $s(z)$ — решение соответствующей задачи Коши для однородного уравнения $s'' = 0$, $v^{(n)}$, $r^{(n)}$ и $s^{(n)}$ суть n -е итерации решений. Уравнение (4.9) решается методом коллокаций с использованием РБФ. Итерационный процесс останавливается при выполнении условия

$$\left| \frac{v^{(n+1)}(z^*) - v^{(n)}(z^*)}{v^{(n)}(z^*)} \right| < \varepsilon,$$

где $\varepsilon > 0$ — заданное число.

Вопросы сходимости предложенного итерационного метода и оценки точности получаемых с его помощью решений задачи (4.3), (4.5), как и верификация алгоритма, представленного в разд. 3, обсуждаются в разд. 5.

Решения ОДУ, представимые в явном виде. Большой интерес, в том числе и для верификации результатов расчетов, представляют те редкие случаи, в которых удается получить решения в виде явных формул. Рассмотрим два из них.

Пусть в уравнении (4.1) $\alpha \neq 0$, $\beta = 1$. Будем искать его решение в виде

$$u(t, \rho) = \sum_{k=1}^n w_k(t) \rho^k,$$

т.е. в виде многочлена по степеням ρ с коэффициентами, зависящими от t . Индукцией по n можно показать, что $w_1(t) = w_3(t) = \dots = w_k(t) = \dots = 0$, $k > 3$. Отличными от нуля могут быть только коэффициенты w_0 и w_2 , которые определяются из системы обыкновенных дифференциальных уравнений

$$w_0'(t) = 4w_0(t)w_2(t) + \alpha w_0(t), \quad w_2'(t) = \left(4 + \frac{4}{\sigma} \right) w_2^2(t) + \alpha w_2(t). \quad (4.10)$$

Общее решение (4.10) имеет вид

$$w_0(t) = \frac{C_2 e^{\alpha t}}{(C - e^{\alpha t})^{\sigma/(\sigma+1)}}, \quad w_2(t) = \frac{\alpha \sigma e^{\alpha t}}{4(\sigma+1)(C - e^{\alpha t})}.$$

Отсюда имеем, что решение уравнения (4.1) искомого вида, удовлетворяющее условию $u(0, R) = 0$, таково:

$$u_1(t, \rho) = \frac{\alpha \sigma e^{\alpha t}}{4(\sigma+1)(C - e^{\alpha t})} \left[\rho^2 - R^2 \left(\frac{C - e^{\alpha t}}{C - 1} \right)^{1/(\sigma+1)} \right]. \quad (4.11)$$

Здесь $C > 1$, $R > 0$ — произвольные константы. Фронт тепловой волны решения (4.11) имеет следующий вид:

$$\rho = a_1(t) = R \left(\frac{C - e^{\alpha t}}{C - 1} \right)^{\delta}, \quad (4.12)$$

где $\delta = 1/(2\sigma + 2)$. Если $\alpha > 0$, то волна движется к началу координат, и при $t = (\ln C)/\alpha$, когда фронт достигает точки $\rho = 0$, происходит blow-up. Если $\alpha < 0$, то волна движется от начала координат, но не может пересечь окружность $\rho = R(C/(C-1))^\delta > R$, и волна быстро (экспоненциально по t) затухает.

Если принять $\alpha = 0$ (т.е. источник/сток отсутствует), то интегрирование системы (4.10) дает другие функции

$$w_0(t) = \frac{C_2}{(C + \eta t)^{\sigma/(\sigma+1)}}, \quad w_2(t) = -\frac{1}{C + \eta t},$$

где $\eta = 4 + 4/\sigma$. Отсюда получаем, что решение уравнения (4.1), удовлетворяющее условию $u(0, R) = 0$, при $\alpha = 0$ имеет следующий вид:

$$u_2(t, \rho) = \frac{1}{C + \eta t} \left[R^2 \left(\frac{C + \eta t}{C} \right)^{2\delta} - \rho^2 \right]. \quad (4.13)$$

Можно видеть, что (4.13) не следует непосредственно из (4.11). Фронт тепловой волны здесь определяется уравнением

$$\rho = a_2(t) = R \left(1 + \frac{\eta t}{C} \right)^\delta. \quad (4.14)$$

Поскольку $0 < \delta < 0.5$, $0 < \sigma/(\sigma + 1) < 1$, то указанный фронт движется достаточно медленно, но тем не менее обладает свойством $a_2(t) \xrightarrow{t \rightarrow +\infty} +\infty$; решение $u_2(t, \rho)$ при всех $\rho > 0$ является ограниченным по t .

Замечание 2. Соотношение (4.13) соответствует следующему квадратичному решению задачи (4.3), (4.5):

$$v_1(z) = -\gamma_1 \sigma \left(\frac{z^2}{2} + z \right), \quad (4.15)$$

которое имеет место при $\alpha = 0$, $\beta = 2 + 1/\sigma$, $\omega = \delta = 1/(2\sigma + 2)$. Ранее решение близкого к (4.13) вида использовалось в работе [22] для верификации расчетов.

5. ПРИМЕРЫ И ОБСУЖДЕНИЕ

Пример 1. Для оценки точности решения задачи (4.3), (4.5) с помощью итерационной процедуры (4.6)–(4.9) в случае, когда фронт тепловой волны задан степенной функцией $a(t) = (c_3 + c_4 t)^\omega$ (см. (4.2)), используем решение (4.15). Для этого нужно, чтобы $a(t)$ имела вид (4.14), поэтому в (4.1)–(4.3) и (4.5) положим $c_3 = R^{1/\delta}$, $c_4 = R^{1/\delta} \eta/C$, $\alpha = 0$, $\beta = 2 + 1/\sigma$.

Поскольку наибольшее отклонение итерационных решений от решения (4.15) наблюдается в точке $z = z^*$, оценивать их точность будем с помощью относительной погрешности

$$D = \left| \frac{v(z^*) - v_1(z^*)}{v_1(z^*)} \right|. \quad (5.1)$$

В табл. 1 представлены погрешности численных решений при $R = 1$, $\varepsilon = 10^{-10}$, различных значениях параметров σ , C , z^* и числа точек коллокации M . Точки коллокации располагались равномерно на отрезке $[z^*, 0]$, за исключением крайней правой, которая была сдвинута из нуля внутрь отрезка ввиду особенности уравнения (4.3). В качестве РБФ здесь и далее использовались мультикватричные функции $f_i = \sqrt{1 + \vartheta^2(z - z_i)^2}$, где z_i , $i = 1, 2, \dots, M$, — точки коллокации, значение

Таблица 1. Погрешности численных решений задачи (4.3), (4.5) при степенном фронте волны

σ	C	z^*	M	D	d	σ	C	z^*	M	D	d
2	5	-0.5	11	$3.6 \cdot 10^{-9}$	$2.2 \cdot 10^{-8}$	5	3	-0.5	11	$6.3 \cdot 10^{-9}$	$6.1 \cdot 10^{-8}$
			21	$6.9 \cdot 10^{-10}$	$9.2 \cdot 10^{-9}$				21	$1.2 \cdot 10^{-9}$	$2.6 \cdot 10^{-8}$
			51	$1.3 \cdot 10^{-10}$	$4.5 \cdot 10^{-9}$				51	$1.9 \cdot 10^{-10}$	$1.2 \cdot 10^{-8}$
		-0.7	11	$7.0 \cdot 10^{-9}$	$2.5 \cdot 10^{-8}$			-0.7	11	$1.2 \cdot 10^{-8}$	$7.2 \cdot 10^{-8}$
			21	$1.1 \cdot 10^{-9}$	$1.2 \cdot 10^{-8}$				21	$2.3 \cdot 10^{-9}$	$3.1 \cdot 10^{-8}$
			51	$1.9 \cdot 10^{-10}$	$5.7 \cdot 10^{-9}$				51	$3.8 \cdot 10^{-10}$	$1.5 \cdot 10^{-8}$
		-0.9	11	$1.5 \cdot 10^{-8}$	$2.9 \cdot 10^{-7}$			-0.9	11	$2.8 \cdot 10^{-8}$	$8.1 \cdot 10^{-8}$
			21	$2.3 \cdot 10^{-9}$	$1.2 \cdot 10^{-8}$				21	$5.1 \cdot 10^{-9}$	$3.2 \cdot 10^{-8}$
			51	$2.9 \cdot 10^{-10}$	$5.9 \cdot 10^{-9}$				51	$8.5 \cdot 10^{-10}$	$1.7 \cdot 10^{-8}$
		-0.99	51	$6.1 \cdot 10^{-10}$	$2.5 \cdot 10^{-8}$			-0.99	51	$1.9 \cdot 10^{-9}$	$8.0 \cdot 10^{-8}$
			101	$1.2 \cdot 10^{-10}$	$3.6 \cdot 10^{-9}$				101	$4.5 \cdot 10^{-10}$	$1.3 \cdot 10^{-8}$

параметра формы ϑ было выбрано, исходя из его влияния на скорость сходимости и точность решения. Лучшие результаты были получены при значениях, близких к $\vartheta = 0.8\sqrt{M}/z^*$, см.[28]. В последнем столбце табл. 1 представлены невязки уравнения (4.3) при подстановке в него численных решений,

$$d = \max_{z \in [z^*, 0]} \left| vv'' + \frac{(v')^2}{\sigma} + \frac{vv'}{z+1} + \gamma_1(z+1)v' + \alpha v^\beta + \gamma_2 v \right|. \tag{5.2}$$

Отметим, что решение получается непрерывным и дважды дифференцируемым.

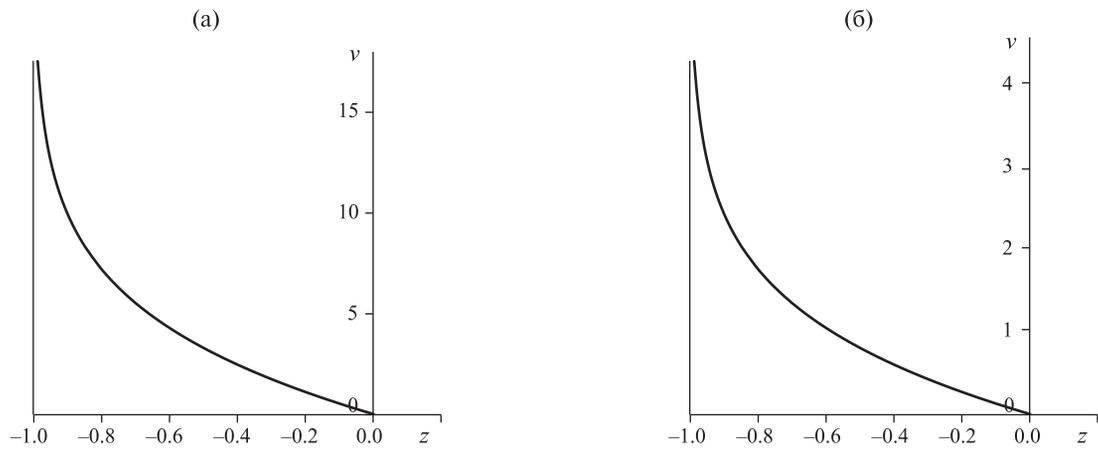
Результаты расчетов позволяют сделать следующие выводы. Погрешность численных решений относительно точного решения (4.15) достаточно мала. С увеличением числа точек коллокации точность повышается, что говорит о сходимости, при этом даже при небольшом их числе порядок погрешности довольно мал.

Числовые значения параметров σ , C незначительно влияют на точность расчетов. Значение z^* , определяющее область решения задачи, также слабо влияет на погрешность до тех пор, пока это значение не приблизится к значению минус один, в котором уравнение имеет особенность. Это происходит, когда интервал времени $t \in [0, T]$, на котором мы решаем уравнение (4.1), неограниченно возрастает. Для примера в табл. 1 был рассмотрен случай $z^* = -0.99$. Тогда, если $M = 11, 21$, итерационный процесс не сходится. Однако достигнуть сходимости и требуемой точности, соответствующей другим значениям z^* , можно, увеличив число точек коллокации.

Невязка (5.2) уменьшается с увеличением числа точек коллокации и коррелирует с относительной погрешностью (5.1). Это дает основание использовать невязку для оценки точности численного решения задачи (4.3), (4.5), когда аналитического тестового решения нет. Этот важный результат мы используем в следующем примере.

Пример 2. Рассмотрим теперь задачу Коши (4.3), (4.5) при $\alpha \neq 0$, $\beta = 1$, $a(t) = c_1 e^{c_2 t}$. Поскольку проинтегрировать ОДУ в квадратурах в этом случае не получается, точность алгоритма (4.6)–(4.9) будем оценивать по величине невязки (5.2). При равномерном расположении точек коллокации невязка в окрестностях концов отрезка $[z^*, 0]$ на несколько порядков больше, чем в остальной части. Особенно возрастает невязка при приближении к точке z^* . Это, очевидно, объясняется тем, что прямая $z = -1$ является вертикальной асимптотой решения (см. фиг. 2). В связи с этим, использовалось сгущение точек коллокации в окрестностях нуля и z^* .

В табл. 2 представлены невязки для решений при $\alpha = 1$, $\varepsilon = 10^{-10}$ и различных значениях параметров σ , c_2 , z^* , M . Вновь наблюдается сходимость относительно числа точек коллокации, но их требуется



Фиг. 2. Графики решений задачи (4.3), (4.5): (а) — для $\sigma = 5$, $c_2 = 1$, (б) — для $\sigma = 3$, $c_2 = 0.5$

Таблица 2. Невязки уравнения (4.3) для численных решений при показательном фронте волны

σ	c_2	z^*	M	d	σ	c_2	z^*	M	d
5	1	-0.5	21	$5.5 \cdot 10^{-7}$	3	0.5	-0.5	21	$9.2 \cdot 10^{-8}$
			51	$1.1 \cdot 10^{-7}$				51	$6.5 \cdot 10^{-9}$
			101	$6.7 \cdot 10^{-9}$				101	$4.3 \cdot 10^{-10}$
		-0.7	51	$2.3 \cdot 10^{-6}$			-0.7	51	$7.1 \cdot 10^{-7}$
			101	$1.2 \cdot 10^{-7}$				101	$1.3 \cdot 10^{-8}$
			201	$1.2 \cdot 10^{-8}$				151	$1.7 \cdot 10^{-9}$
		-0.9	101	$5.2 \cdot 10^{-3}$			-0.9	101	$2.9 \cdot 10^{-4}$
			151	$1.2 \cdot 10^{-4}$				151	$1.8 \cdot 10^{-5}$
			201	$7.2 \cdot 10^{-6}$				201	$4.0 \cdot 10^{-6}$

в разы больше для достижения значений, близких к полученным в предыдущем примере. Точность решения заметно падает с увеличением области решения, поскольку $\lim_{z \rightarrow -1} v(z) = +\infty$. Тем не менее,

лученные результаты демонстрируют хорошую точность решения задачи (4.3), (4.5) и дают основание использовать его для построения решений задачи (2.1), (2.2). Решение, соответствующее выделенным жирным шрифтом строкам табл. 2, использовано ниже для тестирования алгоритма решения задачи (2.1), (2.2) (см. пример 4).

Пример 3. Для тестирования пошагового алгоритма решения задачи (2.1), (2.2), предложенного в разд. 3, используем сначала аналитическое решение (4.11). Пусть уравнение (2.1) имеет вид (4.1), краевое условие (2.2) зададим в виде

$$u|_{\rho=g(t)} = u_1(t, g(t)) \quad (5.3)$$

и рассмотрим три различных варианта функции $g(t)$, определяющей движение фронта в (5.3)

$$g_1(t) = a_1\left(\frac{t}{3}\right), \quad g_2(t) = a_1\left(\frac{t}{2}\right), \quad g_3(t) = a_1\left(\frac{2t}{3}\right), \quad (5.4)$$

где $a_1(t)$ — функция (4.12).

Таблица 3. Погрешности численного решения задачи (4.1), (5.3)

i	t	h	$2N$	M_{in}	Этап 1	Этап 2
1	0.5	0.1	200	50	$9.31 \cdot 10^{-6}$	$1.89 \cdot 10^{-5}$
			350	75	$8.92 \cdot 10^{-6}$	$1.38 \cdot 10^{-5}$
			500	100	$8.86 \cdot 10^{-6}$	$1.25 \cdot 10^{-5}$
		0.01	200	50	$5.44 \cdot 10^{-6}$	$9.71 \cdot 10^{-6}$
			350	75	$4.35 \cdot 10^{-6}$	$8.65 \cdot 10^{-6}$
			500	100	$3.69 \cdot 10^{-6}$	$8.22 \cdot 10^{-6}$
	1.0	0.1	200	50	$1.59 \cdot 10^{-5}$	$3.66 \cdot 10^{-5}$
			350	75	$1.05 \cdot 10^{-5}$	$2.94 \cdot 10^{-5}$
			500	100	$9.94 \cdot 10^{-6}$	$2.87 \cdot 10^{-5}$
		0.01	200	50	$8.97 \cdot 10^{-6}$	$2.63 \cdot 10^{-5}$
			350	75	$8.13 \cdot 10^{-6}$	$2.12 \cdot 10^{-5}$
			500	100	$7.75 \cdot 10^{-6}$	$1.98 \cdot 10^{-5}$
2	0.5	0.1	200	50	$1.64 \cdot 10^{-5}$	$3.31 \cdot 10^{-5}$
			350	75	$1.32 \cdot 10^{-5}$	$2.87 \cdot 10^{-5}$
			500	100	$1.25 \cdot 10^{-5}$	$2.80 \cdot 10^{-5}$
		0.01	200	50	$8.65 \cdot 10^{-6}$	$1.66 \cdot 10^{-5}$
			350	75	$7.53 \cdot 10^{-6}$	$9.53 \cdot 10^{-6}$
			500	100	$7.08 \cdot 10^{-6}$	$9.07 \cdot 10^{-6}$
	1.0	0.1	200	50	$2.35 \cdot 10^{-5}$	$5.59 \cdot 10^{-5}$
			350	75	$1.87 \cdot 10^{-5}$	$4.85 \cdot 10^{-5}$
			500	100	$1.81 \cdot 10^{-5}$	$4.77 \cdot 10^{-5}$
		0.01	200	50	$1.38 \cdot 10^{-5}$	$4.64 \cdot 10^{-5}$
			350	75	$9.72 \cdot 10^{-6}$	$4.28 \cdot 10^{-5}$
			500	100	$9.44 \cdot 10^{-6}$	$4.17 \cdot 10^{-5}$
3	0.5	0.1	200	50	$3.26 \cdot 10^{-5}$	$6.35 \cdot 10^{-5}$
			350	75	$2.94 \cdot 10^{-5}$	$5.97 \cdot 10^{-5}$
			500	100	$2.89 \cdot 10^{-5}$	$5.94 \cdot 10^{-5}$
		0.01	200	50	$1.73 \cdot 10^{-5}$	$3.82 \cdot 10^{-5}$
			350	75	$1.42 \cdot 10^{-5}$	$3.56 \cdot 10^{-5}$
			500	100	$1.36 \cdot 10^{-5}$	$3.54 \cdot 10^{-5}$
	1.0	0.1	200	50	$3.84 \cdot 10^{-5}$	$8.26 \cdot 10^{-5}$
			350	75	$3.47 \cdot 10^{-5}$	$7.94 \cdot 10^{-5}$
			500	100	$3.42 \cdot 10^{-5}$	$7.90 \cdot 10^{-5}$
		0.01	200	50	$2.96 \cdot 10^{-5}$	$7.22 \cdot 10^{-5}$
			350	75	$2.53 \cdot 10^{-5}$	$6.85 \cdot 10^{-5}$
			500	100	$2.48 \cdot 10^{-5}$	$6.84 \cdot 10^{-5}$

Во всех вариантах расчетов были получены симметричные относительно начала координат решения. Поскольку наибольшие отклонения численных решений от точного на этапе 1 наблюдались вдоль определяемого нулевого фронта, оценивать точность на этом этапе будем с помощью

Таблица 4. Погрешности численного решения задачи (4.1), (5.5) при степенном нулевом фронте

i	t	h	$2N$	M_{in}	Этап 1	Этап 2
1	0.5	0.1	200	25	$1.72 \cdot 10^{-4}$	$1.69 \cdot 10^{-3}$
			350	50	$1.61 \cdot 10^{-4}$	$1.36 \cdot 10^{-3}$
			500	75	$1.55 \cdot 10^{-4}$	$1.23 \cdot 10^{-3}$
		0.01	200	25	$1.1 \cdot 10^{-4}$	$9.68 \cdot 10^{-4}$
			350	50	$9.9 \cdot 10^{-5}$	$9.36 \cdot 10^{-4}$
			500	75	$9.4 \cdot 10^{-5}$	$9.14 \cdot 10^{-4}$
	1.0	0.1	200	25	$2.28 \cdot 10^{-4}$	$2.01 \cdot 10^{-3}$
			350	50	$2.16 \cdot 10^{-4}$	$1.78 \cdot 10^{-3}$
			500	75	$2.13 \cdot 10^{-4}$	$1.63 \cdot 10^{-3}$
		0.01	200	25	$1.57 \cdot 10^{-4}$	$1.6 \cdot 10^{-3}$
			350	50	$1.43 \cdot 10^{-4}$	$1.42 \cdot 10^{-3}$
			500	75	$1.37 \cdot 10^{-4}$	$1.36 \cdot 10^{-3}$
2	0.5	0.1	200	25	$1.35 \cdot 10^{-4}$	$9.94 \cdot 10^{-4}$
			350	50	$1.26 \cdot 10^{-4}$	$9.13 \cdot 10^{-4}$
			500	75	$1.18 \cdot 10^{-4}$	$8.84 \cdot 10^{-4}$
		0.01	200	25	$6.9 \cdot 10^{-5}$	$7.97 \cdot 10^{-4}$
			350	50	$6.3 \cdot 10^{-5}$	$6.82 \cdot 10^{-4}$
			500	75	$5.9 \cdot 10^{-5}$	$6.55 \cdot 10^{-4}$
	1.0	0.1	200	25	$1.88 \cdot 10^{-4}$	$1.19 \cdot 10^{-3}$
			350	50	$1.83 \cdot 10^{-4}$	$1.04 \cdot 10^{-3}$
			500	75	$1.81 \cdot 10^{-4}$	$9.6 \cdot 10^{-4}$
		0.01	200	25	$1.13 \cdot 10^{-4}$	$8.73 \cdot 10^{-4}$
			350	50	$1.1 \cdot 10^{-4}$	$8.45 \cdot 10^{-4}$
			500	75	$9.8 \cdot 10^{-5}$	$8.16 \cdot 10^{-4}$
3	0.5	0.1	200	25	$1.12 \cdot 10^{-4}$	$7.42 \cdot 10^{-4}$
			350	50	$9.9 \cdot 10^{-5}$	$7.13 \cdot 10^{-4}$
			500	75	$9.3 \cdot 10^{-5}$	$6.49 \cdot 10^{-4}$
		0.01	200	25	$5.7 \cdot 10^{-5}$	$4.51 \cdot 10^{-4}$
			350	50	$4.8 \cdot 10^{-5}$	$3.81 \cdot 10^{-4}$
			500	75	$4.1 \cdot 10^{-5}$	$3.33 \cdot 10^{-4}$
	1.0	0.1	200	25	$1.65 \cdot 10^{-4}$	$8.73 \cdot 10^{-4}$
			350	50	$1.57 \cdot 10^{-4}$	$8.45 \cdot 10^{-4}$
			500	75	$1.52 \cdot 10^{-4}$	$8.16 \cdot 10^{-4}$
		0.01	200	25	$1.06 \cdot 10^{-4}$	$6.19 \cdot 10^{-4}$
			350	50	$1.04 \cdot 10^{-4}$	$5.84 \cdot 10^{-4}$
			500	75	$9.1 \cdot 10^{-5}$	$5.52 \cdot 10^{-4}$

погрешности найденного фронта относительно $a_1(t)$. Точность решения на этапе 2 оценим погрешностью численного решения относительно точного (4.11) при $\rho = a_1(0)$, где также наблюдалось наибольшее расхождение.

Проиллюстрируем результаты расчетов при $\sigma = 2$, $R = 1$, $C = 2$, $\alpha = -1$. В качестве точек коллокации были приняты $2N$ узлов граничных элементов и M_{in} внутренних точек. В табл. 3 приведены оценки точности на обоих этапах решения в два момента времени при различных h , N и M_{in} . В первом столбце содержатся номера вариантов подвижного фронта $\rho = g(t)$ в соответствии с (5.4). Результаты расчетов показывают, что точность повышается с уменьшением шага по времени и увеличением числа точек коллокации (включая число граничных элементов). Таким образом, алгоритм сходится относительно этих параметров. Расположение подвижного фронта (5.4) относительно нулевого фронта (4.12) влияет на точность (чем эти фронты ближе, тем точнее), но в пределах одного порядка.

Пример 4. Оценим теперь точность численного решения при показательном фронте тепловой волны $a(t) = c_1 e^{c_2 t}$. Для этого используем решение, построенное в примере 2 при $\sigma = 3$, $c_1 = 0.5$, $c_2 = 0.5$. Отметим, что найденные решения задачи (4.3), (4.5) при $M = 51, 101$ совпадали до десятого знака после запятой. Краевое условие (2.2) примем в виде

$$u|_{\rho=g_i(t)} = f_i(t), \quad i = 1, 2, 3, \quad (5.5)$$

где

$$g_1(t) = c_1 e^{\frac{c_2 t}{3}}, \quad g_2(t) = c_1 e^{\frac{c_2 t}{2}}, \quad g_3(t) = c_1 e^{\frac{2c_2 t}{3}},$$

а функции $f_i(t)$ зададим таблично с заданным шагом по времени h в соответствии с численным решением задачи (4.3), (4.5).

В табл. 4 приведены оценки точности в два момента времени при различных h , N и M_{in} , аналогично предыдущему примеру. Выводы также аналогичны примеру 3: наблюдается сходимость относительно шага по времени и числа точек коллокации, увеличение расстояния от заданного подвижного фронта до нулевого фронта незначительно снижает точность. Погрешности в примере 4 заметно больше, чем в примере 3. Очевидно, это связано с тем, что функция $f(t)$ в краевом условии задается в табличном виде. Тем не менее, результаты удовлетворительны для верификации алгоритма.

Пример 5. В завершение вычислительного эксперимента рассмотрим пример решения задачи вида (2.1), (2.2), зависящего от двух пространственных переменных. При этом продемонстрируем также эффективность алгоритма в случае отличного от степенного вида нелинейности уравнения.

Пусть в уравнении (2.1) $F(u) = (\sigma + u)/(\sigma - u)$, что соответствует дробно-линейной зависимости коэффициента теплопроводности от температуры вида $k(T) = \sigma T/(T + \theta)$, $H(u) = u$, и условие (2.2) имеет вид

$$u|_{\rho=1+c_1 t(1+c_2 \cos(\varphi))} = 0, \quad (5.6)$$

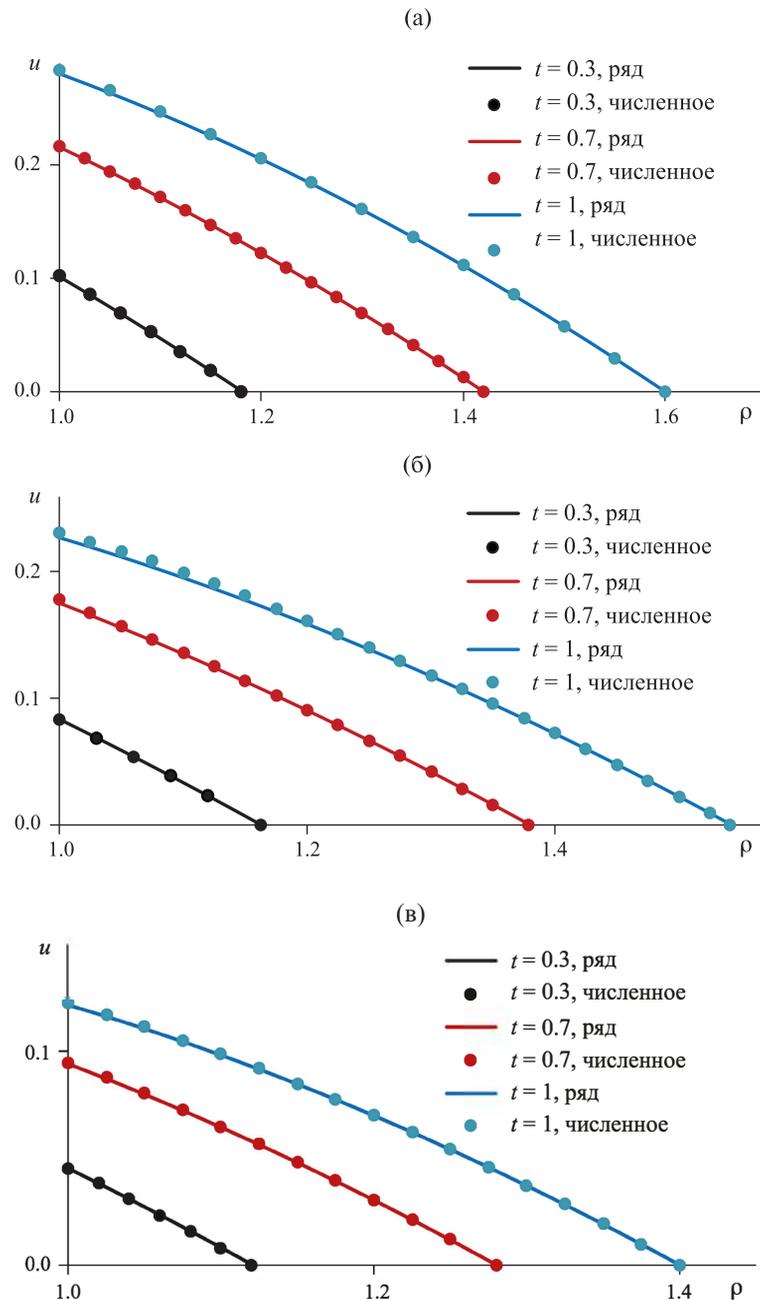
т.е. фронт тепловой волны известен и задан уравнением $\rho = 1 + c_1 t(1 + c_2 \cos(\varphi))$, где $c_1 \neq 0$, $c_2 \neq 0$ — константы. Тогда условие (2.4) в доказательстве теоремы становится однородным, и решение задачи (2.3), (2.4) (которая, напомним, эквивалентна исходной) может быть построено в виде степенного ряда

$$u(t, \rho, \varphi) = \sum_{m=0}^{\infty} u_m(t, \varphi) \frac{r^m}{m!} = \sum_{m=0}^{\infty} u_m(t, \varphi) \frac{[\rho - a(t, \varphi)]^m}{m!}, \quad u_m(t, \varphi) = \left. \frac{\partial^m u}{\partial r^m} \right|_{\rho=a(t, \varphi)}. \quad (5.7)$$

Из условия (2.4) следует, что $u_0 \equiv 0$. Коэффициенты $u_m(t, \varphi)$, $m = 1, 2, \dots$, определяются рекуррентно путем подстановки ряда (5.7) в уравнение (2.3) и приравнивания коэффициентов при соответствующих степенях r . При этом уравнение для $u_1(t, \varphi)$ будет квадратным, а для последующих коэффициентов — линейным. Локальная сходимость ряда следует из теоремы 1 (см. также [23]).

Отрезки ряда (5.7)

$$u_n(t, \rho, \varphi) = \sum_{m=0}^n u_m(t, \varphi) \frac{[\rho - a(t, \varphi)]^m}{m!} \quad (5.8)$$



Фиг. 3. Сравнение численного решения и отрезка ряда: (а) — вдоль луча $\varphi = 0$; (б) — вдоль луча $\varphi = \pi/4$; (в) — вдоль луча $\varphi = \pi/2$

были использованы для верификации численного решения задачи (2.1), (2.2), когда краевое условие неоднородно. Именно, условие (2.2), соответствующее (5.8), было задано на подвижном фронте $g(t, \varphi) = 1 + 0.5c_1 t(1 + c_2 \cos(\varphi))$

$$u|_{\rho=g(t,\varphi)} = u_n(t, g(t, \varphi), \varphi).$$

На фиг. 3 приведены графики численного и аналитического решений вдоль лучей $\varphi = 0$, $\varphi = \pi/4$, $\varphi = \pi/2$ при $\sigma = 3$, $c_1 = 0.4$, $c_2 = 0.5$, $n = 4$, $h = 0.1$, $2N = 500$, $M_{in} = 100$. Результаты показывают, что разработанный в разд. 3 алгоритм позволяет успешно решать двумерные задачи вида (2.1), (2.2) и в случае нелинейности отличной от степенной.

В целом вычислительный эксперимент дал два основных результата: 1) верифицирован предложенный алгоритм решения задачи (2.1), (2.2); 2) обоснована возможность построения референсных решений с помощью численного решения задачи Коши для ОДУ (4.3), (4.5).

ЗАКЛЮЧЕНИЕ И ВЫВОДЫ

В работе выполнено исследование двумерного вырождающегося нелинейного уравнения теплопроводности с источником при нелинейностях общего вида. Рассмотрена проблема отыскания решений, имеющих вид тепловой волны, порождаемой краевым режимом, заданным на движущейся линии, расположенной внутри области определения. Доказана оригинальная теорема существования и единственности таких решений.

Наиболее значимым результатом, по мнению авторов, является то, что формулировка и аппроксимация дополнительных условий на фронте тепловой волны позволили разработать новый пошаговый итерационный алгоритм, использующий совместно метод конечных разностей, метод граничных элементов и метод коллокации, превосходящий для исследуемого типа задач по точности расчетов и скорости сходимости метод двойственной взаимности, применявшийся авторами ранее. Применение радиальных базисных функций, а также специальных замен переменных позволили построить непрерывные по пространственным переменным решения на каждом шаге по времени. В качестве инструмента верификации расчетов были применены точные решения, часть из которых также обладает новизной.

Проведенный комплексный вычислительный эксперимент по сопоставлению результатов применения разработанного алгоритма с аналитическими решениями показал численную сходимость относительно шага по времени и числа точек коллокации, а также устойчивость относительно значений числовых параметров задачи. Это дает возможность использовать в будущем результаты расчетов для изучения свойств исследуемого уравнения.

Дальнейшие исследования могут быть направлены на увеличение количества пространственных переменных до трех, а также на изучение подобных задач для других уравнений параболического типа и их систем.

СПИСОК ЛИТЕРАТУРЫ

1. Зельдович Я. Б., Райзер Ю. П. Физика ударных волн и высокотемпературных гидродинамических явлений. М.: Наука, 1966.
2. Баренблатт Г. И., Ентов В. М., Рыжик В. М. Теория нестационарной фильтрации жидкости и газа. М.: Недра, 1972.
3. Тихонов А. Н., Самарский А. А. Уравнения математической физики. М.: Наука, 2004.
4. Ладыженская О. А., Солонников В. А., Уралцева Н. Н. Линейные и квазилинейные уравнения параболического типа. М.: Наука, 1967.
5. Лыков А. В. Теплообмен. М.: Энергия, 1972.
6. Ханхасаев В. Н., Дармахаев Э. В. О некоторых применениях гиперболического уравнения теплопроводности и методах его решения // Матем. заметки СВФУ. 2018. Т. 25. № 1. С. 98–111.
7. Зельдович Я. Б., Компанец А. С. К теории распространения тепла при теплопроводности, зависящей от температуры // Сб., посвященный 70-летию акад. А. Ф. Иоффе. М.: Изд-во АН СССР, 1950. С. 61–71.
8. Баренблатт Г. И. О некоторых неустановившихся движениях жидкости и газа в пористой среде // Прикл. матем. и механ. 1952. Т. 16. № 1. С. 67–78.
9. Самарский А. А., Галактионов В. А., Курдюмов С. П., Михайлов А. П. Режимы с обострением в задачах для квазилинейных параболических уравнений. М.: Наука, 1987.
10. Сидоров А. Ф. Избранные труды: Математика. Механика. М.: Физматлит, 2001.
11. Степанова Е. В., Шишков А. Е. Начальная эволюция носителей решений квазилинейных параболических уравнений с вырождающимся абсорбционным потенциалом // Матем. сборник. 2013. № 3. С. 79–106.
12. Antontsev S. N., Shmarev S. I. Evolution PDEs with nonstandard growth conditions: Existence, uniqueness, localization, blow-up. Paris: Atlantis Press, 2015.
13. Кудряшов Н. А., Синельщиков Д. И. Аналитические решения нелинейного уравнения конвекции–диффузии с нелинейными источниками // Моделирование и анализ информ. систем. 2016. Т. 23. № 3. С. 309–316.
14. Vazquez J. L. The Porous Medium Equation: Mathematical Theory. Oxford: Clarendon Press, 2007.

15. Казаков А. Л., Сневак Л. Ф. Методы граничных элементов и степенных рядов в одномерных задачах нелинейной фильтрации // Изв. ИГУ. Сер. математика. 2012. Т. 5. № 2. С. 2–17.
16. Казаков А. Л., Кузнецов П. А. Об аналитических решениях одной специальной краевой задачи для нелинейного уравнения теплопроводности в полярных координатах // Сиб. ж. индустриальной матем. 2018. Т. 21. № 2 (74). С. 56–65.
17. Казаков А. Л. О точных решениях краевой задачи о движении тепловой волны для уравнения нелинейной теплопроводности // Сиб. электронные матем. изв. 2019. Т. 16. С. 1057–1068.
18. Partridge P. W., Brebbia C. A., Wrobel L. C. The dual reciprocity boundary element method. Southampton: Comp. Mech. Pub., 1992.
19. Golberg M. A., Chen C. S., Bowman H. Some recent results and proposals for the use of radial basis functions in the BEM // Eng. Anal. Bound. Elem. 1999. V. 23. P. 285–296.
20. AL-Bayati S.A., Wrobel L. C. The dual reciprocity boundary element formulation for convection–diffusion–reaction problems with variable velocity field using different radial basis functions // Int. J. Mech. Sci. 2018. V. 145. P. 367–377.
21. Chen W., Fu Z. — J., Chen C. S. Recent Advances in Radial Basis Function Collocation Methods. Berlin/Heidelberg: Springer, 2013.
22. Казаков А. Л., Неведова О. А., Сневак Л. Ф. Решение задач об иницировании тепловой волны для нелинейного уравнения теплопроводности методом граничных элементов // Ж. вычисл. матем. и матем. физ. 2019. Т. 59. № 6. С. 1047–1062.
23. Kazakov A. L., Spevak L. F., Nefedova O. A., Lempert A. A. On the analytical and numerical study of a two-dimensional nonlinear heat equation with a source term // Symmetry. 2020. V. 12. No 6. P. 921.
24. Сневак Л. Ф., Неведова О. А. Численное решение двумерного нелинейного уравнения теплопроводности с использованием радиальных базисных функций // Комп. иссл. и модел. 2022. Т. 14. № 1. С. 9–22.
25. Brebbia C. A., Telles J. C.F., Wrobel L. C. Boundary element techniques. Theory and applications in engineering. Berlin, Heidelberg: Springer Science & Business Media, 2012.
26. Федотов В. П., Сневак Л. Ф. Аналитическое интегрирование функций влияния для решения задач упругости и теории потенциала методом граничных элементов // Матем. моделирование. 2007. Т. 19. № 2. С. 87–104.
27. Fedotov V. P., Spevak L. F. One approach to the derivation of exact integration formulae in the boundary element method // Eng. Anal. Bound. Elem. 2008. V. 32. No 10. P. 883–888.
28. Franke R. Scattered data interpolation: Tests of some methods // Math. Comput. 1982. V. 38. No 157. P. 181–200.

УДК 517.9

О РАЗРЕШИМОСТИ СУЩЕСТВЕННО НЕЛИНЕЙНОГО ЭЛЛИПТИЧЕСКОГО ДИФФЕРЕНЦИАЛЬНОГО УРАВНЕНИЯ С НЕЛОКАЛЬНЫМИ КРАЕВЫМИ УСЛОВИЯМИ

© 2024 г. О. В. Солонуха^{1, *}¹19333 Москва, ул. Вавилова, 44, кор. 2, ФИЦ ИУ РАН, Россия

*e-mail: solonukha@yandex.ru

Поступила в редакцию 13.09.2023 г.

Переработанный вариант 13.09.2023 г.

Принята к публикации 20.10.2023 г.

Доказаны достаточные условия существования обобщенного решения нелинейного эллиптического дифференциального уравнения с нелокальными краевыми условиями типа Бицадзе–Самарского. Используется условие сильной эллиптичности вспомогательного дифференциально–разностного оператора. При сформулированных условиях дифференциально–разностный оператор является деминепрерывным, коэрцитивным и обладает полуограниченной вариацией, что позволяет применять общую теорию операторов псевдомонотонного типа. Библ. 16.

Ключевые слова: нелокальная задача, дифференциально–разностный оператор, условие сильной эллиптичности, оператор с полуограниченной вариацией.

DOI: 10.31857/S0044466924020097, EDN: YJKVNS

ВВЕДЕНИЕ

Нелокальные эллиптические краевые задачи рассматривались начиная с 30-х гг XX века. В 1969 г. А. В. Бицадзе и А. А. Самарский сформулировали новую нелокальную эллиптическую краевую задачу, возникающую в теории плазмы, см. [1]. Разрешимость задачи в общей постановке была сформулирована как нерешенная задача [2]. В конце 80-х годов была построена общая теория линейных нелокальных эллиптических краевых задач, в рамках которой была решена указанная проблема, см. [3–7]. Нелинейные эллиптические нелокальные задачи с краевым условием типа Бицадзе–Самарского изучались автором ранее, см. [8, 9]. В данной работе исследование продолжено для существенно нелинейных уравнений и возможно несимметричных краевых условий. Будут рассмотрены достаточные условия существования решения существенно нелинейной нелокальной задачи. При исследовании использованы методы теории линейных нелокальных граничных задач [5–7] и теории существенно нелинейных дифференциальных уравнений, см. [10].

В разд. 1 сформулированы свойства разбиения области и границы, а также теорема об изоморфизме функциональных пространств, позволяющая поставить в соответствие нелокальной задаче некое эквивалентное (с точки зрения множества решений) функционально–дифференциальное уравнение, эти результаты доказаны в [11, 12, 8]. В разд. 2 доказаны свойства дифференциально–разностного оператора из эквивалентной задачи. При этом мы полагаем, что дифференциально–разностный оператор удовлетворяет условию сильной эллиптичности. Остальные условия сформулированы для дифференциального оператора и соответствуют условиям, которые рассматривались ранее при изучении дифференциальных уравнений с классическими краевыми условиями. В разд. 3 доказаны теоремы существования решений дифференциально–разностных уравнений с однородным краевым условием типа Дирихле и соответствующих дифференциальных уравнений с нелокальными краевыми условиями.

1. ПОСТАНОВКА ЗАДАЧИ

Для изучения нелинейной краевой задачи нам потребуются некоторые вспомогательные построения, разработанные для линейных задач, см. [5, 6, 7].

Пусть $Q \subset \mathbb{R}^n$ — ограниченная область с границей ∂Q класса C^∞ или $Q = (0, d) \times G$, где $G \subset \mathbb{R}^{n-1}$ — ограниченная область (с границей ∂G класса C^∞ , если $n \geq 3$). В случае $n = 1$ мы полагаем $Q = (0, d)$. Обозначим через $M \subset \mathbb{R}^n$ конечное множество векторов h с целочисленными (или соизмеримыми)

координатами. Через M обозначим аддитивную группу, порожденную множеством M , а через Q_r — открытые связные компоненты множества $Q \setminus (\bigcup_{h \in M} (\partial Q + h))$.

Определение 1. Множество Q_r называется *подобластью*. Семейство \mathcal{R} всех подобластей Q_r ($r = 1, 2, \dots$) называется *разбиением* области Q .

Легко видеть, что множество \mathcal{R} не более чем счетно, при этом $\bigcup_r \bar{Q}_r = \bar{Q}$ и $\bigcup_r \partial Q_r = (\bigcup_{h \in M} (\partial Q + h)) \cap \bar{Q}$.

Известно, что для любой подобласти Q_{r_1} и произвольного вектора $h \in M$ либо найдется подобласть Q_{r_2} такая, что $Q_{r_2} = Q_{r_1} + h$, либо $Q_{r_1} + h \subset \mathbb{R}^n \setminus \bar{Q}$, см. лемму 7.1 из [5]. Таким образом, семейство \mathcal{R} можно разбить на непересекающиеся классы следующим образом: подобласти $Q_{r_1}, Q_{r_2} \in \mathcal{R}$ принадлежат одному классу, если $Q_{r_2} = Q_{r_1} + h$ для некоторого $h \in M$. Будем обозначать подобласти Q_r через Q_{sl} , где s — номер класса, а l — номер подобласти в s -м классе. Очевидно, каждый класс состоит из конечного числа $N = N(s)$ подобластей Q_{sl} . Множество классов может быть конечным или счетным, см. примеры в §7 гл. II из [5]. При этом предполагаются выполненными условия 1 и 2.

Условие 1. Для каждой подобласти Q_{sl} ($s = 1, 2, \dots, l = 1, \dots, N(s)$) и для любого $\varepsilon > 0$ существует открытое множество $G_{sl} \subset Q_{sl}$ с границей $\partial G_{sl} \in C^1$ такое, что $\text{mes}_n(Q_{sl} \setminus G_{sl}) < \varepsilon$, $\text{mes}_{n-1}(\partial G_{sl} \Delta \partial Q_{sl}) < \varepsilon$.

Условие 2. Множество $K = \bigcup_{h_1, h_2 \in M} \{ \bar{Q} \cap (\partial Q + h_1) \cap \overline{(\partial Q + h_2) \setminus (\partial Q + h_1)} \}$ таково, что $\text{mes}_{n-1}(K \cap \partial Q) = 0$.

Кроме разбиения области, необходимо рассмотреть разбиение границы ∂Q , определяемое тем же множеством сдвигов $M \subset \mathbb{R}^n$, см. выше.

Обозначим через Γ_ρ открытые, связные в топологии ∂Q компоненты множества $\partial Q \setminus K$. Множества $\{ \Gamma_\rho + h : \Gamma_\rho + h \subset \bar{Q}, \rho = 1, 2, \dots, h \in M \}$ могут быть разбиты на классы. Множества $\Gamma_{\rho_1} + h_1$ и $\Gamma_{\rho_2} + h_2$ принадлежат одному классу, если

- 1) существует вектор $h \in M$ такой, что $\Gamma_{\rho_1} + h_1 = \Gamma_{\rho_2} + h_2 + h$;
- 2) для любых $\Gamma_{\rho_1} + h_1, \Gamma_{\rho_2} + h_2 \subset \partial Q$ нормали к ∂Q в точках $x \in \Gamma_{\rho_1} + h_1$ и $x - h \in \Gamma_{\rho_2} + h_2$ одинаково направлены.

Обозначим множество $\Gamma_\rho + h$ через Γ_{rj} , где r — номер класса, j — номер элемента в классе ($1 \leq j \leq J = J(r)$). Не нарушая общности, будем считать, что $\Gamma_{r1}, \dots, \Gamma_{rJ_0} \subset Q$, $\Gamma_{r, J_0+1}, \dots, \Gamma_{rJ} \subset \partial Q$ ($0 \leq J_0 = J_0(r) < J(r)$).

Множество всех распределений $u \in \mathcal{D}'(Q)$, являющихся вместе со всеми своими частными производными 1-го порядка функциями из $L_p(Q)$, обозначим через $W_p^1(Q)$. При $p \in (1, \infty)$ пространства Соболева $W_p^1(Q)$ рефлексивны и банаховы относительно нормы $\|u\|_{W_p^1(Q)} = \left\{ \sum_{0 \leq i \leq n} \int_Q |\partial_i u|^p dx \right\}^{1/p}$, здесь

$\partial_0 u \equiv u$. Через $\overset{\circ}{W}_p^1(Q)$ обозначим замыкание множества $\dot{C}^\infty(Q)$ финитных бесконечно дифференцируемых функций в $W_p^1(Q)$, $\|u\|_{\overset{\circ}{W}_p^1(Q)} = \left\{ \sum_{1 \leq i \leq n} \int_Q |\partial_i u|^p dx \right\}^{1/p}$. Также будут рассматриваться сопряженные пространства $L_q(Q) = (L_p(Q))^*$ и $W_q^{-1}(Q) = \left(\overset{\circ}{W}_p^1(Q) \right)^*$, $1/p + 1/q = 1$.

Обозначим $W_{p,\gamma}^1(Q)$ ($\gamma = \{\gamma_{ij}^r\}$) подпространство функций из $W_p^1(Q)$, удовлетворяющих *нелокальным краевым условиям*

$$w|_{\Gamma_{rl}} = \sum_{j=1}^{J_0} \gamma_{lj}^r w|_{\Gamma_{rj}} \quad (r \in B, l = J_0 + 1, \dots, J),$$

$$w|_{\Gamma_{rl}} = 0 \quad (r \notin B, l = 1, \dots, J),$$
(1)

где $J_0 = J_0(r)$, $J = J(r)$, γ_{lj}^r — вещественные числа, $B = \{r : J_0 > 0\}$.

Рассмотрим нелинейное уравнение

$$Aw(x) = - \sum_{1 \leq i \leq n} \partial_i A_i(x, w, \nabla w) + A_0(x, w, \nabla w) = f, \quad x \in Q, \quad (2)$$

с нелокальными краевыми условиями (1).

Определение 2. Пусть $f \in L_q(Q)$, $1/p + 1/q = 1$. Функция $w \in W_{p,\gamma}^1(Q)$ называется *обобщенным решением задачи* (2), (1), если для любого $\xi \in \overset{\circ}{W}_p(Q)$ справедливо интегральное тождество

$$\sum_{1 \leq i \leq n} \int_Q A_i(x, w, \nabla w) \partial_i \xi dx + \int_Q A_0(x, w, \nabla w) \xi dx = \int_Q f \xi dx. \quad (3)$$

Пусть $\Lambda = \{a_h : h \in M\}$ — набор вещественных постоянных коэффициентов. Определим разностный оператор $R : L_p(\mathbb{R}^n) \rightarrow L_p(\mathbb{R}^n)$:

$$Ru(x) = \sum_{h \in M} a_h u(x+h), \quad (4)$$

а также оператор $R_Q = P_Q R I_Q : L_p(Q) \rightarrow L_p(Q)$, здесь $I_Q : L_p(Q) \rightarrow L_p(\mathbb{R}^n)$ — оператор продолжения функций из $L_p(Q)$ нулем в $\mathbb{R}^n \setminus Q$, $P_Q : L_p(\mathbb{R}^n) \rightarrow L_p(Q)$ — оператор сужения функций из $L_p(\mathbb{R}^n)$ на Q . Для исследования свойств оператора R_Q введем матрицы $R_s = \{r_{ij}^s\}_{1 \leq i, j \leq N(s)}$ такие, что

$$r_{ij}^s = \begin{cases} a_h & (h = h_{sj} - h_{si} \in M), \\ 0 & (h_{sj} - h_{si} \notin M), \end{cases} \quad (5)$$

где h_{si} определяется условием $Q_{si} = Q_{s1} + h_{si}$. Из ограниченности области Q и формулы (5) следует, что множество различных матриц R_s конечно. Обозначим эти матрицы через R_{s_v} ($v = 1, \dots, n_1$).

Как известно, см. §7 из [5], для каждого $r = 1, 2, \dots$, найдется единственный номер $s = s(r)$ такой, что $N(s) = J(r)$ и $\Gamma_{rl} \subset \partial Q_{sl}$ ($l = 1, \dots, N$) после перенумерации подобластей s -го класса. Обозначим через $R_{s(r)}$ матрицы, полученные из R_s ($s = s(r)$) путем перенумерования соответствующих столбцов и строк. Пусть e_j^r ($j = 1, \dots, J(r)$) есть j -я строка матрицы размерности $J \times J_0$, полученной путем вычеркивания последних $J - J_0$ столбцов из матрицы $R_{s(r)}$.

Определение 3. Будем говорить, что матрицы R_s соответствуют граничным условиям (1), если выполнено условие 3.

Условие 3. Существует набор Λ такой, что для любого $s = 1, 2, \dots$ матрицы R_s невырожденны, а также для всех $r \in B$ и $s = s(r)$ справедливы соотношения:

$$e_l^r = \sum_{1 \leq j \leq J_0} \gamma_{lj}^r e_j^r \quad (l = J_0 + 1, \dots, J). \quad (6)$$

Кроме того, обозначим через R_{s_0} матрицу порядка $J_0 \times J_0$, полученную из матрицы R_s вычеркиванием последних $N - J_0$ строк и столбцов.

Подробнее (с примерами и иллюстрациями) данное построение приведено в [5] для $p = 2$ и в [8] для $p \in (1, +\infty)$.

Теорема 1 (см. теорему 8.1 из [5] при $p = 2$ и теорему 2.1 из [8] при $p \in (1, +\infty)$). *Предположим, что выполнены условия 1, 2 и 3, а соответствующие матрицы $R_s (s = 1, 2, \dots)$ и $R_{s_0} (s = s(r), r \in B)$ невырождены.*

Тогда существует множество $\gamma = \{\gamma_{ij}^r\}$ такое, что оператор R_Q отображает $\overset{\circ}{W}_p(Q)$ на $W_{p,\gamma}^1(Q)$ непрерывно и взаимнооднозначно.

Из теоремы 1 следует

Теорема 2. *Предположим, что выполнены условия 1, 2 и 3, а соответствующие матрицы $R_s (s = 1, 2, \dots)$ и $R_{s_0} (s = s(r), r \in B)$ невырождены. В этом случае функция $\overset{\circ}{w} \in W_{p,\gamma}^1(Q)$ является обобщенным решением задачи (2), (1) тогда и только тогда, когда $w = R_Q u$, $u \in \overset{\circ}{W}_p(Q)$ и для любого $\xi \in \overset{\circ}{W}_p(Q)$ справедливо интегральное тождество*

$$\sum_{1 \leq i \leq n_Q} \int_Q A_i(x, R_Q u, \nabla R_Q u) \partial_i \xi dx + \int_Q A_0(x, R_Q u, \nabla R_Q u) \xi dx = \int_Q f \xi dx. \tag{7}$$

Таким образом, чтобы найти обобщенное решение дифференциального уравнения (3) с нелокальным и краевыми условиями (2) необходимо найти обобщенное решение дифференциально–разностного уравнения

$$A_R u = A R_Q u = f \tag{8}$$

с краевым условием Дирихле

$$u|_{\partial Q} = 0. \tag{9}$$

Определение 4. Назовем $u \in \overset{\circ}{W}_p(Q)$ обобщенным решением задачи (8), (9), если для любого $\xi \in \overset{\circ}{W}_p(Q)$ справедливо интегральное тождество (7).

2. СВОЙСТВА ДИФФЕРЕНЦИАЛЬНО–РАЗНОСТНОГО ОПЕРАТОРА

Для доказательства разрешимости дифференциально–разностного уравнения (8), (9) необходимо изучить свойства дифференциально–разностного оператора. Важным элементом рассмотрения дифференциально–разностного оператора является изучение его свойств на всех подобластях одного класса разбиения области одновременно.

Определим изоморфизм пространств $U_s : L_p \left(\bigcup_l Q_{sl} \right) \rightarrow L_p^N(Q_{s1})$ по формуле $(U_s u)_l(x) = u(x + h_{sl})$ для всех $x \in Q_{s1}$, где $l = 1, \dots, N = N(s)$, $Q_{s1} + h_{sl} = Q_{sl}$ и $L_p^N(Q_{s1}) = \prod_l L_p(Q_{sl})$. Пространства U_s рефлексивны и банаховы.

Лемма 1 (см. §8 из [5] при $p = 2$ и [11] при $p \in (1, +\infty)$). *Для всех $u \in W_p^1(Q)$ имеем $R_Q u \in W_p^1(Q_{sl})$ и*

$$\|R_Q u\|_{W_p^1(Q_{sl})} \leq c_1 \sum_{j=1}^{N(s)} \|u\|_{W_p^1(Q_{sj})} \quad (s = 1, 2, \dots; l = 1, \dots, N(s)). \tag{10}$$

Причем оператор $R_{Q_s} : L_p^{N(s)}(Q_{s1}) \rightarrow L_p^{N(s)}(Q_{s1})$, заданный соотношением

$$R_{Q_s} = U_s R_Q U_s^{-1}, \tag{11}$$

есть оператор умножения на матрицу R_s порядка $N(s) \times N(s)$, заданную в (5).

Более того, если $\det R_{s_v} \neq 0$ ($v = 1, \dots, n_1$), то существует обратный оператор $R_Q^{-1} : L_p(Q) \rightarrow L_p(Q)$ и для всех $w \in W_p^1(Q)$ справедливы включения $R_Q^{-1} w \in W_p^1(Q_{sl})$, при этом

$$\|R_Q^{-1}w\|_{W_p^1(Q_{sl})} \leq c_2 \sum_{j=1}^{N(s)} \|w\|_{W_p^1(Q_{sj})} \quad (s = 1, 2, \dots; l = 1, \dots, N(s)). \tag{12}$$

Здесь константы $c_1, c_2 > 0$ не зависят от индекса s и функции u , причем

$$R_Q^{-1} = \sum_s U_s^{-1} R_s^{-1} U_s P_s. \tag{13}$$

Будут использованы матрицы $\zeta \in R^{N(s) \times (n+1)}$:

$$\zeta := \begin{pmatrix} \zeta_{10} & \zeta_{11} & \zeta_{12} & \dots & \zeta_{1n} \\ \zeta_{20} & \zeta_{21} & \zeta_{22} & \dots & \zeta_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \zeta_{N(s)0} & \zeta_{N(s)1} & \zeta_{N(s)2} & \dots & \zeta_{N(s)n} \end{pmatrix}.$$

Через ζ_i будем обозначать i -й столбец матрицы ζ , через ζ_l будем обозначать l -ю строку матрицы ζ .

Сформулируем условия на дифференциальные коэффициенты и матрицы R_s . Заметим, что условия роста можно уточнить в зависимости от соотношения степени интегрируемости p и размерности n .

Пусть справедливы следующие условия:

A0) R_s невырожденны;

A1) условие интегрируемости или условие роста: функции $A_i(x, \xi), i = 0, 1, \dots, n$, являются функциями типа Каратеодори (т.е. измеримы по $x \in Q$ и непрерывны по остальным переменным для п.в. $x \in Q$), а также удовлетворяют оценкам роста

$$|A_i(x, \xi)| \leq g_0(x) + c_1 |\xi_0|^{p'-1} + c_1 \sum_{1 \leq j \leq n} |\xi_j|^{p'-1}, \quad i = 1, \dots, n, \tag{14}$$

$$|A_0(x, \xi)| \leq g_0(x) + c_1 \sum_{0 \leq j \leq n} |\xi_j|^{p'-1}, \tag{15}$$

где $c_1 > 0$, $p' \in (1, p)$ и $g_0 \in L_q(Q)$;

A2) условие сильной эллиптичности: для всех s , п.в. $x \in \overline{Q_{s1}}$ и любых $\zeta, \eta \in R^{N(s) \times (n+1)}$ таких, что $\eta_{l0} = \zeta_{l0} \neq 0$ и $\eta \neq \zeta$, существует $\hat{\gamma} > 0$ такая, что

$$\begin{aligned} & \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} (A_i(x + h_{sl}, \zeta_l) - A_i(x + h_{sl}, \eta_l)) (R_s^{-1}(\zeta_i - \eta_i))_l \geq \\ & \geq \hat{\gamma} \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} |\zeta_{li} - \eta_{li}|^p; \end{aligned} \tag{16}$$

A3) условие локальной липшицевости/гёльдеровости функций A_i ($i = 1, \dots, n$) по ξ_0 и функции A_0 по ξ_j ($j = 0, 1, \dots, n$): существуют ρ ($\rho = 1$ при $p \in [2, \infty)$, $\rho \in (1 - 1/p, p - 1]$ при $p \in (1, 2)$) и $\varepsilon > 0$ такие, что для п.в. $x \in Q$ и любых $\xi, \delta \in R^{n+1}$, $|\delta| = \sum_{0 \leq j \leq n} |\delta_j| = |\delta_k| < \varepsilon$,

$$|A_i(x, \xi + \delta) - A_i(x, \xi)| \leq \left(\Psi |\xi_k + \delta_k|^{p-1-\rho} + \Psi \sum_{0 \leq j \leq n} |\xi_j|^{p-1-\rho} + g_\Psi(x) \right) |\delta_k|^p, \tag{17}$$

где $\Psi > 0$, $g_\Psi \in L_{q'}(Q)$, $q' = p / (p - 1 - \rho)$, $k = 0$ при $i = 1, \dots, n$ и $k = 0, 1, \dots, n$ при $i = 0$.

Определение 5. Оператор $A : X \rightarrow Y$ деминепрерывен, если из сходимости $u_m \rightarrow u$ в X следует слабая сходимость $Au_m \rightharpoonup Au$ в Y . Оператор $A : X \rightarrow Y$ ограничен, если образ ограниченного множества из X ограничен в Y ¹.

Лемма 2. Пусть коэффициенты оператора $A : W_p^1(Q) \rightarrow W_q^{-1}(Q)$, заданного в уравнении (2), удовлетворяют условию **A1**). Тогда оператор $A_R : W_p^1(Q) \rightarrow W_q^{-1}(Q)$ деминепрерывен и ограничен.

Доказательство. В силу леммы 1 линейный оператор $R_Q : W_p^1(Q) \rightarrow W_p^1(Q)$ ограничен. Линейный ограниченный оператор непрерывен. В силу условия **A1**) существуют $c_2 > 0$ и $g_1 \in L_q(Q)$ такие, что

$$|A_i(x, \xi)| \leq g_1(x) + c_2 \sum_{0 \leq j \leq n} |\xi_j|^{p-1}, \quad i = 0, 1, \dots, n,$$

т.е. дифференциальный оператор $A : W_p^1(Q) \rightarrow W_q^{-1}(Q)$ деминепрерывен и ограничен, см., например, [13, Гл. 1 §2]. Композиция $A_R = AR_Q$ является деминепрерывным, ограниченным оператором. Лемма доказана.

Определение 6. Оператор $A : \text{Dom}(A) \subset X \rightarrow Y^*$ с линейной (аффинной) областью определения $\text{Dom}(A)$ радиально непрерывен, если для любых $u, v \in \text{Dom}(A)$ и $y \in Y$ функция $\tau \mapsto \langle A(u + \tau v), y \rangle$ непрерывна.

Очевидно, что деминепрерывный оператор является радиально непрерывным.

Определение 7. Оператор $A : X \rightarrow X^*$ называется оператором с полуограниченной вариацией, если существует непрерывная функция C такая, что для всех $u, y \in X$ ($\|u\|_X \leq r_1, \|y\|_X \leq r_1$) справедливо неравенство

$$\langle Au - Ay, u - y \rangle \geq -C \left(r_1; \|u - y\|'_X \right), \quad (18)$$

где $\tau^{-1}C(r_1, \tau r_2) \rightarrow +0$ при $\tau \rightarrow 0$ для всех $r_1, r_2 > 0$, $\|\cdot\|'_X$ — компактная полунорма относительно $\|\cdot\|_X$.

При $X = W_p^1(Q)$ в качестве $\|\cdot\|'_X$ удобно рассматривать $\|\cdot\|_{L_p(Q)}$. $W_p^1(Q) \subset L_p(Q)$ компактно, см. [14, Гл. I, 8, пп. 2].

Операторы с полуограниченной вариацией были рассмотрены Ю. А. Дубинским, см. [15], он же предложил условие сильной эллиптичности для квазилинейных дифференциальных операторов. Ранее автором были исследованы квазилинейные дифференциально-разностные операторы, обладающие свойством полуограниченной вариацией. Как было доказано в 80-е, радиально непрерывные операторы с полуограниченной вариацией являются псевдомонотонными. При исследовании дифференциально-разностных операторов операторы с полуограниченной вариацией обладают следующим преимуществом: для дифференциально-разностных операторов достаточными условиями псевдомонотонности являются условие эллиптичности и условие коэрцитивности, каждое из которых выражено в виде оценки весовых сумм; для доказательства полуограниченной вариации будет использована одна весовая сумма — условие сильной эллиптичности; коэрцитивность дифференциально-разностного оператора будет доказана исходя из условия сильной эллиптичности и условия роста, накладываемого на дифференциальный оператор.

Лемма 3. Пусть справедливы условия **A0)–A3)**. Тогда дифференциально-разностный оператор $A_R : W_p^1(Q) \rightarrow W_q^{-1}(Q)$ обладает полуограниченной вариацией.

Доказательство. Обозначим $w = R_Q u$ и $v = R_Q y$, $u, y \in W_p^1(Q)$, причем в силу невырожденности оператора R_Q существует обратный оператор $R_Q^{-1} : L_p(Q) \rightarrow L_p(Q)$, см. лемму 1. По определению оператора A_R и в силу формулы (13)

¹ в определениях мы рассматриваем абстрактные банаховы пространства X и Y

$$\begin{aligned}
& \langle A_R u - A_R v, u - v \rangle = \\
& = \sum_{0 \leq i \leq n} \int_{\Omega_T} (A_i(x, R_Q u, \nabla R_Q u) - A_i(x, R_Q v, \nabla R_Q v)) \partial_i (u - v) dx = \\
& = \sum_{0 \leq i \leq n} \int_{\Omega_T} (A_i(x, w, \nabla w) - A_i(x, v, \nabla v)) R_Q^{-1} \partial_i (w - v) dx = \\
& = \sum_s \int_{\bigcup_{l=1}^s Q_{sl}} \sum_{0 \leq i \leq n} P_s (A_i(x, w, \nabla w) - A_i(x, v, \nabla v)) U_s^{-1} R_s^{-1} U_s P_s \partial_i (w - v) dx = \\
& = \sum_s \int_{Q_{s1}} \sum_{0 \leq i \leq n} (U_s P_s (A_i(x, w, \nabla w) - A_i(x, v, \nabla v)), R_s^{-1} U_s P_s \partial_i (w - v)) dx = \\
& = \sum_s \int_{Q_{s1}} (I_{s1} + I_{s2} + I_{s3}) dx,
\end{aligned} \tag{19}$$

где (\cdot, \cdot) — скалярное произведение в $R^{N(s)}$. Очевидно, что при $u(x) = v(x)$ для почти всех $x \in Q$ значение данного интеграла неотрицательно. Поэтому, не ограничивая общности, будем считать, что $u(x) \neq v(x)$ для почти всех $x \in Q$. Очевидно, что при этом $w(x) \neq v(x)$ и существует $\lambda \in [0, 1]$ такое, что $\lambda w(x) + (1 - \lambda)v(x) \neq 0$, возможно $\lambda = \lambda(x)$.

Введем матрицы порядка $N(s) \times (n + 1)$

$$\zeta = (U_s P_s w, U_s P_s \partial_1 w, \dots, U_s P_s \partial_n w), \quad \eta = (U_s P_s v, U_s P_s \partial_1 v, \dots, U_s P_s \partial_n v),$$

а также матрицы $\hat{\zeta}$ и $\hat{\eta}$ такие, что

$$\hat{\zeta}_i = \zeta_i, \quad \hat{\eta}_i = \eta_i \quad \forall i = 1, \dots, n, \tag{20}$$

$$\hat{\zeta}_{l0} = \hat{\eta}_{l0} = \lambda \zeta_{l0} + (1 - \lambda) \eta_{l0} \quad \forall l = 1, \dots, N(s). \tag{21}$$

По построению, $\hat{\zeta}_{l0} = \hat{\eta}_{l0} \neq 0$. В то же время $\hat{\zeta}_i - \hat{\eta}_i \equiv \zeta_i - \eta_i \quad \forall i = 1, \dots, n$. Сначала оценим часть подынтегральной суммы правой части (19):

$$\begin{aligned}
I_{s1} + I_{s2} & = \sum_{1 \leq i \leq n} (U_s P_s (A_i(x, w, \nabla w) - A_i(x, v, \nabla v)), R_s^{-1} U_s P_s \partial_i (w - v)) = \\
& = \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} (A_i(x + h_{sl}, \zeta_l) - A_i(x + h_{sl}, \eta_l)) (R_s^{-1} (\zeta_i - \eta_i))_l = \\
& = \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} (A_i(x + h_{sl}, \hat{\zeta}_l) - A_i(x + h_{sl}, \hat{\eta}_l)) (R_s^{-1} (\hat{\zeta}_i - \hat{\eta}_i))_l + \\
& + \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} (A_i(x + h_{sl}, \zeta_l) - A_i(x + h_{sl}, \hat{\zeta}_l)) (R_s^{-1} (\zeta_i - \eta_i))_l + \\
& + \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} (A_i(x + h_{sl}, \hat{\eta}_l) - A_i(x + h_{sl}, \eta_l)) (R_s^{-1} (\zeta_i - \eta_i))_l.
\end{aligned} \tag{22}$$

Первую сумму правой части (22) оценим с помощью условия сильной эллиптичности (16):

$$\begin{aligned}
I_{s1} & = \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} (A_i(x + h_{sl}, \hat{\zeta}_l) - A_i(x + h_{sl}, \hat{\eta}_l)) (R_s^{-1} (\hat{\zeta}_i - \hat{\eta}_i))_l \geq \\
& \geq \hat{Y} \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} |\hat{\zeta}_{li} - \hat{\eta}_{li}|^p = \hat{Y} \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} |\partial_i w(x + h_{sl}) - \partial_i v(x + h_{sl})|^p.
\end{aligned}$$

То есть

$$\sum_s \int_{Q_{s1}} I_{s1} dx \geq \hat{Y} \|w - v\|_{W_p(Q)}^p. \quad (23)$$

Рассмотрим вторую и третью суммы правой части (22)²:

$$\begin{aligned} I_{s2} &\leq \left| \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} \left(A_i(x + h_{sl}, \zeta_l) - A_i(x + h_{sl}, \hat{\zeta}_l) \right) \left(R_s^{-1}(\zeta_i - \eta_i) \right)_l \right| + \\ &+ \left| \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} \left(A_i(x + h_{sl}, \hat{\eta}_l) - A_i(x + h_{sl}, \eta_l) \right) \left(R_s^{-1}(\zeta_i - \eta_i) \right)_l \right| \leq \\ &\leq \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} \left| A_i(x + h_{sl}, \zeta_l) - A_i(x + h_{sl}, \hat{\zeta}_l) \right| \left| \left(R_s^{-1}(\zeta_i - \eta_i) \right)_l \right| + \\ &+ \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} \left| A_i(x + h_{sl}, \hat{\eta}_l) - A_i(x + h_{sl}, \eta_l) \right| \left| \left(R_s^{-1}(\zeta_i - \eta_i) \right)_l \right| \end{aligned} \quad (24)$$

Исходя из условия **A3**, воспользуемся оценкой (17):

$$\begin{aligned} &\left| A_i(x + h_{sl}, \hat{\zeta}_l) - A_i(x + h_{sl}, \zeta_l) \right| \leq \\ &\leq \left(\Psi |\hat{\zeta}_{l0}|^{p-1-p} + \Psi \sum_{0 \leq k \leq n} |\zeta_{lk}|^{p-1-p} + g_\Psi(x + h_{sl}) \right) |\zeta_{l0} - \hat{\zeta}_{l0}|^p. \end{aligned} \quad (25)$$

Аналогично для второй группы слагаемых правой части (24) имеем

$$\begin{aligned} &\left| A_i(x + h_{sl}, \hat{\eta}_l) - A_i(x + h_{sl}, \eta_l) \right| \leq \\ &\leq \left(\Psi |\hat{\eta}_{l0}|^{p-1-p} + \Psi \sum_{0 \leq k \leq n} |\eta_{lk}|^{p-1-p} + g_\Psi(x + h_{sl}) \right) |\eta_{l0} - \hat{\eta}_{l0}|^p. \end{aligned} \quad (26)$$

Подставим (25) и (26) в (24); учитывая, что $\eta_{l0} - \hat{\eta}_{l0} = \lambda(\eta_{l0} - \zeta_{l0})$ и $\zeta_{l0} - \hat{\zeta}_{l0} = (1 - \lambda)(\zeta_{l0} - \eta_{l0})$, см. (21), получаем оценку

$$\begin{aligned} I_{s2} &\leq \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} \left(((\lambda^p + (1 - \lambda)^p) \Psi |\lambda \zeta_{l0} + (1 - \lambda) \eta_{l0}|^{p-1-p} + (1 - \lambda)^p \Psi \sum_{0 \leq k \leq n} |\zeta_{lk}|^{p-1-p} + \right. \\ &+ \lambda^p \Psi \sum_{0 \leq k \leq n} |\eta_{lk}|^{p-1-p} + ((\lambda^p + (1 - \lambda)^p) g_\Psi(x + h_{sl})) |\zeta_{l0} - \eta_{l0}|^p \left. \right) \left| \left(R_s^{-1}(\zeta_i - \eta_i) \right)_l \right| \leq \\ &\leq \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} \left(2^{1-p} \Psi |\lambda \zeta_{l0} + (1 - \lambda) \eta_{l0}|^{p-1-p} + \Psi \sum_{0 \leq k \leq n} |\zeta_{lk}|^{p-1-p} + \right. \\ &+ \Psi \sum_{0 \leq k \leq n} |\eta_{lk}|^{p-1-p} + 2^{1-p} g_\Psi(x + h_{sl}) \left. \right) |\zeta_{l0} - \eta_{l0}|^p \left| \left(R_s^{-1}(\zeta_i - \eta_i) \right)_l \right|. \end{aligned} \quad (27)$$

Здесь учтено, что $\lambda^p + (1 - \lambda)^p \leq 2^{1-p}$ для всех $\lambda \in [0, 1]$, а также $\lambda^p \leq 1$ и $(1 - \lambda)^p \leq 1$. Заметим, что $|\lambda \zeta_{l0} + (1 - \lambda) \eta_{l0}|^\alpha \leq |\zeta_{l0}|^\alpha + |\eta_{l0}|^\alpha$ для всех $\lambda \in [0, 1]$ при $\alpha \geq 0$, а также в силу ограниченности и невырожденности матриц R_s^{-1} имеет вид

² пока будем считать, что $\rho < p - 1$, ситуацию $\rho = p - 1$ поясним в конце доказательства

$$\sum_{1 \leq i \leq n} \left| \left(R_s^{-1}(\zeta_i - \eta_i) \right)_l \right| \leq c_3 \sum_{1 \leq i \leq n} |\zeta_{li} - \eta_{li}|$$

для некоторого $c_3 > 0$. Тогда получим

$$\begin{aligned} I_{s2} \leq & 2^{1-p} c_3 \sum_{1 \leq l \leq N(s)} (\Psi |\zeta_{l0}|^{p-1-p} + \Psi |\eta_{l0}|^{p-1-p} + \Psi \sum_{0 \leq k \leq n} |\zeta_{lk}|^{p-1-p} + \\ & + \Psi \sum_{0 \leq k \leq n} |\eta_{lk}|^{p-1-p} + g_\Psi(x + h_{sl})) |\zeta_{l0} - \eta_{l0}|^p \sum_{1 \leq i \leq n} |\zeta_{li} - \eta_{li}|. \end{aligned} \tag{28}$$

Перейдем к переменным w и v :

$$\begin{aligned} \sum_s \int_{Q_{s1}} I_{s2} dx \leq & c_4 (2\Psi \|w\|_{L_p(Q)}^{p-1-p} + 2\Psi \|v\|_{L_p(Q)}^{p-1-p} + \Psi \sum_{1 \leq k \leq n} \|\partial_i w\|_{L_p(Q)}^{p-1-p} + \\ & + \Psi \sum_{1 \leq k \leq n} \|\partial_i v\|_{L_p(Q)}^{p-1-p} + \|g_\Psi\|_{L_{q'}(Q)}) \|w - v\|_{L_p(Q)}^p \sum_{1 \leq i \leq n} \|\partial_i w - \partial_i v\|_{L_p(Q)} \leq \\ \leq & c_4 \left(c_5 r_1^{p-1-p} + \|g_\Psi\|_{L_{q'}(Q)} \right) \|w - v\|_{L_p(Q)}^p \sum_{1 \leq i \leq n} \|\partial_i w - \partial_i v\|_{L_p(Q)}, \end{aligned} \tag{29}$$

где $\|u\|_{W_p^1(Q)} \leq r_1$ и $\|y\|_{W_p^1(Q)} \leq r_1$. В этой оценке учтено, что в силу ограниченности оператора R_Q $\|w\|_{L_p(Q)} = \|R_Q u\|_{L_p(Q)} \leq c_6 \|u\|_{L_p(Q)}$ и согласно неравенству Фридрихса $\|u\|_{L_p(Q)} \leq c_7 \|u\|_{W_p^1(Q)}$, для $v = R_Q y$ оценки аналогичны. Т.е. $\|w\|_{L_p(Q)} \leq c_6 c_7 \|u\|_{W_p^1(Q)} \leq c_6 c_7 r_1$ и $\|v\|_{L_p(Q)} \leq c_6 c_7 \|y\|_{W_p^1(Q)} \leq c_6 c_7 r_1$, $c_5 = 4\Psi(c_6 c_7)^{p-1-p} + 2\Psi c_6^{p-1-p}$. Для сокращения записи введем функцию $\hat{c}_1(r_1) = c_4 \left(c_5 r_1^{p-2} + \|g_\Psi\|_{L_{q'}(Q)} \right)$. В силу неравенства Юнга из (29) следует, что

$$\begin{aligned} \sum_s \int_{Q_{s1}} I_{s2} dx \leq & \frac{\hat{Y}}{p} \|w - v\|_{W_p^1(Q)}^p + \frac{1}{q} \left(\frac{\hat{c}_1(r_1)}{\hat{Y}} \right)^q \|w - v\|_{L_p(Q)}^{pq} \leq \\ \leq & \frac{\hat{Y}}{p} \|w - v\|_{W_p^1(Q)}^p + \frac{c_6^{pq}}{q} \left(\frac{\hat{c}_1(r_1)}{\hat{Y}} \right)^q \|u - y\|_{L_p(Q)}^{pq}. \end{aligned} \tag{30}$$

Осталось оценить слагаемое при $i = 0$ в подынтегральной сумме правой части (19):

$$\begin{aligned} I_{s3} \leq & \left| (U_s P_s (A_0(x, w, \nabla w) - A_0(x, v, \nabla v)), R_s^{-1} U_s P_s (w - v)) \right| = \\ = & \left| \sum_{1 \leq l \leq N(s)} (A_0(x + h_{sl}, \zeta_l) - A_0(x + h_{sl}, \eta_l)) \left(R_s^{-1}(\zeta_0 - \eta_0) \right)_l \right| \leq \\ \leq & \sum_{1 \leq l \leq N(s)} |A_0(x + h_{sl}, \zeta_l) - A_0(x + h_{sl}, \eta_l)| \left| \left(R_s^{-1}(\zeta_0 - \eta_0) \right)_l \right|. \end{aligned} \tag{31}$$

Исходя из условия А3), подставим оценку (17) в (31) покоординатно, аналогично выводу формулы (28) из (26). Так как

$$\begin{aligned} & |A_0(x + h_{sl}, \zeta_l) - A_0(x + h_{sl}, \eta_l)| \leq \\ & \leq |A_0(x + h_{sl}, \zeta_l) - A_0(x + h_{sl}, \eta_{l0}, \zeta_{l1}, \dots, \zeta_{ln})| + \\ & + |A_0(x + h_{sl}, \eta_{l0}, \zeta_{l1}, \dots, \zeta_{ln}) - A_0(x + h_{sl}, \eta_{l0}, \eta_{l1}, \zeta_{l2}, \dots, \zeta_{ln})| + \\ & + \dots + |A_0(x + h_{sl}, \eta_{l0}, \dots, \eta_{l,n-1}, \zeta_{ln}) - A_0(x + h_{sl}, \eta_l)| \leq \\ & \leq \sum_{0 \leq k \leq n} \left(\Psi \sum_{0 \leq j \leq k} |\eta_{lj}|^{p-1-p} + \Psi \sum_{k \leq j \leq n} |\zeta_{lj}|^{p-1-p} + g_\Psi \right) |\zeta_{lk} - \eta_{lk}|^p \leq \\ & \leq \left(\Psi \sum_{0 \leq j \leq n} |\zeta_{lj}|^{p-1-p} + \Psi \sum_{0 \leq j \leq n} |\eta_{lj}|^{p-1-p} + g_\Psi \right) \sum_{0 \leq k \leq n} |\zeta_{lk} - \eta_{lk}|^p, \end{aligned}$$

то

$$\begin{aligned} I_{s3} & \leq \sum_{1 \leq l \leq N(s)} \left(\widehat{\Psi} \sum_{0 \leq j \leq n} (|\zeta_{lj}|^{p-1-p} + |\eta_{lj}|^{p-1-p}) + g_\Psi(x + h_{sl}) \right) \times \\ & \times \sum_{0 \leq k \leq n} |\zeta_{lk} - \eta_{lk}|^p \left| (R_s^{-1}(\zeta_{.0} - \eta_{.0}))_l \right|. \end{aligned}$$

Вернемся к функциям w, v, u и y и воспользуемся неравенством Гёльдера:

$$\begin{aligned} & \sum_s \int_{Q_{s1}} I_{s3} dx \leq \sum_s \sum_{1 \leq l \leq N(s)} \sum_{0 \leq k \leq n} \int_{Q_{s1}} \Psi \left(\sum_{0 \leq j \leq n} |\partial_j v(x + h_{sl})|^{p-1-p} + \right. \\ & + \sum_{0 \leq j \leq n} |\partial_j w(x + h_{sl})|^{p-1-p} + g_\Psi(x + h_{sl}) \left. \right) |\partial_k w(x + h_{sl}) - \partial_k v(x + h_{sl})|^p \times \\ & \times |u(x + h_{sl}) - y(x + h_{sl})| dx = \\ & = \sum_{0 \leq k \leq n} \int_Q \left(\Psi \sum_{0 \leq j \leq n} |\partial_j v|^{p-1-p} + \Psi \sum_{0 \leq j \leq n} |\partial_j w|^{p-1-p} + g_\Psi \right) |\partial_k w - \partial_k v|^p |u - y| dx \leq \\ & \leq \left\| \Psi \sum_{0 \leq j \leq n} (|\partial_j v|^{p-1-p} + |\partial_j w|^{p-1-p}) + g_\Psi \right\|_{L_{q'}(Q)} \sum_{0 \leq k \leq n} \|\partial_k(w - v)\|_{L_p(Q)}^p \|u - y\|_{L_p(Q)}. \end{aligned}$$

Заметим, что

$$\begin{aligned} & \left\| \Psi \sum_{0 \leq j \leq n} (|\partial_j v|^{p-1-p} + |\partial_j w|^{p-1-p}) + g_\Psi \right\|_{L_{q'}(Q)} \leq \\ & \leq \Psi \left(\|v\|_{W_p^1(Q)}^{p-1-p} + \|w\|_{W_p^1(Q)}^{p-1-p} \right) + \|g_\Psi\|_{L_{q'}(Q)} \leq c_8 r_1^{p-1-p} + \|g_\Psi\|_{L_{q'}(Q)} =: \hat{c}_2(r_1). \end{aligned}$$

Здесь, как и выше, были учтены оценки $\|w\|_{L_p(Q)} \leq c_6 c_7, \|u\|_{W_p^1(Q)} \leq c_6 c_7 r_1$ и $\|v\|_{L_p(Q)} \leq c_6 c_7, \|y\|_{W_p^1(Q)} \leq c_6 c_7 r_1$, т.е. $c_8 = 2\Psi((c_6 c_7)^{p-1-p} + c_6^{p-1-p})$. Осталось воспользоваться известным не-

равенством $ab \leq \frac{(\varepsilon a)^p}{p} + \frac{1}{q} \left(\frac{b}{\varepsilon}\right)^q$:

$$\begin{aligned}
 & \sum_s \int_{Q_{s1}} I_{s3} dx \leq \hat{c}_2(r_1) \|u - y\|_{L_p(Q)} \sum_{0 \leq j \leq n} \|\partial_j w - \partial_j v\|_{L_p(Q)}^p = \\
 & = \hat{c}_2(r_1) \|u - y\|_{L_p(Q)} \|w - v\|_{L_p(Q)}^p + \hat{c}_2(r) \|u - y\|_{L_p(Q)} \sum_{1 \leq j \leq n} \|\partial_j w - \partial_j v\|_{L_p(Q)}^p \leq \\
 & \leq c_6^p \hat{c}_2(r_1) \|u - y\|_{L_p(Q)}^{1+\rho} + \frac{\hat{\Upsilon} \rho}{p} \|w - v\|_{W_p^1(Q)}^p + \frac{p - \rho}{p} \left(\frac{\hat{c}_2(r_1)}{\hat{\Upsilon}} \right)^{\frac{p}{p-\rho}} \|u - y\|_{L_p(Q)}^{\frac{p}{p-\rho}}. \tag{32}
 \end{aligned}$$

Подставляя (23), (30) и (32) в (19), получим

$$\begin{aligned}
 \langle A_R u - A_R y, u - y \rangle & \geq \sum_s \int_{Q_{s1}} |I_{s1} - |I_{s2}| - |I_{s3}| dx \geq \\
 & \geq \left(\hat{\Upsilon} - \frac{\hat{\Upsilon}}{p} - \frac{\rho \hat{\Upsilon}}{p} \right) \|w - v\|_{W_p^1(Q)}^p - \frac{c_6^{\rho q} \left(\hat{c}_1(r_1) \right)^q}{q \left(\hat{\Upsilon} \right)^q} \|u - y\|_{L_p(Q)}^{\rho q} - \\
 & - c_6^p \hat{c}_2(r_1) \|u - y\|_{L_p(Q)}^{1+\rho} - \frac{p - \rho}{p} \left(\frac{\hat{c}_2(r_1)}{\hat{\Upsilon}} \right)^{\frac{p}{p-\rho}} \|u - y\|_{L_p(Q)}^{\frac{p}{p-\rho}} = \\
 & = \frac{\hat{\Upsilon}(p - 1 - \rho)}{p} \|w - v\|_{W_p^1(Q)}^p - C(r_1; \|u - y\|_{L_p(Q)}). \tag{33}
 \end{aligned}$$

Поскольку $\left(1 - \frac{1 + \rho}{p}\right) \hat{\Upsilon} \geq 0$, $\rho q = \frac{\rho p}{p - 1} > 1$, $1 + \rho > 1$, $\frac{p}{p - \rho} > 1$ и $\overset{\circ}{W}_p(Q) \subset L_p(Q)$ компактно, то A_R — оператор с полуограниченной вариацией.

Заметим, что при $\rho = p - 1$, что возможно при $p \in (1, 2]$, доказательство существенно упрощается, поскольку из **A3**) следует, что $g_\Psi \in L_\infty(Q)$, а функции A_i удовлетворяют оценке

$$|A_i(x, \xi + \delta) - A_i(x, \xi)| \leq \hat{g}_\Psi(x) |\delta_k|^{p-1},$$

где $|\hat{g}_\Psi(x)| \leq |g_\Psi(x)| + 2\Psi(n + 1)$, т.е. $\hat{g}_\Psi \in L_\infty(Q)$. Следовательно, в оценках (29) и (32) мы получим

$$\begin{aligned}
 \sum_s \int_{Q_{s1}} I_{s2} dx & \leq 2^{2-p} \|\hat{g}_\Psi\|_{L_\infty(Q)} \|w - v\|_{L_p(Q)}^{p-1} \sum_{1 \leq i \leq n} \|\partial_i w - \partial_i v\|_{L_p(Q)}, \\
 \sum_s \int_{Q_{s1}} I_{s3} dx & \leq 2^{2-p} \|\hat{g}_\Psi\|_{L_\infty(Q)} \|u - y\|_{L_p(Q)} \sum_{0 \leq i \leq n} \|\partial_i w - \partial_i v\|_{L_p(Q)}^{p-1}.
 \end{aligned}$$

Таким образом,

$$\begin{aligned}
 \langle A_R u - A_R y, u - y \rangle & \geq \sum_s \int_{Q_{s1}} |I_{s1} - |I_{s2}| - |I_{s3}| dx \geq \\
 & \geq \left(\hat{\Upsilon} - \frac{\hat{\Upsilon}}{p} - \frac{\hat{\Upsilon}}{q} \right) \|w - v\|_{\overset{\circ}{W}_p(Q)}^p - \frac{c_6^p \left(2^{2-p} \|\hat{g}_\Psi\|_{L_\infty(Q)} \right)^q}{\hat{\Upsilon}^q} \|u - y\|_{L_p(Q)}^p - \\
 & - c_6^{p-1} 2^{2-p} \|\hat{g}_\Psi\|_{L_\infty(Q)} \|u - y\|_{L_p(Q)}^p - \frac{1}{q} \left(\frac{2^{2-p} \|\hat{g}_\Psi\|_{L_\infty(Q)}}{\hat{\Upsilon}} \right)^q \|u - y\|_{L_p(Q)}^p = \\
 & = - \left(\frac{c_6^p + 1}{q} \left(\frac{2^{2-p} \|\hat{g}_\Psi\|_{L_\infty(Q)}}{\hat{\Upsilon}} \right)^q + c_6^{p-1} 2^{2-p} \|\hat{g}_\Psi\|_{L_\infty(Q)} \right) \|u - y\|_{L_p(Q)}^p
 \end{aligned}$$

$$= -C(r_1; \|u - y\|_{L_p(Q)}).$$
(34)

Так как $p > 1$ и $\overset{\circ}{W}_p^1(Q) \subset L_p(Q)$ компактно, то A_R — оператор с полуограниченной вариацией.

Замечание 1. При $p > 2$ оценка (33) имеет вид

$$\begin{aligned} \langle A_R u - A_R y, u - y \rangle &\geq \left(1 - \frac{2}{p}\right) \hat{\Upsilon} \|w - v\|_{\overset{\circ}{W}_p^1(Q)}^p - \frac{1}{q} \left(\frac{c_6 \hat{c}_1(r_1)}{\hat{\Upsilon}}\right)^q \|u - y\|_{L_p(Q)}^q - \\ &\quad - c_6 \hat{c}_2(r_1) \|u - y\|_{L_p(Q)}^2 - \frac{1}{q} \left(\frac{\hat{c}_2(r_1)}{\hat{\Upsilon}}\right)^q \|u - y\|_{L_p(Q)}^q \geq \\ &\geq c_9 \frac{\hat{\Upsilon}(p-2)}{p} \|u - y\|_{\overset{\circ}{W}_p^1(Q)}^p - C(r_1; \|u - y\|_{L_p(Q)}), \end{aligned}$$

где использована оценка $\|R_Q u\|_{\overset{\circ}{W}_p^1(Q)}^p \geq c_9 \|u\|_{\overset{\circ}{W}_p^1(Q)}^p$, справедливая в силу невырожденности линейного оператора R_Q .

В литературе, посвященной дифференциальным уравнениям в частных производных, более исследованным является случай, когда $p \geq 2$ и справедливо условие локальной липшицевости ($\rho = 1$) или дифференцируемости. При $p \in (1, 2)$ необходимо установить $\rho \in (0, 1)$ для интегрируемости рассматриваемых функций. Логично назвать это условие условием локальной гёльдеровости аналогично условию локальной липшицевости. Покажем, что из классического определения локальной липшицевости следует оценка (15).

Лемма 4. Пусть $p \in [2, \infty)$ и существуют функция типа Каратеодори $\tilde{\Psi}(x, \xi)$ и $\varepsilon > 0$ такие, что для п.в. $x \in Q$ и любых $\xi, \delta \in R^{n+1}$, $|\delta| = \sum_{0 \leq j \leq n} |\delta_j| = |\delta_k| < \varepsilon$,

$$|A_i(x, \xi + \delta) - A_i(x, \xi)| \leq \tilde{\Psi}(x, \xi) |\delta_k|, \tag{35}$$

$$\tilde{\Psi}(x, \xi) \leq g_\Psi(x) + \Psi \sum_{0 \leq j \leq n} |\xi_j|^{p-2}, \tag{36}$$

где $\Psi > 0$, $g_\Psi \in L_{q'}(Q)$, $q' = p / (p - 2)$. Тогда справедлива оценка (15) с $\rho = 1$.

Доказательство. Очевидно, что существуют $n_2 \in N$ и $\delta \in R^{n+1}$ ($|\delta| < \varepsilon$) такие, что $\xi = \hat{\xi} + n_2 \delta$. Тогда

$$\begin{aligned} |A_i(x, \hat{\xi}) - A_i(x, \xi)| &= \left| \sum_{1 \leq k \leq n_2} (A_i(x, \hat{\xi} + k\delta) - A_i(x, \hat{\xi} + (k-1)\delta)) \right| \leq \\ &\leq \sum_{1 \leq k \leq n_2} |A_i(x, \hat{\xi} + k\delta) - A_i(x, \hat{\xi} + (k-1)\delta)|. \end{aligned}$$

Воспользуемся (35):

$$\begin{aligned} |A_i(x, \hat{\xi} + k\delta) - A_i(x, \hat{\xi} + (k-1)\delta)| &\leq \tilde{\Psi}(x, \hat{\xi} + (k-1)\delta) |\delta| = \\ &= \tilde{\Psi}\left(x, \hat{\xi} + \frac{k-1}{n_2}(\xi - \hat{\xi})\right) \frac{|\xi - \hat{\xi}|}{n_2}. \end{aligned}$$

То есть

$$|A_i(x, \hat{\xi}) - A_i(x, \xi)| \leq \frac{1}{n_2} \sum_{1 \leq k \leq n_2} \tilde{\Psi}\left(x, \hat{\xi} + \frac{k-1}{n_2}(\xi - \hat{\xi})\right) |\xi - \hat{\xi}|.$$

Переходя к пределу при $n_2 \rightarrow \infty$, в силу непрерывности функции Ψ получаем, что

$$|A_i(x, \hat{\xi}) - A_i(x, \xi)| \leq \int_0^1 \tilde{\Psi}(x, \hat{\xi} + \tau(\xi - \hat{\xi})) d\tau |\xi - \hat{\xi}|. \quad (37)$$

Оценим теперь интеграл по τ из формулы (37). Для этого подставим в (37) оценку (36) и используем известное неравенство $\int_0^1 |a + \tau(b - a)|^\alpha d\tau \leq |a|^\alpha + |b|^\alpha$ при $\alpha \geq 0$. Поскольку мы считаем, что ξ и $\hat{\xi}$ отличаются только в одной координате j , то

$$\begin{aligned} \int_0^1 \tilde{\Psi}(x, \hat{\xi} + \tau(\xi - \hat{\xi})) d\tau &\leq \int_0^1 \left(\Psi \sum_{0 \leq k \leq n} |\hat{\xi}_k + \tau(\xi_k - \hat{\xi}_k)|^{p-2} + g_\Psi(x) \right) d\tau = \\ &= \Psi \int_0^1 |\hat{\xi}_j + \tau(\xi_j - \hat{\xi}_j)|^{p-2} d\tau + \Psi \sum_{k \neq j} |\xi_k|^{p-2} + g_\Psi(x) \leq \\ &\leq \Psi |\hat{\xi}_j|^{p-2} + \Psi \sum_{0 \leq k \leq n} |\xi_k|^{p-2} + g_\Psi(x). \end{aligned}$$

Подставив это выражение в (37), получаем оценку (15).

Определение 8. Оператор $A : X \rightarrow X^*$ называется *коэрцитивным*, если существует $u_0 \in X$ такая, что

$$\|u\|_X^{-1} \langle Au, u - u_0 \rangle \rightarrow \infty \text{ при } \|u\|_X \rightarrow \infty. \quad (38)$$

Лемма 5. Пусть справедливы условия **A0)–A2)**. Тогда дифференциально–разностный оператор $A_R : W_p^1(Q) \rightarrow W_q^{-1}(Q)$ коэрцитивен.

Доказательство. Пусть $w = R_Q u, u \in W_p^1(Q)$, причем в силу невырожденности оператора R_Q существует обратный оператор $R_Q^{-1} : L_p(Q) \rightarrow L_p(Q)$, см. лемму 1. По определению оператора A_R и в силу формулы (13)

$$\begin{aligned} \langle A_R u, u \rangle &= \sum_{0 \leq i \leq n} \int_{\Omega_T} A_i(x, R_Q u, \nabla R_Q u) \partial_i u dx = \sum_{0 \leq i \leq n} \int_{\Omega_T} A_i(x, w, \nabla w) R_Q^{-1} \partial_i w dx = \\ &= \sum_s \int_{\bigcup_l Q_{sl}} \sum_{0 \leq i \leq n} P_s A_i(x, w, \nabla w) U_s^{-1} R_s^{-1} U_s P_s \partial_i w dx = \\ &= \sum_s \int_{Q_{s1}} \sum_{1 \leq i \leq n} (U_s P_s (A_i(x, w, \nabla w) - A_i(x, w, \nabla 0)), R_s^{-1} U_s P_s \partial_i (w - 0)) dx + \\ &\quad + \sum_s \int_{Q_{s1}} \sum_{1 \leq i \leq n} (U_s P_s A_i(x, w, \nabla 0), R_s^{-1} U_s P_s \partial_i w) dx + \\ &+ \sum_s \int_{Q_{s1}} (U_s P_s A_0(x, w, \nabla w), R_s^{-1} U_s P_s \partial_i w) dx = \sum_s \int_{Q_{s1}} (I_{s1} + I_{s2} + I_{s3}) dx, \end{aligned} \quad (39)$$

где (\cdot, \cdot) — скалярное произведение в $R^{N(s)}$. Поскольку нас интересует поведение данного интеграла при $\|u\|_{0,1}^{W_p(Q)} \rightarrow \infty$, то, не ограничивая общности, будем считать, что $u(x) \neq 0$ для почти всех $x \in Q$, т.е. $w(x) \neq 0$ для почти всех $x \in Q$.

Введем матрицу порядка $N(s) \times (n + 1)$

$$\zeta = (U_s P_s w, U_s P_s \partial_1 w \dots, U_s P_s \partial_n w),$$

а также матрицу η такую, что $\eta_0 = \zeta_0$ и $\eta_i = 0$ для всех $i = 1, \dots, n$.

Первую сумму правой части (22) оценим с помощью условия сильной эллиптичности (16):

$$\begin{aligned} I_{s1} &= \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} (A_i(x + h_{sl}, \zeta_l) - A_i(x + h_{sl}, \eta_l)) (R_s^{-1} \zeta_i)_l \geq \\ &\geq \hat{Y} \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} |\zeta_{li}|^p = \hat{Y} \sum_{1 \leq l \leq N(s)} \sum_{1 \leq i \leq n} |\partial_i w(x + h_{sl})|^p. \end{aligned}$$

То есть

$$\sum_s \int_{Q_{s1}} I_{s1} dx \geq \hat{Y} \|w\|_{W_p(Q)}^p = \hat{Y} \|R_Q u\|_{W_p(Q)}^p \geq c_9 \hat{Y} \|u\|_{W_p(Q)}^p, \quad (40)$$

где $\|R_Q u\|_{W_p(Q)}^p \geq c_9 \|u\|_{W_p(Q)}^p$ в силу невырожденности линейного оператора R_Q .

Рассмотрим вторую сумму правой части (39) с учетом оценки роста (15) и ограниченности оператора R_Q :

$$\begin{aligned} \sum_s \int_{Q_{s1}} |I_{s2}| dx &\leq \sum_{1 \leq i \leq n} \int_Q |A_i(x, w, \nabla 0) \partial_i u| dx \leq \\ &\leq \sum_{1 \leq i \leq n} \int_Q (g_0(x) + c_1 |w(x)|^{p'-1}) |\partial_i u(x)| dx \leq \left(\|g_0\|_{L_q(Q)} + c_1 \|w\|_{L_p(Q)}^{p'-1} \right) \|u\|_{W_p^1(Q)} \leq \\ &\leq \|g_0\|_{L_q(Q)} \|u\|_{W_p^1(Q)} + c_1 c_6^{p'-1} \|u\|_{L_p(Q)}^{p'-1} \|u\|_{W_p^1(Q)}. \end{aligned} \quad (41)$$

Аналогично для третьего слагаемого правой части (39) имеем

$$\begin{aligned} \sum_s \int_{Q_{s1}} |I_{s3}| dx &\leq \int_Q |A_0(x, w, \nabla w) u| dx \leq \\ &\leq \int_Q \left(g_0(x) + c_1 |w(x)|^{p'-1} + c_1 \sum_{1 \leq j \leq n} |\partial_j w(x)|^{p'-1} \right) |u(x)| dx \leq \\ &\leq \left(\|g_0\|_{L_q(Q)} + c_1 \|w\|_{L_p(Q)}^{p'-1} + c_1 \|w\|_{W_p^1(Q)}^{p'-1} \right) \|u\|_{L_p(Q)} \leq \\ &\leq \|g_0\|_{L_q(Q)} \|u\|_{W_p^1(Q)} + c_1 c_6^{p'-1} \|u\|_{L_p(Q)}^{p'-1} \|u\|_{W_p^1(Q)} + c_1 c_6^{p'-1} \|u\|_{W_p^1(Q)}^{p'-1} \|u\|_{L_p(Q)}. \end{aligned} \quad (42)$$

Таким образом, в правой части (39) мы имеем три слагаемых. Первое слагаемое строго положительно и имеет степень роста $p > 1$, см. (40). Второе и третье слагаемые имеют степень роста $p' < p$, см. (41) и (42). Следовательно, $\langle A_R u, u \rangle \geq c_{10} \|u\|_{W_p^1(Q)}^p - c_{11}$ при достаточно больших значениях $\|u\|_{W_p^1(Q)}$, где $c_{10}, c_{11} > 0$. Коэрцитивность доказана.

3. СУЩЕСТВОВАНИЕ РЕШЕНИЙ

Теорема 3. Пусть выполнены условия **A0)–A3)**. Тогда в пространстве $W_p^1(Q)$ для любого $f \in W_q^{-1}(Q)$ существует непустое, ограниченное, слабозамкнутое множество обобщенных решений задачи (8), (9).

Доказательство. Из условий теоремы следует, что A_R — деминепрерывный, см. лемму 2, оператор с полуограниченной вариацией, см. лемму 3, причем A_R коэрцитивен, см. лемму 5. Следовательно, обобщенное решение задачи (8), (9) существует, см. теорему 3.1 из [15, Гл. 2, §3], множество решений ограничено в силу коэрцитивности A_R и слабо замкнуто, см. следствие 4.1. из [16, Гл. 1, §4].

Теорема 4. Пусть выполнены условия 1–3 из разд. 1, а соответствующие матрицы R_s и R_{s_0} невырождены, а также справедливы условия **A1)–A3)** из разд. 2. Тогда в пространстве $W_{p,\gamma}^1(Q)$ для любого $f \in W_q^{-1}(Q)$ существует непустое, ограниченное, слабозамкнутое множество обобщенных решений задачи (2), (1).

Доказательство. Из условий **A0)–A3)** из разд. 2 следует, что существует непустое, ограниченное, слабозамкнутое множество обобщенных решений задачи (8), (9), см. теорему 3. При этом, если

$u \in \overset{\circ}{W}_p(Q)$ — решение (8), (9), то в силу условий 1–3 из разд. 1 и невырожденности матриц R_s и R_{s_0} $w = R_Q u \in W_{p,\gamma}^1(Q)$ — обобщенное решение задачи (2), (1), см. теорему 2. Поскольку линейный оператор R_Q непрерывен и ограничен, то множество обобщенных решений задачи (2), (1) непусто, ограничено и слабозамкнуто в пространстве $W_{p,\gamma}^1(Q)$.

Пример 1. Пусть $p \in (1, \infty)$, $Q = (0, 2) \times (0, 1)$, $f \in L_q(Q)$, $1/p + 1/q = 1$. Рассмотрим задачу с краевыми условиями Бицадзе-Самарского:

$$Aw(x) = f \quad (x \in Q), \tag{43}$$

$$\begin{aligned} w(x_1, 0) = w(x_1, 1) = 0 \quad (0 \leq x_1 \leq 2), \\ w(0, x_2) = \gamma_1 w(1, x_2) \quad (0 \leq x_2 \leq 1), \\ w(2, x_2) = \gamma_{-1} w(1, x_2) \quad (0 \leq x_2 \leq 1). \end{aligned} \tag{44}$$

Пусть $\gamma_1 \gamma_{-1} \neq 1$, $W_{p,\gamma}^1(Q)$ содержит функции из $W_p^1(Q)$, удовлетворяющие краевым условиям (44). Легко видно, что разностный оператор

$$Ru(x) = u(x_1, x_2) + \gamma_1 u(x_1 + 1, x_2) + \gamma_{-1} u(x_1 - 1, x_2)$$

таков, что $w = R_Q u \in W_{p,\gamma}^1(Q)$ для любого $u \in \overset{\circ}{W}_p(Q)$. Действие оператора определяется матрицей

$$R_1 = \begin{pmatrix} 1 & \gamma_1 \\ \gamma_{-1} & 1 \end{pmatrix}. \text{ Если } \gamma_1 \gamma_{-1} \neq 1, \text{ то } R_1 \text{ невырождена. В то же время } R_{10} = (1) \text{ также невырождена. Обратной к } R_1 \text{ является матрица } R_1^{-1} = \frac{1}{1 - \gamma_1 \gamma_{-1}} \begin{pmatrix} 1 & -\gamma_1 \\ -\gamma_{-1} & 1 \end{pmatrix}. \text{ Пусть } \hat{\gamma}_0 := 1 / (1 - \gamma_1 \gamma_{-1}), \hat{\gamma}_k = -\gamma_k / (1 - \gamma_1 \gamma_{-1}).$$

Сформулируем условия, достаточные для существования обобщенного решения $w \in W_{p,\gamma}^1(Q)$ задачи (43), (44) для любого $f \in W_q^{-1}(Q)$:

1) условие роста: для п.в. $x \in [0, 2] \times [0, 1]$ и всех $\xi \in \mathbb{R}^3$

$$|A_i(x, \xi)| \leq g_0(x) + c_1 |\xi_0|^{p'-1} + c_1 \sum_{j=1,2} |\xi_j|^{p'-1}, \quad i = 1, 2,$$

$$|A_0(x, \xi)| \leq g_0(x) + c_1 \sum_{0 \leq j \leq 2} |\xi_j|^{p'-1},$$

где $c_1 > 0$, $p' \in (1, p)$ и $g_0 \in L_q(Q)$;

2) условие сильной эллиптичности: для п.в. $x \in [0, 1] \times [0, 1]$ и всех $\zeta, \eta \in \mathbb{R}^{2 \times 3}$ таких, что $\eta \neq \zeta$ и $\eta_{l0} = \zeta_{l0}$, для некоторого $\hat{\gamma} > 0$

$$\begin{aligned} \sum_{l,m=1,2} \sum_{i=1,2} \hat{\gamma}_{m-l} (A_i(x_1 + l - 1, x_2, \zeta_l) - A_i(x_1 + l - 1, x_2, \eta_l)) (\zeta_{mi} - \eta_{mi}) \geq \\ \geq \hat{\gamma} \sum_{l,i=1,2} |\zeta_{li} - \eta_{li}|^p; \end{aligned}$$

3) условие локальной липшицевости/гёльдеровости функций A_i ($i = 1, 2$) по ξ_0 и функции A_0 по ξ_j ($j = 0, 1, 2$): существуют ρ ($\rho = 1$ при $p \in [2, \infty)$), $\rho \in (1 - 1/p, p - 1]$ при $p \in (1, 2)$) и $\varepsilon > 0$ такие, что для п.в. $x \in [0, 2] \times [0, 1]$ и любых $\xi, \delta \in R^3$, $|\delta| = \sum_{0 \leq j \leq n} |\delta_j| = |\delta_k| < \varepsilon$,

$$|A_i(x, \xi_0 + \delta_0, \xi_1, \xi_2) - A_i(x, \xi)| \leq \left(\Psi |\xi_0 + \delta_0|^{p-1-\rho} + \Psi \sum_{0 \leq j \leq 2} |\xi_j|^{p-1-\rho} + g_\Psi(x) \right) |\delta_0|^\rho,$$

$$|A_0(x, \xi + \delta) - A_0(x, \xi)| \leq \left(\Psi |\xi_k + \delta_k|^{p-1-\rho} + \Psi \sum_{0 \leq j \leq 2} |\xi_j|^{p-1-\rho} + g_\Psi(x) \right) |\delta_k|^\rho,$$

где $\Psi > 0$, $g_\Psi \in L_{q'}(Q)$, $q' = p / (p - 1 - \rho)$, $i = 1, 2$, $k = 0, 1, 2$.

Пример 2. Пусть $p \in (1, \infty)$, $n \geq 3$, $Q = (0, 3) \times G$, где $G \subset \mathbb{R}^{n-1}$ — ограниченная область с границей ∂G класса C^∞ , $f \in W_q^{-1}(Q)$, $1/p + 1/q = 1$. Пусть $x' = (x_2, \dots, x_n)$. Рассмотрим задачу с нелокальными краевыми условиями типа Бицадзе-Самарского:

$$Aw(x) = f \quad (x \in Q), \tag{45}$$

$$\begin{aligned} w(x_1, x') = w(x_1, x') = 0 & \quad (0 \leq x_1 \leq 3, x' \in \partial G), \\ w(0, x') = -w(1, x') & \quad (x' \in G), \\ w(3, x') = 0 & \quad (x' \in G). \end{aligned} \tag{46}$$

$W_{p,\gamma}^1(Q)$ содержит функции из $W_p^1(Q)$, удовлетворяющие краевым условиям (46). Покажем, что разностный оператор

$$Ru(x) = u(x_1, x') - u(x_1 + 1, x') + u(x_1 + 2, x')$$

таков, что $w = R_Q u \in W_{p,\gamma}^1(Q)$ для любого $u \in W_p^1(Q)$:

$$w(0, x') = R_Q u(0, x') = u(0, x') - u(1, x') + u(2, x') = -u(1, x') + u(2, x'); \tag{47}$$

$$w(1, x') = R_Q u(1, x') = u(1, x') - u(2, x') + u(3, x') = u(1, x') - u(2, x'); \tag{48}$$

$$w(2, x') = R_Q u(2, x') = u(2, x') - u(3, x') = u(2, x'); \quad w(3, x') = u(3, x') = 0.$$

То есть, $w(0, x') = -w(1, x')$, см. (47), (48), $w(x_1, x') = R_Q u(x_1, x') = 0$, w удовлетворяет краевым условиям (46). Действие оператора R_Q определяется матрицей

$$R_1 = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \text{ причём } R_{10} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \text{ и } R_1^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

R_1 и R_{10} невырождены. Заметим также, что обратный оператор R_0^{-1} в рассматриваемом примере является разностным, поскольку его действие определяется матрицей R_1^{-1} . В общем случае обратный к разностному оператору не обязан быть разностным.

Для существования обобщенного решения $w \in W_{p,\gamma}^1(Q)$ задачи (45), (46) для любого $f \in W_q^{-1}(Q)$ достаточно выполнения условий **A1**, **A3** и условия сильной эллиптичности для дифференциально–разностного оператора, для нашего конкретного разностного оператора оценка (16) имеет вид

$$\begin{aligned} & \sum_{1 \leq l \leq 3} \sum_{1 \leq i \leq n} (A_i(x_1 + l - 1, x', \zeta_l) - A_i(x_1 + l - 1, x', \eta_l)) (\zeta_{li} - \eta_{li}) + \\ & + \sum_{1 \leq l \leq 2} \sum_{1 \leq i \leq n} (A_i(x_1 + l - 1, x', \zeta_l) - A_i(x_1 + l - 1, x', \eta_l)) (\zeta_{l+1,i} - \eta_{l+1,i}) \geq \\ & \geq \hat{\gamma} \sum_{1 \leq l \leq 3} \sum_{1 \leq i \leq n} |\zeta_{li} - \eta_{li}| \quad \text{для п.в. } x_1 \in [0, 1], x' \in \bar{G}. \end{aligned}$$

Заметим, что для разрешимости уравнения (45) с краевыми условиями Дирихле достаточно выполнения условий **A1**, **A3** и условия сильной эллиптичности для дифференциального оператора:

$$\sum_{1 \leq i \leq n} (A_i(x, \xi) - A_i(x, \mu)) (\xi_i - \mu_i) \geq \hat{\gamma} \sum_{1 \leq i \leq n} |\xi_i - \mu_i|^p \quad \text{для п.в. } x \in [0, 1] \times \bar{G}.$$

В условии сильной эллиптичности дифференциально–разностного оператора обязательно появляются слагаемые, соответствующие связям значений рассматриваемых функций в точках различных подобластей одного класса разбиения области.

СПИСОК ЛИТЕРАТУРЫ

1. Бицадзе А. В., Самарский А. А. О некоторых простейших обобщениях линейных эллиптических задач // Докл. АН СССР. 1969. Т. 185. № 4. С. 739–740.
2. Самарский А. А. О некоторых проблемах теории дифференциальных уравнений // Дифференц. уравнения. 1980. Т. 16. № 1. С. 1925–1935.
3. Скубачевский А. Л. Нелокальные эллиптические задачи с параметром // Матем. сборник. 1983. Т. 121. № 6. С. 201–210.
4. Скубачевский А. Л. Эллиптические задачи с нелокальными условиями вблизи границы // Матем. сборник. 1986. Т. 129. № 2. С. 279–302.
5. Skubachevskii A. L. Elliptic Functional Differential Equations and Applications. Basel–Boston–Berlin: Birkhäuser, 1997.
6. Скубачевский А. Л. Неклассические краевые задачи. I // СМФН, РУДН. 2007. Т. 26. С. 3–132.
7. Скубачевский А. Л. Неклассические краевые задачи. II // СМФН, РУДН. 2009. Т. 33. С. 3–179.
8. Солонуха О. В. Об одной нелинейной нелокальной задаче эллиптического типа // Ж. вычисл. матем. и матем. физ. 2017. Т. 57. № 3. С. 60–72.
9. Солонуха О. В. Обобщенные решения квазилинейных эллиптических дифференциально–разностных уравнений // Ж. вычисл. матем. и матем. физ. 2020. Т. 60. № 12. С. 2085–2097.
10. Солонуха О. В. О существовании решений нелинейных параболических вариационных неравенств с односторонними ограничениями // Матем. заметки. 2005. Т. 77. № 3. С. 460–476.
11. Солонуха О. В. Об одном классе существенно нелинейных эллиптических дифференциально–разностных уравнений // Труды МИАН. 2013. Т. 283. С. 226–244.
12. Solonukha O. V. On nonlinear and quasilinear elliptic functional–differential equations // Discrete and Continuous Dynamic Systems, Seria S. 2016. V. 9. N 3. P. 847–868.

13. *Красносельский М. А.* Топологические методы в теории нелинейных интегральных уравнений. М.: Гостехиздат, 1956.
14. *Соболев С. Л.* Некоторые применения функционального анализа в математической физике. 3-е изд., перераб. и доп. / Под ред. О. О. Олейник. М.: Наука. 1988.
15. *Дубинский Ю. А.* Нелинейные эллиптические и параболические уравнения // Итоги науки и техники: ВИНТИ. Современные проблемы математики, 1976. № 9. С. 5–130.
16. *Иваненко В. И., Мельник В. С.* Вариационные методы в задачах управления для систем с распределенными параметрами. Киев: Наук. думка, 1988.

ЛОКАЛЬНАЯ ВЕЙВЛЕТНАЯ АДАПТАЦИЯ ДЕКАРТОВЫХ СЕТОК В ВЫЧИСЛИТЕЛЬНЫХ ЗАДАЧАХ ГАЗОВОЙ ДИНАМИКИ¹⁾

© 2024 г. А. Л. Афендигов^{1,*}, В. С. Никитин^{1,**}

¹125047 Москва, Миусская пл., 4, ИПМ им. М. В. Келдыша РАН, Россия

*e-mail: andre@keldysh.ru

**e-mail: zogorlol@gmail.com

Поступила в редакцию 28.06.2023 г.

Переработанный вариант 12.10.2023 г.

Принята к публикации 19.10.2023 г.

Представлен метод динамической локальной адаптации градуированных декартовых деревьев для численного решения задач газовой динамики. Локальный вейвлетный анализ газодинамического поля на базе неравномерных В-сплайнов применяется независимо к каждой ячейке расчетной сетки и позволяет выделить негладкие или существенно нелинейные участки решения (или наоборот, достаточно гладкие и линейные) и модифицировать сетку для расчета следующего шага по времени так, чтобы у разномасштабных особенностей течения было адекватное сеточное разрешение. В комбинации с другими методами вычислительной газовой динамики, такими как метод свободной границы, представленный метод позволяет эффективно решать нестационарные задачи с обтеканием движущихся тел. На ряде таких задач продемонстрирована работа предложенного варианта вейвлетной адаптации. Библ. 37. Фиг. 5.

Ключевые слова: вычислительная газовая динамика, вейвлетная адаптация, вейвлетный анализ, локально адаптивные декартовые сетки.

DOI: 10.31857/S0044466924020103, EDN: YJKTI

1. ВВЕДЕНИЕ

Проблема адаптации сеток для увеличения точности определяемых численно решений начально-краевых задач возникла вместе с появлением современных численных методов и проведением расчетов на первых ЭВМ. В вычислительной газовой динамике идеи адаптации, связанные с выделением сильных ударных волн и сгущением сетки без изменения числа узлов (ячеек) “пропорционально градиенту” или с использованием диффузионной прибавки (см. [1], [2]), были развиты в разных направлениях (см., например, [3]–[16]). Тем не менее, до сих пор даже для стационарных задач нет общепринятого определения оптимальности сетки, хотя давно понята связь этого вопроса с такими характеристиками как величины ε -энтропии и асимптотики поперечников функциональных компактов, в которых лежат разыскиваемые решения (см. [17], [18]).

В задачах газовой динамики трудность состоит еще и в том, что нет ясности с выбором функциональных компактов, в которых разыскиваются решения. В результате, в некоторых пакетах машинное время, затрачиваемое на адаптацию сетки, становится сравнимым со временем расчета газовой динамики. Сложность задачи усугубляется необходимостью строить алгоритмы, учитывающие архитектуру суперкомпьютеров, например, структуру памяти или наличие графических ускорителей. В силу появления все новых модификаций параллельных суперкомпьютеров эта тематика продолжает оставаться актуальной. Кроме того, отметим, что большинство известных адаптивных алгоритмов применяются для решения линейных стационарных задач (см. обсуждение этого вопроса в [19]). Нестационарность и нелинейность задачи ограничивают возможность использования апостериорного анализа решения, требуют адекватного учета связи временных и пространственных масштабов и, наконец, балансировки (т.е. как часто проводить) процессов измельчения и огрубления сетки.

¹⁾ Работа выполнена при поддержке Московского центра фундаментальной и прикладной математики, Соглашение с Минобрнауки РФ (код 075-15-2019-1623).

Для учета граничных условий есть ряд подходов, позволяющих моделировать внутренние граничные условия модификацией, определяющей системы уравнений. Одним из таких подходов является метод свободной границы (free boundary method, см. [20]–[22]). В нем для учета влияния обтекаемых твердых тел вводится компенсационный поток — специальная поправка, делающая определяющую систему уравнений неоднородной. Появление таких подходов привело к возобновлению интереса к использованию декартовых сеток, разных вариантов их структуризации и адаптации. Алгоритмы на таких сетках устроены проще, а простота и универсальность построения таких сеток хорошо подходит для исследования динамики в нестационарных задачах с изменяющимися внутренними границами.

В настоящей работе описан и опробован новый алгоритм вейвлетной адаптации сетки, основанный на идеях использования градуированных декартовых деревьев и многомасштабных разложений (см. [13]–[16] и имеющиеся там ссылки). В нем применяется вариант преобразования вейвлетного типа, основанный на использовании неравномерных В-сплайнов и метода лифтинга (см. [23]–[26]). Этот подход позволяет эффективно решать такие задачи, как, например, задача о сверхзвуковом вылете пеллета (см. [27]) и задачи о свободном движении системы тел под действием набегающей ударной волны (см. [28]).

В работах [29]–[31] аналогичные подходы применяются для глобального многомасштабного вейвлетного анализа, который используется для контроля ошибки, перехода между разными масштабами и для вычисления пространственных производных. Для обобщения на многомерный случай используются тензорные произведения, а для внутренних граничных условий — метод фиктивных областей (см. [32]–[34]) (другая вариация метода погруженной границы). Ключевым отличием подхода, используемого в настоящей работе, является локальность вейвлетного анализа. Анализируется не вычислительная сетка целиком, а каждая ячейка в отдельности, причем для многомерного обобщения каждая ось тоже анализируется отдельно. Локальность вейвлетного анализа позволяет сделать весь метод предельно локальным, что необходимо для эффективной параллелизации. Кроме того, предложенный метод адаптации является модульным, т.е. не используется другими компонентами вычислительного алгоритма, и легковесным, т.е. не требует хранения в памяти дополнительных конструкций, связанных с разными масштабами.

2. ГРАДУИРОВАННЫЕ ДЕРЕВЬЯ И ИЕРАРХИИ ДЕКАРТОВЫХ СЕТОК

В версии метода декартовых сеток, развиваемой для решения нестационарных задач газовой динамики в [20], [22] и используемой в настоящей работе, начально-краевая задача в области со сложной — многосвязной и зависящей от времени границей, сводится к неоднородной задаче в стандартной прямоугольной области с помощью метода свободной границы. Метод свободной границы удобен тем, что позволяет путем введения в уравнения движения сил реакции стенок, рассматривая задачу в фиксированной прямоугольной области. Поэтому при использовании этого метода без ограничения общности задача об адаптации сетки может быть сведена заменой переменных к задаче в единичном кубе $\Omega \subset R^d$ независимо от числа движущихся в заданном объеме тел. Идея построения градуированных декартовых деревьев возникла в гармоническом анализе, а затем была четко реализована в работе [18], где они называются «элементами класса элементарных продолжений разбиений единичного куба». Далее слова «градуированное декартово» опускаются там, где это не может привести к недоразумению. Полученные в этой статье результаты опираются на одну специальную теорему о субаддитивных функциях множеств, а доказательство основано на подробном исследовании некоторого конкретного алгоритма разбиений куба $\Omega^d \subset R^d$. Затем для заданной функции $u(x)$ из пространства $W_p^q(\Omega^d)$ на построенном для нее разбиении (дереве) определяется аппроксимирующий агрегат — кусочно-полиномиальная функция. В применении к теории аппроксимации этот нелинейный алгоритм порождает аппарат приближений, тесно увязанный с особенностями конкретной функции, что позволяет добиться хорошей скорости приближений. В этой статье параметром является число кубов в разбиении (листьев в дереве), а субаддитивной функцией — интегральная норма. Полученные в этом направлении результаты позволили получить оценки некоторых поперечников и ε -энтропии компактов в $W_p^q(\Omega^d)$.

Задачи, решаемые в настоящей работе, и методы исследования — иные в силу наличия сильных разрывов и того, что в расчетах приходится работать со значениями вектора газодинамических переменных в узлах, а не с интегральными средними. Кроме того, в выделенной фиксированной области течение зависит от времени и может существенно усложняться или упрощаться. Например, может появиться, а затем исчезнуть сложная конфигурация ударных волн, контактных разрывов и волн разрежения. При этом вне описанных сингулярностей решение может оставаться гладким.

Если предположить, что разрыв размазан на несколько ячеек, т.е. вектор-функция газодинамических величин сглажена, что в реальности и происходит при использовании методов типа конечного объема, то цель исследования состоит в построении зависящей от времени «квазиоптимальной» последовательности

деревьев, на каждом из которых и происходит расчет шага по времени. То есть рассматривается задача о введении в численных алгоритмах для решения эволюционных нелинейных задач более подробной сетки вблизи особенностей решения и адекватном укрупнении ячеек в областях высокой гладкости. На каждом временном шаге считается заданным двоичное (можно рассматривать и q -ичные) разбиение $\{V_{l,k}\}$. Здесь индекс $0 \leq l \leq L$ отвечает уровню разбиения, а мультииндекс k — расположению ячейки на этом уровне, т.е.

$$V_{l,k} = \left(2^{-l}k_1, 2^{-l}(k_1 + 1)\right) \times \dots \times \left(2^{-l}k_d, 2^{-l}(k_d + 1)\right),$$

где d — размерность пространства, а

$$k_i \in I_l, \quad I_l = \{0, 1, \dots, 2^l - 1\}.$$

Определим множество целочисленных векторов $J = \{0, 1\}^d$. Тогда замыкание каждой ячейки $V_{l,k}$, $l < L$, может быть представлено в виде

$$\bar{V}_{l,k} = \bigcup_{m \in M_{l,k}} \bar{V}_{l+1,m}, \quad (2.1)$$

где $M_{l,k}$ — множество индексов, определяющих разбиение ячейки $V_{l,k}$ на ячейки, принадлежащие уровню $(l + 1)$

$$M_{l,k} = \{2k + j; j \in J\}.$$

Назовем $M_{l,k}$ уточняющим набором индексов, а разложение (2.1) — уточняющим разбиением ячейки $V_{l,k}$. Тем самым, каждая ячейка $V_{l+1,m}$, принадлежащая уровню $l + 1$, входит в уточняющее разбиение некоторой ячейки $V_{l,k}$ уровня l .

В этом разложении ячейка $V_{l+1,m}$ называется потомком ячейки $V_{l,k}$, а $V_{l,k}$, в свою очередь, называется предком $V_{l+1,m}$ уровня l . По индукции для каждой $V_{l+1,m}$ определяются предки всех уровней до $l = 0$. Ячейки $V_{l,k}$, $V_{r,m}$, для которых $(l,k) \neq (r,m)$ и $\dim(\bar{V}_{l,k} \cap \bar{V}_{r,m}) = d - 1$, назовем прямыми соседями.

Определение. Назовем градуированным декартовым деревом порядка такое разбиение куба

$$\bar{\Omega} = \bigcup_{(l,k) \in \mathcal{L}^n} \bar{V}_{l,k},$$

$$\mathcal{L}^n \subset \{(l,k), \quad l = 0, \dots, L; k \in I_l^d\},$$

что

- 1) $V_{l,k} \cap V_{r,m} \neq \emptyset \Leftrightarrow (l,k) = (r,m)$;
- 2) если ячейки $V_{l,k}$ и $V_{r,m}$ — прямые соседи, то для уровней l, r выполняется $|l - r| \leq b$;
- 3) если $(l,k) \in \mathcal{L}^n$ и $k \in M_{l-1,m}$, то для любого $k_1 \in M_{l-1,m}, (l,k_1) \in \mathcal{L}^n$.

Другими словами, каждая ячейка уровня l из этой иерархической структуры входит в дерево только вместе со всеми элементами уточняющего разбиения предка этой ячейки предшествующего уровня, а уровни ячеек, имеющих пересекающиеся грани, отличаются не более, чем на b .

В алгоритме [4] фиксируется число конечных ячеек (листьев в дереве), а величина b произвольна, однако при расчете временного шага сильная неоднородность сетки может вести к потере точности, и поэтому в настоящей работе ограничимся.

Замечание 1. Из определения $M_{l,k}$ следует, что при проведении вычислений удобно записывать координаты $k = (k_1, \dots, k_d)$ ячейки $V_{l,k}$ в двоичной системе счисления и, приписывать нули перед значащими цифрами так, чтобы число знаков в двоичной записи составляло l . Тогда набор цифр, стоящих на s -ых позициях двоичных записей индексов k_i , определяет соответствующий вектор j в уточняющем разложении предка данной ячейки на уровне $s - 1$. Указанное свойство позволяет находить соседей и предков данной ячейки, пользуясь двоичной арифметикой.

Рассмотрим для наглядности двумерный случай. В памяти компьютера градуированное декартовое дерево задается связанной структурой, где каждому элементу структуры поставлена в соответствие

некоторая ячейка из множества $\{V_{l,k}\}$. В каждом элементе этой структуры хранятся уровень разбиения l , векторный индекс ячейки k и флаг F ее принадлежности дереву. При этом $F = 1 \Leftrightarrow (l,k) \in \mathcal{L}^n$. Если ячейка принадлежит дереву (такие ячейки назовем листовыми), то в ней дополнительно хранится вся необходимая для вычислений информация о газодинамическом поле в этой ячейке. Если же $F = 0$, $(l,k) \notin \mathcal{L}^n$, т.е. эта ячейка не входит в дерево, то в ней дополнительно хранятся адреса элементов, соответствующих потомкам данной ячейки. Эти адреса индексируются уточняющим вектором j из разложения (2.1). При этом в упомянутой структуре содержатся только элементы, соответствующие таким нелистовым ячейкам, у которых есть листовые потомки. В результате, если начать с корневой ячейки, последовательными операциями прохождения по адресу можно прийти до любой листовой ячейки.

Построение эффективного алгоритма поиска соседних ячеек в дереве является ключевым моментом и для работы солвера, и для адаптации сетки. Опишем, как в рамках предложенной структуры для каждой ячейки указываются прямые соседи.

По каждой оси $e_a, a = 1, \dots, d$, соседями (необязательно листовыми) ячейки $V_{l,k}$ будут ячейки $V_{l,k_{left}}$, $V_{l,k_{right}}$, у которых соответствующий индекс отличается на единицу: $k_a^{left} = k_a - 1$, $k_a^{right} = k_a + 1$. Приведем алгоритм подъема по дереву для получения ячейки $V_{r,m}$ на псевдокоде:

```

V = V0,0 — поиск начинается с корневой ячейки
while r>0 and F(V)=0 — поиск останавливается, либо когда достигнут необходимый уровень разбиения,
либо когда текущая ячейка оказалась листовой
  r=r-1
  j=(m>>r)&1 — получаем i-е биты вектора m, определяющие уточняющий вектор
  m=m>>1 — т.е. индекс ячейки целочисленно делится пополам
  V=Childj(V) — переход по указателю на потомка с индексом j
end while
return V
    
```

С помощью алгоритма поиска находятся ячейки $V_{l,k_{left}}$, $V_{l,k_{right}}$. Если получившиеся ячейки нелистовые, то левыми (правыми) соседями исходной ячейки будут все листовые потомки получившейся ячейки с уточняющим индексом $k_a = 1$ (соответственно $k_a = 0$). Повторяя операцию для каждой из осей, получим всех прямых соседей данной ячейки.

3. О РЕШЕНИИ УРАВНЕНИЙ ГАЗОВОЙ ДИНАМИКИ НА ГРАДУИРОВАННЫХ ДЕКАРТОВЫХ ДЕРЕВЬЯХ

Рассмотрим движение газа, описываемое системой Эйлера, которая в стандартной консервативной форме имеет вид

$$\frac{\partial \mathbf{q}}{\partial t} + \frac{\partial \mathbf{f}_k}{\partial x_k} = 0,$$

с обычным предположением о суммирование по повторяющимся индексам. Здесь $\mathbf{q} = (\rho, \rho \mathbf{u}, \rho E)^T$ — вектор консервативных переменных, $\mathbf{f}_k = (\rho u_k, \rho u_k u_m + \delta_{km} p, \rho u_k H)^T$ — векторы потоков, $m = 1, 2, 3$, а δ_{km} — символ Кронекера; ρ , u_k , p — плотность, компоненты вектора скорости и давление соответственно, $H = e + p/\rho$ — энтальпия, e — удельная внутренняя энергия, а $E = e + 0.5 u_k u_k$ — полная удельная энергия. Газ предполагается идеальным совершенным с показателем адиабаты γ , т.е. $p = (\gamma - 1)\rho e$. В области Ω с набором твердых включений Γ и условием непротекания на них $(\mathbf{u}, \mathbf{n})_{\Gamma} = 0$, где \mathbf{n} — вектор нормали к поверхности, ставится стандартная начально-краевая задача для уравнений Эйлера. Краевые условия на $\partial\Omega$ устанавливаются в зависимости от стоящей задачи и, например, для задачи обтекания моделируют постоянный поток вдали от включений. Рассмотрим однородную задачу Коши, которая решается в единичном кубе Ω методом конечного объема. Детали вычисления компенсационных потоков для неоднородной задачи могут быть найдены в [20], [27]. Задача о распаде разрыва может решаться приближенно, например, по Русанову или точно по Годунову (см. [2]), сама конструкция реализуется на градуированном декартовом дереве.

Если расчетная сетка состоит из ячеек разных уровней, то она является неконформной, и при расчете следующего шага суммарные потоки через заданную ячейку вычисляются по формуле

$$\mathbf{Q}_{j,k}^{n+1} = \mathbf{Q}_{j,k}^n - \frac{\Delta t}{V_{j,k}} \sum_{\sigma} s_{\sigma} \mathbf{F}_{\sigma}, \quad (3.1)$$

где $\mathbf{Q}_{j,k}^n$ — среднее значение вектора консервативных переменных в ячейке $V_{j,k}$ на временном шаге с номером n (объем ячейки $V_{j,k}$), s_{σ} — площадь грани ячейки, а \mathbf{F}_{σ} — численный поток через грань.

Потоки по каждому из направлений, умноженные на площади соответствующих частей границы, суммируются по границе ячейки. Устойчивость оператора послойного перехода в (3.1) обеспечивается переменным шагом по времени Δt с выполнением условия Куранта—Фридрихса—Леви (см. [2]). Для увеличения порядка аппроксимации схемы можно применить подсеточную реконструкцию газодинамических параметров более высокого порядка точности, например, кусочно-линейное восполнение, обобщающее на неравномерные сетки интерполяцию схемы MUSCL (Monotone Upstream-Centered Scheme for Conservation Laws). Схемы второго порядка точности по пространству и выше являются немонотонными и приводят к паразитическим осцилляциям в численных решениях вблизи поверхностей сильных разрывов. Для подавления этих осцилляций используются ограничители производных minmod и Ван Альбада (G.D. van Albada, B. van Leer, W. Roberts). Детали указанных конструкций и соответствующие ссылки имеются, например, в [20].

В результате расчета $(n + 1)$ шага по времени каждой ячейке $V_{j,k}$ дерева $\{V_{j,k}\}^n$ оказывается поставлено в соответствие новое значение вектора газодинамических переменных, отнесенное к центрам ячеек. В разд. 5 описан алгоритм вейвлетного анализа, с помощью которого по этим векторам строится дерево $\{V_{j,k}\}^{n+1}$, адаптированное под решение на новом шаге. Обычно достаточно использовать не весь вектор, а только одно наиболее характерное поле. Как правило это поле плотности, либо давления. Но сначала в разд. 4 рассмотрим вейвлетное преобразование функции на последовательности неравномерно расположенных узлов, на основе которого будет строиться вейвлетный анализ в ячейках дерева.

4. ОДНОМЕРНОЕ ДИСКРЕТНОЕ ВЕЙВЛЕТНОЕ ПРЕОБРАЗОВАНИЕ НА НЕРАВНОМЕРНОЙ СЕТКЕ

Рассмотрим сначала вейвлетный анализ скалярной функции $\tilde{f}(x)$ на узлах отрезка $[a, b]$. Считаем заданным разбиение

$$a = x_0 < x_1 < \dots < x_{n+k}, \quad x_0 = a, \quad x_{n+k} = b, \quad (4.1)$$

где узлы x_j соответствуют центрам ячеек. Для решения вопроса о соответствии функции и ее особенностей разбиению на отрезке предлагается использовать преобразование вейвлетного типа, основанное на теории B -сплайнов (см., например, [17], [25]). Для удобства напомним некоторые общеизвестные факты. Для сетки (4.1) введем B -сплайны степени k :

$$B_i^k(x) = [x_i, \dots, x_{i+k+1}] (y - x)_+^k, \\ i = 0, \dots, n - 1,$$

где через $[x_i, \dots, x_{i+k+1}]g(y)$ обозначена разделенная разность функции $g(y)$, а

$$(y - x)_+^k = \begin{cases} (y - x)^k, & y \geq x, \\ 0, & y < x. \end{cases}$$

Известно, что $B_i(x)$, $i = 0, \dots, n - 1$, являются сплайнами дефекта 1 с минимальным носителем $\text{supp} B_i(x) = [x_i, x_{i+k+1}]$ и линейно независимы. Кроме того, они удовлетворяют соотношениям

$$B_i^k(x) = \frac{x - x_i}{x_{i+k+1} - x_i} B_i^{k-1}(x) + \frac{x_{i+k+1} - x}{x_{i+k+1} - x_i} B_{i+1}^{k-1}(x)$$

и

$$\int_{x_i}^{x_{i+k+1}} B_i^k(s) g^{(k+1)}(s) ds = k! [x_i, \dots, x_{i+k+1}] g(s).$$

Из последнего равенства следует, что

$$\int_{x_i}^{x_{i+k+1}} B_i^k(s) s^l ds = \frac{k!l!}{(k+l+1)!} [x_i, \dots, x_{i+k+1}] s^{k+l+1} \tag{4.2}$$

и, в частности,

$$\int_{x_i}^{x_{i+k+1}} B_i^k(s) ds = \frac{1}{(k+1)}. \tag{4.3}$$

В дальнейшем, в тех случаях, когда это не может привести к недоразумениям, верхний индекс в обозначениях сплайнов будет опускаться.

Для сплайнов на равномерных сетках давно известно тождество

$$B^k(x) = \sum_{m=0}^{k+1} 2^{-k} C_k^m B^k(2x - m),$$

называемое двухмасштабным соотношением и лежащее в основе теории сплайн-вейвлетов (см., например, [35]). Аналогичное соотношение справедливо и для неравномерных сеток (см. [25]). Рассмотрим, например, при четном $n+k$ сетку, полученную из (4.1) отбрасыванием каждого второго узла

$$\left\{ \tilde{x}_0 = x_0, \tilde{x}_1 = x_2, \dots, \tilde{x}_{\frac{n+k}{2}} = x_{n+k} \right\}. \tag{4.4}$$

Через $\tilde{B}_i(x)$ обозначим сплайны, построенные для этой сетки:

$$\begin{aligned} \tilde{B}_i(x) &= [\tilde{x}_i, \dots, \tilde{x}_{i+k+1}] (y-x)_+^k, \\ i &= 0, \dots, N-1; \quad N = \frac{n+k}{2}. \end{aligned}$$

Заметим, что $\tilde{x}_i = x_{2i}$.

Обозначим через V линейное пространство $V = \text{span}\{B_0(x), \dots, B_{n-1}(x)\}$ и $\tilde{V} = \text{span}\{\tilde{B}_0(x), \dots, \tilde{B}_{N-1}(x)\}$.

Лемма 1. $\tilde{V} \subset V$.

Это утверждение непосредственно следует из (4.1), (4.4) и линейной независимости B -сплайнов.

Отсюда следует, что существует подпространство W , являющееся дополнением \tilde{V} до прямой суммы

$$V = \tilde{V} + W, \quad \tilde{V} \cap W = \{0\},$$

и при этом $\dim W = \frac{n+k}{2}$.

Выберем в W базис $\{\xi_0, \dots, \xi_{n-N}\}$, состоящий из функций, имеющих p нулевых моментов. Обозначим через $Z = \{\zeta_j, j = 0, \dots, n-1\}$ базис биортогональный к $\{B_0(x), \dots, B_{n-1}(x)\}$, а через $\Sigma = \{\sigma_i, i = 0, \dots, N-1\}$, $\Theta = \{\theta_i, i = 0, \dots, n-N\}$ — базисы \tilde{V} и W , составленные из векторов базиса, биортогонального к базису $\{\tilde{B}_0(x), \dots, \tilde{B}_{N-1}(x), \xi_0, \dots, \xi_{n-N}\}$.

Пусть $\tilde{f}(x) = f(x) + \varepsilon(x)$, где $|\varepsilon(x)| < \varepsilon_0, \varepsilon_0 \ll 1$, а $f(x) \in \text{span}\{B_i(x), i = 0, \dots, n-1\}$, и

$$f(x) = \sum_{j=0}^{n-1} b_j \zeta_j. \tag{4.5}$$

Тем самым, ключевым для построения адаптивных сеток предлагаемым методом является предположение о том, что разыскиваемое решение лежит в пространстве V или хотя бы с достаточной точностью приближается элементами этого пространства.

Коэффициенты разложения функции $f(x)$ по базисам Σ в \tilde{V} и Θ в W зададут гладкие $c_j = (f, \tilde{B}_j(x))$ и вейвлетные $w_j = (f(x), \xi_j(x))$ компоненты вейвлетного разложения. Алгоритм для вычисления вейвлетного разложения описан ниже.

Из леммы 1 следует, что

$$\tilde{B}_i^k(x) = \sum_{j=2i}^{2i+k+1} r_{ij} B_j^k(x).$$

Компоненты низкочастотного фильтра r_{ij} легко найти, пользуясь линейной независимостью сплайнов $B_j^k(x)$. Для этого обозначим

$$y_l = B(x_l), \quad z_l^i = B_{2i}(x_l), \quad l = 2i + 1, \dots, 2i + k.$$

Тогда r_{ij} определяются последовательно;

$$\begin{aligned} r_{i,2i} &= y_{2i+1} / z_{2i+1}^{2i}, \dots, r_{i,2i+s} = \\ &= \left(y_{2i+s+1} - \sum_{j=2i}^{2i+s-1} r_{ij} z_j^{2i} \right) / z_{2i+s+1}^{2i}. \end{aligned}$$

Так, например, для $k = 1$, учитывая условие (4.3), имеем

$$\begin{aligned} \tilde{B}_i(x) &= \frac{x_{2i+1} - x_{2i}}{x_{2i+4} - x_{2i}} B_{2i}(x) + \\ &+ \frac{x_{2i+3} - x_{2i+1}}{x_{2i+4} - x_{2i}} B_{2i+1}(x) + \\ &+ \frac{x_{2i+4} - x_{2i+3}}{x_{2i+4} - x_{2i}} B_{2i+2}(x). \end{aligned} \tag{4.6}$$

Описанная процедура вычислительно устойчива и более проста, чем предложенная в [26].

Для выбора базиса в подпространстве W из матрицы r_{ij} нужно выбрать невырожденный минор ρ_{ij} порядка N , который всегда существует в силу линейной независимости сплайнов $\{\tilde{B}_0(x), \dots, \tilde{B}_{N-1}(x)\}$. Такой минор не единственен и имеется свобода, которой можно распорядиться так, чтобы минимизировать носитель $\text{supp} \xi_l(x)$. Естественно взять элементы базиса $\{\hat{\xi}_0, \dots, \hat{\xi}_{n-N}\}$ в W из функций $B_j(x) \in \{B_0(x), \dots, B_{n-1}(x)\}$. Этот «пробный» базис можно подправить, положив

$$\xi_l = \hat{\xi}_l - \sum_{j=j_{\min}(l)}^{j_{\min}(l)+p-1} \alpha_{lj} \tilde{B}_j,$$

нормированный условием $\|\xi_l\|_{L_1} = 1$ так, чтобы удовлетворить условиям на моменты

$$(\xi_l, x^s) = 0, \quad s = 0, \dots, p - 1. \tag{4.7}$$

Число $j_{\min}(l)$ определено из условия

$$\tilde{x}_{j_{\min}(l)} = \max \{ \tilde{x}_j \mid \tilde{x}_j \leq x_{j_l} \}.$$

Здесь через j_l , $l = 1, \dots, n - N$, обозначены номера столбцов, не попавших в выбранный минор.

Положим $\hat{\xi}_l = B_{j_l}$; тогда в силу (4.7) вектор α_{lj} , $j = 0, \dots, p - 1$, однозначно определяется из линейной системы с невырожденной матрицей

$$\begin{aligned} (B_{j_l}, x^s) &= \sum_{j=0}^{p-1} \alpha_{lj} (\tilde{B}_{j_{\min}(l)+j}, x^s), \\ j, s &= 0, \dots, p - 1, \end{aligned}$$

где невырожденность матрицы непосредственно следует из (4.2).

Тем самым, определено дискретное вейвлетное преобразование с

$$c_i = \sum_{j=1}^N r_{ij} b_j$$

для сглаженных компонент и

$$w_l = b_{j_l} - \sum_{m=0}^{p-1} \alpha_{lm} c_{j_{\min(l)+m}}$$

для вейвлетных компонент.

Отметим, что для анализатора с $(k, p) = (1, 2)$ необходим шаблон с семью узлами, а матрица r_{ij} , $i = 0, 1, j = 0, \dots, 4$, определена соотношениями (4.6). Положим

$$\rho = \begin{Bmatrix} r_{00} & r_{04} \\ r_{01} & r_{14} \end{Bmatrix}.$$

В этом случае несложные вычисления для $j_l = l + 1$ дают

$$\begin{aligned} \alpha_{00} &= \frac{x_6 + x_4 - x_3 - x_1}{x_6 - x_0}, \alpha_{01} = \frac{x_3 + x_1 - x_4 - x_0}{x_6 - x_0}, \\ \alpha_{10} &= \frac{x_6 - x_3}{x_6 - x_0}, \alpha_{11} = \frac{x_3 - x_0}{x_6 - x_0}; \\ \alpha_{20} &= \frac{x_6 + x_2 - x_5 - x_3}{x_6 - x_0}, \alpha_{21} = \frac{x_5 + x_3 - x_0 - x_2}{x_6 - x_0}. \end{aligned}$$

Итого, результатом вейвлетного анализа функции будет величина

$$\tilde{w} = \max_l \left(w_l \frac{\|\xi_l\|}{x_{n+k} - x_0} \right). \tag{4.8}$$

Отношение в (4.8) — нормировочный множитель, нужный чтобы соотносить величины вейвлетных коэффициентов, полученные на разных разбиениях.

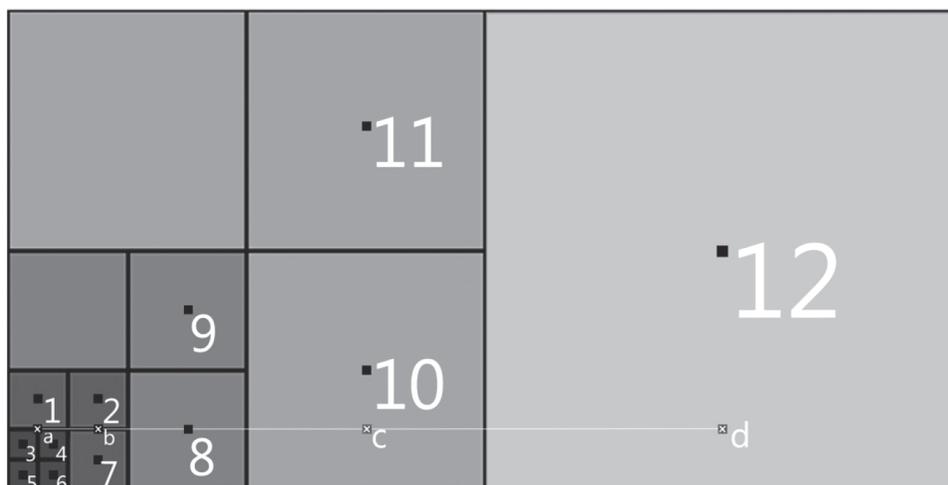
Вольно говоря, эта величина характеризует потерю информации при переходе от функции, заданной на сетке (4.1), к функции, заданной на прореженной сетке (4.4). По величине этих коэффициентов и производится решение об укрупнении или измельчении сетки. Так, например, при $k = 1, p = 2$ для линейных функций $f(x)$ и только для них имеем $\tilde{w} = 0$.

5. ЛОКАЛЬНАЯ ВЕЙВЛЕТНАЯ АДАПТАЦИЯ ГРАДУИРОВАННЫХ ДЕКАРТОВЫХ ДЕРЕВЬЕВ

Локальная адаптация деревьев состоит в том, чтобы по дереву $\{V_{j,k}\}^n$ построить дерево $\{V_{j,k}\}^{n+1}$, принимая для каждой отдельной листовой ячейки одно из трех решений: разбить ячейку на 2^d потомков большего уровня, объединить с другими потомками предка предыдущего уровня в одну ячейку меньшего уровня, либо оставить без изменений. Рассмотрим сначала одномерные деревья.

Для их вейвлетного анализа ведется поиск соседей слева и справа от исследуемой ячейки. Центры исследуемой ячейки и ее соседей используются в качестве узлов разбиения (4.1) для вейвлетного преобразования. Для эффективной работы алгоритма на суперкомпьютерах с распределенной памятью необходимо, чтобы эта процедура была как можно более локальной, т.е. использовала как можно меньше соседних ячеек. Минимальный шаблон для проведения анализа состоит из $2p + 2k + 1$ узлов, т.е. $n = 2p + k$.

В качестве $f(x)$ для вейвлетного преобразования возьмем функцию, совпадающую в центрах ячеек с выбранным компонентом газодинамического поля на $(l + 1)$ -м шаге. Значения функции в промежутках между узлами нужны для вычисления скалярных произведений $(f(x), B_j(x))$. Для приближенного вычисления этих интегралов без интерполяции данных в промежуточных точках воспользуемся квадратурными формулами. В рамках анализа предположим, что $f(x)$ в промежуточных точках является интерполяционным сплайном степени $p - 1$, тогда произведение $f(x)B_j(x)$ будет сплайном степени $p + k - 1$, и скалярные произведения в (4.5) находятся точно с помощью квадратурных формул



Фиг. 1. Фрагмент градуированного декартового дерева первого порядка.

Гаусса точности $p + k - 1$. В случае $p = 2, k = 1$ для проведения вычислений достаточно воспользоваться формулой Симпсона. Если, кроме того, сетка (4.1) равномерна, а функция $f(x)$ — линейна на $\text{supp } B_j(x)$, то формула Симпсона сводится к $(f(x), B_j(x))_{L_2} = \frac{1}{2}f(x_{j+1})$. Этим в известной мере обосновывается обычный прием в численных приложениях теории вейвлетов, когда значения коэффициентов в разложениях функции по ортогональным или биортогональным системам заменяются на значения функции в соответствующих узлах. Воспользовавшись этим приближением для крайних узлов с индексами $j = 0, j = n - 1$ можно уменьшить необходимый размер шаблона на 2, что значительно увеличивает локальность анализа.

При программной реализации метода выгодно учесть, что если фиксирован порядок b декартового дерева и размер шаблона, то несложно описать все возможные конфигурации сетки и вычислить все возможные коэффициенты r_{ij}, α_j и нормы до компиляции программы. Например, если шаблон состоит из пяти ячеек, число различных конфигураций $K = 57$. Таким образом, в качестве входных данных алгоритма, вычисляющего $w(x_j)$ помимо значений в центрах ячеек $f(x)$, требуются только разницы уровней соседних ячеек.

Для многомерных градуированных декартовых деревьев, предлагается квазиодномерный подход, позволяющий избежать вычислительно трудоемких конструкций, связанных с построением тензорных произведений сплайнов как, например, в [29]. Для этого для каждой из осей так же, как в случае одномерных деревьев, ищутся соседи слева и справа от исследуемой ячейки по каждой из осей. Центры этих ячеек на неравномерном участке сетки не обязаны лежать на одной прямой. Поэтому значения выбранного для анализа газодинамического поля проецируются из центров найденных ячеек на прямую, параллельную оси и проходящую через центр исследуемой ячейки. Получив узлы разбиения и значения функции в них, мы свели задачу к случаю одномерных деревьев.

Ниже приведены типичные приемы для построения разбиений типа (4.1) и интерполяции газодинамического поля на прямую на примере двумерной сетки и пятиточечного шаблона.

На фиг. 1 представлен участок сетки, на котором нужно вычислить данные для анализа ячейки 8 по горизонтальной оси для пятиточечного шаблона. Крестами обозначены точки, значения в которых необходимо вычислить для одномерного вейвлетного анализа. Черными точками обозначены центры ячеек, значения в которых будут использованы для построения значений поля в крестовых точках — a, b, c, d . Для проведения анализа с привлечением ячеек уровня выше, чем исследуемая, как, например, в случае с точками a и b , проводится осреднение:

$$f_a = 0.5(f_1 + 0.25(f_3 + f_4 + f_5 + f_6)),$$

$$f_b = 0.5(f_2 + f_7).$$

Если же используемые ячейки имеют уровень меньший, чем исследуемая, проводится экстраполяция:

$$f_c = f_{10} - 0.25(h_y)_{10} (f'_y)_{10},$$

$$f_d = f_{12} - 0.375(h_y)_{12} (f'_y)_{12}.$$

Производные по осям отличных от оси, по которой производится анализ, приближенно вычисляются с помощью соседней ячейки того же уровня, которая всегда существует для градуированных декартовых деревьев порядка 1:

$$(f'_y)_{10} = 0.5(f_9 - f_8) / (h_y)_{10}, \quad (5.1)$$

$$(f'_y)_{12} = 0.5(f_{11} - f_{10}) / (h_y)_{12}. \quad (5.2)$$

Порядок декартового дерева также ограничивает число шаблонов для подобных интерполяций.

Замечание 2. Описанный метод сведения анализа на деревьях к анализу на одномерной сетке подходит для расчетов с любым порядком аппроксимации, в том числе, первым. Если же в случае вычислений с высоким порядком в ячейках уже хранятся градиенты исследуемого поля, то можно использовать их для реконструкции значений вместо приближений (5.1), (5.2).

Результатом вейвлетного анализа для ячейки многомерного дерева будет значение

$$\widehat{w}(V_{l,k}) = \max_{0 < i \leq d} \widetilde{w}_i,$$

где d — размерность пространства, а \widetilde{w}_i — результат одномерного вейвлетного анализа вдоль оси i . Значение \widehat{w} далее сравнивается с заданными пороговыми величинами C_{\max} и C_{\min} :

- 1) если $\widehat{w}(V_{l,k}) > C_{\max}$, то соответствующая ячейка измельчается;
- 2) если $C_{\min} < \widehat{w}(V_{l,k}) < C_{\max}$, то соответствующая ячейка остается без изменений;
- 3) если $\widehat{w}(V_{l,k_i}) < C_{\min}$ для каждого потомка общего предка предыдущего уровня, то ячейки объединяются. В противном случае потомок остается без изменений.

Замечание 3. Для восстановления порядка дерева, т.е. для выполнения условия отличия уровней соседних ячеек не более чем на единицу, необходимо дополнительно измельчить некоторые ячейки для балансировки. Эта процедура находится вне рамок этой работы.

Замечание 4. В процессе адаптации сетки очень важно не испортить решение, для чего следует применять гладкие, физически осмысленные и ненарушающие законы сохранения схемы интерполяции. В работах [27], [28] использовалась WENO-интерполяция (см. [36], [37]).

6. РЕЗУЛЬТАТЫ ЧИСЛЕННЫХ ЭКСПЕРИМЕНТОВ

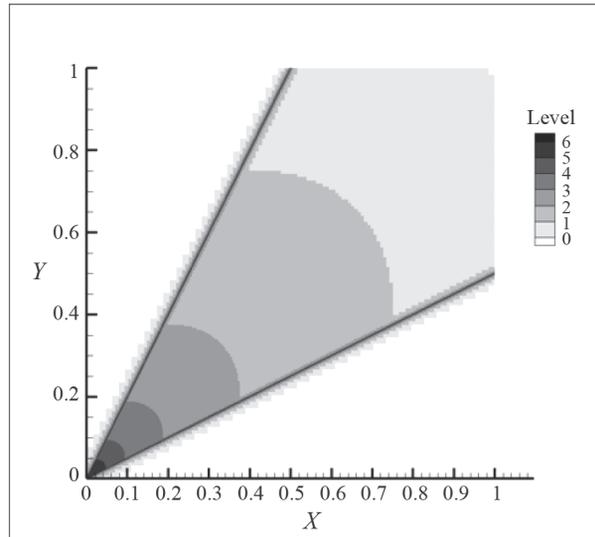
Ниже мы рассмотрим работу предложенного алгоритма вейвлетной адаптации в вычислительных экспериментах. В дальнейших примерах используется вейвлетный анализатор с порядком вейвлетов $k = 1$ и числом нулевых моментов $p = 2$.

6.1. Аналитически заданное поле

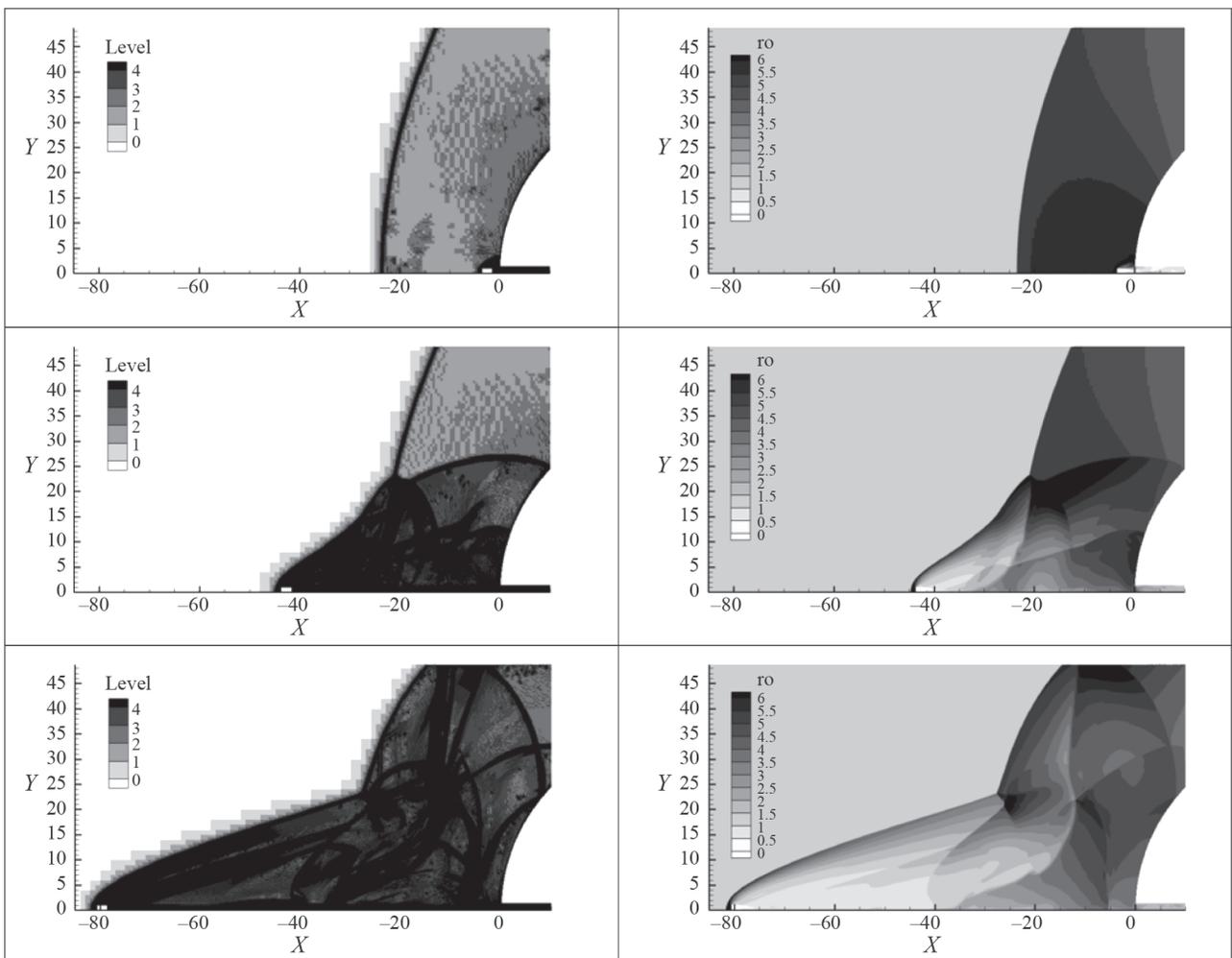
Работу вейвлетного анализатора можно наглядно продемонстрировать на адаптации сетки для аналитически заданного поля, моделирующего волну разряжения:

$$f(x, y) = \begin{cases} \operatorname{arctg}(0.5), & y < 0.5x, \\ \operatorname{arctg}(2), & y > 2x, \\ \operatorname{arctg}(y/x), & 0.5x < y < 2x. \end{cases}$$

Область $[0, 1] \times [0, 1]$ изначально разбита на 64×64 ячеек. Все ячейки области l раз подвергаются вейвлетной адаптации, где $l = 6$ — максимальный уровень разбиения базовой ячейки.



Фиг. 2. Результат адаптации аналитически заданного поля.



Фиг. 3. Уровни адаптации ячеек (слева) в промежуточных моментах времени и поля плотности (справа) в те же моменты в задаче о вылете pellets (фрагмент области).

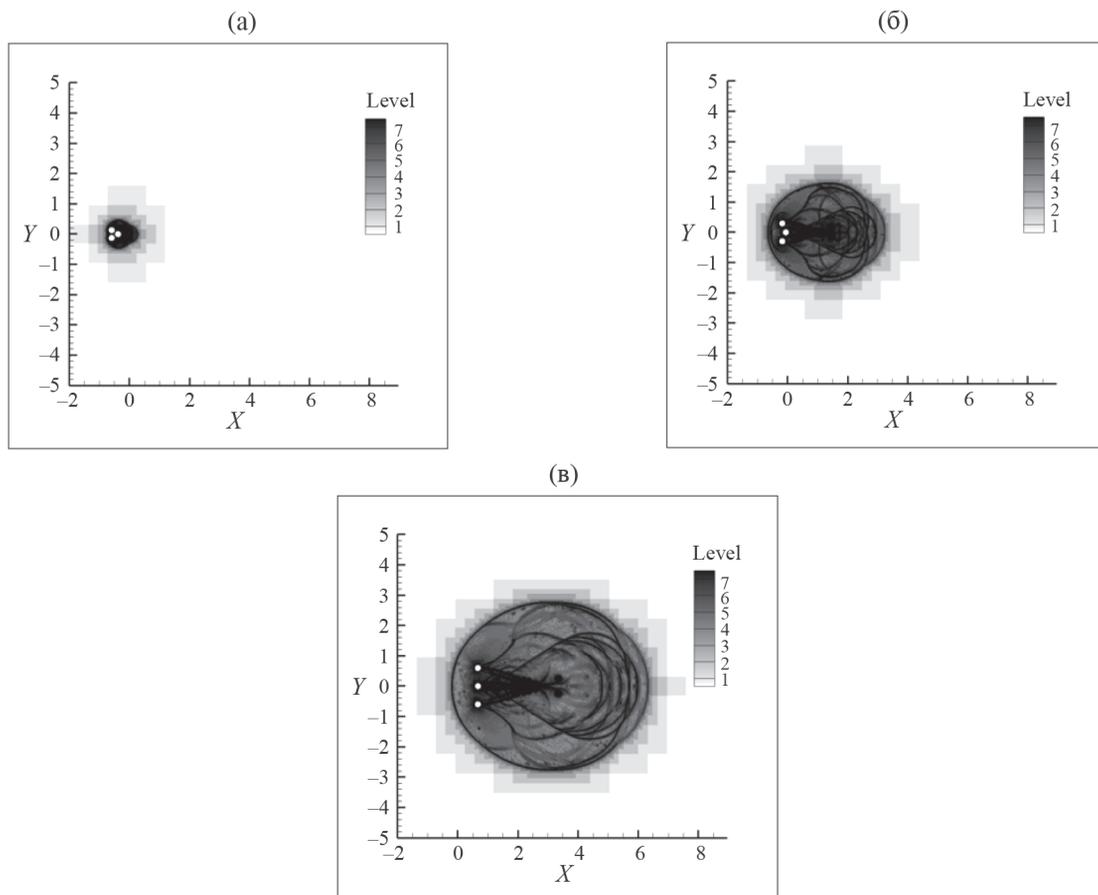
На фиг. 2 цветом обозначается уровень разбиения ячейки. Видно, что в области, где f не является константой, по мере приближения к началу координат увеличивается сложность функции, а вместе с ней — подробность сетки.

6.2. Адаптация декартовых сеток в газодинамических задачах

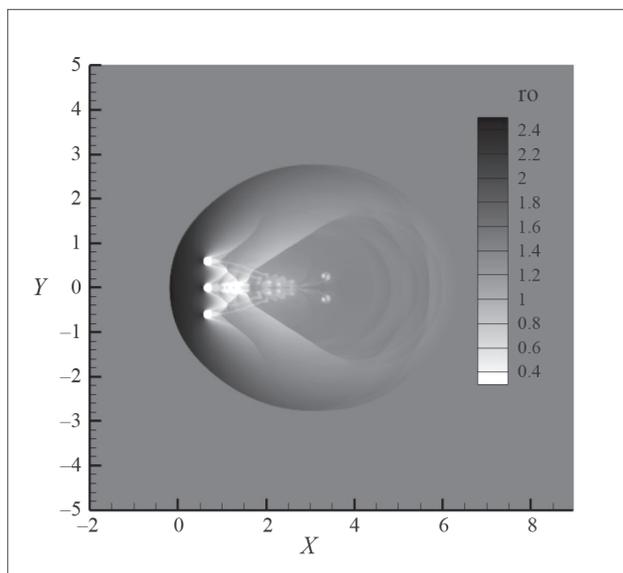
Теперь продемонстрируем работу предложенного метода на двух содержательных газодинамических задачах. В качестве функции f для вейвлетного анализа использовалось поле плотности. В обеих задачах использовался идеальный газ с показателем адиабаты $\gamma = 1.4$ и посчитаны по схеме второго порядка аппроксимации. В первой задаче тело движется по заданному закону с постоянной скоростью, во второй система тел движется свободно под действием давления газа. Дополнительно проводилась адаптация до максимального уровня вдоль подвижных границ обтекаемых тел.

Задача 1. Эта задача о вылете пеллета подробно рассмотрена в [27]. В ней цилиндр малого размера $2\text{мм} \times 2\text{мм}$ вылетает из канала неподвижного затупленного цилиндра $100 \times 70\text{мм}$ со сферическим концом. Канал на оси симметрии цилиндров имел диаметр 3мм . Система тел обтекалась сверхзвуковым потоком идеального газа с числом Маха $M = 3$ а малый цилиндр двигался в противоположном потоку направлении с числом Маха $M = 4$. Расчетная область $[-300, 100] \times [0, 130]$ (в мм). Пороговые величины вейвлетной адаптации $C_{\max} = 0.0005$, $C_{\min} = 0.0001$.

Задача 2. Задача о разлете тел при сверхзвуковом обтекании в плоской постановке, опубликованная в [28], была пересчитана в настоящей работе в большей области и с большим числом уровней, поэтому опишем ее подробнее. Три цилиндра с осями, перпендикулярными расчетной области радиуса $R = 0.1$ и массой $M = 1$, покоятся в начальный момент времени. Центральные оси цилиндров расположены в вершинах правильного треугольника с координатами $\mathbf{X}_1(0) = (-0.6, 0.125)$, $\mathbf{X}_2(0) = (-0.6, -0.125)$, $\mathbf{X}_3(0) = (0.125\sqrt{3} - 0.6, 0)$. Расчетная область $[-2.14] \times [-8.8]$.



Фиг. 4. Эволюция расчетной сетки в задаче о разлете цилиндров. Начальная стадия расчета (а), средняя стадия (б) и финальный момент времени (в). Цветом обозначен уровень ячейки.



Фиг. 5. Поле плотности в финальный момент времени в задаче о разлете цилиндров.

На нижней, верхней и правой границах задано условие свободного вытекания. На левой границе задан постоянный вектор, совпадающий с начальными данными $(\rho, u_x, u_y, p)^T = (1.4, 2, 0, 1)^T$. Размер базовой сетки 25×25 с девятью уровнями адаптации. Пороговые величины вейвлетной адаптации $C_{\max} = 0.0005$, $C_{\min} = 0.0001$.

Цилиндры взаимодействуют с потоком, получая импульс под действием поверхностных сил. Расчет производился до выхода цилиндров на одну прямую $x = 1.813$ в момент времени $t = 2.3525$. Все величины даны в безразмерном виде. О методике расчета подробнее можно прочитать в исходной публикации [28].

На фиг. 3 для задачи 1 и на фиг. 4, 5 для задачи 2 видно, как сложные конфигурации разномасштабных особенностей течения, включающих головные ударные волны, волны разряжения и более мелкие скачки уплотнения эффективно динамически улавливаются вейвлетным анализатором, локально измельчающим сетку в соответствующих областях, а затем огрубляющим эти ячейки на более поздних шагах, если области с негладким и существенно нелинейным решением переместились в другие ячейки.

7.ЗАКЛЮЧЕНИЕ

Представлен эффективный метод локальной адаптации декартовых сеток, использующий одномерное вейвлетное преобразование на базе неравномерных В-сплайнов. Метод действует локально, выделяя необходимую ему информацию о газодинамическом поле из минимально возможного количества соседей исследуемой ячейки. В комбинации с другими приведенными в работе методами, такими как метод свободной границы, его можно применять для решения широкого класса задач газовой динамики, в том числе, существенно нестационарных, с обтеканием движущихся тел и активными перестройками течения.

Способность локального анализатора динамически выделять разномасштабные особенности газодинамических полей на текущем расчетном шаге и адекватно подстраивать сетку под решение продемонстрирована на ряде численных экспериментов.

СПИСОК ЛИТЕРАТУРЫ

1. Годунов С. К. и др. Численное решение многомерных задач газовой динамики. М.: Наука, 1976.
2. Теоретические основы и конструирование вычислительных алгоритмов. Ред. К. И. Бабенко, М.: Наука, 1978.

3. Young D. P., Melvin R. G., Bieterman M. B., Johnson F. T., Samant S. S., and Bussoletti J. E. A locally refined rectangular grid finite element method: Application to computational fluid dynamics and computational physics // J. Comput. Phys. 1991. V. 92. N 1. P. 1.
4. Khokhlov A. M. Fully threaded tree algorithms for adaptive refinement fluid dynamics simulations // J. Comput. Phys. 1998. V. 143. P. 519–543.
5. Aftosmis M. J., Berger M. J., Melton J. E. Adaptive cartesian mesh generation // Chapter 22. Handbook of grid generation, ed. J. Thompson, B. Soni, N. Weatherill. CRC Press, 1999.
6. Pember R. B., Bell J. B., Colella P., Crutchfield W. Y., and Welcome M. L. An Adaptive Cartesian Grid Method for Unsteady Compressible Flow in Complex Geometries // AIAA Paper 93-3385-CP (1993).
7. Бреславский П. В., Мажукин В. И. Метод динамической адаптации в задачах газовой динамики // Матем. моделирование. 1995. Т. 7. № 12. С. 48–78.
8. Бреславский П. В., Мажукин В. И. Динамически адаптирующиеся сетки для взаимодействующих разрывных решений // Ж. вычисл. матем. и матем. физ. 2007. Т. 47. № 4. С. 717–737.
9. Афендииков А. Л., Луцкий А. Е., Плёнкин А. В. Вейвлетный анализ локализованных структур в идеальной и вязкой моделях // Матем. моделирование. 2011. Т. 23. № 1. С. 41–50.
10. Афендииков А. Л. и др. Локализация разрывов в полях газодинамических функций с помощью вейвлет анализа // Матем. моделирование. 2008. № 7. С. 65–84.
11. Афендииков А. Л., Луцкий А. Е., Плёнкин А. В. Локализация особенностей газодинамических полей и адаптация расчетной сетки к положению разрывов // Матем. моделирование. 2012. Т. 24. № 12. С. 49–54.
12. Harten A. Multiresolution algorithms for the numerical solution of hyperbolic conservation laws // Comm. Pure Appl. Math. 1995. V. 48. N 12. P. 1305–1342.
13. Harten A. Multiresolution representation of data: A general framework // SIAM J. Numer. Anal. 1996. V. 33. N 3. P. 1205–1256.
14. Zumbusch G. Parallel multilevel methods. Adaptive mesh refinement and loadbalancing // Adv. Numeric. Math. Teubner, Wiesbaden, 2003.
15. Bramkamp F., Lamby P., Mueller S. An adaptive multiscale finite volume solver for unsteady and steady state flow computations // J. Comput. Phys. 2004. V. 197. P. 460–490.
16. Hartmann D., Meinke M., Schroder W. A strictly conservative Cartesian cut-cell method for compressible viscous flows on adaptive grids // Comput. Meth. Appl. Mech. Engineer. 2011. V. 200. P. 1038–1052.
17. Бабенко К. И. Основы численного анализа. М.: Наука, 1998.
18. Бирман М. Ш., Соломяк М. З. Кусочно-полиномиальные приближения функций классов W_p^α // Матем. сб. 1967. Т. 73. № 115. С. 331–355.
19. Василевский Ю. В. и др. Автоматизированные технологии построения неструктурированных расчетных сеток. М.: Физматлит, 2016.
20. Меньшов И. С., Корнев М. А. Метод свободной границы для численного решения уравнений газовой динамики в областях с изменяющейся геометрией // Матем. моделирование. 2014. Т. 26. № 5. С. 99–112. англ. пер.: I. S. Menshov, M. A. Kornev. Free-boundary method for the numerical solution of gasdynamic equations in domains with varying geometry // Math. Mod. Comput. Simulat. 2014. V. 6. N 6. P. 612–621.
21. Меньшов И. С., Павлухин П. В. Эффективный параллельный метод сквозного счета задач аэродинамики на несвязных декартовых сетках // Ж. вычисл. матем. и матем. физ. 2016. Т. 56. № 9. С. 1677–1691.
22. Афендииков А. Л. и др. Адаптивные вейвлетные алгоритмы на декартовых сетках. М.: ИПМ им. М. В. Келдыша РАН, 2016. 232 с.
23. de Boor C. Splines as linear combinations of B-splines. A Survey // in APPROXIMATION THEORY, II G. G. Lorentz, C. K. Chui, and L. L. Schumaker (ed.), Academic Press (New York), 1976, 1–47 (emended version, 1986).
24. Sweldens W. The Lifting Scheme: A custom-design construction of biorthogonal wavelets // Appl. Comput. Harmonic Anal. 1996. V. 3. P. 186–200.
25. Ford J. M., Oseledets I. V., Tyrtysnikov E. E. Matrix approximations and solvers using tensor products and non-standard wavelet transforms related to irregular grids // Russ. J. Numer. Anal. Math. Model. 2004. V. 19. N 2. P. 185–204.
26. Демьянович Ю. К. Калибровочное соотношение для B-сплайнов на неравномерной сетке // Матем. моделирование. 2001. Т. 13. № 9. С. 98–100.
27. Афендииков А. Л., Луцкий А. Е., Меньшов И. С., Никитин В. С., Ханхасаева Я. В. (ИПМ им. М. В. Келдыша РАН) Численное моделирование вылета пеллета из затупленного тела. М.: ИПМ, РАН, 2017.

28. *Афендииков А. Л., Никитин В. С.* Численное моделирование на адаптивных сетках свободного движения системы тел в сверхзвуковом потоке газа // Матем. моделирование. 2020. Т. 32. № 12. С. 55–64; *Math. Models Comput. Simul.* 2021. V. 13. N 4. P. 667–673.
29. *Regele J. D., Vasilyev O. V.* An adaptive wavelet-collocation method for shock computations // *Inter. J. Comput. Fluid Dynamic.* 2009. V. 23. N 7. P. 503–518.
30. *Vasilyev O. V., Bowman C.* Second generation wavelet collocation method for the solution of partial differential equations // *J. Comput. Phys.* 2000. V. 165. P. 660–693.
31. *Vasilyev O. V., Kevlahan N. K.R., Goldstein D. E., Vezolainen A. V., Regele J., Nejadmalayeri A., Reckinger S., BrownDymkoski E., Rogoz E.* Adaptive wavelet environment for in silico universal multiscale modeling (AWESUMM) (2019).
32. *Dymkoski E. Brown, Kasimov N., Vasilyev O. V.* A characteristic based volume penalization method for general evolution problems applied to compressible viscous flows // *J. Comput. Phys.* 2014. V. 262. P. 344–357.
33. *Dymkoski E. Brown, Kasimov N., Vasilyev O. V.* Characteristic based volume penalization method for arbitrary Mach flows around solid obstacles // *Direct and Large Eddy Simulation IX* / ed. by J. Frohlich, H. Kuerten, B. Geurts, V. Armenio. Springer, 2015. P. 10–115.
34. *Vasilyev O. V., Kevlahan N. K. R.* Hybrid wavelet collocation Brinkman penalization method for complex geometry flows // *Inter. J. Numer. Meth. Fluid.* 2002. V. 40. P. 531–538.
35. *Chui C. K.* An Introduction to wavelets. Elsevier, 2014. P. 266.
36. *Chi-Wang Shu,* High Order ENO and WENO Schemes for Computational Fluid Dynamics // *High-order methods for computational physics*, Springer, Heidelberg, 1999, P. 438–480.
37. *Тишкин В. Ф., Пескова Е. Е., Жалнин Р. В., Горюнов В. А.* О построении WENO-схем для гиперболических систем уравнений на неструктурированных сетках // *Изв. вузов. Приволжский регион. Физ.-матем. науки.* 2014. № 1 (29).

УДК 519.642

НОВЫЕ КОМПЬЮТЕРНО-ЭКОНОМИЧНЫЕ АППРОКСИМАЦИИ СЛУЧАЙНЫХ ФУНКЦИЙ ДЛЯ РЕШЕНИЯ СТОХАСТИЧЕСКИХ ЗАДАЧ ТЕОРИИ ПЕРЕНОСА¹⁾

© 2024 г. Г. А. Михайлов^{1,2,*}, И. Н. Медведев^{1,2,**}

¹630090 Новосибирск, пр-т Акад. Лаврентьева, 6, ИВМ и МГ СО РАН, Россия

²630090 Новосибирск, ул. Пирогова, 1, НГУ, Россия

*e-mail: gam@sscc.ru

**e-mail: min@osmf.sccc.ru

Поступила в редакцию 28.04.2023 г.

Переработанный вариант 28.04.2023 г.

Принята к публикации 19.10.2023 г.

Разработана новая сеточная аппроксимация однородного изотропного случайного поля с заданной средней корреляционной длиной. Эта аппроксимация строится путем разбиения координатного пространства на ансамбль кубиков, размер которых воспроизводит среднюю корреляционную длину при независимом выборе значения поля из заданного одномерного распределения в каждом элементе разбиения. Сформулирован также недавно предложенный авторами метод корреляционно-рандомизированного моделирования переноса частиц через случайную среду. Проведено сравнение точности и трудоемкости соответствующих алгоритмов метода Монте-Карло для решения задач о переносе гамма-квантов через случайную среду типа мозаики Вороного. Для проверки гипотезы о существенном влиянии одномерного распределения и корреляционного радиуса оптической плотности среды на перенос излучения были также проведены дополнительные расчеты для случайного пуассоновского “поля воздушных шаров” в воде. Дано обобщение сеточной аппроксимации на неизотропные случайные поля. Библ. 19. Фиг. 3. Табл. 4.

Ключевые слова: метод максимального сечения, стохастическая среда, сеточная аппроксимация, корреляционно-рандомизированные алгоритмы, корреляционная длина, перенос гамма-квантов, трудоемкость алгоритма.

DOI: 10.31857/S0044466924020118, EDN: YJCQUB

1. ВВЕДЕНИЕ

1.1. Рассматривается процесс переноса частиц (квантов излучения), математическая модель которого определяется кинетическим уравнением типа (см. [1, 2])

$$(\omega, \text{grad}\Phi) + \Sigma(\mathbf{r}, \mathbf{v})\Phi(\mathbf{r}, \mathbf{v}) = \int \Sigma_s(\mathbf{r}, \mathbf{v}')w(\mathbf{v}, \mathbf{v}'; \mathbf{r})\Phi(\mathbf{r}, \mathbf{v}')d\mathbf{v}' + \Phi_0(\mathbf{r}, \mathbf{v}). \quad (1)$$

Здесь $\mathbf{r} \in R^3$, $\mathbf{v} = v\omega$, ω — единичный вектор направления скорости, $v = |\mathbf{v}|$; $\Phi(\mathbf{r}, \mathbf{v})$ — плотность потока частиц-квантов излучения; $\Sigma(\mathbf{r}, v)$ — полное макроскопическое сечение ослабления; $\Sigma_s(\mathbf{r}, v)$ — макроскопическое сечение рассеяния; $\Phi_0(\mathbf{r}, \mathbf{v})$ — плотность распределения частиц в источнике; $w(\mathbf{v}, \mathbf{v}'; \mathbf{r})$ — индикатриса рассеяния кванта в точке $(\mathbf{r}, \mathbf{v}')$. Для решения задач переноса численно-статистически моделируется [3],[2] цепь Маркова столкновений частицы с элементами вещества, свободный пробег l между которыми распределен с плотностью

$$f(l) = \Sigma(\mathbf{r}(l), v)\exp(-\tau(l)),$$

где $\mathbf{r}(l) = \mathbf{r}' + l\omega$, а $\tau(l) = \int_0^l \Sigma(\mathbf{r}(t), v)dt$ (см. [3],[2]). Если $\Sigma \equiv \text{const}$, то пробег можно моделировать по формуле $l = (-\ln \alpha) / \Sigma$, где α — случайное число, равномерно распределенное в интервале $(0, 1)$. Если

¹⁾ Работа выполнена в рамках государственного задания ИВМиМГ СО РАН № 0251-2022-0002.

же плотность среды существенно меняется, то может быть полезным метод максимального сечения (выравнивания, дельта-рассеяния) (см. [4]–[6]). Этот метод основан на том, что в уравнении (1) $\Sigma(\mathbf{r}, v)$ заменяется на $\Sigma_{\max}(v) \geq \Sigma(\mathbf{r}, v)$ и $\Sigma_s(v', \mathbf{r})w(\mathbf{v}, \mathbf{v}'; \mathbf{r})$ на

$$\Sigma_s(\mathbf{r}, v')w(\mathbf{v}, \mathbf{v}'; \mathbf{r}) + [\Sigma_{\max}(v') - \Sigma(\mathbf{r}, v')] \delta(\mathbf{v} - \mathbf{v}').$$

В соответствующей стохастической модели длины пробегов моделируются для коэффициента ослабления $\Sigma_{\max}(v)$, в точке столкновения \mathbf{r} с вероятностью $(\Sigma_{\max}(v) - \Sigma(\mathbf{r}, v)) / \Sigma_{\max}(v)$ скорость не изменяется, а с вероятностью

$$\Sigma(\mathbf{r}, v) / \Sigma_{\max}(v) \quad (2)$$

реализуется физическое столкновение с условными вероятностями Σ_s / Σ , $(\Sigma - \Sigma_s) / \Sigma$ рассеяния и поглощения. Адекватность метода следует также из известного сохранения свойства пуассоновости при случайном прореживании пуассоновского потока столкновений [6].

1.2. Метод максимального сечения часто применяется для решения стохастических задач теории переноса со случайным полем плотности среды $\rho(\mathbf{r})$, которое для простоты изложения в основном будем предполагать изотропным с нормированной корреляционной функцией $k(t)$ и корреляционным радиусом

$$L = \int_0^{\infty} K(\mathbf{r}, \mathbf{r}'(l)) dl = \frac{1}{4\pi} \int_{\Omega} \int_0^{\infty} K(\mathbf{r}, \mathbf{r} + \omega l) d\omega dl, \quad (3)$$

где $K(\mathbf{r}, \mathbf{r}'(l))$ — коэффициент корреляции между $\rho(\mathbf{r})$ и $\rho(\mathbf{r}'(l))$ для $|\mathbf{r} - \mathbf{r}'(l)| = l$, ω — изотропный единичный вектор.

При этом

$$\Sigma(\mathbf{r}, v) = \rho(\mathbf{r})\sigma(\mathbf{r}, v) \text{ и } \Sigma_s(\mathbf{r}, v) = \rho(\mathbf{r})\sigma_s(\mathbf{r}, v),$$

где $\sigma(\mathbf{r}, v)$, $\sigma_s(\mathbf{r}, v)$ — эффективные микроскопические сечения ослабления и рассеяния, которым соответствует нормированная индикатриса рассеяния $w(\mathbf{v}, \mathbf{v}'; \mathbf{r})$. Для решения стохастических задач теории переноса эффективен “метод двойной рандомизации” [7], в котором для каждой реализации среды моделируется одна или несколько (для уменьшения трудоемкости) траекторий частиц и на этой основе строятся несмещенные оценки линейных функционалов, например осредненной по реализациям среды вероятности прохождения. Трудоемкость (сложность) такого алгоритма может быть очень большой и даже неограниченной при $L \rightarrow 0$, например, для “мозаичных” моделей $\rho(\mathbf{r})$, пуассоновских, Вороного и т.п. (см., например, [8]–[10]). В связи с этим в работе [11] был сформулирован несложный корреляционно рандомизированный алгоритм (КР-алгоритм) для оценки влияния стохастичности поля ρ на перенос частиц, эвристически основанный на том факте, что вероятность прохождения в значительной степени определяются корреляционной длиной L и одномерным распределением поля ρ [8].

1.3. Формулировка КР-алгоритма. Сформулируем теперь КР-алгоритм, который определяется пунктами $\mathcal{CR}1$, $\mathcal{CR}2$.

$\mathcal{CR}1$) В процессе переноса пробег из \mathbf{r}' в \mathbf{r} моделируется по формуле $l = -\ln \alpha / \Sigma_{\max}(v)$.

$\mathcal{CR}2$) Если $l = |\mathbf{r} - \mathbf{r}'| < L$, то в алгоритме максимального сечения фиксируется $\rho(\mathbf{r}) = \rho(\mathbf{r}')$, иначе значение $\rho(\mathbf{r})$ выбирается случайно из одномерного распределения поля ρ .

Таким образом, КР-алгоритм строится путем рандомизации алгоритма максимального сечения с учетом значения L . Эвристически ясно, что этот алгоритм в случае анизотропного рассеяния в значительной степени соответствует корреляционной функции, близкой к “ступеньке” $k_0(t) = 1$ при $0 \leq t \leq L$. Как показывают приведенные выше эвристические соображения, а также модельные исследования и численные эксперименты, такой алгоритм дает удовлетворительные результаты для анизотропного рассеяния при достаточно малых значениях L ($L < 0.4 / \Sigma_{\max}$), когда стандартное моделирование слишком трудоемко. При этом значение L в алгоритме можно корректировать (особенно в односкоростном случае) на основе упрощенных тестовых задач [12]. Ясно, что трудоемкость КР-алгоритма ограничена при $L \rightarrow 0$ и его можно использовать для оценки соответствующей скорости сходимости в стохастических задачах переноса даже в неограниченных областях. Отметим, что в случае неанизотропного однородного поля ρ

можно также использовать КР-алгоритм, заменив L на $L(\omega), \omega = (\mathbf{r} - \mathbf{r}') / |\mathbf{r} - \mathbf{r}'|$. В заключение разд. 2 заметим, что КР-алгоритм при $L = 0$ реализует моделирование траекторий в осредненной среде, т.е. для $\rho(\mathbf{r}) \equiv E\rho$. Это получается непосредственно из (2) по формуле полной вероятности.

1.4. Усовершенствованный КР-алгоритм (ФКРА). Далее предполагается, что нормированная корреляционная функция $k(t)$ поля $\rho(\mathbf{r})$ неотрицательна: $k(t) \geq 0$. Можно полагать, что КР-алгоритм реализует выполнение равенства $\rho(\mathbf{r}) = \rho(\mathbf{r}')$ в алгоритме выравнивания с вероятностью $k_0(l) = 1$ при $l \leq L$, где $l = |\mathbf{r} - \mathbf{r}'|$. Это дает основание для рассмотрения функционального корреляционно-рандомизированного алгоритма (ФКРА), который по аналогии с КР-алгоритмом (КРА) строится следующим образом:

FCR1) Пробег l из \mathbf{r}' в \mathbf{r} моделируется по формуле $l = -\ln \alpha / \Sigma_{\max}$.

FCR2) Реализуется случайное число α_1 ; если $\alpha_1 < k(l)$, то $\rho(\mathbf{r}) = \rho(\mathbf{r}')$, иначе $\rho(\mathbf{r})$ выбирается из одномерного распределения поля ρ .

Однако для построения ФКРА требуется оценка корреляционной функции $k(t)$ в целом, а для КРА достаточно каким-либо образом определить лишь значение L , т.е. корреляционный масштаб среды.

Для ФКРА выполняется соотношение

$$\begin{aligned} E[\rho(\mathbf{r}')\rho(\mathbf{r}(l))|l] &= \\ &= k(l)E\rho^2 + [1 - k(l)](E\rho)^2 = k(l) \cdot D\rho. \end{aligned}$$

Следовательно, коэффициент корреляции между \mathbf{r}' и $\mathbf{r}(l)$ при фиксированном l равен $k(l)$. По-видимому, это означает, что ФКРА обеспечивает хорошую точность оценки среднего числа столкновений в заданной области. Впрочем, это предположение требует дополнительного исследования.

В заключение этого раздела отметим, что корреляционно-рандомизированные алгоритмы, в отличие от алгоритмов двойной рандомизации, дают возможность исследовать прохождение частиц через бесконечно протяженный слой вещества и тем самым погрешность оценок, возникающую при замене такого слоя на “брусек” (см. далее разд. 4).

2. НОВАЯ КОМПЬЮТЕРНО-ЭКОНОМИЧНАЯ СЕТОЧНАЯ АППРОКСИМАЦИЯ $\rho_h(\mathbf{r})$ ОДНОРОДНОГО СЛУЧАЙНОГО ПОЛЯ $\rho(\mathbf{r})$

2.1. Основанный на рандомизации алгоритма максимального сечения КР-алгоритм реализует вычислительную аппроксимацию процесса переноса. Он весьма экономичен, так как не требует реализации поля $\rho(\mathbf{r})$ в целом и двойной рандомизации (см. разд. 1). Однако этот алгоритм в основном эффективен для существенно анизотропного рассеяния при условии $L < 0.4 / \Sigma_{\max}$ [12]. Недостатком КРА является также невозможность оценки физически значимой величины дисперсии $D(\zeta | \rho)$ за исключением бернуллиевой оценки с малым параметром $p = E\zeta$ (см. далее п. св. 5.3).

Эти соображения определили мотивацию разработки компьютерно-экономичной сеточной аппроксимации поля плотности $\rho_h(\mathbf{r})$, которая фактически представляет геометрически детерминированную стохастическую мозаику, аналогичную полю Вороного и в определенном далее смысле аппроксимирующую его. Мозаика Вороного строится на основе базового пуассоновского точечного потока $\{\mathbf{r}_i\}, i = 0, 1, 2, 3, \dots$, интенсивности λ_p , определяющего разбиение пространства R^3 на ячейки $\{S\}$, каждая из которых является множеством точек, наиболее близких к одной из точек потока. Геометрические свойства такого разбиения детально изучены, например, в [9]. Элементы разбиения являются выпуклыми многогранниками, и поэтому величина (3) вследствие однородности поля ρ преобразуется к виду

$$L = \frac{1}{4\pi} \int l_S(\mathbf{r}, \omega) d\omega \approx 0.459\lambda_p^{-1/3}, \tag{4}$$

где $\forall \mathbf{r} \in S, l_S(\mathbf{r}, \omega)$ — расстояние от точки \mathbf{r} до границы элемента S вдоль направления ω .

Отметим, что в работе [13] было показано и в дальнейшем дополнительно проверено расчетами, что нормированная корреляционная функция $k_L(t)$ поля Вороного хорошо аппроксимируется экспонентой

$$k_L(t) \approx e^{-t/L}. \tag{5}$$

Формула (5) позволяет строить приближенный ФКР-алгоритм, уточняющий КРА (см. далее п. св. 5.2)

2.2. Новая сеточная аппроксимация (сеточная мозаика) $\rho_h(\mathbf{r})$ строится на основе разбиения пространства R^3 прямоугольной сеткой, т.е. на кубики $\{S_h\}$ со стороной h . Для каждого элемента S_h

независимо реализуется случайное значение ρ соответственно заданному одномерному распределению поля $\rho(\mathbf{r})$. Эффективное значение $L = l_h$ поля $\rho_h(\mathbf{r})$ строится по аналогии с (4) путем равномерного осреднения корреляционной длины в этом поле:

$$l_h = \frac{h^{-3}}{4\pi} \int_{S_h} \int l_{S_h}(\mathbf{r}, \omega) d\mathbf{r} d\omega. \quad (6)$$

Ясно, что выполняются равенства

$$L = l_h = h l_1, \quad h = L / l_1, \quad (7)$$

причем достаточно точные вычисления дали значение

$$l_1 = 0.44831$$

с погрешностью, не превышающей 0.00001.

Трудоёмкость (сложность) мозаики $\rho_h(\mathbf{r})$ существенно меньше трудоёмкости мозаики Вороного, так как для определения значения $\rho_h(\mathbf{r})$ не требуется перебора значений ρ , определенных на начальном этапе двойной рандомизации для всех элементов сетки. Номер элемента прямоугольной сетки, в которой находится точка $\mathbf{r} = (x_1, x_2, x_3)$ определяется очевидным образом с помощью стандартной встроеной функции вычисления целой части (например, $\text{floor}(x_i / h)$ в C++).

3. МИНИМИЗАЦИЯ ТРУДОЕМКОСТИ АЛГОРИТМА И ОЦЕНКА ПРАТИЧЕСКИ ЗНАЧИМЫХ ЗНАЧЕНИЙ ПОГРЕШНОСТИ

Известно (см., например, [2]), что трудоёмкость алгоритма метода Монте-Карло определяется величиной $S(\zeta) = T_\zeta D\zeta$, где T_ζ — среднее время моделирования на ЭВМ для получения одного выборочного значения ζ . Представленные ниже расчеты для стохастической среды с мозаикой Вороного и ее сеточной аппроксимации были проведены простым вариантом “метода двойной рандомизации” (см. п. св. 1.2), в котором для каждой реализации плотности $\rho(\mathbf{r})$ моделировалась одна траектория частицы. Если для каждой реализации плотности моделировать m условно-независимых траекторий частиц, то согласно “формуле полной дисперсии” (см., например, [2]), величину $D\zeta_m$ можно представить в виде

$$D = D\zeta_m = D_0 + \frac{D_1}{m}, \quad (8)$$

$$D_0 = DE(\zeta | \rho), \quad D_1 = ED(\zeta | \rho),$$

где D_0 — физически значимая дисперсия оценки по флуктуациям поля ρ , а D_1 — среднее значение условной дисперсии по траекториям Ω для фиксированной реализации поля ρ . Отметим, что здесь D — “элементарная” дисперсия (на одну траекторию). Выполняется очевидное соотношение для среднего времени реализации оценки ζ_m

$$T_{\zeta_m} = T_0 + m T_1, \quad (9)$$

где T_0 — среднее время, затрачиваемое на реализацию случайного поля, а T_1 — среднее время, затрачиваемое на моделирование одной случайной траектории ζ . Величина трудоёмкости $S(\zeta_m)$ в этом случае, как функция от m , имеет минимум в точке (см., например [2])

$$m_{\text{opt}} = \sqrt{\frac{T_0 D_1}{T_1 D_0}} \quad \text{и} \quad S(\zeta_{m_{\text{opt}}}) = (\sqrt{T_0 D_0} + \sqrt{T_1 D_1})^2. \quad (10)$$

Если ζ — “бернуллиева” оценка для малой вероятности $p = E\zeta$, то

$$D_1 = ED(\zeta | \rho) = Ep(\rho)[1 - p(\rho)] \approx p = \\ = E\zeta \text{ и } D_0 \approx D\zeta_m - E\zeta / m.$$

Значение m_{opt} можно оценить, сделав два предварительных расчета для $m = 1$ и $m > 1$, подставив полученные значения для $D\zeta_1, D\zeta_m$ в (8) и найдя решение системы из 2-х линейных уравнений для D_0, D_1 . В этой системе можно использовать в качестве D дисперсию среднего значения ζ_m , так как в (9) фигурирует отношение D_1 / D_0 . Аналогичным способом определяются значения T_0, T_1 из (9) для полученных из расчетов значений T_{ζ_1}, T_{ζ_k} . Отметим еще раз, что D_0 — физически значимая дисперсия при условии, что D элементарна.

4. ТЕСТОВАЯ ЗАДАЧА О ПРОХОЖДЕНИИ ГАММА-КВАНТОВ СКВОЗЬ СТОХАСТИЧЕСКУЮ СРЕДУ

4.1. В качестве исходной стохастической среды рассмотрим мозаику Вороного в “бруске”

$$\mathbf{V}_H = [0, X_{max}] \times [0, Y_{max}] \times [0, H] \in R^3$$

и будем полагать, что элементы такого разбиения (выпуклые многогранники) случайно заполнены водой с заданной вероятностью ρ или воздухом с заданной вероятностью $1 - \rho$. Известно (см., например, [8]), что при моделировании пуассоновского потока точек в ограниченной области случайное поле типа мозаики Вороного не является однородным и изотропным, поэтому пуассоновский поток следует моделировать в области большего размера, чтобы мозаичное поле было близко к однородному и изотропному. На основе предварительных расчетов [8] было решено использовать брусок объема:

$$\mathbf{V}_{H,2L} = [-2L, X_{max} + 2L] \times [-2L, Y_{max} + 2L] \times [-2L, H + 2L].$$

Использование “бруска” связано с невозможностью реализации поля $\rho(r)$ в неограниченной области. Отметим, что сравнение вероятностей прохождения квантов через слой и “брусок” было реализовано в [12] с помощью КР-алгоритма.

Моделирование мозаичного поля Вороного в $\mathbf{V}_{H,2L}$ строилось в два этапа. Сначала реализовывалось пуассоновское число η случайных точек $\{\mathbf{r}_i\}$ согласно параметру $\lambda_p (X_{max} + 4L)(Y_{max} + 4L)(H + 4L)$, где λ_p — интенсивность пуассоновского точечного потока. Далее координаты точек $\{\mathbf{r}_i\}, i = 1, 2, \dots, \eta$, выбирались независимо и равномерно в области $\mathbf{V}_{H,2L}$, а случайный выбор вещества (вода или воздух) в каждой ячейке S_i , построенной около точки \mathbf{r}_i , независимо разыгрывался с вероятностью ρ и $1 - \rho$ соответственно. Здесь стоит отметить, что вероятность ρ определяет среднюю “геометрическую” плотность смеси в области $\mathbf{V}_{H,2L}$.

Пусть в центре основания бруска \mathbf{V}_H в точке $(50, 50, 0)$ расположен точечный источник, испускающий гамма-кванты в направлении ω_0 , перпендикулярном плоскости XY , с начальной энергией 1 МэВ. Полагаем, что вне бруска \mathbf{V}_H находится абсолютное “черное” подпространство, поглощающее выходящее гамма-излучение. Определим в качестве основных функционалов -

- вероятность $P(\text{up})$ прохождения гамма-квантов через верхнюю площадку бруска \mathbf{V}_{100} ($z = H = 100$ см);
- вероятность $P(\text{down})$ вылета гамма-квантов через нижнюю площадку бруска \mathbf{V}_{100} ($z = 0$ см).

Дополнительно в проведенных расчетах подсчитывались средние значения энергии $E(\text{up}_{Det}), E(\text{down}_{Det})$ частиц регистрируемых в плоских детекторах, расположенных на верхней ($z = H = 100$ см) и нижней ($z = 0$ см) площадках бруска \mathbf{V}_H в квадратах $-10 \leq x \leq 10, -10 \leq y \leq 10$.

4.2. Остановимся теперь на использовании физической модели переноса гамма-квантов с учетом комптоновского рассеяния и поглощения в результате фотоэффекта в среде M (здесь вода или воздух). Микроскопическое сечение комптоновского рассеяния определяются выражением [14], [15]

$$\tilde{\Sigma}(\tilde{E}) = 2\pi r_e^2 \left\{ \frac{1 + \tilde{E}}{\tilde{E}^3} \left[\frac{2\tilde{E}(1 + \tilde{E})}{1 + 2\tilde{E}} - \ln(1 + 2\tilde{E}) \right] + \frac{1}{2\tilde{E}} \ln(1 + 2\tilde{E}) - \frac{1 + 3\tilde{E}}{(1 + 2\tilde{E})^2} \right\},$$

где r_e — классический радиус электрона. Здесь \tilde{E} — безразмерная энергия гамма-кванта, т.е. энергия в электрон-вольтах, поделенная на $m_e c^2$, где m_e — масса покоя электрона, а c^2 — скорость света. Макроскопическое сечение комптоновского рассеяния задается соотношением

$$\Sigma_k(E) = \rho_m N_A \frac{Z}{A} \tilde{\Sigma}(\tilde{E}), \quad (11)$$

где ρ_m — плотность среды, N_A — число Авогадро, Z и A — число электронов и молекулярный вес среды. Значения макроскопического сечения поглощения $\sigma_c(E)$ в процессе расчета можно вычислять на основе табличных данных (см., например [14]) с использованием линейной интерполяции. Отметим, что введение в число параметров процесса переноса энергии E формально меняет уравнение 1.

Траектория гамма-кванта реализуется по следующей схеме.

- 0) Выбор начальной точки соответственно плотности распределения источника $\pi(x)$.
- 1) Моделирование случайной длины свободного пробега l согласно по формуле $l = (\ln \alpha) / \Sigma(E)$.
- 2) Проверка вылета частицы из среды M .
- 3) Пересчет координат радиус-вектора \mathbf{r} нового положения частицы $\mathbf{r}' = \mathbf{r} + \omega l$.
- 4) Вычисление нового значения энергии E' . Известно [15], что если гамма-квант с безразмерной энергией E рассеивается в результате комптон-эффекта, то индикатриса рассеяния $g(E, E')$ пропорциональна функции

$$f(E, E') = \frac{E'}{E} + \frac{E}{E'} + \left(\frac{1}{E} - \frac{1}{E'} \right) \left(2 + \frac{1}{E} - \frac{1}{E'} \right),$$

$$E(1 + 2E)^{-1} < E' < E.$$

Это выражение непосредственно следует из формулы Клейна-Нишины для дифференциального сечения комптоновского рассеяния. Здесь и далее символом α мы будем обозначать случайное число, равномерно распределенное в интервале $(0, 1)$. Случайную величину E' здесь можно моделировать с помощью простого варианта метода исключения [15]:

а) реализуем $E_0 = E(1 + 2E\alpha_1)(1 + 2E)^{-1}$;

б) если $\alpha_2[1 + 2E + (1 + 2E)^{-1}] > f(E_0, E)$, то снова выполняется п. а) и т.д., иначе $E' = E_0$.

Этот алгоритм вытекает из неравенства $f(E, E') \leq 1 + 2E + (1 + 2E)^{-1}$.

5) Выбор случайного типа столкновения: комптоновское рассеяние или поглощение с вероятностями $q_k(E') = \Sigma_k(E') / \Sigma(E')$ и $\Sigma_c(E') / \Sigma(E')$ соответственно.

б) В случае комптоновского рассеяния частицы — пересчет величины косинуса угла рассеяния между старым $\omega = (\omega_x, \omega_y, \omega_z)$ и новым $\omega' = (\omega'_x, \omega'_y, \omega'_z)$ направлениями пробега [16], [15]

$$\mu(\omega, \omega') = 1 - \frac{1}{E'} + \frac{1}{E}.$$

Далее реализуется моделирование азимутального угла рассеяния θ по формуле $\theta = 2\pi\alpha$ и пересчет координат нового вектора направления ω' по стандартным формулам (см., например, [2], [15]).

7) Переход на п. 1 для дальнейшего моделирования.

5. РЕЗУЛЬТАТЫ РАСЧЕТОВ ДЛЯ РЕШЕНИЯ ТЕСТОВОЙ ЗАДАЧИ

5.1. Для заданных значений

$$X_{\max} = Y_{\max} = H = 100 \text{ см}, \rho = 0.9 \text{ и } \lambda_p = 0.00005, 0.0001, 0.001, 0.01, 0.015$$

были реализованы три схемы моделирования переноса гамма излучения через брусок \mathbf{V}_{100} . Первое и наиболее трудоемкое прямое моделирование (Direct) случайной траектории гамма-кванта в рассматриваемой стохастически неоднородной среде для построения “бернулиевой” несмещенной оценки осуществлялась в два этапа. Сначала строилась реализация поля Вороного по изложенному выше рандомизированному алгоритму в области $\mathbf{V}_{H, 2L}$, а потом, для такой реализации поля моделировалась случайная траектория гамма-кванта по изложенной выше общей схеме моделирования (см. п. св. 4.2) с сечениями

из [14]. Поскольку в задачах переноса гамма-излучения сечения среды и примесей зависят от энергии E , то здесь и далее мы полагаем $\Sigma = \Sigma(E)$. В рассматриваемой стохастической задаче переноса гамма-квантов со случайным полем плотности среды для моделирования случайной длины свободного пробега использовался метод максимального сечения (см.п. св. 1.2), а после очередного случайного пробега проводилась проверка, в какую из ячеек попала частица. Именно на такую проверку и реализацию поля Вороного тратилось существенное время прямого моделирования для каждой траектории уже при $\lambda_p \geq 0.001$ (см. табл. 1).

Во второй схеме была реализована новая сеточная аппроксимация (сеточная мозаика) области (см.п. 2.2). Исходный “брусоч” V_H был разбит на кубики S_h так, что количество кубиков для каждой реализации составляло (см. соотношение (7))

$$\frac{X_{\max} Y_{\max} H}{h^3} = \frac{X_{\max} Y_{\max} H l_1^3}{L^3} = 0.93 \lambda_p X_{\max} Y_{\max} H,$$

а случайный выбор типа вещества в каждом кубике S_h независимо “разыгрывался” с вероятностью p (вода) и $1 - p$ (воздух) соответственно. Пусть N_V это среднее число центров ячеек мозаики Вороного внутри V_H . В данном случае $X_{\max} = Y_{\max} = H = 100$ см и имеет место соответствие, представленное в табл. 1

Таблица 1. Значения корреляционного радиуса L , среднего числа центров ячеек N_V мозаики Вороного и качества кубиков сеточной мозаики $0.93 \lambda_p X_{\max} Y_{\max} H_{\max}$ в зависимости от λ_p .

λ_p	0.00005	0.0001	0.001	0.01	0.015
L	12.46	9.89	4.59	2.13	1.86
N_V	168.19	271.79	1658.12	12780.6	18605.7
$0.93 \lambda_p X_{\max} Y_{\max} H_{\max}$	46.5	93	930	9300	13950

В третьей схеме моделирования перенос гамма-квантов в среде определялся КР-алгоритмом (см. п. 1.3), в котором значение L определялось выражением (4).

5.2. Введем следующие обозначения для относительных погрешностей $\delta_{\text{Grid}} G(\cdot)$, $\delta_{\text{CR}} G(\cdot)$ оценок функционала $G(\cdot)$ при использовании сеточной аппроксимации и КР-алгоритма соответственно:

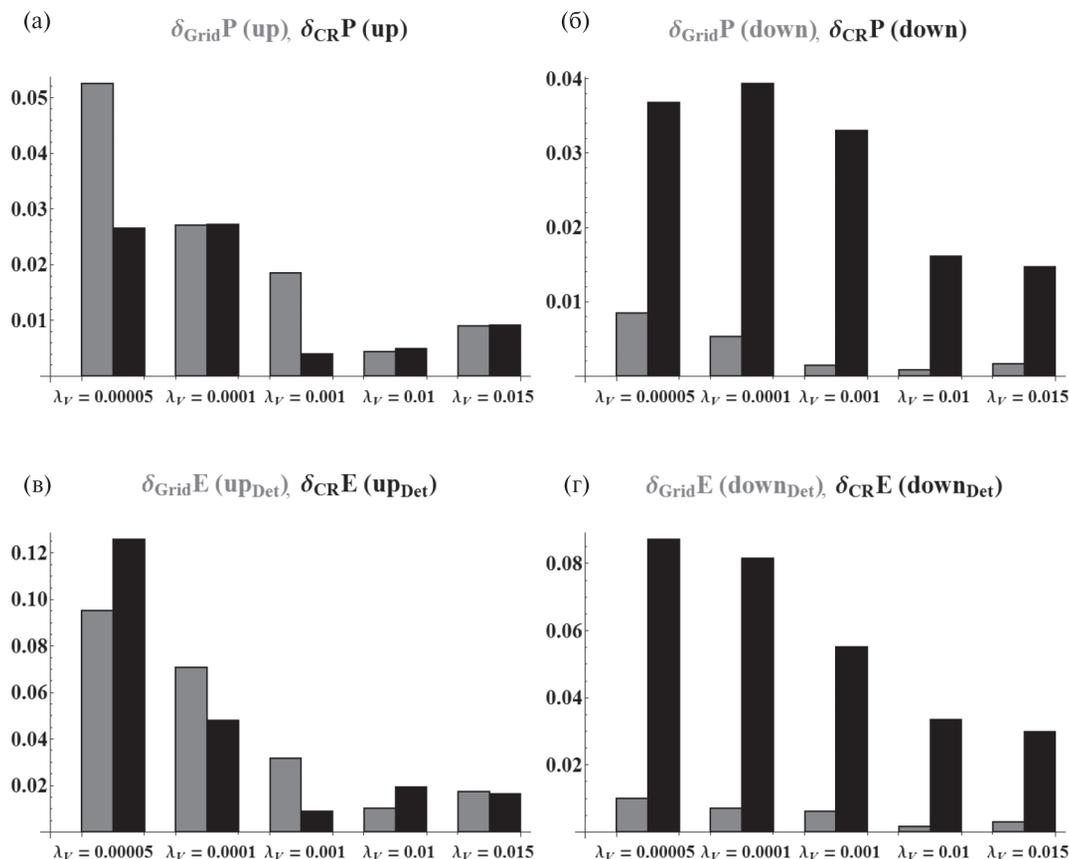
$$\delta_{\text{Grid}} G(\cdot) = \frac{|G_{\text{Direct}}(\cdot) - G_{\text{Grid}}(\cdot)|}{G_{\text{Direct}}(\cdot)},$$

$$\delta_{\text{CR}} G(\cdot) = \frac{|G_{\text{Direct}}(\cdot) - G_{\text{CR}}(\cdot)|}{G_{\text{Direct}}(\cdot)},$$

относительно $G_{\text{Direct}}(\cdot)$ — “точной” оценки, полученной для стохастической среды с мозаикой Вороного. На фиг. 1 представлены графически значения относительных погрешностей функционалов

$$P(\text{up}), P(\text{down}), E(\text{up}_{\text{Det}}), E(\text{down}_{\text{Det}}) \tag{12}$$

для $\lambda_p = 0.00005, 0.0001, 0.001, 0.01, 0.015$ и начальной энергии 1 Мэв. Для представленных оценок моделировалось 10^8 траекторий гамма-квантов, причем относительные среднеквадратические статистические погрешности определенных функционалов (12), полученные с использованием алгоритмов Direct, Grid, CR составили менее 0.005, что дает крайне незначительную погрешность в масштабе столбцов фиг. 1. Серым цветом отмечены относительные погрешности функционалов, полученные с использованием сеточной мозаики (Grid), а черным — КР-алгоритмом (CR). Как видно из фиг. 1, при использовании сеточной аппроксимации (Grid) погрешности оценок функционалов $P(\text{down}), E(\text{down}_{\text{Det}})$ меньше 1% и в 4–10 раз меньше соответствующих значений погрешностей функционалов, полученных при использовании КР-алгоритма (CR). При этом значения погрешностей оценок функционалов $P(\text{up}), E(\text{up}_{\text{Det}})$, полученных при использовании сеточной аппроксимации и КР-алгоритма несущественно отличаются друг от друга и, следовательно, КР-алгоритм существенно выигрывает по трудоемкости согласно табл. 3 (Точечный источник с начальной энергией 1 Мэв).

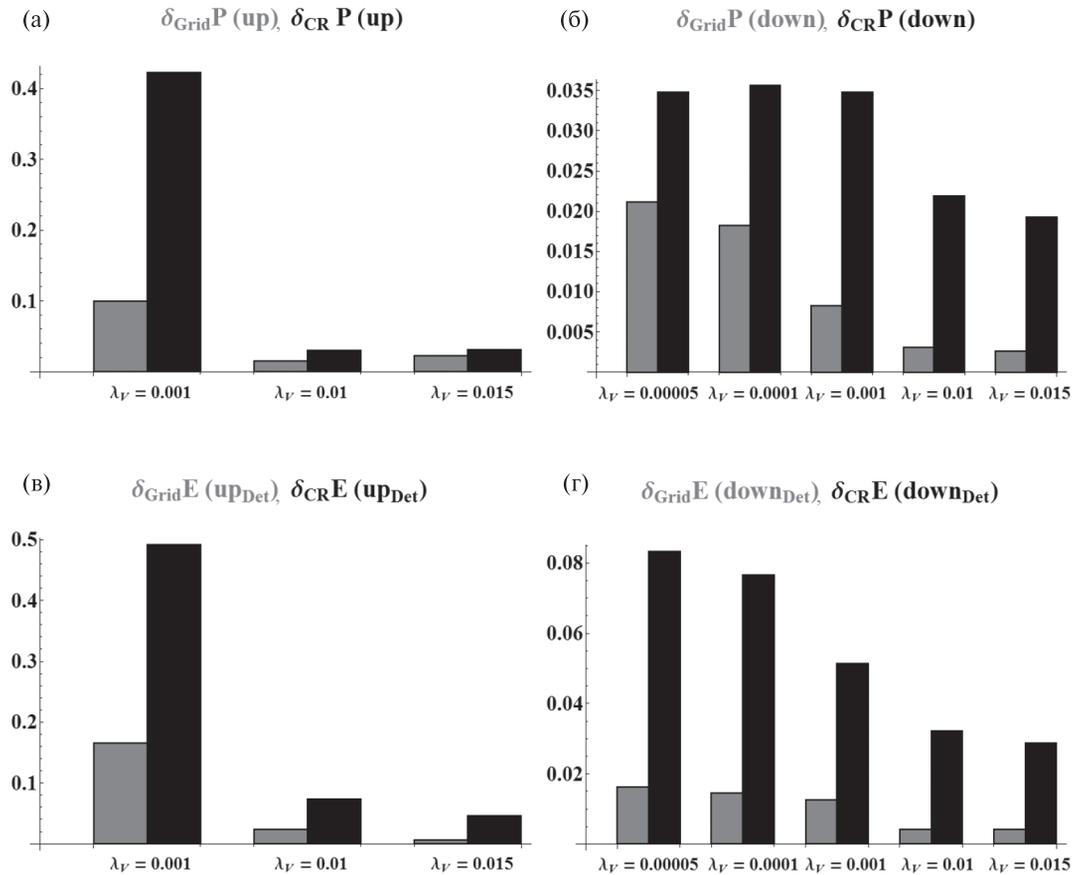


Фиг 1. Значения относительных погрешностей функционалов $P(\text{up})$, $P(\text{down})$, $E(\text{up}_{\text{Det}})$, $E(\text{down}_{\text{Det}})$, полученные с использованием приближенных алгоритмов (Grid, CR). Точечный источник с начальной энергией 1 МэВ.

Таблица 2. Значения относительных погрешностей оценок функционалов $P(\text{up})$, $P(\text{down})$, полученные с использованием приближенных алгоритмов (Grid, CR) при $\lambda_p = 0.00005, 0.0001, 0.001, 0.01, 0.015$. Точечный источник с начальной энергией 0.1 МэВ.

λ_p	$\delta_{\text{Grid}}P(\text{up})$	$\delta_{\text{CR}}P(\text{up})$	$\delta_{\text{FCR}}P(\text{up})$	$\delta_{\text{Grid}}P(\text{down})$	$\delta_{\text{CR}}P(\text{down})$	$\delta_{\text{FCR}}P(\text{down})$
0.00005	0.509	5.63	0.078	0.023	0.021	0.078
0.0001	0.285	5.75	0.025	0.018	0.035	0.041
0.001	0.099	0.423	0.078	0.008	0.008	0.008
0.01	0.015	0.024	0.030	0.003	0.021	0.019
0.015	0.023	0.031	0.005	0.002	0.019	0.069

Были проведены также аналогичные расчеты для источника с начальной энергией 0.1 МэВ. Полученные результаты представлены на фиг. 2 и в табл. 2. Как видно из фиг. 2, при использовании сеточной аппроксимации (Grid) полученные значения относительных погрешностей оценок функционалов $P(\text{up})$, $P(\text{down})$, $E(\text{up}_{\text{Det}})$, $E(\text{down}_{\text{Det}})$ существенно меньше соответствующих значений относительных погрешностей функционалов, полученных при использовании КР-алгоритма (CR). Более того, при $\lambda_p = 0.00005, 0.0001$ значения $\delta_{\text{Grid}}P(\text{up})$ оказались в 11–20 раз меньше соответствующих значений $\delta_{\text{CR}}P(\text{up})$ (выделенные жирным шрифтом в табл. 2). Как и в модельной задаче из [12] это связано с существенным нарушением условия $L < 0.4 / \Sigma_{\text{max}}(E)$ применимости КР-алгоритма (КРА) в данном варианте расчетов. Поэтому для приближенной оценки функционалов $P(\text{up})$, $P(\text{down})$ здесь был также реализован ФКР алгоритм (FCR) из п. св. 1.2 с корреляционной функцией 5. Как видно из табл. 2, при использовании ФКР алгоритма полученные значения относительных погрешностей оценок функционалов $P(\text{up})$, $P(\text{down})$ в целом существенно меньше соответствующих значений $\delta_{\text{CR}}P(\text{up})$, $\delta_{\text{CR}}P(\text{down})$



Фиг. 2. Значения относительных погрешностей оценок функционалов $P(\text{up}), P(\text{down}), E(\text{up}_{det}), E(\text{down}_{Det})$, полученные с использованием приближенных алгоритмов (Grid, CR). Точечный источник с начальной энергией 0.1 МэВ.

Таблица 3. Время (мин), затраченное на моделирование 10^8 траекторий гамма-квантов с использованием алгоритмов Direct, Grid, CR.

Точечный источник с начальной энергией 0.1 МэВ					
λ_p	0.00005	0.0001	0.001	0.01	0.015
Direct	30.1	45.4	255.6	1935.4	2791.6
Grid	6.4	7.2	28.9	265.3	385.3
CR	3.4	3.4	3.5	3.4	3.4

Точечный источник с начальной энергией 1 МэВ					
λ_p	0.00005	0.0001	0.001	0.01	0.015
Direct	40.1	60.4	334.5	2502.3	3590.2
Grid	9.1	9.8	33.5	278.3	403.5
CR	5.8	5.6	5.7	5.8	5.9

и сопоставимы со значениями $\delta_{Grid}P(\text{up}), \delta_{Grid}P(\text{down})$. Это связано с использованием в ФКР корреляционной функции, которая не требуется для построения сеточной аппроксимации. Интересно отметить, что в варианте с $E_0 = 1$ МэВ для функционалов $P(\cdot)$ при переходе от КРА к ФКРА погрешность практически не изменилась, так как здесь условие применимости КРА в основном выполнено.

Как было замечено ранее, при получении “точной” оценки для исходной стохастической среды с мозаикой Вороного существенное время моделирования траектории занимает реализация такого поля и проверка, в какую из ячеек и с какой примесью попала частица. Время реализации сеточной мозаики и определение номера прямоугольной сетки существенно меньше, а в КР-алгоритме приближенное случайное поле реализуется в каждой точке столкновения из одномерного распределения ρ . Это наглядно продемонстрировано в табл. 1, где представлено время (в минутах), затраченное на моделирование 10^8 траекторий гамма-квантов с использованием алгоритмов Direct, Grid, CR для оценки значений относительных погрешностей функционалов $P(\text{up}), P(\text{down}), E(\text{up}_{\text{Det}}), E(\text{down}_{\text{Det}})$ (см. фиг. 1). Например, при $\lambda = 0.015 (L = 1.86)$ и мощности источника 0.1 Мэв, значения $\delta_{\text{Grid}} P(\text{up}), \delta_{\text{CR}} P(\text{up})$ практически совпадают, а трудоемкость алгоритма Direct (время расчета) в 7.2 раза больше соответствующего значения трудоемкости алгоритма Grid, которое, в свою очередь, в 113 раз больше значения трудоемкости алгоритма CR.

Замечание. В проведенных расчетах реализация сеточной мозаики была проведена согласно п. 2.2. При такой реализации для достаточно мелкой сетки существенное время моделирования тратится на реализацию независимых случайных начальных значений ρ во всех кубиках S_h . В таком случае можно значительно уменьшить время моделирования путем реализации случайных значений ρ лишь в тех кубиках, которые “посещает” моделируемая частица. Дополнительные расчеты показали, что время расчетов в алгоритме Grid при этом всего на 5–10% больше соответствующих значений времени для алгоритма CR (см. табл. 1) и практически ограничено при уменьшении корреляционного масштаба среды.

5.3. Рассмотрим теперь в качестве примера оптимизацию трудоемкости приближенной оценки $P_{\text{Grid}}(\text{down})$ при использовании сеточной аппроксимации для источника с начальной энергией 0.1 Мэв. Для $\lambda_p = 0.001$ были проведены предварительные расчеты путем моделирования 10^7 траекторий для $m = 1, 10$ и получены из (8) и (9) следующие оценки:

$$D_0 \approx 0.0052, \quad D_1 \approx 0.2297, \quad T_0 \approx 17.5156, \\ T_1 \approx 150.91, \quad m_{\text{opt}} = 17.515.$$

Поскольку полученное значение m_{opt} не является целым, то целесообразно воспользоваться несмещенной модификацией оценки ζ_m с ветвлением траектории, в которой после реализации поля моделируется $m = 17$ или $m = 18$ траекторий с вероятностями 1 – 0.515 и 0.515 соответственно. Именно такое распределение числа траекторий гарантирует минимальную среднеквадратичную погрешность в классе произвольных целочисленных распределений с фиксированным средним значением 17.515 (см. [17]). В разд. 3 была указана также приближенная: оценка $D_1 \approx p = P_{\text{Grid}}(\text{down}) = E\zeta_1 = 0.234996$; соответствующее приближенное оптимальное значение m_{opt} здесь составило 18.67.

Таблица 4. Значение величины трудоемкости $S(\zeta_m)$ для $m = 1, 10, 17.515, 50, 100$ ($m_{\text{opt}} = 17.515$) для оценки $P_{\text{Grid}}(\text{down})$. Точечный источник с начальной энергией 0.1 Мэв.

m	1	10	17.515	50	100
$D\zeta_m$	$2.349 \cdot 10^{-1}$	$2.818 \cdot 10^{-1}$	$1.873 \cdot 10^{-2}$	$9.799 \cdot 10^{-3}$	$7.503 \cdot 10^{-3}$
T_{ζ_m} (сек)	172	368	518	1218	2284
$S(\zeta_m)$	40.56	10.37	9.71	11.94	17.13

В табл. 4 представлены полученные в расчетах значения трудоемкости при $\lambda_p = 0.001$ для оценки $P(\text{down})$ с использованием сеточной аппроксимации при $m = 1, 10, 17.515, 50, 100$; видно, что минимальное значение $S(\zeta_m)$ составило **9.71** при $m = m_{\text{opt}} = 17.515$. Данные из табл. 1 показывают, что здесь использование оценок сеточной аппроксимации (CR) сравнительно с моделированием поля Вороного (Direct) сокращает трудоемкость примерно в 40 раз.

6. ДОПОЛНИТЕЛЬНЫЕ СРАВНИТЕЛЬНЫЕ РАСЧЕТЫ ДЛЯ СЛУЧАЙНЫХ ПОЛЕЙ РАЗНОГО ТИПА

В работе [8] было замечено, что вероятность прохождения в значительной степени определяются корреляционной длиной L и одномерным распределением поля ρ . Рассмотрим следующую стохастическую модель среды, описанную в [18]. Предположим, что “брусоч” V_H заполнен смесью вещества W (вода)

и ансамблем шаров из воздуха $\mathbf{B}_{r_b}(r_i)$ радиуса r_b с центрами в точках $r_i, i = 0, 1, 2, \dots, \eta$. Таким образом, $\mathbf{V}_H = \mathbf{V}_W \cup \mathbf{V}_{\text{air}}$, где \mathbf{V}_{air} — часть области \mathbf{V}_H , заполненная, возможно пересекающимися, шарами из воздуха. Предположим, что координаты центров шаров образуют однородный пуассоновский ансамбль $R_\eta(\mathbf{V}_{H,r_b})$ в области

$$\mathbf{V}_{H,r_b} = [-r_b, X_{\max} + r_b] \times [-r_b, Y_{\max} + r_b] \times [-r_b, H + r_b]$$

с мерой интенсивности $\Lambda(\mathbf{V}_{H,r_b}) = \lambda_p * \text{Vol}(\mathbf{V}_{H,r_b})$, где $\text{Vol}(\mathbf{V}_{H,r_b}) = (X_{\max} + 2r_b)(Y_{\max} + 2r_b)(H + 2r_b)$. Для этой модели стохастической среды средняя "геометрическая" плотность $\rho(\mathbf{V}_H)$ смеси в бруске \mathbf{V}_H определяется через вероятность $P(\mathbf{r} \in \mathbf{V}_W)$ следующим образом [10]:

$$\begin{aligned} \rho(\mathbf{V}_H) &= P(\mathbf{r} \in W) = P\left(R_\eta(\mathbf{V}_{H,r_b}) \subset \mathbf{V}_{H,r_b} \setminus \mathbf{B}_{r_b}\right) = \\ &= e^{-\Lambda(\mathbf{V}_{H,r_b}) \frac{4}{3}\pi r_b^3 / \text{Vol}(\mathbf{V}_{H,r_b})}. \end{aligned} \tag{13}$$

Для такого пуассоновского ансамбля шаров в бруске \mathbf{V}_{H,r_B} в работе [11] было найдено следующее выражение для корреляционной функции:

$$\begin{aligned} k_b(t) &= e^{-\Lambda V_b} \left(e^{V_d} - 1 \right) / \left(1 - e^{-\Lambda V_b} \right), \\ 0 \leq t = |\mathbf{r} - \mathbf{r}'| &\leq 2r_b, \end{aligned} \tag{14}$$

где

$$V_b = \frac{4}{3}\pi r_b^3, \text{ а } V_d = 2\frac{\pi d^2}{3}(3r_b - d), d = (2r_b - t) / 2,$$

есть объем общей части пересекающихся шаров радиуса r_b с центрами, отстоящими на расстоянии $t \leq 2r_b$, а $\Lambda = \Lambda(\mathbf{V}_{H,r_b}) / \text{Vol}(\mathbf{V}_{H,r_b})$. Полагая $k_b(t) = 0$ при $t > 2r_b$ для описанной модели с шарами можно определить корреляционную длину как

$$L_{\text{ball}}(r_b) = \int_0^\infty k_b(t) dt. \tag{15}$$

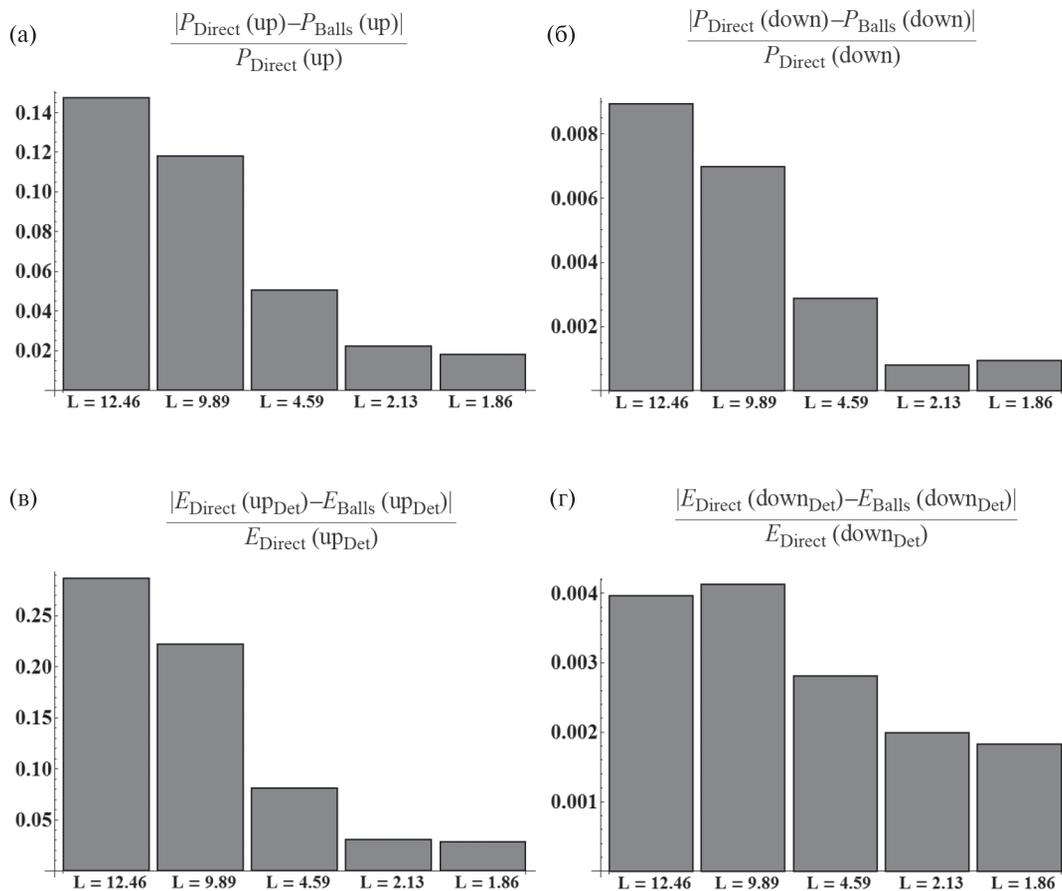
Была решена задача подбора параметров описанной выше стохастической модели среды с шарами так, чтобы величина $L_{\text{ball}}(r_b)$ совпала с корреляционным радиусом L (см. (4)) поля Вороного для заданных значений $\lambda_p = \lambda_{p_0}$ и $\rho(\mathbf{V}_H) = \rho_0$. Для этого было найдено численное значение $r_b = r_b(\lambda_{p_0})$ из равенства

$$0.459\lambda_p^{-1/3} = \int_0^{2r_b} k_b(t) dt,$$

а потом, используя полученное значение $r_b(\lambda_{p_0})$, определено значение меры интенсивности $\Lambda(\mathbf{V}_{H,r_b})(r_b(\lambda_{p_0}))$ (параметр распределения Пуассона) из соотношения

$$\Lambda(\mathbf{V}_{H,r_b})(r_b(\lambda_{p_0})) = -\frac{\ln \rho * \text{Vol}(\mathbf{V}_{H,r_b}(\lambda_{p_0}))}{\frac{4}{3}\pi r_b^3(\lambda_{p_0})}.$$

На фиг. 3 графически представлены относительные разности функционалов $P(\text{up}), P(\text{down}), E(\text{up}_{\text{Det}}), E(\text{down}_{\text{Det}})$ полученные для поля Вороного (Direct) и стохастической модели среды с шарами (Balls) при $L = L_{\text{ball}}(r_b)$ для $L = 12.46, 9.89, 4.59, 2.13, 1.86$ путем моделирования 10^8 траекторий гамма-квантов. Из фиг. 3 видно, что для поля Вороного и соответствующей ($L = L_{\text{ball}}(r_b)$) модели среды с шарами относительные разности функционалов $P(\text{up}), P(\text{down}), E(\text{up}_{\text{Det}}), E(\text{down}_{\text{Det}})$ уменьшаются с ростом L и меньше 6% при $L \leq 4.59$.



Фиг. 3. Значения относительных разностей функционалов $P(\text{up})$, $P(\text{down})$, $E(\text{up}_{\text{Det}})$, $E(\text{down}_{\text{Det}})$, полученные для поля Вороного и стохастической модели среды с шарами при $L = L_{\text{ball}}(r_b)$ для $\lambda_p = 0.00005, 0.0001, 0.001, 0.01, 0.015$. Точечный источник с начальной энергией 1 Мэв.

Рассмотрим теперь вкратце сеточную аппроксимацию неизотропного однородного случайного поля $\rho(\mathbf{r})$ с заданными корреляционными длинами L_x, L_y, L_z путем разбиения пространства R^3 на параллелепипеды со сторонами $(a_x, a_y, a_z) = \mathbf{a}$. Конкретную приближенную корреляционную длину, в частности, по оси Z при этом целесообразно определить формулой вида (4) с заменой $l_S(\mathbf{r}, \omega)$ на $l_S(\mathbf{r}, \omega) |\omega_z|$ (см., например, [19]). Остальные соотношения и алгоритм построения соответствующей оценки $\tilde{L}_z(\mathbf{a})$ обобщаются соответственно. В результате получается система из трех уравнений типа $\tilde{L}_z(\mathbf{a}) = L_z$, в качестве начального приближения решения которой естественно использовать $\mathbf{a}_0 = (L_x, L_y, L_z)$. Отметим, что если $L_z = \infty$, то и \tilde{L}_z бесконечно.

СПИСОК ЛИТЕРАТУРЫ

1. Дэвисон Б. Теория переноса нейтронов. М.: Атомиздат, 1960.
2. Марчук Г. И., Михайлов Г. А., Назаралиев М. А. и др. Метод Монте-Карло в атмосферной оптике. Новосибирск: Наука, 1976. [Engl. transl.: Springer-Verlag, 1980].
3. Спанье Дж., Гелбард З. Метод Монте-Карло и задачи теории переноса нейтронов. М.: Атомиздат, 1972.
4. Coleman W. A. Mathematical verification of a certain Monte Carlo sampling technique and applications of the techniques to radiation transport problems // J. Nucl. Sci. and Engng. 1968. V. 32. N 1. P. 76–81.
5. Woodcock E., Murphy T., Hemmings P., Longworth S. Techniques used in the GEM code for Monte Carlo neutronics calculations in reactors and other systems of complex geometry // Proc. Conf. Applications of Computing Methods to Reactor Problems. 1965. V. 557. P. 2.

6. Михайлов Г. А., Аверина Т. А. Алгоритм “максимального сечения” в методе Монте-Карло // Докл. АН. 2009. Т. 428. № 2. С. 163–165.
7. Михайлов Г. А. Рандомизированные алгоритмы метода Монте-Карло для задач со случайными параметрами (метод “двойной рандомизации”) // Сиб. ж. вычисл. матем. 2019. Т. 22. № 2. С. 187–200.
8. Ambos A. Y., Mikhailov G. A. Numerically statistical simulation of the intensity field of the radiation transmitted through a random medium // Russian Journal of Numerical Analysis and Math. Modelling. 2018. V. 33. N 3. P. 161–171.
9. Larmier C., Zoia A., Malvagi F., Dumonteil E., Mazzolo A. Monte Carlo particle transport in random media: The effects of mixing statistics // J. of Quantitative Spectroscopy and Radiative Transfer. 2017. V. 196. P. 270–286.
10. Глазов Г. Н., Титов Г. А. Статистические характеристики коэффициента ослабления в разорванной облачности. I. Модель с шарами одинакового радиуса // Вопросы лазерного зондирования атмосферы. Новосибирск, 1976. С. 126–139.
11. Михайлов Г. А., Медведев И. Н. Новый корреляционно рандомизированный алгоритм оценки влияния стохастичности среды на перенос частиц // Доклады АН. Математика, информатика, процессы управления. 2021. Т. 498. № 1. С. 55–58.
12. Medvedev I. N., Mikhailov G. A. New correlative randomized algorithms for statistical modelling of radiation transfer in stochastic medium // Russian Journal of Numerical Analysis and Math. Modelling. 2021. V. 36. N 4. P. 219–225.
13. Амбос А. Ю. Вычислительные модели мозаичных однородных изотропных случайных полей и задачи переноса излучения // Сиб. журн. вычисл. матем. 2016. Т. 19. № 1. С. 19–32.
14. Сторм Э., Исраэль Х. Сечения взаимодействия гамма-излучения (для энергий 0,001–100 МэВ и элементов с 1 по 100). М.: Атомиздат, 1973. 256 с.
15. Ермаков С. М., Михайлов Г. А. Статистическое моделирование. М.: Наука, 1982. 296 с.
16. Фано У., Спенсер Л., Бергер М. Перенос гамма-излучения. М.: Госатомиздат, 1963. 284 с.
17. Бреднихин С. А., Медведев И. Н., Михайлов Г. А. Оценка параметров критичности ветвящихся процессов методом Монте-Карло // Журн. вычисл. матем. и матем. физ. 2010. Т. 50. № 2. С. 362–374.
18. Medvedev I. N., Mikhailov G. A. Randomized exponential transformation algorithm for solving the stochastic problems of gamma-ray transport theory // Russian Journal of Numerical Analysis and Math. Modelling. 2020. V. 35. N 3. P. 153–162.
19. Medvedev Iia N. On the efficiency of using correlative randomized algorithms for solving problems of gamma radiation transfer in stochastic medium // Russian Journal of Numerical Analysis and Mathem. Modelling. 2022. V. 37. N 4. P. 231–240.

УДК 519.635

ИССЛЕДОВАНИЕ ПРИМЕНЕНИЯ ЯВНО-ИТЕРАЦИОННОЙ СХЕМЫ К МОДЕЛИРОВАНИЮ ДОЗВУКОВЫХ РЕАГИРУЮЩИХ ГАЗОВЫХ ПОТОКОВ¹⁾

© 2024 г. Е. Е. Пескова^{1, *}, О. С. Язовцева^{1, **}

¹430005 Саранск, ул. Большевистская, 68,

Национальный исследовательский Мордовский гос. ун-т, Республика Мордовия, Россия

*e-mail: e.peskova@math.mrsu.ru

**e-mail: yaos@math.mrsu.ru

Поступила в редакцию 28.06.2023 г.

Поступила в редакцию 02.09.2023 г.

Поступила в редакцию 20.10.2023 г.

Статья посвящена исследованию возможности применения явно-итерационной схемы ЛИ–М для расчета диссипативных членов к решению задач дозвуковых реагирующих потоков с радикально-цепными реакциями, активными диффузионными процессами, значительной теплопередачей и энергопоглощением. Моделирование подобных течений характеризуется ограничением на шаг интегрирования по времени, связанным, в первую очередь, с преобладанием диффузионных процессов над конвективными и наличием быстрых химических реакций. Математическая модель описана с использованием многокомпонентных уравнений Навье–Стокса. Совокупность разномасштабных процессов в модели обусловила использование принципа расщепления по физическим процессам: химическая кинетика проинтегрирована методом Радо с адаптивным шагом по времени; конвективный поток рассчитывается с использованием потока Русанова и WENO схемы; диссипативные потоки с помощью явно-итерационной схемы ЛИ–М. В результате разработаны численный алгоритм и код для исследования дозвуковых реагирующих течений в осесимметричной геометрии и проведен ряд вычислительных экспериментов. Для тестирования реализованного алгоритма было решено одномерное нестационарное неоднородное уравнение. Показано, что применение схемы ЛИ–М к расчету диссипативной части позволяет избавиться от диффузионного ограничения на шаг интегрирования по времени. Было проведено численное моделирование процесса высокотемпературной конверсии метана в одномерной постановке. Данный процесс характеризуется быстрыми химическими реакциями, значительными локальными изменениями температуры, плотности газа и теплофизических характеристик, что накладывает серьезные ограничения на шаг интегрирования по времени. Показано, что алгоритм позволяет проводить расчеты с шагом, превышающим диффузионные ограничения на шаг по времени. Проведено сравнение расчетов с расчетами по ранее верифицированному алгоритму, показано хорошее совпадение результатов со значительным выигрышем по времени выполнения программы. Проведено численное моделирование течения газа в цилиндрической трубе, полученные результаты верифицированы путем сеточной сходимости. Библ. 29. Фиг. 7.

Ключевые слова: математическое моделирование, уравнения Навье–Стокса, расщепление по физическим процессам, чебышёвская явно-итерационная схема, осесимметричные течения, конверсия метана.

DOI: 10.31857/S0044466924020126, EDN: YJAXQP

1. ВВЕДЕНИЕ

Задачи разработки эффективных и высокоточных вычислительных алгоритмов и программ для моделирования многокомпонентных реагирующих газовых течений широко известны и являются актуальными для современной науки и техники. Поскольку реагирующие газовые потоки характеризуются многомасштабностью протекающих физико-химических процессов, возникают жесткие ограничения на шаг интегрирования по времени. Эффективным средством решения задач, включающими в себя столь разномасштабные процессы, является использование устойчивых неявных разностных схем [Lawrence,

¹⁾ Работа выполнена при финансовой поддержке РФФ (код проекта № 23-21-00202), <https://rscf.ru/project/23-21-00202/>.

1989, Жалнин, 2022]. Однако нелинейность источниковых членов и теплофизических характеристик газовой смеси приводит к увеличению времени неявного интегрирования. Использование параллельных технологий и распределенных вычислений для явных разностных схем способствует их активному развитию в настоящее время. Основным их недостатком является сохранение устойчивости алгоритма для сложных систем только при крайне малом шаге интегрирования по времени, что затрудняет проведение вычислительных экспериментов. Активно развиваемым подходом к моделированию реагирующих потоков с помощью явных разностных схем, является расщепление многокомпонентной системы уравнений Навье–Стокса на несколько этапов по типам рассматриваемых процессов [Борисов, 2022, Жуков, 2020]. Данные работы ориентированы на исследование сверхзвуковых реагирующих течений. В настоящей статье предложена разработка эффективного алгоритма для дозвуковых газовых потоков, которые актуальны при моделировании процессов в химических реакторах [Снытников, 2021], в гемодинамике [Василевский, 2022], задач гидродинамики в водоемах [Сухинов, 2023].

Численное решение системы уравнений Навье–Стокса при малых числах Маха, дополненной уравнениями для компонент смеси, сопряжено с трудностями, связанными со значительно различающимися временами конвективного переноса и распространения акустических возмущений, с малыми изменениями давления с одновременным изменением в разы скорости потока, с быстрыми скоростями протекания химических процессов, с преобладанием диффузионных процессов над конвективным переносом. Данные трудности широко известны и решаются посредством расщепления уравнений по физическим процессам [Марчук, 1988, Day, 2000]. Решение системы уравнений химической кинетики, как правило, проводят с собственным адаптивным шагом в рамках общего газодинамического шага. Проблемы малого изменения давления в системе решаются с использованием идеи SIMPLE-подобных алгоритмов [Флетчер, 1991, Van Doormaal, 1984, Борисов, 2015], в которых расчет давления выделяется в отдельный шаг и решается уравнение эллиптического типа. Однако, в настоящее время не решена проблема преобладания скорости диссипативных процессов над конвективными, которая возникает по причине активных химических превращений. В силу изменения компонентного состава и температуры смеси возрастают коэффициенты теплопроводности и диффузии с уменьшением общего расчетного шага на порядки. Как следствие, расчет системы уравнений химической кинетики, которая может включать в себя десятки уравнений, и использование схем повышенного порядка аппроксимации при пространственной дискретизации приводят к неприемлемым временам расчета задачи в целом.

Перспективный подход к преодолению ограничения на шаг интегрирования по времени уравнений Навье–Стокса, связанного с диссипативными процессами, описан в цикле работ В. Т. Жукова [Жуков, 2010, Жуков, 2020]. В этих работах предлагается вести расчет с шагом интегрирования по времени, определенным из условия устойчивости явной схемы для гиперболической системы, но расчет диссипативных членов предлагается вести по схеме локальных итераций. Схема основана на явном итерационном процессе, параметрами в котором являются корни многочленов Чебышёва. В отличие от неявных схем она не предполагает дополнительного предобуславливания или расчета необходимого количества итераций, что существенно упрощает ее применение для многофакторных задач. Сохранение достаточно крупного шага по времени при устойчивом счете делает ее эффективнее большинства явных схем при легкости применения к ней параллельных или распределенных вычислений. Данный алгоритм построен для классических уравнений Навье–Стокса для сверхзвуковых течений и успешно прошел валидацию и верификацию [Жуков, 2020] на задачах с известными аналитическими и численными решениями.

Настоящая работа посвящена исследованию возможности применения явно-итерационной схемы ЛИ–М для расчета диссипативных членов [Жуков, 2010] с целью увеличения шага интегрирования по времени к решению задач дозвуковых реагирующих потоков с радикально-цепными реакциями, активными диффузионными процессами, значительной теплопередачей и энергопоглощением.

Вычислительный алгоритм построен на основе ранее разработанного алгоритма для исследования дозвуковых реагирующих потоков, в котором система уравнений Навье–Стокса расщеплялась на химическую кинетику, газодинамические уравнения (включающую конвективные и диссипативные члены) и уравнение для расчета давления [Жалнин, 2017, Gubaydullin, 2021]. Общий расчетный шаг в этом алгоритме определялся величиной диссипативных членов. Согласно алгоритму, предложенному в [Жуков, 2020], газодинамический блок в настоящей работе расщеплен на конвективный и диссипативный этапы, диссипативный этап рассчитывается по явно-итерационной схеме с чебышёвскими параметрами, что позволяет вести расчет с общим шагом по времени, превышающим диффузионные ограничения. Для решения задач кинетики применен метод Радо с адаптивным шагом по времени [Haiger, 1996]. Расчет конвективных членов осуществлен с использованием потоков Русанова [Русанов, 1961] с реконструкцией газодинамических параметров на границах ячеек WENO-схемой [Shu, 1997]. Расчет диссипативного этапа проводится с применением явно-итерационной чебышёвской схемы и включает в себя 3 шага:

диффузионный, вязкий и теплопроводный. Поскольку расчет конвективных потоков проводился без учета давления, на последнем этапе необходимо рассчитать поправку к давлению и скорректировать вектор скорости. Решение эллиптического уравнения для поправки проводится методом Якоби, он является достаточно экономичным, поскольку изменения давления в области незначительны и за начальное приближение принимается давление на предыдущем шаге по времени.

Для подтверждения заявленных свойств схемы ЛИ–М и тестирования построенного алгоритма было решено одномерное неоднородное нестационарное уравнение. Вычислительные эксперименты проводились на задаче конверсии метана в одномерной и осесимметричной геометриях [Лашина, 2023]. Поскольку процесс расщепления метана возможен только под воздействием высоких температур (около 1200°C) и характеризуется быстрыми химическими реакциями, значительным энергопоглощением и преобладанием диффузии, данная задача представляется хорошим кандидатом для валидации и верификации построенного алгоритма.

2. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

В настоящей работе для описания математической модели используется модификация уравнений Навье–Стокса в приближении малых чисел Маха, в основе которой лежит модель горения с малыми числами Маха [Rehm, 1978], скорректированная с использованием средств асимптотического анализа в [Majda, 1985]. Использование в модели уравнения для энтальпии смеси позволяет учитывать влияние химических реакций на энергию системы, поскольку энтальпия определяется с учетом энтальпии образования каждого вещества в ходе химической реакции. При достаточно малом числе Маха (в случае решаемых нами задач $M \approx 10^{-3}$) давление p может быть представлено в виде $p = p_0 + \pi$, где p_0 — термодинамическое давление, постоянное в области, π — динамическая составляющая давления, удовлетворяющая условию $\pi / p_0 \sim O(M^2)$. Характер существенно дозвуковых течений ($M \ll 1$) позволяет опустить учет диссипативных членов в уравнении энтальпии, вклад диссипативных сил оценивается величиной второго порядка от числа Маха [Majda, 1985]. Однако даже при таком малом числе Маха, среда не может быть признана несжимаемой, поскольку процесс характеризуется существенными локальными изменениями концентраций и термодинамических характеристик, что является следствием активной генерации радикалов.

Таким образом, система уравнений может быть записана в виде:

$$\frac{\partial \rho Y_m}{\partial t} = -\nabla \cdot (\rho Y_m \vec{v}) - \nabla \cdot \vec{J}_m + R_m, \quad (1)$$

$$\frac{\partial \rho \vec{v}}{\partial t} = -\nabla \cdot (\rho \vec{v} \vec{v}) - \nabla \pi + \nabla \cdot \bar{\tau}, \quad (2)$$

$$\frac{\partial \rho h}{\partial t} = -\nabla \cdot (\rho h \vec{v}) - \nabla \cdot \vec{q}, \quad (3)$$

Здесь $m = \overline{1, M}$, M — количество компонент в газовой смеси, ρ — плотность смеси, Y_m — массовая доля m -й компоненты смеси, \vec{v} — вектор скорости, h — энтальпия смеси, \vec{J}_m — вектор диффузионного потока, R_m — скорость образования или расхода m -й компоненты смеси, $\bar{\tau}$ — тензор вязких напряжений, \vec{q} — вектор потока тепла.

Для термодинамической составляющей давления записывается уравнение состояния газовой смеси:

$$p_0 = \rho R T \sum_m \frac{Y_m}{M_{wm}}, \quad (4)$$

где T — температура, M_{wm} — молекулярная масса m -й компоненты смеси, R — универсальная газовая постоянная.

Плотность смеси определяется из выражения:

$$\rho = \sum_m \rho Y_m. \quad (5)$$

Вектор диффузионного потока компонента и вектор потока тепла для смеси определяются с использованием модели средних по смеси значений:

$$\vec{J}_m = -\rho D_{m,mix} \nabla Y_m, \quad (6)$$

$$\vec{q} = -\lambda \nabla T - \sum_m h_m \rho D_{m,mix} \nabla Y_m. \quad (7)$$

Тензор вязких напряжений имеет вид:

$$\bar{\tau} = \mu (\nabla \vec{v} + (\nabla \vec{v})^T) - \frac{2}{3} \mu (\nabla \cdot \vec{v}) I, \quad (8)$$

где μ — коэффициент динамической вязкости смеси, I — единичная матрица.

Коэффициенты диффузии, вязкости, теплопроводности являются функциями, зависящими от температуры и концентраций компонент смеси [Stadnichenko, 2016].

Энтальпия реакционной смеси представляется как средневзвешенное энтальпий компонент

$$h(T, Y_m) = \sum_m Y_m h_m(T), \quad h_m = \int_{T_0}^T C_{pm}(T) dT + h_m^0. \quad (9)$$

Здесь h_m^0 — энтальпия образования m -й компоненты при стандартной температуре $T_0 = 298.15$ К, $C_{pm}(T)$ — удельная теплоемкость m -й компоненты при постоянном давлении.

Химические превращения в этой системе уравнений описываются кинетическими схемами химических реакций. Скорость образования или расхода компоненты смеси находится из выражения:

$$R_m = M_{wm} \sum_{n=1}^N \nu_{mn} w_n. \quad (10)$$

Здесь M_{wm} — молекулярная масса компоненты m , стехиометрический коэффициент компоненты m в стадии реакции n , w_n — скорость протекания стадии n , N — количество элементарных стадий.

Для замыкания систему уравнений (1)–(3) необходимо дополнить выражением для нахождения динамической составляющей давления π . Для ее определения по методу проекции [Флетчер, 1991] интегрирование уравнения (2) разделено на два этапа: предварительный расчет вектора скорости без учета динамической составляющей давления; коррекция вектора скорости на учет давления. Из последнего этапа можно получить уравнение Пуассона для динамической составляющей давления. Исследуемые течения в областях активных химических превращений характеризуются значительным изменением объема и, следовательно, плотности, что определяет ненулевое значение дивергенции вектора скорости. Для нахождения условия на дивергенцию вектора скорости используется соотношение для термодинамического давления (4), уравнение неразрывности в недивергентном виде и уравнение (3), записанное относительно температуры. Подробный вывод условия на дивергенцию вектора скорости для модели (1)–(3) представлен в работах [Борисов, 2015, Снытников, 2023]:

$$\begin{aligned} S \equiv \nabla \cdot \vec{v} = & \frac{1}{\rho C_p T} \left(\nabla \cdot \lambda \nabla T + \sum_m \rho D_{m,mix} \nabla Y_m \nabla h_m \right) + \\ & + \frac{1}{\rho} \sum_m \frac{M_w}{M_{wm}} (\nabla \cdot \rho D_{m,mix} \nabla Y_m) + \\ & + \frac{1}{\rho} \sum_m \left(\frac{M_w}{M_{wm}} - \frac{h_m}{C_p T} \right) R_m. \end{aligned} \quad (11)$$

Здесь $C_p(T)$ — теплоемкость смеси при постоянном давлении, $\lambda(T)$ — коэффициент теплопроводности смеси, $D_{m,mix}(T)$ — коэффициент диффузии m -й компоненты смеси, $h_m(T)$ — энтальпия m -й компоненты смеси, M_w — молекулярная масса смеси.

Таким образом, основными уравнениями модели являются (1)–(3), (11).

3. ВЫЧИСЛИТЕЛЬНЫЙ АЛГОРИТМ

Для исследования возможности использования явно-итерационной схемы с чебышёвскими параметрами [Жуков, 2020] для моделирования реагирующих многокомпонентных потоков с малыми числами Маха будем рассматривать систему уравнений (1)–(3) в цилиндрической системе координат для осесимметрического течения:

$$\frac{\partial U}{\partial t} + \frac{\partial(F^1(U) - H^1(U))}{\partial z} + \frac{1}{r} \frac{\partial(r(F^2(U) - H^2(U)))}{\partial r} = W(U). \quad (12)$$

Здесь U — результирующий вектор основных переменных среды, $F^k(U) = F^{k,\text{conv}}(U) + F^{k,p}(U)$ — вектор конвективных потоков, $H^k(U) = H^{k,D}(U) + H^{k,\mu}(U) + H^{k,\lambda}(U)$ — вектор диффузионных, вязких и тепловых потоков, $W(U)$ — вектор источниковых членов, отвечающий за химические реакции:

$$U = \begin{pmatrix} \rho Y_m \\ \rho u_z \\ \rho u_r \\ \rho h \end{pmatrix}, F^{1,\text{conv}}(U) = \begin{pmatrix} \rho u_z Y_m \\ \rho u_z^2 \\ \rho u_z u_r \\ \rho h u_z \end{pmatrix}, F^{1,p}(U) = \begin{pmatrix} 0 \\ \pi \\ 0 \\ 0 \end{pmatrix},$$

$$F^{2,\text{conv}}(U) = \begin{pmatrix} \rho u_r Y_m \\ \rho u_z u_r \\ \rho u_r^2 \\ \rho h u_r \end{pmatrix}, F^{2,p}(U) = \begin{pmatrix} 0 \\ 0 \\ \pi \\ 0 \end{pmatrix},$$

$$H^{1,D}(U) = \begin{pmatrix} J_{mz} \\ 0 \\ 0 \\ 0 \end{pmatrix}, H^{1,\mu}(U) = \begin{pmatrix} 0 \\ \tau_{zz} \\ \tau_{zr} \\ 0 \end{pmatrix}, H^{1,\lambda}(U) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ q_z \end{pmatrix},$$

$$H^{2,D}(U) = \begin{pmatrix} J_{mr} \\ 0 \\ 0 \\ 0 \end{pmatrix}, H^{2,\mu}(U) = \begin{pmatrix} 0 \\ \tau_{rz} \\ \tau_{rr} \\ 0 \end{pmatrix}, H^{2,\lambda}(U) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ q_r \end{pmatrix}, W = \begin{pmatrix} R_m \\ 0 \\ 0 \\ 0 \end{pmatrix};$$

$$J_{mz} = \rho D_{m,\text{mix}} \frac{\partial Y_m}{\partial z}, J_{mr} = \rho D_{m,\text{mix}} \frac{\partial Y_m}{\partial r},$$

$$q_z = \lambda \frac{\partial T}{\partial z} + \sum_{m=1}^M h_m \rho D_{m,\text{mix}} \frac{\partial Y_m}{\partial z}, q_r = \lambda \frac{\partial T}{\partial r} + \sum_{m=1}^M h_m \rho D_{m,\text{mix}} \frac{\partial Y_m}{\partial r},$$

$$\tau_{zz} = \mu \left(2 \frac{\partial u_z}{\partial z} - \frac{2}{3} \left(\frac{\partial u_z}{\partial z} + \frac{1}{r} \frac{\partial r u_r}{\partial r} \right) \right), \tau_{zr} = \mu \left(2 \frac{\partial u_r}{\partial r} - \frac{2}{3} \left(\frac{\partial u_z}{\partial z} + \frac{1}{r} \frac{\partial r u_r}{\partial r} \right) \right),$$

$$\tau_{zr} = \tau_{rz} = \mu \left(\frac{\partial u}{\partial r} + \frac{\partial v}{\partial z} \right),$$

$m = \overline{1, M}$, M — количество компонент в газовой смеси, u_z , u_r — проекции вектора скорости \vec{u} на оси z и r .

Численный алгоритм для описанной системы будет релизован в рамках метода конечных объемов. Использована равномерная по каждому направлению сетка прямоугольников $\Omega_h = \{z_i = ih_z, r_j = jh_r; i = \overline{1, N_z}, j = \overline{1, N_r}; N_z h_z = L_z, N_r h_r = L_r\}$, $L_{z(r)}$ — размеры расчетной области вдоль осей $\{z, r\}$. В каждой ячейке сетки заданы значения газодинамических параметров и концентраций компонент смеси. Систему уравнений (12) необходимо дополнить граничными условиями, для расчетов представленной модели задаются условия втекания и вытекания потока, условия прилипания с температурой стенок, на оси симметрии задаются условия симметрии.

Интегрированием по ячейкам сетки строится разностная схема:

$$\frac{U^{n+1} - U^n}{\Delta t} + L_h \left(F^{\text{conv}}(U) + F^p(U) - H^D(U) - H^\mu(U) - H^\lambda(U) \right) = W(U). \quad (13)$$

Здесь $L_h F^{\text{conv}}(U)$, $L_h F^p(U)$, $L_h H^D(U)$, $L_h H^\mu(U)$, $L_h H^\lambda(U)$ — сеточные аналоги конвективных, диффузионных, вязких и тепловых потоков, $W(U)$ — дискретный аналог скорости образования или расхода каждой компоненты смеси.

Переход с шага по времени t^n на шаг $t^{n+1} = t^n + \Delta t$ по схеме (13) строится на основе расщепления на решение уравнений химической кинетики, на гиперболическую задачу без учета давления, параболическую задачу и эллиптическую задачу для расчета давления. В свою очередь параболическая задача делится на три этапа: расчет диффузионных, вязких и теплопроводных членов.

Опишем каждый этап построения схемы. Пусть на шаге t^n имеем значения вектора $U^n = (\rho Y_m, \rho u_z, \rho u_r, \rho h)$, $m = \overline{1, M}$. Температуру смеси T находим с использованием метода Ньютона из равенства (9).

Первым этапом интегрирования является решение уравнений химической кинетики:

$$\frac{U^{\text{chem}} - U^n}{\Delta t} = W(U^n). \quad (14)$$

Для решения использован подключаемый модуль RADAU5 [Hairer, 1996], реализующий трехстадийный метод Рунге-Кутты пятого порядка точности с адаптивным шагом по времени. Результатом расчетов является обновленный вектор $U^{\text{chem}} = ((\rho Y_m)^{\text{chem}}, \rho u_z, \rho u_r, \rho h)^T$, $m = \overline{1, M}$, в котором последние три строки остаются неизменными после этого этапа.

На втором этапе проведен расчет конвективных потоков без учета давления.

Данный этап реализован явно с нахождением значений потоков на границе ячейки с использованием потоков Русанова [Русанов, 1961].

$$\frac{U^{\text{conv}} - U^{\text{chem}}}{\Delta t} + L_h F^{\text{conv}}(U^n) = 0. \quad (15)$$

В обновленный вектор $U^{\text{conv}} = ((\rho Y_m)^{\text{conv}}, (\rho u_z)^{\text{conv}}, (\rho u_r)^{\text{conv}}, (\rho h)^{\text{conv}})^T$, $m = \overline{1, M}$, записывается результат расчета.

Здесь

$$L_h F^{\text{conv}}(U^n) = \frac{F_{i+1/2j}^{1,\text{conv}} - F_{i-1/2j}^{1,\text{conv}}}{h_z} + \frac{r_{ij+1/2} \cdot F_{ij+1/2}^{2,\text{conv}} - r_{ij-1/2} \cdot F_{ij-1/2}^{2,\text{conv}}}{r \cdot h_r},$$

$$F_{i+1/2j}^{1,\text{conv}} = 0.5 \left(F_{i+1/2j}^{1,\text{conv}}(U_{i+1/2j}^+) + F_{i+1/2j}^{1,\text{conv}}(U_{i+1/2j}^-) - \alpha_1 (U_{i+1/2j}^+ - U_{i+1/2j}^-) \right),$$

$$F_{ij+1/2}^{2,\text{conv}} = 0.5 \left(F_{ij+1/2}^{2,\text{conv}}(U_{ij+1/2}^+) + F_{ij+1/2}^{2,\text{conv}}(U_{ij+1/2}^-) - \alpha_2 (U_{ij+1/2}^+ - U_{ij+1/2}^-) \right),$$

$$\alpha_1 = \max(|u_{z,i+1/2j}^+|, |u_{z,i+1/2j}^-|),$$

$$\alpha_2 = \max(|u_{r,ij+1/2}^+|, |u_{r,ij+1/2}^-|),$$

где $U_{i+1/2j}^+$, $U_{i+1/2j}^-$ — значения вектора переменных U^{chem} слева и справа от границы между i и $i+1$ ячейками, рассчитываются по WENO схеме 5-го порядка [Shu, 1997].

Третий этап алгоритма — диссипативный.

Он состоит из трех явно-итерационных процедур, примененных последовательно к уравнениям законов сохранения массы, импульса и энтальпии без учета конвективных членов. Каждую из подзадач можно записать в виде:

$$\frac{\partial H}{\partial t} - D_h V = f_h. \quad (16)$$

Для случая решения системы уравнений сохранения массы

$$\begin{aligned} H &= (\rho Y_m)^\top, \quad V = (Y_m)^\top, \\ D_h V &= \left(\nabla \cdot (\rho D_{m,mix} \nabla Y_m) \right)^\top, \\ f_h &= 0, \quad m = \overline{1, M}; \end{aligned}$$

уравнения сохранения импульса:

$$H = (\rho u_z, \rho u_r)^\top, \quad V = (u_z, u_r)^\top, \quad D_h V = \nabla \cdot (\bar{\tau}), \quad f_h = 0;$$

уравнения сохранения энергии:

$$\begin{aligned} H &= \rho h, \quad V = T, \quad D_h V = \nabla \cdot (\lambda \nabla T), \\ f_h &= \nabla \cdot \left(\sum_{m=1}^M h_m \rho D_{m,mix} \nabla Y_m \right). \end{aligned}$$

Здесь ∇ — оператор дивергенции в цилиндрической системе координат.

Для каждой из подзадач применена следующая итерационная процедура ($l = \overline{1, 2p-1}$):

$$H^l = \frac{H^{\text{conv}} + \Delta t b_l H^{l-1} + \Delta t (D_h V^{l-1} + f_h)}{1 + \Delta t b_l}. \quad (17)$$

Алгоритм нахождения итерационных параметров b_l приведен в следующем разделе.

Применив процедуру (17) к уравнениям сохранения массы, где $H^{\text{conv}} = ((\rho Y_m)^{\text{conv}})^\top$, $H^0 = ((\rho Y_m)^n)^\top$, $m = \overline{1, M}$, — векторы, составленные из концентраций компонент газовой смеси, получим вектор

$$H^l = H^{n+1} = ((\rho Y_m)^{n+1})^\top. \quad (18)$$

Из H^{n+1} находим переменные $\rho^{n+1} = \sum_m (\rho Y_m)^{n+1}$, $Y_m^{n+1} = (\rho Y_m)^{n+1} / \rho^{n+1}$, $m = \overline{1, M}$.

После применения процедуры (17) к уравнениям сохранения импульса с $H^{\text{conv}} = ((\rho u_z)^{\text{conv}}, (\rho u_r)^{\text{conv}})^\top$, $H^0 = ((\rho u_z)^n, (\rho u_r)^n)^\top$ получим

$$H^l = H^{n+1} = \left((\rho u_z^*)^{n+1}, (\rho u_r^*)^{n+1} \right)^\top, \quad (19)$$

откуда с помощью H^{n+1} находим $u_z^* = \frac{(\rho u_z^*)^{n+1}}{\rho^{n+1}}$, $u_r^* = \frac{(\rho u_r^*)^{n+1}}{\rho^{n+1}}$. Это значение скорости предварительное, поскольку в расчетах пока не учтено влияние давления.

Применим процедуру (17) к уравнениям сохранения энергии в энтальпийной форме при $H^{\text{conv}} = (\rho h)^{\text{conv}}$, $H^0 = (\rho h)^n$, $f_h = \nabla \cdot \left(\sum_{m=1}^M h_m \rho D_{m,mix} \nabla Y_m \right)$. Значения переменной f_h рассчитаны из значений на временном слое n . В результате получим

$$H^l = H^{n+1} = (\rho h)^{n+1}. \quad (20)$$

С помощью H^{n+1} найдем $h^{n+1} = \frac{(\rho h)^{n+1}}{\rho^{n+1}}$.

Четвертый этап заключается в коррекции вектора скорости и расчете динамической составляющей давления.

Особенностью течений с малыми числами Маха является крайне малое изменение давления при одновременном увеличении скорости в разы за счет химических реакций и диссипативных процессов, поэтому для учета влияния давления на скорость система уравнений (12) решена без учета потока $F^p(U)$. Скорость должна быть скорректирована после вычисления давления согласно выражению, полученному из уравнения сохранения импульса, преобразованного с учетом уравнения неразрывности:

$$\frac{\vec{u}^{n+1} - \vec{u}^*}{\Delta t} = -\frac{1}{\rho^n} \nabla \pi^{n+1}. \quad (21)$$

Возьмем дивергенцию от обеих частей выражения (21) и, согласно условию на дивергенцию вектора скорости (11), получим уравнение Пуассона для определения динамической составляющей давления π^{n+1} :

$$\nabla \cdot \frac{1}{\rho^n} \nabla \pi^{n+1} = \frac{1}{\Delta t} (\nabla \cdot \vec{u}^* - S^{n+1}). \quad (22)$$

В этом уравнении S^{n+1} рассчитано с использованием параметров на шаге $n + 1$, найденных на предыдущих этапах схемы. Полученное уравнение решается итерационным методом Якоби. Поскольку изменения давления в области незначительны и за начальное приближение принимается давление на предыдущем шаге по времени, достаточно одной-двух итераций данного метода. После нахождения динамической составляющей давления π^{n+1} проводим коррекцию вектора скорости с использованием выражения (21), в результате получаем u_z^{n+1}, u_r^{n+1} .

4. АЛГОРИТМ НАХОЖДЕНИЯ ИТЕРАЦИОННЫХ ПАРАМЕТРОВ

Алгоритм схемы локальных итераций ЛИ–М подробно описан в работах [Жуков, 2010, Жуков, 2020]. В данном разделе приведены расчетные формулы для степени чебышевского многочлена и итерационных параметров.

Схема ЛИ–М является двухслойной по времени, и алгоритм перехода от слоя t^n к слою t^{n+1} — явно-итерационный. Число итераций l определяется из условия устойчивости и рассчитывается для каждого слоя n . В данной работе мы используем схему ЛИ–М точности $O(\Delta t)$.

Степень чебышёвского многочлена p задается формулой:

$$p = \langle 0.25\pi\sqrt{\Delta t\lambda_{\max}} + 1 \rangle, \quad (23)$$

где $\langle s \rangle$ — наименьшее целое, большее или равное s , $\lambda_{\max} = \lambda_{\max}^D + \lambda_{\max}^u + \lambda_{\max}^\lambda$ — сумма значений верхних границ спектров дискретных операторов, отвечающих аппроксимациям процессов диффузионного переноса за счет диффузии, вязкости и теплопроводности. В соответствии с теоремой Гершгорина о кругах [29] величины $\lambda_{\max}^D, \lambda_{\max}^u, \lambda_{\max}^\lambda$ можно рассчитать, вычисляя для каждого узла сетки сумму модулей коэффициентов разностных аппроксимаций диссипативных членов.

Рассчитываем множество нулей многочлена Чебышёва и упорядочиваем их для устойчивости [26]:

$$K_p = \left\{ \cos \frac{2i-1}{2p} \pi, i = \overline{1, p} \right\}. \quad (24)$$

По упорядоченному множеству $\beta_m \in K_p$, $m = \overline{1, p}$, рассчитываем параметры

$$\alpha_m = \frac{\lambda_{\max}}{1 + z_1} (z_1 - \beta_m), \quad m = \overline{1, p}, \quad (25)$$

где корню $\beta_1 = z_1 = \cos \frac{0.5\pi}{p}$ соответствует параметр $a_1 = 0$.

Окончательно набор итерационных параметров состоит из $l = 2p - 1$ чисел $\{b_1, \dots, b_l\} = \{a_p, \dots, a_2, a_p, \dots, a_2, a_1\}$.

5. ЧИСЛЕННЫЕ ЭКСПЕРИМЕНТЫ

Настоящий раздел посвящен вычислительным экспериментам для одномерного неоднородного нестационарного уравнения, имеющего аналитическое решение, с целью сопоставления полученных результатов расчетов. Также проиллюстрированы заявленные свойства вычислительного алгоритма: возможность проведения расчетов с шагом, превышающим диффузионные ограничения на шаг по времени. Во второй части раздела представлены расчеты пространственно одномерной задачи дозвукового течения реагирующей газовой смеси с радикальной кинетикой углеводородов. Проведен сравнительный анализ описанного в настоящей статье алгоритма на основе схемы ЛИ–М и разработанного ранее вычислительного алгоритма, верифицированного на экспериментальных данных, численных и аналитических решениях. В третьей части раздела представлены расчеты практической задачи высокотемпературной конверсии метана в осесимметричной геометрии течения, проведена демонстрация сеточной сходимости построенного алгоритма.

5.1. Решение одномерного неоднородного уравнения

Для тестирования алгоритма рассмотрим одномерное неоднородное нестационарное уравнение [Зайцев, 1996, Снытников, 2001]

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = \frac{\partial^2 u}{\partial x^2} + f(u), \quad (26)$$

где $f(u) = -\theta + \gamma e^{-2u}$, имеющее аналитическое решение

$$u(x, t) = \ln \left(\pm \sqrt{\frac{\gamma}{\theta}} + c \cdot \exp(\pm \sqrt{\theta} x - t(\theta \pm \sqrt{\theta})) \right). \quad (27)$$

Выбор данного уравнения обусловлен наличием конвективного и диссипативного переноса, и неоднородной части, схожей с уравнениями химической кинетики. В расчетах принимаем $\theta = 1$, $\gamma = 1$, $x \in [0, 2]$, $t \in [0, 2]$.

Нами проведены расчеты с использованием алгоритма из п. 3 с расщеплением на нелинейную правую часть (выражение (14)), расчет конвективной части (выражение (15)) и расчет диссипативной части (выражение (16)) с отношением $\frac{\Delta t}{h} = 0.1$. Также проведен расчет по схеме без расщепления на конвективную и диссипативную составляющую, последняя рассчитывалась по схеме с центральными разностями. Шаг по времени в данном случае ограничен условием $\frac{\Delta t}{h^2} \leq 0.5$, в расчетах принималось $\frac{\Delta t}{h^2} = 0.2$.

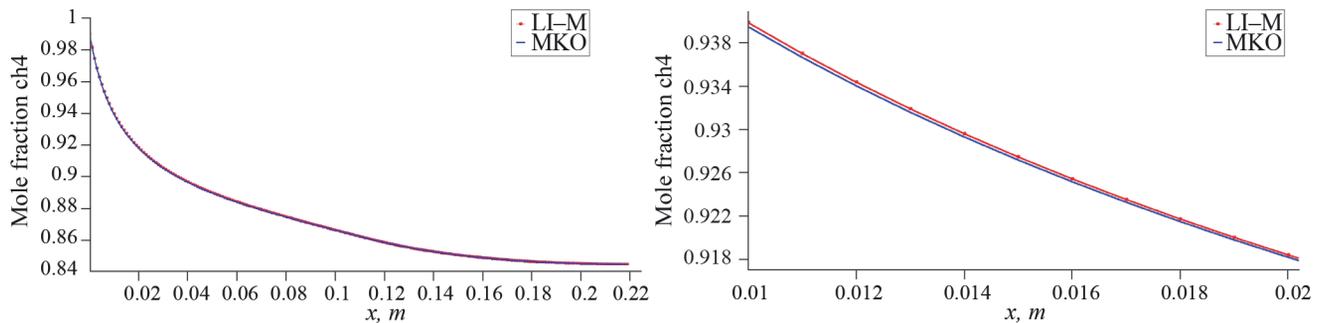
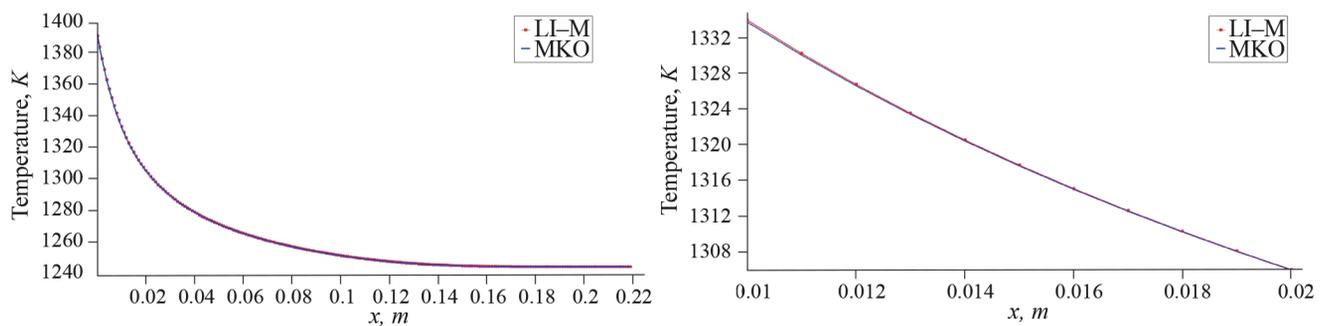
Для сетки размерности $N = 100$, $h = 0.02$ схема ЛИ–М делает 1000 шагов по времени, но на каждом шаге происходит пересчет диссипативного потока $l = 2p - 1$ раз, степень чебышёвского многочлена $p = 4$. На этой же сетке и временном интервале явная схема делает 25000 шагов по времени. Замеряемое расчетное время данной одномерной задачи показало, что алгоритм, основанный на схеме ЛИ–М требует в 20 раз меньше расчетного времени. Максимальное отклонение численного решения по схеме ЛИ–М от точного при $h = 0.02$ составило $\Delta_{\max} = 7.2 \cdot 10^{-4}$. Мы не приводим в этом разделе сопоставления графиков точного решения и расчетов по явной схеме и схеме ЛИ–М, а также на последовательности сгущающихся сеток, поскольку все решения дают результаты с очень малым различием.

5.2. Решение одномерных дозвуковых уравнений Навье–Стокса с радикальной кинетикой углеводородов

С использованием построенного вычислительного алгоритма решены уравнения Навье–Стокса в приближении малых чисел Маха с радикальной кинетикой углеводородов в пространственно-одномерной расчетной области. Очевидно, что вычислительный алгоритм для пространственно одномерной системы естественным образом вытекает из представленной осесимметричной схемы (разд. 3) обнулением члена, отвечающего производной по координате, сонаправленной с цилиндрическим радиусом. Проведено сравнение новых результатов с результатами ранее разработанного кода для исследования дозвуковых реагирующих газовых течений, конвективные и диффузионные потоки в котором рассчитывались без расщепления по явной схеме [Жалнин, 2017]. Алгоритм без расщепления был протестирован в двумерной и осесимметричной постановке на экспериментальных данных и аналитических решениях [Жалнин, 2017, Gubaydullin, 2021], что говорит о правомерности сравнения с ним.

Таблица 1. Сравнение эффективности алгоритмов

	Алгоритм на основе ЛИ–М	Алгоритм на основе МКО [Жалнин, 2017]
Шаг по времени Δt , сек	$0.5 \cdot 10^{-3}$	10^{-4}
Число шагов	2000	10000
Расчетное время, сек	350	1668

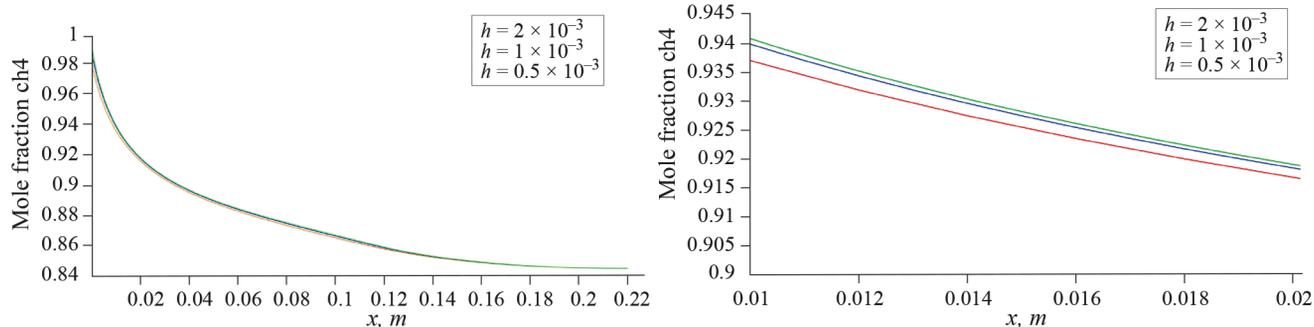
**Фиг. 1.** Распределение метана по длине реактора: красный пунктир – схема ЛИ–М, синяя линия – схема [Жалнин, 2017].**Фиг. 2.** Распределение температуры по длине реактора: красный пунктир – схема ЛИ–М, синяя линия – схема [Жалнин, 2017].

Вычислительные эксперименты проводились в области $\Omega = [0 \text{ м}, 0.22 \text{ м}]$. В начальный момент времени область заполнена метаном (CH_4), скорость 0.1 м/сек, давление 101325 Па. На входе слева подается газ, аналогичный исходной смеси. Справа задано условие вытекания, давление 101325 Па. Варьировались температуры газа на входе и в области. Расчеты проводились до 1 секунды.

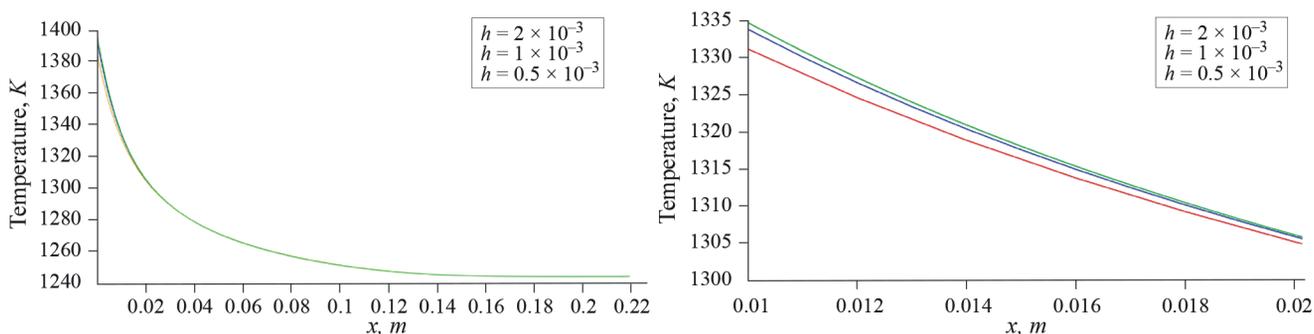
Для описания схемы химических превращений принята радикальная кинетическая схема [Лашина, 2023], включающая $M = 15$ компонент смеси.

В расчетах принимаем равномерную сетку отрезков с шагом $h = 10^{-3}$ м. Шаг интегрирования по времени задавался следующим образом. Для конвективных процессов шаг по времени определяется из условия Куранта–Фридрихса–Леви $\frac{a\tau_{\text{conv}}}{h} \leq 1$. Максимально допустимый шаг для исследуемой задачи в случае отсутствия диссипативных членов определен как $\tau_{\text{conv}} \approx 0.5 \cdot 10^{-2}$ сек, поскольку скорость распространения возмущений приближенно равна $a \approx 0.15$ м/сек. Как известно, на практике значение шага интегрирования зачастую принимается ниже расчетного. Возьмем шаг интегрирования равный

$\Delta t = 0.5 \cdot 10^{-3}$ сек. Диффузионный шаг по времени определяется как $\Delta t_{\text{dif}} = \frac{1}{\lambda_{\text{max}}}$ посредством вычисления верхних границ спектров дискретных операторов, отвечающих аппроксимациям процессов



Фиг. 3. Распределение метана: расчет на сгущающихся сетках.



Фиг. 4. Распределение температуры: расчет на сгущающихся сетках.

диффузионного переноса, вязкости и теплопроводности. В рассматриваемом в настоящей работе диапазоне температур $\Delta t_{\text{dif}} \approx 10^{-4}$ сек.

Ниже представлено сравнение быстродействия алгоритмов на основе ЛИ-М (3–5 итераций для диссипативных членов внутри общего расчетного шага) и метода конечных объемов [Жалнин, 2017] при начальной температуре реакционной смеси 1400 К.

Из табл. 1 следует, что расчет ускорен более, чем в четыре раза, несмотря на высокую трудоемкость решения уравнений химической кинетики.

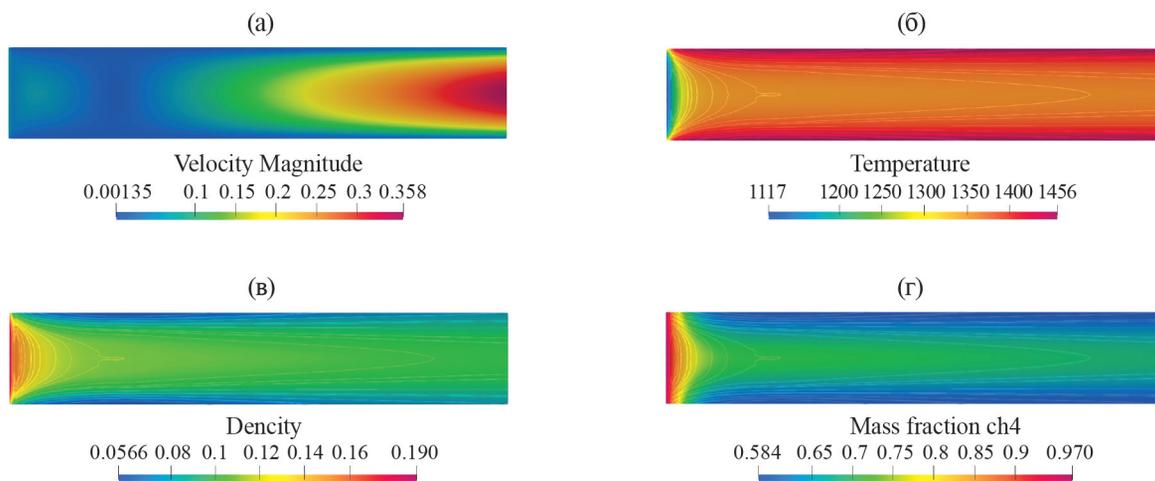
На фиг. 1, 2 представлены профили мольной доли метана и температуры газовой смеси по реактору для двух расчетных схем, справа — фрагменты графиков в области резких изменений газодинамических характеристик и концентраций веществ ($[0.01\text{ м}, 0.02\text{ м}]$). Из графиков очевидна согласованность результатов расчетов различными методами.

Исследование сходимости построенного алгоритма было проведено на последовательности сгущающихся сеток ($h = 0.5 \cdot 10^{-3}$ м, $h = 10^{-3}$ м, $h = 2 \cdot 10^{-3}$ м) с сохранением отношения $\frac{\Delta t}{h} = 0.5$. Результаты

расчетов представлены на фиг. 3, 4. Во всей расчетной области Ω решения практически неразличимы (левый график), поэтому справа выделены фрагменты графиков, где отличия заметны — в области падения температуры и мольной доли метана.

5.3. Моделирование конверсии метана в осесимметричной геометрии

Вычислительные эксперименты проводились в реакторе цилиндрической формы длиной 220 мм, шириной 40 мм. В начальный момент времени область заполнена метаном (CH_4), температура 1073 К, давление 101325 Па. На входе слева по всему диаметру трубы подается метан температурой 1073 К с расходом 40 л/ч. Справа по всему диаметру трубы задано условие вытекания, давление 101325 Па. Температура стенок 1473 К. Данные условия приводят к значительным перепадам температуры, плотности и концентрации в расчетной области в начальные моменты времени. Для описания схемы химических превращений принята схема [Лашина, 2023] с 15 компонентами смеси, как в разделе выше.

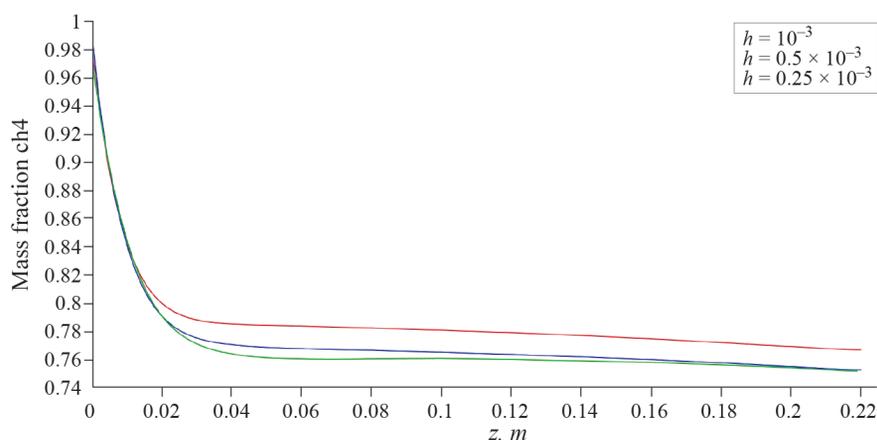


Фиг. 5. Распределение газодинамических характеристик и метана по трубе: (а) скорость; (б) температура; (в) плотность; (г) массовая доля метана.

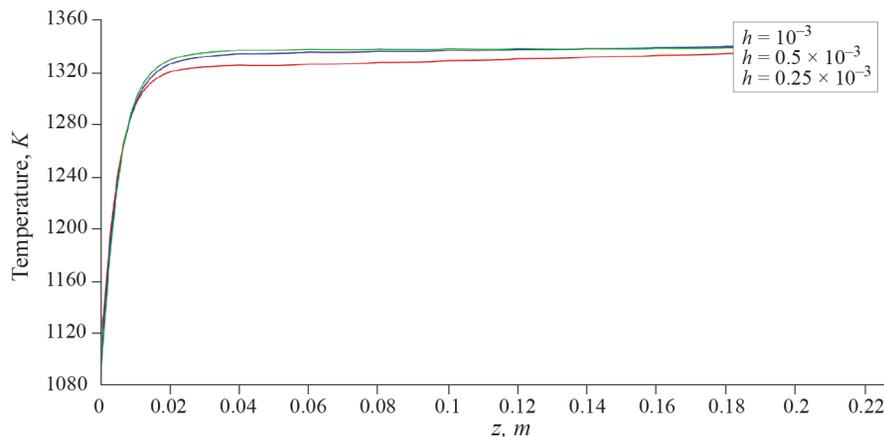
В расчетах принимаем равномерную сетку прямоугольников с шагом $h_z = h_r = 10^{-3}$ м. Поскольку скорость распространения возмущений для данной задачи может превышать 1 м/с, шаг интегрирования по времени возьмем $\Delta t = 10^{-4}$ сек. Отметим, что расчетный диффузионный шаг для данной задачи $\Delta t_{\text{dif}} \approx 10^{-5}$.

На фиг. 5 представлены результаты вычислительного эксперимента в момент времени 0.5 сек. Из графиков видно, что газодинамические параметры внутри области отличаются на порядки, что оказывает влияние на значения коэффициентов диффузии. Скорость смеси (фиг. 5а) теряет свое начальное распределение во второй части трубы, в которой происходит значительное ускорение потока. Около стенок формируется зона активных химических превращений с уменьшением доли метана (фиг. 5г) за счет нагрева стенок реактора (фиг. 5б). В центральной части реактора температура газовой смеси ниже, поскольку здесь происходят эндотермические реакции с поглощением энергии. Поскольку давление в реакторе равно атмосферному, нагрев приводит к разрежению газа с минимальным значением плотности у стенок (фиг. 5в). Максимальные значения плотности наблюдаются у входа газовой смеси, поскольку поступает относительно холодный метан.

Было проведено исследование сходимости на последовательности сгущающихся сеток (220×20 , 440×40 , 880×80) с сохранением отношения $\frac{\Delta t}{h} = 0.1$. Результаты расчетов вдоль оси реактора представлены на фиг. 6, 7. При дальнейшем увеличении количества расчетных ячеек графики неразличимы с сеткой размерности 880×80 . Можно сделать вывод о сеточной сходимости алгоритма для случая осесимметричного течения.



Фиг. 6. Распределение метана: расчет на сгущающихся сетках.



Фиг. 7. Распределение температуры: расчет на сгущающихся сетках.

6. ЗАКЛЮЧЕНИЕ

В статье исследована возможность применения явно-итерационной схемы ЛИ–М для расчета диссипативных членов при исследовании дозвуковых реагирующих газовых потоков с резкими локальными изменениями газодинамических и термодинамических параметров за счет химических реакций и, как следствие, преобладанием диффузионных процессов над конвективными. Для достижения поставленной цели разработаны вычислительный алгоритм и код, проведен ряд вычислительных экспериментов.

Выполнены расчеты одномерного неоднородного нестационарного уравнения, получено хорошее соответствие расчетных и аналитических решений, а также показана возможность избежать диффузионного ограничения на шаг по времени. Алгоритм применен для решения одномерных дозвуковых уравнений Навье–Стокса с радикальной кинетикой углеводородов, получено хорошее согласование с ранее разработанным верифицированным алгоритмом. Проведено моделирование конверсии метана в осесимметричной геометрии течения, получены картины распределения газодинамических характеристик. Расчет на последовательности сгущающихся сеток показал сходимость построенного алгоритма. Сделаны выводы о возможности проведения расчетов системы уравнений Навье–Стокса с дозвуковым характером течения с шагом по времени большим, чем для явного счета процессов диффузии, что значительно сокращает расчетное время даже в присутствии жестких систем уравнений химической кинетики и схем повышенного порядка аппроксимации.

Авторы выражают благодарность В. Н. Снытникову за предоставление экспериментального материала и продуктивное обсуждение результатов.

СПИСОК ЛИТЕРАТУРЫ

1. Lawrence S. L., Tannehill J. C., Chaussee D. S. Upwind Algorithm for the Parabolized Navier–Stokes Equations // AIAA Journal. 1989. V. 27. N 9. P. 1175–1183.
2. Жалнин Р. В., Масыгин В. Ф., Тишкин В. Ф. Решение двумерных задач газовой динамики с использованием неявной схемы для метода Галеркина с разрывными базисными функциями на неструктурированных треугольных сетках // Сиб. журнал вычисл. матем.. 2022. Т. 25. № 1. С. 19–32.
3. Борисов В. Е., Критский Б. В., Рыков Ю. Г. Программный модуль MCFL–Chem для расчета высокоскоростных течений смеси реагирующих газов // Препринты ИПМ им. М. В. Келдыша. 2022. № 21. 40 с.
4. Жуков В. Т., Феодоритова О. Б., Новикова Н. Д., Дубень А. П. Явно-итерационная схема для интегрирования по времени системы уравнений Навье–Стокса // Матем. моделирование. 2020. Т. 32. № 4. С. 57–74.
5. Снытников В. Н., Снытников Вл. Н., Масюк Н. С., Маркелова Т. В., Пармон В. Н. Приборы и техника эксперимента // Combustion Theory and Modelling. 2021. № 3. С. 129–137.
6. Василевский Ю. В., Симаков С. С., Гамилов Т. М., Саламатова В. Ю., Добросердова Т. К., Копытов Г. В., Богданов О. Н., Данилов А. А., Дергачев М. А., Добровольский Д. Д., Косухин О. Н., Ларина Е. В., Мелешкина А. В., Мычка Е. Ю., Харин В. Ю., Чеснокова К. В., Шипилов А. А. Персонализация математических моделей в кардиологии: трудности и перспективы // Компьютерные исследования и моделирование. 2022. Т. 14. № 4. С. 911–930.

7. *Сухинов А. И., Чистяков А. Е., Никитина А. В., Атаян А. М., Литвинов В. Н.* Метод решения сеточных уравнений для задач гидродинамики в плоских областях // Матем. моделирование. 2023. Т. 35. № 3. С. 35–58.
8. *Марчук Г. И.* Методы расщепления. М.: Наука, 1988.
9. *Day M. S., Bell J. B.* Numerical simulation of laminar reacting flows with complex chemistry // Combustion Theory and Modelling. 2000. V. 4. N 4. P. 535–556.
10. *Флетчер К.* Вычислительные методы в динамике жидкости. М.: Мир, 1991.
11. *Van Doormaal J. R., Raithby G. D.* Enhancement of the SIMPLE method for predicting incompressible fluid flow // Heat Transfer. 1984. V. 7. P. 147–163.
12. *Борисов В. Е., Кулешов А. А., Савенков Е. Б., Якуш С. Е.* Программный комплекс TCS3D: вычислительная модель // Препринты ИПМ им. М. В. Келдыша. 2015. № 110. 20 с.
13. *Жуков В. Т.* О явных методах численного интегрирования для параболических уравнений // Математическое моделирование. 2010. Т. 22. № 10. С. 127–158.
14. *Жуков В. Т., Новикова Н. Д., Феодоритова О. Б.* Об одном подходе к интегрированию по времени системы уравнений Навье–Стокса // Ж. вычисл. матем. и матем. физ. 2020. Т. 60. № 2. С. 267–280.
15. *Жалнин Р. В., Пескова Е. Е., Стадниченко О. А., Тишкин В. Ф.* Моделирование течения многокомпонентного реагирующего газа с использованием алгоритмов высокого порядка точности // Вестн. Удмуртского университета. Математика. Механика. Компьютерные науки. 2017. Т. 27. № 1. С. 608–617.
16. *Gubaydullin I. M., Zhalnin R. V., Masyagin V. F., Peskova E. E., Tishkin V. F.* Simulation of propane pyrolysis in a flow-through chemical reactor under constant external heating // Mathematical Models and Computer Simulations. 2021. V. 13. N 3. P. 437–444.
17. *Hairer E., Wanner G.* Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems. Springer-Verlag, Berlin, 1996.
18. *Русанов В. В.* Расчет взаимодействия нестационарных ударных волн с препятствиями // Ж. вычисл. матем. и матем. физ. 1961. Т. 1. № 2. С. 267–279.
19. *Shu C. W.* Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws // ICASE Report. 1997. № 97–65. P. 79.
20. *Лашина Е. А., Пескова Е. Е., Снытников В. Н.* Математическое моделирование нестационарной температурной конверсии метан-этановых смесей в широком диапазоне температур // Химия в интересах устойчивого развития. 2023. Т. 31. № 3. С. 288–296.
21. *Rehm R. G., Baum H. R.* The equation of motion for thermally driven, buoyant flows // J. Research NBS. 1978. V. 83. № 3. P. 297–308.
22. *Majda A., Sethian J.* The derivation and numerical solution of the equations for zero Mach number combustion. Combustion Science and Technology. 1986. V. 42. P. 185–205.
23. *Stadnichenko O. A., Snytnikov V. N., Snytnikov V. N., Masyuk N. S.* Mathematical modeling of ethane pyrolysis in a flow reactor with allowance for laser radiation effects // Chemical Engineering Research and Design. 2016. V. 109. P. 405–413.
24. *Борисов В. Е., Кулешов А. А., Савенков Е. Б., Якуш С. Е.* Программный комплекс TCS3D: математическая модель // Препринты ИПМ им. М. В. Келдыша. 2015. № 6. 20 с.
25. *Снытников В. Н., Пескова Е. Е., Стояновская О. П.* Модель двухтемпературной среды газ-твердые наночастицы с лазерным пиролизом метана // Математическое моделирование. 2023. Т. 35. № 4. С. 24–50.
26. *Лебедев В. И., Финогенов С. А.* О порядке выбора итерационных параметров в чебышёвском циклическом методе // Ж. вычисл. матем. и матем. физ. 1971. Т. 11. № 2. С. 425–438.
27. *Зайцев В. Ф., Полянин А. Д.* Дифференциальные уравнения с частными производными: Справочник. М.: Международный. программа, 1996.
28. *Снытников В. Н., Юрченко Е. М.* Схема расщепления для задач фильтрации газа с химическими реакциями // Вычисл. технологии. 2001. Т. 6. № 5. С. 95–105.
29. *Гантмахер Ф. Р.* Теория матриц. М.: Наука, 1966.

УДК 519.642.2, 532.59

*Работа посвящена памяти академика
Владимира Евгеньевича Захарова
(1939–2023 гг.), внесшего
огромный вклад в развитие теории нелинейного
взаимодействия волн в системах гидродинамического типа.*

АЛГОРИТМ РЕШЕНИЯ ЧЕТЫРЕХВОЛНОВОГО КИНЕТИЧЕСКОГО УРАВНЕНИЯ В ЗАДАЧАХ ВОЛНОВОЙ ТУРБУЛЕНТНОСТИ¹⁾

© 2024 г. Б. В. Семисалов^{1,2,3,*}, С. Б. Медведев^{2,3},
С. В. Назаренко⁴, М. П. Федорук^{3,2}

¹630090 Новосибирск, пр-т Акад. Коптюга, 4,

Институт математики им. С. Л. Соболева СО РАН, Россия

²630090 Новосибирск, пр-т Акад. Лаврентьева, 6,

Федеральный исследовательский центр информационных и вычислительных технологий,

³630090 Новосибирск, ул. Пирогова, 1, Новосибирский государственный университет,

⁴Université Côte d'Azur, CNRS, Institut de Physique de Nice (INPHYNI),

17 rue Julien Lauprêtre 06200 Nice, France

*e-mail: vibis87@gmail.com

Поступила в редакцию 29.06.2023 г.

Переработанный вариант 15.09.2023 г.

Принята к публикации 20.10.2023 г.

Предложен метод численного решения четырехволновых кинетических уравнений, возникающих в задачах волновой (слабой) турбулентности при описании однородного изотропного взаимодействия волн. Для расчета интеграла столкновений разработаны быстросходящиеся кубатурные формулы, позволяющие адаптировать алгоритм к особенностям решений и ядер интегралов. Проведены эксперименты на сходимость в задачах интегрирования из реальных приложений. Для учета существенной разномасштабности задач турбулентности в алгоритме реализованы и протестированы дробно-рациональные приближения решений и новая схема итераций по времени. Эффективность разработанного алгоритма продемонстрирована при моделировании обратного каскада частиц бозе-газа при формировании конденсата Бозе–Эйнштейна. Библ. 51. Фиг. 10. Табл. 1.

Ключевые слова: волновая турбулентность, кинетическое уравнение, нелинейное взаимодействие волн, расчет интеграла столкновений, особая точка, кубатурная формула, экспоненциальная сходимость, дробно-рациональное приближение, адаптивный метод, метод коллокаций, метод установления, нелинейное уравнение Шрёдингера, конденсация Бозе–Эйнштейна, волны на глубокой воде.

DOI: 10.31857/S0044466924020136, EDN: YIZWIJ

1. ВВЕДЕНИЕ И ПОСТАНОВКА ЗАДАЧИ

Выявление механизмов возникновения и развития турбулентных течений в жидкостях, газах и плазме остается одной из открытых проблем современной физики. Важную роль в этой проблеме играют процессы передачи энергии и других инвариантов между разномасштабными возмущениями. Такие процессы, называемые каскадами, возникают во многих нелинейных физических системах и характеризуются стационарными потоками инвариантов между масштабами, на которых происходит накачка и диссипация этих инвариантов [1, 2, 3].

Дальнейшие исследования показали, что физика случайного нелинейного взаимодействия волн не ограничивается каскадными процессами, она намного богаче. Значительный интерес представляет сама эволюция спектров энергии и других инвариантов в пространстве Фурье. Здесь обнаруживается множество эффектов: изотропная и анизотропная эволюция, возникновение неустойчивых стационарных

¹⁾ Работа выполнена при финансовой поддержке РФФИ (грант № 22-11-00287).

спектров, автомодельные режимы первого и второго рода, взрывной рост значений спектра и многое другое. Попытки математического исследования этих эффектов в последние десятилетия привели к бурному развитию теории волновой турбулентности (ТВТ) [4, 5], центральным объектом которой является кинетическое уравнение (КУ) с интегралом столкновений в правой части, описывающее волновые взаимодействия в гамильтоновых нелинейных физических системах [6].

К настоящему времени для КУ получено немного строгих результатов. Наиболее известными из них являются спектры Колмогорова–Захарова, представляющие точные стационарные решения КУ и описывающие каскадные процессы. Основным инструментом исследования эволюции решения КУ остается численное моделирование. Однако и здесь возникают существенные сложности, связанные с нелинейностью КУ, свойствами ядер интегралов столкновения, неограниченностью областей интегрирования, и самое главное с тем, что частотные диапазоны, в которых ищутся решения КУ (отличия частот накачки и диссипации энергии), могут достигать десяти декад, то есть градиенты решений могут варьироваться в пределах десяти порядков. Наибольший интерес для специалистов в области экспериментальной и теоретической физики состоит как раз в том, чтобы сделать частотный диапазон максимально широким и наблюдать в нем осредненное (сглаженное) поведение спектра.

В данной работе предложен один подход к решению *четырёхволновых КУ*, которые описывают появление двух новых волн при взаимодействии двух других волн. Как правило, такие КУ возникают в физических системах с кубической нелинейностью (например, в системах, описываемых *нелинейным уравнением Шрёдингера, НУШ*, см. [7, 8]). Четырёхволновые КУ также могут быть получены для систем с квадратичной нелинейностью, в которых трёхволновые резонансы невозможны (см. [9] и более простой способ исключения нерезонансных членов в приложении А.3 [5]). Кроме того, подобные КУ могут быть выведены из уравнения Больцмана при соответствующем виде функции распределения частиц, [10].

В [4] (см. гл. 6) сформулирован общий подход к выводу КУ в предположении слабой нелинейности волновых взаимодействий. В рамках этого подхода считается, что амплитуды и фазы волн являются независимыми случайными величинами, и, кроме того, фазы имеют равномерное распределение. На основе указанных предположений получен общий вид четырёхволновых КУ:

$$\begin{aligned} \frac{dn_{\mathbf{k}}}{dt} = & \int_{\mathbb{R}^{3d}} |W_{12}^{\mathbf{k}^3}|^2 \delta_{12}^{\mathbf{k}^3} \delta(\omega_{12}^{\mathbf{k}^3}) n_1 n_2 n_{\mathbf{k}} \times \\ & \times \left[\frac{1}{n_{\mathbf{k}}} + \frac{1}{n_3} - \frac{1}{n_1} - \frac{1}{n_2} \right] d\mathbf{k}_1 d\mathbf{k}_2 d\mathbf{k}_3, \end{aligned} \tag{1.1}$$

где $d \in \{2, 3\}$ — размерность физической системы, $n_{\mathbf{k}} = n(\mathbf{k}, t)$ — спектр волнового действия, т.е. спектр инварианта физической системы, \mathbf{k} — волновой вектор в d -мерном пространстве Фурье, $t > 0$ — время, $W_{12}^{\mathbf{k}^3} = W(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ — коэффициент взаимодействия волн $\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3$, возникающий при записи гамильтониана в пространстве Фурье (см. [4], п. 6.8); $n_i = n(\mathbf{k}_i, t)$, $i = 1, 2, 3$, $\delta_{12}^{\mathbf{k}^3} = \delta(\mathbf{k} + \mathbf{k}_3 - \mathbf{k}_1 - \mathbf{k}_2)$ — дельта функция Дирака, $\delta(\omega_{12}^{\mathbf{k}^3}) = \delta(\omega + \omega_3 - \omega_1 - \omega_2)$, ω_i — частота волны \mathbf{k}_i .

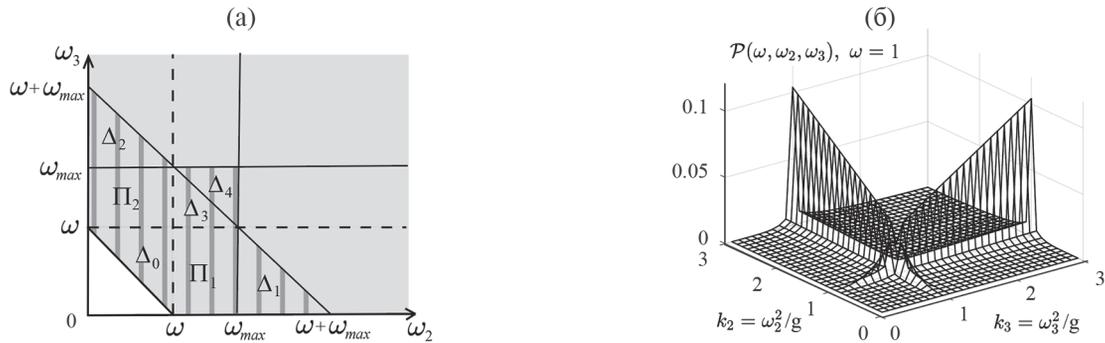
Существенный интерес представляет *однородный изотропный случай*. В таком случае решение КУ зависит только от модуля волнового вектора $n_{\mathbf{k}} = n(k, t)$, $k = |\mathbf{k}|$, и уравнение (1.1) допускает осреднение по угловым координатам в пространстве Фурье. В результате в правой части возникают новые переменные интегрирования $k_i = |\mathbf{k}_i|$, $i = 1, 2, 3$. Далее на основе дисперсионного соотношения волновое число k выражается через частоту волны ω , и используется сдвиговое свойство дельта функции $\delta(\omega_{12}^{\mathbf{k}^3})$. В итоге получается уравнение

$$\begin{aligned} \frac{dn_{\omega}}{dt} = St_{\omega}(n) = & \int_{\Delta_{\omega}} \mathcal{P}(\omega, \omega_2, \omega_3) \times \\ & \times [n_{\delta} n_2 n_3 + n_{\omega} (n_2 n_3 - n_{\delta} n_3 - n_{\delta} n_2)] d\omega_2 d\omega_3, \end{aligned} \tag{1.2}$$

где $n_{\omega} = n(\omega, t)$ — спектр волнового действия, зависящий от частоты волны ω , $n_i = n(\omega_i, t)$, $i = 2, 3$, $n_{\delta} = n(\omega_{\delta}, t)$, $\omega_{\delta} = \omega_2 + \omega_3 - \omega$,

$$\Delta_{\omega} = \{(\omega_2, \omega_3) : \omega_2 \geq 0, \omega_3 \geq 0, \omega_2 + \omega_3 \geq \omega\}$$

есть двумерная область интегрирования (см. серую область на фиг. 1а). Ядро $\mathcal{P}(\omega, \omega_2, \omega_3)$ интеграла столкновений, как правило, является неограниченной функцией, содержащей существенные особенности.



Фиг. 1. Область интегрирования Δ_ω : неограниченная серая область при $\omega \in [0, \infty)$, ограниченная заштрихованная область при $\omega \in [0, \omega_{\max})$ (а); ядро интеграла столкновений, описывающее взаимодействия поверхностных волн на глубокой воде в однородном изотропном случае (б).

Приведем два примера приложений, где возникают уравнения вида (1.2) (другие приложения обсуждаются в [5], гл. 5).

Пример 1. Физические системы, описываемые уравнением Шредингера с кубической нелинейностью, возникающие в оптике, космологии, динамике сверхтекучих жидкостей и бозе-газа (см. [11, 12, 13]). В таких системах коэффициент взаимодействия $W_{12}^{k^3} \equiv 1$ и при отсутствии стабильных когерентных структур (в неакустическом режиме) дисперсионное соотношение есть $\omega = k^2$. Кроме того, в трехмерном случае удается вывести явное выражение для $\mathcal{P}(\omega, \omega_2, \omega_3)$ (см. приложение А из [8]):

$$\mathcal{P}(\omega, \omega_2, \omega_3) = \frac{4\pi^3}{\sqrt{\omega}} \min[\sqrt{\omega}, \sqrt{\omega_2}, \sqrt{\omega_3}, \sqrt{\omega_2 + \omega_3 - \omega}]. \tag{1.3}$$

Пример 2. Динамика волн на поверхности глубокой воды. В этом приложении $W_{12}^{k^3}$ имеет сложный вид [14] (см. также некоторые приближенные формулы из [15]), и явную формулу для $\mathcal{P}(\omega, \omega_2, \omega_3)$ в общем случае получить не удастся, имеется лишь интегральное выражение, представляющее осреднение по угловым координатам в пространстве Фурье. Из него с использованием квадратурных формул удается вычислить \mathcal{P} при различных $\omega, \omega_2, \omega_3$, см. фиг. 1б. Здесь дисперсионное соотношение $\omega = \sqrt{gk}$, g — ускорение свободного падения.

Отметим, что в рассмотренных примерах ядро \mathcal{P} содержит особенности (изломы и/или большие градиенты) в окрестности линий $\omega_2 = \omega, \omega_3 = \omega$ (штриховые линии на фиг. 1а) и в окрестности границы $\Delta_\omega: \omega_2 = 0, \omega_3 = 0$ и $\omega_2 + \omega_3 = \omega$. Это типичная ситуация для четырехволновых КУ.

Характерное поведение решения (1.2) — степенная функция. В частности точные решения (1.2) с ядром (1.3), спектры Колмогорова–Захарова, есть

$$n_\omega \sim \omega^{-7/6}, \quad n_\omega \sim \omega^{-3/2} \ln^{-1/3} \frac{\omega}{\omega_f}, \quad \omega \gg \omega_f, \tag{1.4}$$

где ω_f — частота накачки. О выводе, верификации и валидации таких решений подробно написано в [16]. Как отмечено выше, подобные решения реализуются, когда в системе имеется накачка и диссипация инвариантов, действующие на существенно разных масштабах (дистанция между ними в спектральном пространстве может достигать десяти декад). Для моделирования накачки и диссипации к уравнению (1.2) нужно добавить соответствующие члены:

$$\begin{aligned} \frac{dn_\omega}{dt} = & \underbrace{St_\omega(n) + \frac{c_f}{\sqrt{2\pi\sigma}} \exp\left[-\frac{1}{2}\left(\frac{\omega - \omega_f}{\sigma}\right)^2\right]}_{\text{накачка}} - \\ & \underbrace{\left[(\omega / \omega_L)^{-p} + (\omega / \omega_R)^q \right] n_\omega}_{\text{диссипация}}, \end{aligned} \tag{1.5}$$

где $c_f, \sigma, \omega_f, \omega_L, \omega_R, p, q$ — положительные параметры: ω_L, ω_R — характерные значения частот низкочастотной диссипации (связанной, например, с потерей устойчивости) и высокочастотной диссипации (например, вязкой), $\omega_L \ll \omega_f \ll \omega_R$, числа p, q определяют крутизну спектра в окрестности диссипации и соответствуют порядкам производных диссипативных членов в уравнениях динамики рассматриваемой среды; c_f, σ характеризуют градиент спектра накачки. Отметим, что градиенты членов, описывающих накачку и диссипацию, могут отличаться на несколько порядков.

Численное решение уравнений (1.2) и (1.5) предполагает постановку задач Коши для них на ограниченном частотном интервале $\omega \in [0, \omega_{\max}]$ с условием $\omega_L, \omega_f, \omega_R \in [0, \omega_{\max}]$. В таком случае, как правило, предполагается, что значения n_ω при $\omega > \omega_{\max}$ пренебрежимо малы (см., например, приложение В из [8]), а область интегрирования Δ_ω становится ограниченной областью сложной формы, которую далее мы будем представлять в виде совокупности подобластей $\Delta_0 - \Delta_4, \Pi_1, \Pi_2$, см. заштрихованную часть на фиг. 1а. В заключении этого параграфа сформулируем две важные задачи, возникающие при разработке вычислительных алгоритмов решения уравнений (1.2) и (1.5) в указанной постановке.

Задача 1. Разработка быстросходящихся кубатурных формул в подобластях на фиг. 1а. Ключевой здесь является проблема расчета интегралов вида

$$I^g = I^g[f] = \int_{-1}^1 g(x)f(x)dx, \tag{1.6}$$

где $g(x)$ имеет точки ветвления или другие особенности на границах отрезка $[-1, 1]$, $f(x)$ — функция высокого порядка гладкости. Действительно, подстановка степенного решения из (1.4) в правую часть (1.2) и анализ задачи интегрирования вдоль переменной ω_2 в подобласти Π_2 на фиг. 1а (или вдоль переменной ω_3 в подобласти Π_1) приводит к выражению (1.6), в котором $g(x) = (1+x)^\beta$, $\beta > -1$. Причем, при поиске нестационарных решений (при эволюции спектра к стационарному режиму) β является априори неизвестной величиной и может изменяться со временем.

С другой стороны, при наличии особенностей у ядра \mathcal{P} в окрестности линий $\omega_2 = \omega, \omega_3 = \omega$ при интегрировании вдоль переменной ω_2 в подобласти Π_2 (вдоль переменной ω_3 в подобласти Π_1 , а также при интегрировании по Δ_0) мы также можем прийти к выражению (1.6), где $g(x)$ имеет особые точки в малых окрестностях $x = -1$ или $x = 1$, тип которых (полюс, точка ветвления, существенно особая точка) неизвестен.

Задача 2. Высокоточная аппроксимация решения $n(\omega, t)$ на отрезке $\omega \in [0, \omega_{\max}]$, градиенты которого могут отличаться на несколько порядков. Отметим, что помимо адаптации приближения к большим градиентам по переменной ω (градиентам функций, задающих накачку и диссипацию в (1.5), или градиентам спектра при взрывном росте его значений, по поводу последнего см. [19, 18, 17]) возникает проблема построения устойчивой схемы для итераций по времени с шагом, существенно превышающим минимальное расстояние между узлами сетки по переменной ω . По аналогии с условиями Куранта–Фридрихса–Леви явные схемы здесь не будут эффективными. Адаптивные схемы на основе методов Рунге–Кутты, апробированные авторами при получении решений из [16], также привели к неприемлемому уменьшению шага по t и большим временным затратам.

Целью настоящей работы является конструирование, обоснование и реализация алгоритмов решения задач 1, 2 и применение этих алгоритмов для численного моделирования волновой кинетики на основе уравнений (1.2), (1.5).

2. МЕТОДЫ РАСЧЕТА ИНТЕГРАЛА СТОЛКНОВЕНИЙ

Прежде всего отметим, что для перехода к расчету интеграла вида (1.6) при интегрировании по треугольным подобластям $\Delta_0 - \Delta_4$ необходимо применить отображения стандартного квадрата $Sq = \{(x_1, x_2) : -1 \leq x_1, x_2 \leq 1\}$ на $\Delta_0 - \Delta_4$ [20]. К примеру, такое отображение для области Δ_0 и соответствующий Якобиан имеют вид:

$$\omega_2(x_1, x_2) = -\frac{\omega}{8}[3x_1 - x_2(x_1 + 1) - 1] + \frac{\omega}{2}, \tag{2.1}$$

$$\omega_3(x_1, x_2) = -\frac{\omega}{8}[3x_2 - x_1(x_2 + 1) - 1] + \frac{\omega}{2},$$

$$J(x_1, x_2) = \left(\frac{\omega^2}{16}\right)(2 - x_1 - x_2).$$

Таким образом, если $\varphi(\omega_2, \omega_3)$ — подынтегральная функция (1.2), то интеграл по Δ_0 от $\varphi(\omega_2, \omega_3)$ сводится к интегралу по Sq от функции $\tilde{\varphi}(x_1, x_2) = \varphi(\omega_2(x_1, x_2), \omega_3(x_1, x_2))$:

$$\begin{aligned} & \int_{\Delta_0} \varphi(\omega_2, \omega_3) d\omega_2 d\omega_3 = \\ & = \frac{\omega^2}{16} \int_{-1}^1 \int_{-1}^1 \tilde{\varphi}(x_1, x_2) (2 - x_1 - x_2) dx_1 dx_2. \end{aligned}$$

Важным обстоятельством для нас является то, что при указанной замене граница треугольника переходит в границу Sq , поэтому особенности $\tilde{\varphi}(x_1, x_2)$ лежат на сторонах Sq , т.е. в точках $x_1 = \pm 1$ при $x_2 \in [-1, 1]$ и в точках $x_2 = \pm 1$ при $x_1 \in [-1, 1]$.

Интегралы по подобластям Π_1, Π_2 сводятся к интегралам по Sq с помощью линейных замен переменных. Задачу интегрирования по Sq с помощью теоремы Фубини легко свести к нахождению повторных интегралов вида (1.6) по стандартному отрезку $[-1, 1]$ с указанными особенностями в точках $x = \pm 1$. Для решения проблемы интегрирования используем фундаментальные результаты теории ортогональных функций (квадратурные формулы максимального алгебраического порядка точности), теории меры (обобщенное неравенство Маркова–Стилтьеса) и теории разностных аппроксимаций (формула Эйлера–Маклорена). Конкретно, нам потребуются нижеследующие теоретические результаты.

1. Квадратурные формулы Гаусса–Якоби. Пусть в формуле (1.6) $g(x) = \gamma(x) = (1-x)^\alpha (1+x)^\beta$, $\alpha > -1$, $\beta > -1$. Построим последовательность многочленов Якоби, ортонормированных на отрезке $[-1, 1]$ с весом $\gamma(x)$: $p_0(x), p_1(x), \dots$, таких что

$$\int_{-1}^1 \gamma(x) p_n(x) p_m(x) dx = \delta_{mn},$$

где δ_{mn} — символ Кронекера. Пусть $\lambda_n > 0$ — коэффициент при старшей степени полинома p_n , а x_1, \dots, x_n — корни p_n , которые являются попарно различными, принадлежащими отрезку $[-1, 1]$ числами (о свойствах ортогональных многочленов и их корней см. [21], гл. 1, 2). Тогда для расчета интеграла (1.6) имеет место квадратурная формула:

$$\begin{aligned} I^g[f] & \approx I_n^\gamma = I_n^\gamma[f] = \sum_{k=1}^n w_k f(x_k), \\ w_k & = -\frac{\lambda_{n+1}}{\lambda_n} \frac{1}{p_{n+1}(x_k) p_n'(x_k)}. \end{aligned} \quad (2.2)$$

Числа x_k и w_k называются соответственно узлами и весами квадратурной формулы I_n^γ , $k = 1, \dots, n$.

Для весов имеет место более удобное представление: $w_j = [\sum_{k=0}^n (p_k(x_j))^2]^{-1}$, см. [22], с. 48. Имеются так-

же алгоритмы быстрого вычисления узлов и весов с помощью построения матрицы Якоби (см. [23], п. 2.7.5 и [24], гл. 12, gauss.m).

Рассмотрим функционал погрешности формулы (2.2): $E_n^\gamma[f] = I^g - I_n^\gamma$. Для $E_n^\gamma[f]$ имеется классическая оценка через супремум норму $\|f^{(2n)}\|$, см. [23], стр. 98. Однако в случае ограниченной гладкости $f(x)$ или при наличии у ее аналитического продолжения особых точек в комплексной плоскости, величина $\|f^{(2n)}\|$ может не существовать, либо расти слишком быстро, чтобы гарантировать сходимость. В связи с этим целесообразно применять более общую оценку погрешности.

Теорема 1. Для функционала погрешности квадратуры (2.2) имеет место оценка (см. [25], гл. 5, § 3, п. 3)

$$|E_n^\gamma[f]| \leq 2E_{2n}^b[f] \int_{-1}^1 \gamma(x) dx, \quad (2.3)$$

где $E_n^b[f]$ — погрешность наилучшего приближения функции f полиномом степени не выше n . Скорость убывания величины $E_n^b[f]$ связана со свойствами гладкости f . Конкретно, имеют место следующие оценки (см. [26], гл. 7, 8):

1. Пусть $v \geq 1$, $f(x) \in C^{v-1}[-1,1]$, $f^{(v)} \in L[-1,1]$ имеет ограниченную вариацию: $\int_{-1}^1 |f^{(v)}(x)| dx \leq V$, тогда при любых $n > v$

$$|E_n^b[f]| \leq \frac{2V}{\pi v(n-v)^v}. \tag{2.4}$$

2. Пусть $f(x)$ — аналитична на отрезке $[-1,1]$ и имеет аналитическое продолжение $f(z)$ в эллипс Бернштейна E_ρ с суммой полюсов $\rho > 1$, $z \in E_\rho \subset \mathbb{C}$. Пусть на контуре эллипса $|f(z)| \leq M$. Тогда

$$|E_n^b[f]| \leq \frac{2M\rho^{-n}}{\rho-1}. \tag{2.5}$$

Отметим важный нюанс: оценки (2.4), (2.5) доказаны для приближений Фурье-Чебышёва. Однако, как показано в [27], такие приближения являются ненасыщаемыми, то есть при любой гладкости $f(x)$ порядок их сходимости совпадает с порядком сходимости наилучшего полиномиального приближения.

2. Формулы Гаусса для функций с разрывами второго рода. Квадратурная формула, описанная выше, при поиске интегралов (1.6) с функцией $f(x)$ высокой гладкости является оптимальной (имеет наивысшую алгебраическую степень точности), если известен вид функции $g(x) = (1-x)^\alpha(1+x)^\beta$. Однако, как отмечено во введении, числа α, β могут меняться, а также возможны случаи, когда вид функции $g(x)$ неизвестен, но известно, что $g(x)$, либо ее производные терпят разрыв второго рода в окрестностях точек ± 1 . В таком случае полезным является следующий результат, полученный с применением обобщенного неравенства Маркова–Стилтьеса (см. лемму 6.1 в [28]). Здесь и далее будем использовать обозначение $\psi(x) = g(x)f(x)$.

Лемма 1. Пусть интегранд (1.6), $\psi(x)$, и все его производные до порядка $2n$ существуют и непрерывны на интервале $(-1,1)$, и $\psi^{(j)}(x) \geq 0$ при $x \in (-1,1)$, $j = 0,1,\dots,2n$ (далее множество таких функций будем обозначать $M_{2n}^+(-1,1)$). Построим классическую квадратурную формулу Гаусса (КФГ) I_n^1 (формулу (2.2) при $\gamma(x) \equiv 1$) для приближения интеграла (1.6) и обозначим $E_n^1[\psi] = I_n^g[f] - I_n^1[\psi]$. Тогда

$$\max \left\{ \int_{x_n}^1 \psi(x) dx - w_n \psi(x_n), 0 \right\} \leq E_n^1[\psi] \leq \int_{x_n}^1 \psi(x) dx. \tag{2.6}$$

Если же $\psi^{(j)}(x) \leq 0$ при $x \in (-1,1)$, $j = 0,1,\dots,2n$ (далее множество таких функций обозначено $M_{2n}^-(-1,1)$), то

$$\max \left\{ \int_{-1}^{x_1} \psi(x) dx - w_1 \psi(x_1), 0 \right\} \leq E_n^1[\psi] \leq \int_{-1}^{x_1} \psi(x) dx. \tag{2.7}$$

Из приведенного утверждения можно получить следующий результат.

Лемма 2. Если в (1.6) $f(x) \in C^{v-1}[-1,1]$, $v \geq 1$, $g(x) = (1-x)^\alpha$, α — нецелое число, $-1 < \alpha < v/2 - 1$, и $\psi \in M_{2n}^+(-1,1)$, то

$$|E_n^1[\psi]| \leq M \frac{j_1^{2(\alpha+1)}}{2^{\alpha+1}(1+\alpha)} n^{-2(1+\alpha)} \text{ при всех } n = 1,2,\dots, \tag{2.8}$$

где $M = \|f\|$, $j_1 \approx 2.4048$ — первый положительный ноль функции Бесселя нулевого порядка. Кроме того, если $M > \pi(1+\alpha)$, то при $n \rightarrow \infty$

$$\left| E_n^1[\Psi] \right| \sim C_{\alpha, M} n^{-2(1+\alpha)}, \text{ где } \frac{j_1^{2(\alpha+1)}}{2^{\alpha+1}} \left(\frac{M}{1+\alpha} - \pi \right) \leq C_{\alpha, M} \leq \frac{M j_1^{2(\alpha+1)}}{2^{\alpha+1}(1+\alpha)}. \quad (2.9)$$

Если в условиях леммы $\nu > \alpha/2 - 1$, то имеют место оценки вида (2.3), (2.4) при $\gamma(x) = 1$.

На практике достаточно требовать монотонность подынтегральной функции и ее производных в малой окрестности точки $x = 1$, т.е. нужно чтобы $\psi \in \mathcal{M}_{2n}^+(1 - \delta, 1)$, где $\delta > 0$ — малое число. При этом, если в точке $x = 1 - \delta$ условие монотонности нарушается, то область интегрирования можно разбить на два отрезка $[-1, 1] = [-1, 1 - \delta] \cup [1 - \delta, 1]$, и в каждом отрезке строить КФГ.

Результат аналогичный лемме 2 имеет место при $g(x) = (1+x)^\beta$, где $-1 < \beta$ — нецелое число.

В этой работе предложена и проверена численно идея об улучшении оценок (2.8), (2.9) за счет замены переменной интегрирования в (1.6) вида $x = h(y)$, где $h(y)$ конформно отображает отрезок $[-1, 1]$ в себя, $h(-1) = -1$, $h(1) = 1$. Такая замена приводит к интегралу

$$I^g[f] = \int_{-1}^1 g(h(y)) f(h(y)) h'(y) dy. \quad (2.10)$$

Далее для расчета (2.10) будем применять КФГ с узлами y_1, \dots, y_n и рассмотрим три вида функций $h(y)$.

Теорема 2. Пусть $f(x) \in C^{\nu-1}[-1, 1]$, $\nu \geq 1$, интегранд (1.6) имеет особенность в одной из точек $x = \pm 1$, $g(x) = (1 \mp x)^\alpha$, α — нецелое число, $-1 < \alpha < \nu/2 - 1$, и $\psi \in \mathcal{M}_{2n}^\pm(-1, 1)$.

1. Зададим

$$h(y) = h_\pm(y) = \pm 1 + \varepsilon \sinh \left[\frac{y \mp 1}{2} \sinh^{-1} \frac{2}{\varepsilon} \right], \quad (2.11)$$

где $\varepsilon > 0$ — малое число, \sinh^{-1} — обратная функция к \sinh . Тогда погрешность КФГ при вычислении (2.10) удовлетворяет оценке:

$$\left| E_n^1[\Psi] \right| \leq \frac{M \tilde{\varepsilon}}{2^{\alpha+1}} \frac{j_1^{2(\alpha+1)}}{2^{\alpha+1}(1+\alpha)} n^{-2(1+\alpha)}, \quad (2.12)$$

$$\tilde{\varepsilon} \sim \varepsilon \ln[1/\varepsilon] \text{ при } \varepsilon \rightarrow 0.$$

2. Пусть

$$h(y) = \sin(\pi y / 2). \quad (2.13)$$

Тогда погрешность КФГ при вычислении (2.10) удовлетворяет оценке:

$$\left| E_n^1[\Psi] \right| \leq M \frac{(j_1^2 \pi)^{2(\alpha+1)}}{2^{5(\alpha+1)}(1+\alpha)} n^{-4(1+\alpha)}. \quad (2.14)$$

3. Если функция $f(x)$ аналитична на отрезке $[-1, 1]$ и

$$h(y) = \tan \left[A \sinh \frac{2By}{1-y^2} \right], \quad (2.15)$$

то скорость сходимости КФГ выше степенной и ниже экспоненциальной. Здесь A, B — положительные параметры, их рекомендуемые значения $A = \pi/2$, $B = \pi/4$ (см. [29] и [23], п. 2.9.2).

Замечание 1. Замены переменной (2.13), (2.15) приводят к симметричному сгущению узлов в окрестности точек ± 1 , поэтому могут быть использованы, когда интегранд имеет особенности в окрестностях обеих точек $x = -1$ и $x = 1$. Замену (2.11) можно обобщить для учета двух таких особенностей:

$$h(y) = \left[\left(h_+^{-1} + h_-^{-1} \right) / 2 \right]^{-1}, \tag{2.16}$$

где верхний индекс “−1” обозначает обратную функцию. Причем, за счет наличия параметра ϵ замена (2.16) обеспечивает более гибкую адаптацию к особенностям, чем (2.13), (2.15). Замена переменной (2.11), (2.16) являются частным случаем отображений (4.3), (4.4), которые обсуждаются в п. 4.1. Отметим также, что поскольку, как правило, $\psi = gf \notin \mathcal{M}_{2n}^+(-1,1) \cap \mathcal{M}_{2n}^-(-1,1)$, то применение леммы 1 для анализа сходимости при наличии двух указанных особенностей невозможно.

Замечание 2. Оценка погрешности для (2.13) основана на равенствах $h'(-1) = h'(1) = 0$. В качестве обобщения можно рассмотреть функцию $h(y)$, удовлетворяющую условиям $h^{(j)}(-1) = h^{(j)}(1) = 0$ при $j = 1, \dots, r - 1, r \geq 2$. При использовании такой замены получаем оценку погрешности

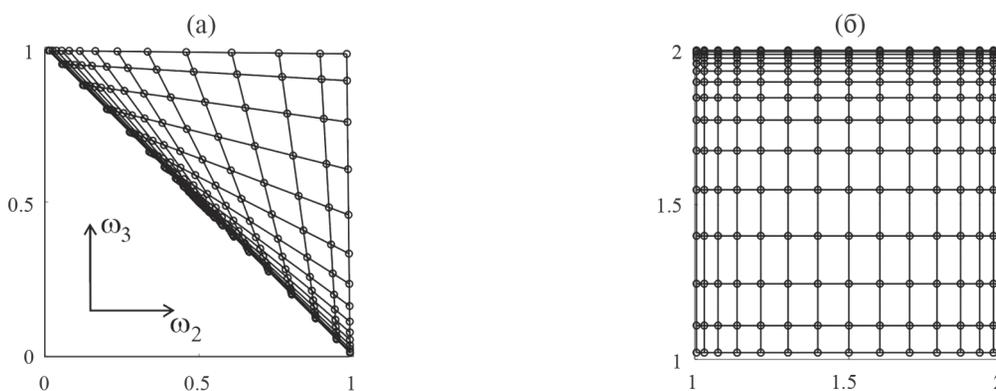
$$\left| E_n^1[\Psi] \right| \sim \tilde{C}_{\alpha, M} n^{-2r(1+\alpha)} \text{ при } n \rightarrow \infty,$$

где $\tilde{C}_{\alpha, M}$ зависит только от M и α . Из этой оценки видно, что порядок сходимости увеличивается пропорционально r . Указанные обобщения можно получить, если например, возвести функцию (2.13) в степень r , или использовать в качестве $h(y)$ полиномиальную функцию определенного вида, см. [31].

Замечание 3. Тот факт, что при использовании (2.15) скорость сходимости становится выше степенной связан со свойством $h^{(j)}(-1) = h^{(j)}(1) = 0$, где $j \in \mathbb{N}$. Замена переменной $h(y)$, удовлетворяющая таким требованиям, была впервые предложена в [32]. Однако в этой работе мы воспользовались известным отображением из [29].

3. Формулы Мори для функций с особенностями в точках $x = \pm 1$. Применение замен переменной $h(y)$, таких что $h^{(j)}(\pm 1) = 0$ при $j = 1, \dots, r - 1$ и формулы трапеций с равномерным распределением на отрезке $[-1, 1]$ по переменной y дает высокоточную квадратурную формулу для расчета интегралов с особенностями в окрестностях точек ± 1 . Оценки погрешности таких приближений строятся на основе формулы Эйлера–Маклорена (общий подход описан, например, в [23], п. 2.9). В работе [29] на основе такой идеи и замены (2.15) предложена так называемая двуэкспоненциальная квадратурная формула “double exponential formula”:

$$I_n^{DE}[gf] = \frac{2}{n+1} \sum_{k=1}^n g(h(-1+ks)) \times \\ \times f(h(-1+ks)) h'(-1+ks), \quad s = \frac{2}{n+1}.$$



Фиг. 2. Сетки из 15×15 узлов в подобластях Δ_0 (а) и Π_1 (б) при $\omega = 1, \omega_{\max} = 2$. Координаты узлов — образы нулей полиномов Лежандра, НПЛ, под действием (2.1) и замены (2.11) по обоим координатам (а); координаты узлов — НПЛ $[-1, 1]$ на $[1, 2]$ по направлению ω_2 и образы НПЛ под действием (2.11) по направлению ω_3 (б). Здесь использована замена $h_1(y)$ из (2.11) при $\epsilon = 10^{-1}$.

Теорема 3 (см. [29,3]). При $g(x) = (1-x^2)^\alpha$ и $f(x) \equiv 1$. Асимптотика убывания погрешности I_n^{DE} носит экспоненциальный характер:

$$\begin{aligned} E_n^{DE}[\Psi] &= E_n^{DE}[gf] = I^s[f] - I_n^{DE}[\Psi] \sim \\ &\sim C_1 \exp\left(-C_2 \frac{n}{\log^2 n}\right), \text{ при } n \rightarrow \infty, \end{aligned}$$

где C_1 и C_2 — положительные постоянные.

В заключение раздела покажем для примера, как распределяются узлы описанных методов интегрирования в областях Δ_0 и Π_1 . На фиг. 2 показаны узлы кубатурной формулы, построенной с использованием (2.1), (2.11).

3. ТЕСТОВЫЕ РАСЧЕТЫ ИНТЕГРАЛОВ ИЗ ПРИЛОЖЕНИЙ

Для анализа эффективности предложенных методов расчета интегралов с особенностями проведем экспериментальное исследование сходимости формул, описанных в разд. 2. При этом будем исходить из того, что числа α и β или даже вид функции $g(x)$ нам неизвестен. Рассмотрим следующие тестовые задачи, имеющие непосредственное отношение к приложениям.

Тест 1. Интеграл для случая обратного каскада в НУШ при наличии накачки экспоненциального вида, см. уравнение (1.5), расположенной на высоких частотах $\omega_f \sim \omega_{\max}$:

$$I_1(\omega_f, \sigma) = \int_{-1}^1 (1-\omega^2)^{-2/3} \exp\left[-\frac{(\omega-\omega_f)^2}{2\sigma^2}\right] d\omega. \quad (3.1)$$

Рассмотрим два варианта этой тестовой задачи: вариант 1 — расчет $I_1(0,0.1)$. Интегралы такого вида возникают, когда правая часть $St_\omega(n)$ рассчитывается при $\omega \sim \omega_f \sim \omega_{\max}$. Вариант 2 — расчет $I_1(0.9998, 10^{-4})$. Такие задачи возникают, когда правая часть $St_\omega(n)$ рассчитывается при $\omega \ll \omega_f \sim \omega_{\max}$:

$$I_1(0,0.1) = I_{1,1} = \frac{\sqrt{\pi} \Gamma(1/3) {}_1F_1(1/2, 5/6; -1/2\sigma^2)}{\Gamma(5/6)} \approx 0.252377677219947773,$$

$$I_1(0.9998, 10^{-4}) = I_{1,2} \approx 0.058481158415646.$$

Здесь Γ и ${}_1F_1$ означают гамма-функцию Эйлера и гипергеометрическую функцию соответственно.

Тест 2. Интеграл для случая прямого каскада в НУШ (см. второй спектр в (1.4)), возникающий при расчете правой части (1.2) (в прямом каскаде, как правило, $0 \sim \omega_f \ll \omega_{\max}$, поэтому верхний предел по ω зададим в 1000 раз большим ω_f):

$$I_2 = \int_{\omega_f}^{1000\omega_f} \omega^{-1} \ln^{-1/3} \frac{\omega}{\omega_f} d\omega = \frac{3}{2} (\log 1000)^{2/3}. \quad (3.2)$$

Тест 3. Интеграл для случая волн на поверхности глубокой воды:

$$I_3(p) = \int_0^1 \omega^p \exp\left[\frac{1}{\log(1-\omega)}\right] d\omega. \quad (3.3)$$

Здесь для моделирования свойств ядра \mathcal{P} , использована функция $\exp\left[\frac{1}{\log(1-\omega)}\right]$ с существенно особыми точками на границах интервала интегрирования, скорость убывания которой в окрестности $\omega = 0$ быстрее экспоненциальной. Рассмотрим два варианта этой тестовой задачи: вариант 1, расчет $I_3(0)$, — задача, возникающая при вычислении интеграла $St_\omega(n)$ на постоянном спектре, представляющем глобальное термодинамическое равновесие; вариант 2, расчет $I_3(-8)$. Этот интеграл возникает при вычислении

правой части КУ на спектре Колмогорова–Захарова в прямом каскаде энергии $n_{\omega} \sim 1 / \omega^8$ (такое решение было впервые опубликовано в [33], где символ ω использован для обозначения не частотной, а волновой переменной, см. также [14]):

$$I_3(0) = I_{3,1} = 2K_1(2) \approx 0.2797317636330448,$$

$$I_3(-8) = I_{3,2} \approx 1205.63842587396,$$

где $K_1(2)$ — модифицированная функция Бесселя II рода.

Для организации расчетов в тестах 1–3, там где необходимо, с помощью линейных замен переменной перейдем к стандартному отрезку интегрирования $[-1, 1]$ и используем методы, предложенные в разд. 2: формулы Гаусса–Якоби (для краткости обозначим их GJ), классические формулы Гаусса (G), формулы Гаусса с заменами переменных (2.11), (2.16) (G_{\sinh}), (2.13) (G_{\sin}), (2.15) (G_{\tanh}) и двухэкспоненциальные формулы (DE). Пусть $I_{n,a}$ — приближенное значение интеграла, a — номер теста. В экспериментах будем рассчитывать относительную погрешность

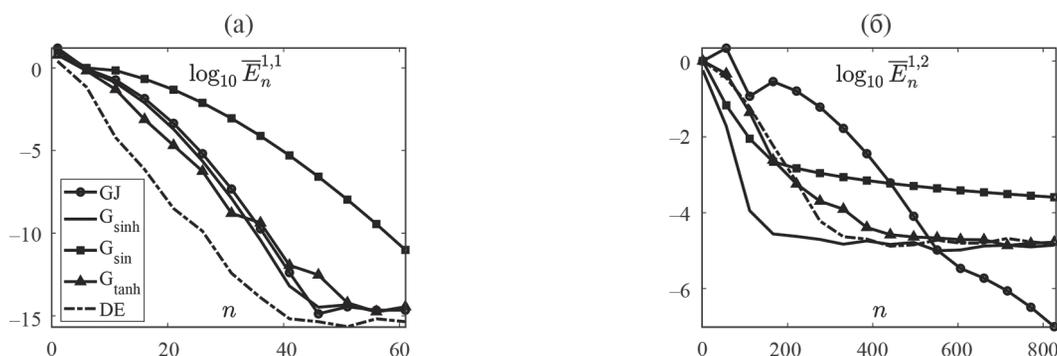
$$\bar{E}_n^a = |I_a - I_{n,a}| / |I_a|.$$

В расчетах замены (2.11), (2.16) и параметры ϵ и A, B варьировались для достижения максимальной скорости сходимости (максимального угла наклона графика $\log_{10} \bar{E}_n^a$).

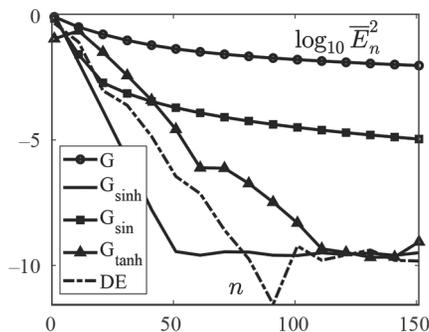
На фиг. 3 показаны графики десятичных логарифмов погрешностей расчета $I_{1,1}$ (график а) и $I_{1,2}$ (график б). При вычислении $I_{1,1}$ в методе G_{\sinh} задано $\epsilon = 10^{0.3}$ и использовано сгущение (2.16), в методах G_{\tanh} и DE заданы параметры $A = 0.6\pi / 2, B = 0.6\pi / 4$. При расчете $I_{1,2}$ в методе G_{\sinh} задано $\epsilon = 10^{-12}$ и использована замена $h_+(y)$ из (2.11), в методах G_{\tanh} и DE заданы $A = 1.2\pi / 2, B = 1.2\pi / 4$.

Важное и неожиданное заключение, следующее из фиг. 3, состоит в том, что даже при известных значениях параметров α и β формула GJ, имеющая максимальную алгебраическую степень точности, не обеспечивает максимальную скорость убывания погрешности: на графике а видно, что методы DE, G_{\sinh} и G_{\tanh} сходятся быстрее. Рекордные результаты демонстрирует DE (для достижения точности 0.1% DE требует примерно в 2 раза меньше узлов, чем GJ). На фиг 3б видно, что при $n < 400$ погрешность GJ больше, чем у всех других методов. Однако только GJ позволят получить повышенную точность (погрешность 10^{-7} и ниже) при большом числе узлов. Погрешность других методов, достигнув значений порядка 10^{-5} , не уменьшается при увеличении n . Отметим, что наилучшие результаты при расчете $I_{1,2}$ демонстрирует G_{\sinh} , позволяя получить максимальную точность уже на 200 узлах. В целом, тест $I_{1,2}$ показывает, насколько ресурсоемкой является задача расчета интеграла столкновений на практике — видно, что в подобластях $\Delta_0 - \Delta_4, \Pi_1, \Pi_2$ требуется задавать сетки не менее, чем из 200×200 узлов и адаптировать дополнительно параметры методов интегрирования. Если же адаптировать параметры не удастся, могут потребоваться сетки из 500×500 узлов и более.

На фиг. 4 приведены значения погрешностей, полученные при расчете I_2 . В методе G_{\sinh} использована замена переменной $h_-(y)$ из (2.11) и задано $\epsilon = 10^{-15}$; в методах G_{\tanh} и DE заданы $A = 1.5\pi / 2,$



Фиг. 3. Логарифмы относительных погрешностей при расчете $I_{1,1}$ (а), $I_{1,2}$ (б).



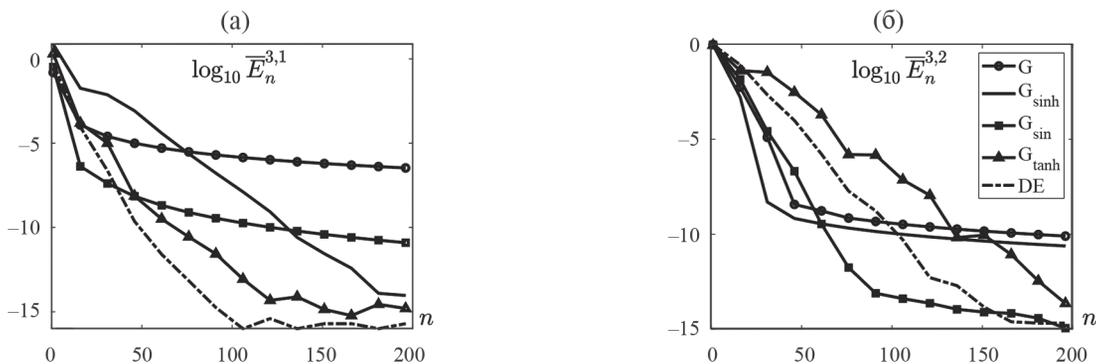
Фиг. 4. Логарифм относительной погрешности при расчете интеграла I_2 .

$B = 1.5\pi / 4$. Подынтегральная функция, включающая логарифм, имеет точки ветвления на границе отрезка, $\omega = \omega_f$, и в малой окрестности границы, $\omega = 0$. Формула GJ в таком случае неприменима, поэтому для сравнения используется классическая формула G, которая демонстрирует самое медленное убывание погрешности с ростом n . По скорости сходимости лидирует метод G_{\sinh} . Хорошие результаты демонстрирует DE.

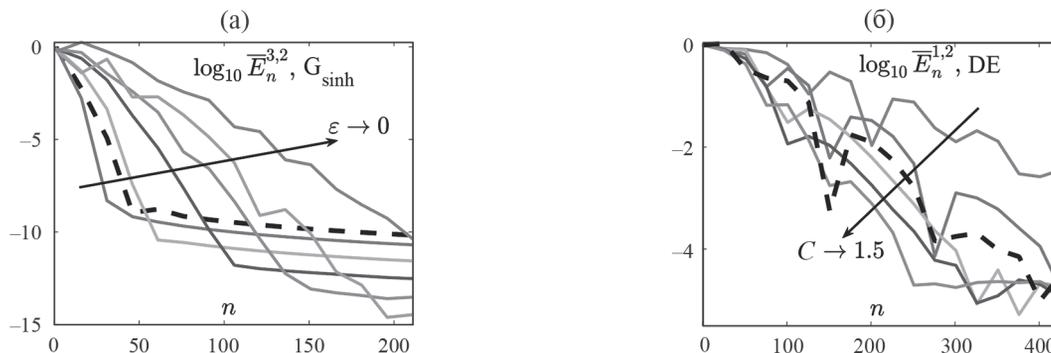
Отметим, что здесь и в других расчетах методы DE и G_{\tanh} , основанные на применении одной и той же замены переменной (2.15), имеют близкую скорость сходимости. Однако DE, как правило, дает погрешность на 2–4 порядка меньше, чем G_{\tanh} . Это связано с тем, что сетка G_{\tanh} сильно разрежена в центре отрезка интегрирования, что приводит к снижению точности аппроксимации при наличии минимальных градиентов интегранта внутри отрезка.

На фиг. 5 показаны графики десятичных логарифмов погрешностей расчета $I_{3,1}$ (график а) и $I_{3,2}$ (график б). В обоих экспериментах при реализации метода G_{\sinh} использована замена переменной (2.16). При этом в расчете $I_{3,1}$ задано $\varepsilon = 10^{-8}$, в расчете $I_{3,2}$ — $\varepsilon = 10^{-1}$. В обоих экспериментах при реализации G_{\tanh} и DE заданы $A = 1.2\pi / 2$, $B = 1.2\pi / 4$.

Результаты на фиг. 5 наглядно показывают, что при изменении параметра подынтегральной функции скорость сходимости любого метода в сравнении с другими может сильно измениться. Действительно, при расчете $I_3(0)$ методы DE и G_{\tanh} сходятся существенно быстрее других, а метод G_{\sinh} существенно отстает от них. При расчете $I_3(-8)$ максимальную скорость сходимости при малых n демонстрирует G_{\sinh} ; методы G и G_{\sin} близки по скорости сходимости к G_{\sinh} . При этом G_{\sin} дает повышенную точность (погрешность порядка 10^{-13} при $n = 100$). Методы DE и G_{\tanh} в этой задаче сходятся медленнее других. Отсюда можно сделать вывод, что моделирование того или иного режима эволюции решения КУ требует предварительных расчетов на сходимость для выбора подходящего метода интегрирования и настройки его параметров.



Фиг. 5. Логарифмы относительных погрешностей при расчете интегралов $I_{3,1}$ (а), $I_{3,2}$ (б).



Фиг. 6. Логарифмы относительных погрешностей (а) при вариации параметра ϵ метода G_{\sinh} при вычислении $I_{3,2}$, сплошные линии по направлению стрелки соответствуют $\epsilon = 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-7}$, штриховая линия — $\epsilon = 10^1$; (б) при вариации коэффициента C в значениях параметров $A = C\pi/2, B = C\pi/4$ метода DE при вычислении $I_{1,2}$, сплошные линии по направлению стрелки соответствуют $C = 0.6, 0.8, 1, 1.2, 1.4$, штриховая линия — $C = 3$.

Следует заметить, что наблюдаемый характер убывания погрешностей предложенных методов соответствует теоретическим оценкам, представленным в предыдущем разделе. Методы G, G_{\sinh} при больших n демонстрируют степенной характер убывания погрешности (см. фиг. 4, 5), причем порядок сходимости G_{\sinh} примерно в два раза выше G . Эти результаты соответствуют оценкам (2.8), (2.14). На фиг. 3б, 4, 5б погрешность метода G_{\sinh} быстро убывает при увеличении n от единицы (такое почти скачкообразное падение погрешности возникает из-за малого значения множителя $\tilde{\epsilon}$ в оценке (2.12)), затем характер убывания меняется на степенной, и асимптотика сходимости G_{\sinh} совпадает с асимптотикой G , что соответствует оценкам (2.8), (2.12).

Отметим, что при использовании G_{\sinh} уменьшение ϵ ниже некоторого порогового значения приводит к тому, что скорость сходимости замедляется, однако при этом возрастает максимальная точность, которую позволяет достичь метод. Эта ситуация продемонстрирована на фиг. 6, а. Здесь изображен график логарифма погрешности метода G_{\sinh} при вычислении $I_{3,2}$. Видно, что при уменьшении ϵ от 10^1 сначала угол наклона графика увеличивается, затем, достигая максимальных значений в районе $\epsilon^* = \epsilon = 10^{-1}$, начинает уменьшаться. Оптимальные значения ϵ^* зависят от свойств подынтегральной функции.

Из рисунка видно, что применение метода G_{\sinh} ведет к альтернативе: либо мы получаем предельную точность 10^{-8} на малом числе узлов ($n \sim 50$) и далее с ростом n погрешность убывает медленно (это реализуется при относительно больших ϵ , см. линию, соответствующую $\epsilon \sim 10^{-1}$), либо получаем точность машинной арифметики 10^{-15} на значительно большем числе узлов ($n \sim 200$) при $\epsilon < 10^{-5}$. По нашему мнению эффект замедления скорости сходимости при $\epsilon \rightarrow 0$ связан с тем, что сетка узлов квадратурной формулы становится более разреженной в центре отрезка интегрирования.

Характерное изменение скорости сходимости метода DE при изменении параметров A, B изображено на фиг. 6б. Здесь мы ввели дополнительный коэффициент $C: A = C\pi/2, B = C\pi/4$ и привели логарифм погрешности метода DE при вычислении $I_1(0.9998, 10^{-4})$. На графике видно, что увеличение C от значений $C = 0.6$ сначала приводит к увеличению скорости сходимости, но после достижения некоторого оптимального значения C^* (в этом эксперименте $C^* \approx 1.5$) скорость сходимости с ростом C начинает уменьшаться. Как и в случае метода G_{\sinh} , оптимальные значения C^* зависят от свойств подынтегральной функции. Отметим однако, что во всех проведенных экспериментах эти значения отличаются от $C = 1$ не более, чем в 2 раза.

Подводя итог этого раздела, сформулируем следующие основные выводы: 1) формулы GJ в рассмотренных задачах не являются оптимальными даже при известных α и β ; 2) при отсутствии оптимизации параметров рекомендуется использовать метод DE с параметрами $A = \pi/2, B = \pi/4$, если есть возможность оптимизировать метод под задачу, лучше использовать G_{\sinh} ; 3) для обоснованного выбора метода интегрирования в каждой конкретной задаче необходимы предварительные эксперименты на сходимость.

4. НОВЫЙ МЕТОД РЕШЕНИЯ ЭВОЛЮЦИОННОЙ ЗАДАЧИ

Далее для уравнения (1.5) будем рассматривать задачу Коши

$$\begin{aligned} \frac{\partial n_\omega}{\partial t} &= \mathcal{S}t_\omega(n) + \mathfrak{F}(\omega) - \mathfrak{D}(\omega)n_\omega, \\ n_\omega(0) &= n_I(\omega), \end{aligned} \quad (4.1)$$

где

$$\begin{aligned} \mathfrak{F}(\omega) &= \frac{c_f}{\sqrt{2\pi\sigma}} \exp\left[-\frac{1}{2}\left(\frac{\omega - \omega_f}{\sigma}\right)^2\right], \\ \mathfrak{D}(\omega) &= \left[(\omega / \omega_L)^{-p} + (\omega / \omega_R)^q\right] n_\omega \end{aligned}$$

суть члены, задающие накачку и диссипацию соответственно. Поскольку модуль правой части (4.1) и сама неизвестная функция могут обращаться в бесконечность в точке $\omega = 0$, при разработке численного метода целесообразно ввести малое число ω_{\min} , $0 < \omega_{\min} \leq \omega_L$, и искать решение (4.1) на отрезке $\omega \in \omega = [\omega_{\min}, \omega_{\max}]$, а для расчета интеграла $\mathcal{S}t_\omega(n)$ использовать аналитическое или непрерывное продолжение решения в область $[0, \omega_{\min}]$ (о таком алгоритме подробнее см. в [17]). Кроме того, зачастую полагается, что при $\omega > \omega_{\max}$ решение пренебрежимо мало [8].

Далее зададим в качестве начального данного $n_I(\omega) \equiv 0$ и опишем метод поиска решения задачи (4.1) — функции $n_\omega(t)$ при $\omega \in \omega$, $t \geq 0$.

4.1. Приближение решения по переменной ω

Как отмечено во введении, при наличии накачки и диссипации, действующих на разных частотных масштабах, градиенты решения КУ по переменной ω могут отличаться на несколько (вплоть до десяти) порядков. Для аппроксимации функции с такими сильными особенностями будем использовать дробно-рациональную барицентрическую интерполяционную формулу [35]:

$$\begin{aligned} n(\omega, t) \approx r_M [n](\omega, t) &= \sum_{m=1}^M \frac{\xi_m n(\tilde{\omega}_m, t)}{\mathcal{L}(\omega) - \tilde{y}_m} \Big/ \sum_{m=1}^M \frac{\xi_m}{\mathcal{L}(\omega) - \tilde{y}_m}, \\ \forall t &\geq 0, \end{aligned} \quad (4.2)$$

где $\xi_m = M / T'_M(y_m) = (-1)^{m-1} \sin(\frac{2m-1}{2M}\pi)$ — веса интерполяции; $T_M(y)$ — полином Чебышёва степени M ; $\tilde{y}_m = \chi(y_m)$, y_m — нули $T_M(y)$, $m = 1, \dots, M$; $\mathcal{L}(\omega)$ — линейное отображение отрезка $\bar{\omega}$ на $[-1, 1]$; $\tilde{\omega}_m = \mathcal{L}^{-1}(\tilde{y}_m)$ — узлы интерполяции.

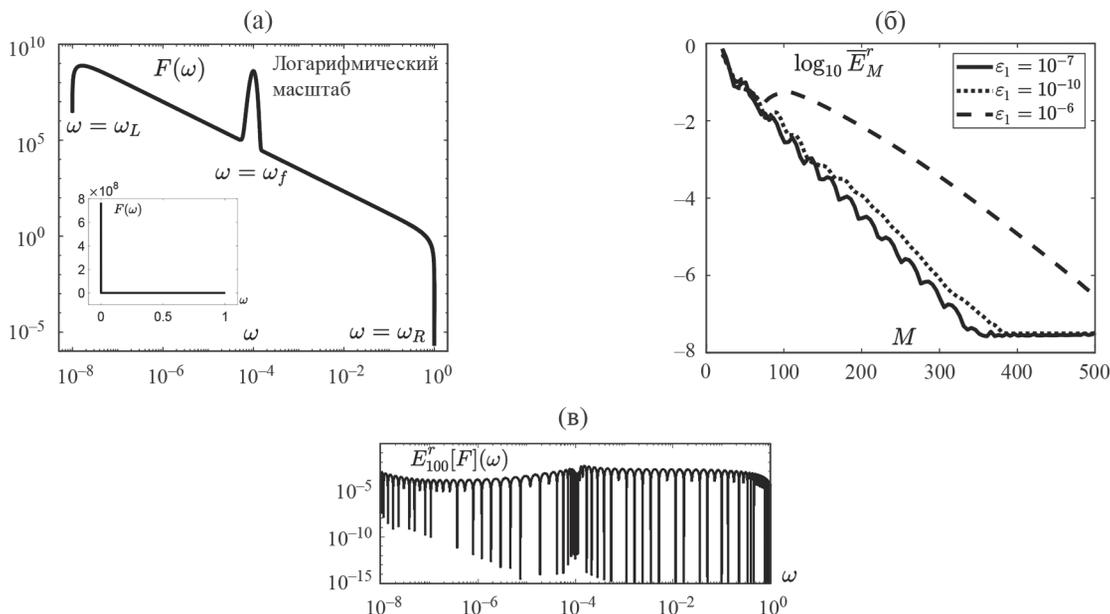
Функция $\chi(y)$ — конформное отображение отрезка $[-1, 1]$ в себя с условиями $\chi(-1) = -1$, $\chi(1) = 1$. Если $\chi = \text{id}$ — тождественное отображение, то приближение (4.2) совпадает с интерполяционным полиномом с узлами Чебышёва [34]. В [35] заложена идея построения функции $\chi(y)$, позволяющая существенно повысить точность интерполяции (4.2) при наличии у функции n_ω изолированных особых точек в комплексной плоскости \mathbb{C} , лежащих близко к отрезку $\bar{\omega}$: отображение $\chi(y)$ должно “уносить” ближайшую к отрезку особую точку на большое расстояние в \mathbb{C} . Таким образом, в силу теоремы Бернштейна, скорость сходимости (4.2) существенно возрастет по сравнению с классическим случаем $\chi(y) = \text{id}$.

В [30] показано, что в качестве $\chi(y)$ можно выбрать функцию

$$\chi(y) = \delta + \varepsilon \sinh \left\{ \left(\sinh^{-1} \frac{1-\delta}{\varepsilon} + \sinh^{-1} \frac{1+\delta}{\varepsilon} \right) \frac{y-1}{2} + \sinh^{-1} \frac{1-\delta}{\varepsilon} \right\}, \quad (4.3)$$

в которой (δ, ε) — координата особой точки аналитического продолжения n_ω , как функции ω , в \mathbb{C} . Отметим, что замены (2.11) являются частным случаем (4.3) при $\delta = \pm 1$.

Важная проблема, касающаяся обобщения (4.2) на случай, когда у аналитического продолжения $n(\mathcal{L}^{-1}(y), t)$ имеется несколько особых точек: $y_1^*, \dots, y_S^* \in \mathbb{C}$, обсуждалась в [36]–[38]. Было показано, что для построения отображения $\chi(y)$ в таком случае можно применять формулу:



Фиг. 7. Функция $F(\omega)$ в логарифмическом масштабе и в линейном масштабе на внутреннем графике (а); логарифмы относительных погрешностей \overline{E}_M^r при различных значениях ε_1 (б); график отклонения $r_M[F](\omega)$ от $F(\omega)$ при $M = 100$, $\varepsilon_1 = 10^{-7}$ в логарифмическом масштабе (в).

$$\chi = \left(\frac{1}{S} \sum_{s=1}^S \chi_s^{-1} \right)^{-1}, \tag{4.4}$$

где индекс “ -1 ” означает обратную функцию; χ_1, \dots, χ_S имеют вид (4.3), где $\delta = \delta_1, \dots, \delta_S$, $\varepsilon = \varepsilon_1, \dots, \varepsilon_S$ являются действительной и мнимой координатами особых точек y_1^*, \dots, y_S^* соответственно.

Более того, эксперименты показывают, что отображения вида (4.3), (4.4) можно использовать при построении приближений (4.2) как для целых функций с экспоненциальными пиками (моделирование накачки), так и для функций со степенными погранслоями на отрезке $\overline{\omega}$ (моделирование диссипации). При этом в качестве δ выбирается $\omega = \omega_f$ (для накачки) и $\omega = 0$, $\omega = \omega_{\max}$ (для диссипации), а в качестве ε — малое число, значение которого можно связать с градиентами членов накачки и диссипации.

Для демонстрации высокой точности (4.2)–(4.4) рассмотрим задачу о приближении функции $F(\omega) = n_\omega + \mathfrak{F}(\omega) - \mathfrak{D}(\omega)n_\omega$ при $n_\omega = \omega^{-7/6}$ (спектр Колмогорова–Захарова из (1.6)), $\omega_L = 10^{-8}$, $\omega_R = 1$, $\omega_f = 10^{-4}$, $\sigma = 10^{-5}$, $c_f = 10^4$, $p = q = 2$ на отрезке $\overline{\omega} = [10^{-8}, 1]$.

Для приближения $F(\omega)$ используем формулу (4.2), в которой зависимость от t отсутствует: $r_M[F](\omega)$. Для построения отображения $\chi(y)$ используем (4.4) и учтем, что $F(\omega)$ имеет особенности в окрестности точек $\omega = \omega_{\min}, \omega_f, \omega_{\max}$ (см. фиг. 7а), поэтому положим $S = 3$, $\delta_1 = \mathcal{L}(\omega_{\min}) = -1$, $\delta_2 = \mathcal{L}(\omega_f)$, $\delta_3 = \mathcal{L}(\omega_{\max}) = 1$. Значения $\varepsilon_1 = 10^{-7}$, $\varepsilon_2 = 10^{-5}$, $\varepsilon_3 = 10^{-1}$, обеспечивающие максимальную скорость убывания относительной погрешности приближения (4.2) с ростом M , были подобраны экспериментально. Наибольшее влияние на убывание погрешности имеет параметр ε_1 , см. фиг. 7б.

Относительная погрешность в этом эксперименте определялась по формуле:

$$\overline{E}_M^r = \max_{\omega \in \Omega_M} E_M^r[F](\omega),$$

$$E_M^r[F](\omega) = |F(\omega) - r_M[F](\omega)| / \mu, \quad \mu = \max_{\omega \in \overline{\omega}} |F(\omega)|.$$

Здесь множество точек $\Omega_M = \bigcup_{m=1}^{M-1} \{\tilde{\omega}_m + k(\tilde{\omega}_{m+1} - \tilde{\omega}_m) / K, k = 1, \dots, K\}$, $K = 50$, т.е. Ω_M есть объединение равномерных подсеток на интервалах $[\tilde{\omega}_m, \tilde{\omega}_{m+1}]$, $m = 1, \dots, M - 1$.

Из фиг. 7а видно, что в линейном масштабе правая часть и решение КУ визуализируются в виде скачка в нуле. Однако реальная картина поведения решения, которую видно в логарифмическом масштабе, намного богаче. Для аппроксимации такого поведения классические методы не годятся, а формулы (4.2)–(4.4) дают высокую точность на достаточно малом количестве узлов, см. фиг. 7 б, в.

4.2. Приближение решения по времени

Для приближения решения по времени создадим новый алгоритм. Пусть требуется вычислить эволюцию решения $n(\omega, t)$ на отрезке $t \in [0, T]$ с начальными данными $n(\omega, 0) = n_I(\omega)$. Используем модификацию интерполяционного полинома с узлами Чебышева при фиксированном ω и $t \in [0, T]$:

$$n(\omega, t) \approx P_K[n](\omega, t) = \frac{1}{K} \sum_{k=1}^K \zeta_k(t) \frac{\xi_k T_K(\Lambda(t))}{\Lambda(t) - \theta_k} \times n(\omega, t_k) + n_I(\omega), \quad \zeta_k(t) = \frac{1 + \Lambda(t)}{1 + \theta_k}, \tag{4.5}$$

где $\Lambda(t)$ — линейное отображение отрезка $[0, T]$ на отрезок $[-1, 1]$; $t_k = \Lambda^{-1}(\theta_k)$ — узлы интерполяции, $k = 1, \dots, K$; θ_k — нули T_K , ξ_k — веса интерполяционной формулы. Множитель $\zeta_k(t)$ позволяет автоматически учесть начальное условие (условие на левой границе временного отрезка). Подобные приближения для решения краевых задач были предложены независимо в [25] и [39], их свойства обсуждались в [40]. Здесь мы обобщаем эту технику для решения задачи Коши.

Задачу Коши (4.1) будем решать с помощью метода коллокаций с узлами $(\tilde{\omega}_m, t_k)$, расположенными в прямоугольнике $\Omega = [\omega_{\min}, \omega_{\max}] \times [0, T]$. Для приближения решения в Ω используем прямое (тензорное) произведение интерполяций (4.2), (4.5):

$$n(\omega, t) \approx r_M \times P_K = \frac{1}{K} \sum_{m=1}^M \sum_{k=1}^K \zeta_k(t) \frac{\xi_m \xi_k T_K(\Lambda(t)) n(\tilde{\omega}_m, t_k)}{(\mathcal{L}(\omega) - \tilde{y}_m)(\Lambda(t) - \theta_k)} \Big/ \sum_{m=1}^M \frac{\xi_m}{\mathcal{L}(\omega) - \tilde{y}_m} + n_I(\omega). \tag{4.6}$$

Для приближения производной по времени в (4.1) найдем производную по t от (4.6) и запишем ее в узлах $(\tilde{\omega}_m, t_k)$. Используя основное свойство интерполяции, $r_M[n](\tilde{\omega}_i, t) = n(\tilde{\omega}_i, t)$, и обозначая $n_{ij} = n(\tilde{\omega}_i, t_j)$, $(n_t)_{ij} = \lim_{\omega \rightarrow \tilde{\omega}_i, t \rightarrow t_j} \partial n / \partial t$, получаем

$$(n_t)_{ij} \approx \sum_{k=1, k \neq j}^K a_{jk} n_{ik} + v_j n_{ij},$$

$$a_{jk} = \frac{2}{T} \left(\frac{(-1)^{j+k} (1 + \theta_j) s_k}{s_j (1 + \theta_k) (\theta_j - \theta_k)} \right), \tag{4.7}$$

$$v_j = \frac{2}{T} \left(\frac{1}{1 + \theta_j} + \frac{\theta_j}{2s_j^2} \right),$$

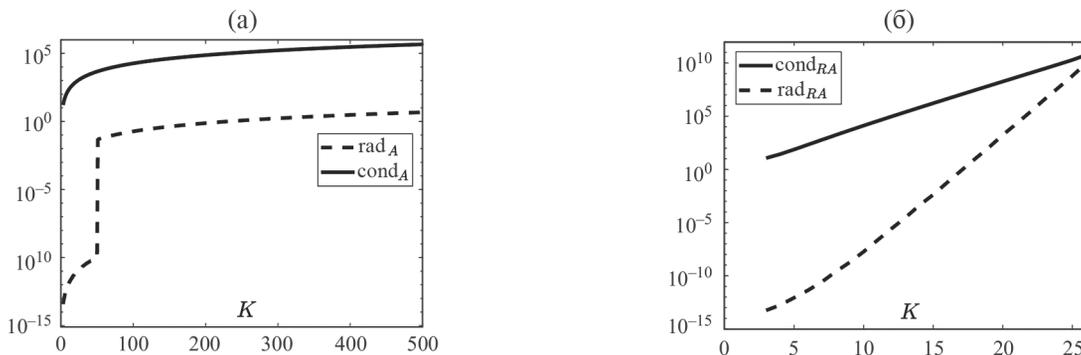
где $s_j = \sqrt{1 - \theta_j^2}$, $j = 1, \dots, K$. В рамках п. 4.2 полагаем $T=2$.

Пусть \mathcal{N} , $\mathcal{N}_t - M \times K$ матрицы, составленные из значений n_{ij} и $(n_t)_{ij}$ соответственно; $A - K \times K$ матрица, содержащая v_j на диагонали и a_{jk} вне диагонали, $j \neq k$. Тогда приближение (4.7) можно записать в матричном виде:

$$\mathcal{N}_t \approx \mathcal{N} A^T. \tag{4.8}$$

Для построения алгоритма нам понадобится спектральное разложение матрицы A :

$$A = R_A D_A R_A^{-1}, \tag{4.9}$$



Фиг. 8. Зависимости обусловленностей (сплошная линия) и радиусов интервалов, в которых гарантированно лежат значения обусловленностей (штриховая линия), от числа узлов K для матриц A (а) и R_A (б).

где R_A — матрица собственных векторов A , $D_A = \text{diag}(d_1, \dots, d_K)$ — матрица собственных значений.

Проведем анализ собственных чисел матрицы A и чисел обусловленности матрицы R_A с помощью интервальных методов [41, 42].

Пакет INTLAB позволяет проводить расчеты с гарантией точности, иными словами вычислять радиусы интервалов, в которые гарантированно попадает искомое значение собственного числа или значения числа обусловленности матрицы. В табл. 1 приведены значения центров и радиусов интервалов, в которые гарантированно попадают значения собственных чисел матрицы A при $K = 10$.

Таблица 1. Значения собственных чисел d_k матрицы A при $K = 10$, и радиусы интервалов rad_k , гарантированно содержащих собственные числа d_k

k	d_k	rad_k
1	58.651955985441816	7.11e-15
2,3	$2.999445156055363 \pm 7.96046273340461i$	9.55e-15
4,5	$4.26941311045503 \pm 6.062577061672625i$	1.49e-14
6,7	$5.028785289672468 \pm 3.933803674349112i$	3.43e-14
8	5.695873617901074	1.34e-13
9,10	$5.52844164214539 \pm 1.9302196040513i$	7.96e-14

В значениях d_k символом “ i ” обозначена мнимая единица. На фиг. 8 в логарифмической шкале приведены значения верхних границ обусловленностей cond_A и cond_{R_A} матриц A и R_A , посчитанных во второй норме, и радиусы интервалов rad_A и rad_{R_A} , в которых гарантированно лежат значения обусловленностей.

Благодаря проведенному анализу удастся сделать следующие выводы:

- Все собственные числа матрицы A достаточно далеко отстоят друг от друга и от нуля и лежат в комплексной плоскости.
- Обусловленность матрицы A с ростом числа узлов растет достаточно медленно, что обеспечивает численную устойчивость алгоритмов, основанных на обращении матрицы A .
- Обусловленность матрицы R_A с ростом числа узлов растет быстро и при $K = 25$ превышает 10^{10} . Таким образом, алгоритм, основанный на спектральном разложении (4.9), будет численно устойчивым только при небольших K .

Дальнейший анализ (4.9), основанный на методах дихотомии матричного спектра [43] показал, что при больших K используемой шестнадцатизначной арифметики `double` недостаточно, чтобы локализовать с высокой точностью собственные числа матрицы A . Это связано с наличием собственных

подпространств очень близких к вырожденным. Эксперименты показали, что в расчетах указанный эффект начинает проявляться при $K > 20$.

4.3. Аппроксимация и решение задачи (4.1)

Для решения нелинейной задачи (4.1) используем метод установления. Введем дополнительную переменную ϑ , положим $n = n(\omega, t, \vartheta)$, и будем искать решение (4.1) в области Ω как предел решений задачи

$$\begin{aligned} B_{\vartheta} n &= \frac{\partial n}{\partial t} - (St_{\omega}(n) + \mathfrak{F}(\omega) - \mathfrak{D}(\omega)n), \\ n(t, \omega, 0) &= n_I(\omega), (t, \omega) \in \Omega \end{aligned} \quad (4.10)$$

при $\vartheta \rightarrow \infty$, где B_{ϑ} — специальный оператор, содержащий производную по ϑ . Далее используем простой оператор $B_{\vartheta}^1 = \frac{\partial}{\partial \vartheta}$ и более сложный оператор $B_{\vartheta}^2 = (k_1 - k_2 \frac{\partial}{\partial t}) \frac{\partial}{\partial \vartheta}$, где k_1, k_2 — положительные параметры.

Для реализации итерационного процесса введем по переменной ϑ равномерную сетку с шагом Δ и узлами $\vartheta_s = s\Delta$, $s = 1, 2, \dots$; аппроксимируем производную по ϑ разностным отношением: $\partial n / \partial \vartheta \approx (n^{s+1} - n^s) / \Delta$, где $n^s = n(\omega, t, \vartheta_s)$; и получим схемы:

$$\begin{aligned} &\left(1 - \Delta \mathfrak{D}(\omega) - \Delta \frac{\partial}{\partial t}\right) n^{s+1} = \\ &= n^s - \Delta (St_{\omega}(n^s) + \mathfrak{F}(\omega)) = f_1(n^s, \omega), \end{aligned} \quad (4.11)$$

$$\begin{aligned} &\left(k_1 - \Delta \mathfrak{D}(\omega) - (k_2 + \Delta) \frac{\partial}{\partial t}\right) n^{s+1} = \\ &= \left(k_1 - k_2 \frac{\partial}{\partial t}\right) n^s - \Delta (St_{\omega}(n^s) + \mathfrak{F}(\omega)) = f_2(n^s, \omega), \end{aligned} \quad (4.12)$$

соответствующие применению операторов $B_{\vartheta}^1, B_{\vartheta}^2$.

Для сведения (4.11), (4.12) к задачам линейной алгебры используем приближения, описанные в пун. 4.2. Пусть \mathcal{N}^s — матрица значений решения $n(\omega, t, \vartheta_s)$ в узлах $(\tilde{\omega}_m, t_k)$; $F_{1,2}(\mathcal{N}^s)$ — матрицы значений $f_{1,2}(n^s, \omega)$ в тех же узлах; D_{ω} — диагональная матрица со значениями $\mathfrak{D}(\tilde{\omega}_1), \dots, \mathfrak{D}(\tilde{\omega}_M)$ на диагонали. Подставляя в (4.11), (4.12) приближение (4.6) и используя метод коллокаций с учетом формулы (4.8), получаем соответствующие задачи линейной алгебры для перехода с предыдущего, s -го, на следующий, $(s+1)$ -й, шаг метода установления

$$\mathcal{N}^{s+1} - \Delta (D_{\omega} \mathcal{N}^{s+1} + \mathcal{N}^{s+1} A^T) = F_1(\mathcal{N}^s), \quad (4.13)$$

$$k_1 \mathcal{N}^{s+1} - \Delta D_{\omega} \mathcal{N}^{s+1} - (k_2 + \Delta) \mathcal{N}^{s+1} A^T = F_2(\mathcal{N}^s). \quad (4.14)$$

Используем теперь спектральное разложение (4.9), умножим (4.13), (4.14) на матрицу R_A^{-1} справа и введем обозначения: $V^{s+1} = \mathcal{N}^{s+1} R_A^{-1}$, $G_{1,2}^s = F_{1,2}(\mathcal{N}^s) R_A^{-1}$. В итоге получаем системы с диагональными матрицами

$$V^{s+1} - \Delta (D_{\omega} V^{s+1} + V^{s+1} D_A) = G_1^s, \quad (4.15)$$

$$k_1 V^{s+1} - \Delta D_{\omega} V^{s+1} - (k_2 + \Delta) V^{s+1} D_A = G_2^s, \quad (4.16)$$

решения которых даются элементарными формулами:

$$v_{ij} = \frac{g_{1ij}}{1 - \Delta [\mathfrak{D}(\tilde{\omega}_i) + d_j]},$$

$$v_{ij} = \frac{g_{2ij}}{k_1 - \Delta \mathfrak{D}(\tilde{\omega}_i) - (k_2 + \Delta)d_j},$$

соответственно. Здесь v_{ij} — элементы матрицы V^{s+1} ; g_{1ij} , g_{2ij} — элементы матриц G_1^s , G_2^s . Параметры метода установления Δ , k_1 , k_2 следует выбирать таким образом, чтобы для всех i и j имели место соотношения:

$$\mathfrak{D}(\tilde{\omega}_i) + d_j \neq 1 / \Delta, \quad \Delta \mathfrak{D}(\tilde{\omega}_i) + (k_2 + \Delta)d_j \neq k_1.$$

Решение на $(s + 1)$ -м шаге метода установления дается формулой $\mathcal{N}^{s+1} = V^{s+1}R_A$. Затем, используя те же формулы, по значениям \mathcal{N}^{s+1} находим решение \mathcal{N}^{s+2} , и т.д. до тех пор, пока решение не установится. Критерий остановки итерационного процесса:

$$\|\mathcal{N}^{s+1} - \mathcal{N}^s\| / \Delta \leq \epsilon_R,$$

где $\|\cdot\|$ — максимальный элемент матрицы, ϵ_R — малое число.

Анализ сходимости метода установления в одной прикладной задаче дан в [44]. Исследование сходимости алгоритма, схожего с описанным выше, проведено для линейной задачи в [40]. Отметим также, что при вычислении v_{ij} в знаменателе стоит комбинация, включающая собственные числа нестационарного оператора d_j , значения диссипативной функции $\mathfrak{D}(\tilde{\omega}_j)$, задающей большой градиент в спектральном пространстве, и параметры метода установления Δ , k_1 , k_2 . Балансируя эти параметры, удается снизить влияние больших градиентов на точность схемы по времени и получить решение в тех задачах, где методы Рунге–Кутты и многоточечные методы решения жестких систем уравнений не сработали.

Ограничения на K , описанные в конце п. 4.2, приводили к тому, что разрешение (4.5) было недостаточным для аппроксимации решения на большом временном интервале. В связи с этим, значения T следует выбирать исходя из значений градиентов и частоты осцилляций решения как функции переменной t . Эти значения следует подбирать экспериментально так, чтобы метод установления сходился. Значения решения при $t = T$ можно использовать в качестве начальных данных для расчета решения в области $[\omega_{\min}, \omega_{\max}] \times [T, 2T]$. Таким образом, повторяя описанные операции, можно найти решение $n(\omega, t)$ при $(\omega, t) \in [\omega_{\min}, \omega_{\max}] \times [0, lT]$, где l — сколь угодно большое натуральное число.

5. МОДЕЛИРОВАНИЕ КОНДЕНСАЦИИ БОЗЕ–ЭЙНШТЕЙНА

В [16] проведен анализ КУ вида (1.2) с ядром (1.3), описывающего случайное нелинейное взаимодействие волн в бозе-газе, приводящее к формированию конденсата Бозе–Эйнштейна. Динамика бозе-газа моделируется на основе уравнения Гросса–Питаевского, которое представляет собой уравнение Шредингера с кубической нелинейностью.

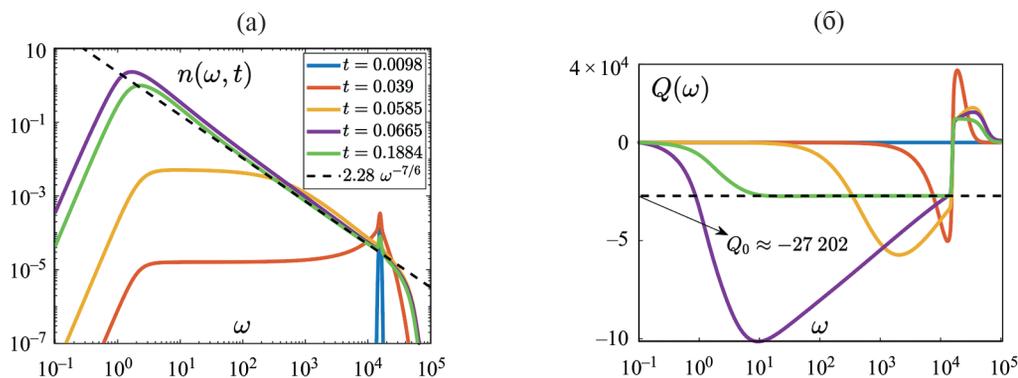
В рамках теории волновой турбулентности конденсация Бозе–Эйнштейна объясняется возникновением обратного каскада частиц бозе-газа, который описывается точным стационарным решением КУ (1.2), (1.3) — спектром Колмогорова–Захарова. Это решение, найденное в [16] с точностью до постоянного множителя, имеет вид

$$n_\omega = C_i |Q_0|^{1/3} \omega^{-7/6}, \quad C_i \approx 7.5774045 \times 10^{-2}. \tag{5.1}$$

Здесь Q_0 — постоянный поток частиц в инерционном частотном диапазоне, расположенном между частотами накачки и диссипации. Точное выражение для C_i , включающее гипергеометрические функции и являющееся достаточно громоздким, дано в [16].

Важное свойство решения (5.1) в том, что оно реализуется, когда в системе присутствует накачка и диссипация, действующие на существенно разных частотных масштабах: накачка — на высоких частотах, диссипация — на низких. Таким образом, для получения (5.1) в численных экспериментах нужно решать уравнение (1.5). При этом в области между накачкой и диссипацией (в инерционном диапазоне) возникает постоянный поток числа частиц, который задается формулой:

$$Q(\omega) = -2\pi \int_0^\omega \sqrt{\tilde{\omega}} St_{\tilde{\omega}} d\tilde{\omega}. \tag{5.2}$$



Фиг. 9. Установление решения КУ (1.5), (1.3) к стационарному спектру Колмогорова–Захарова (штриховая линия) (а), установление потока частиц в инерционном диапазоне (б).

Нужно отметить, что существенное значение, как для теории, так и для практики представляет сам процесс установления спектра к решению (5.1). Важные вопросы, связанные с этим: является ли решение (5.1) устойчивым, как быстро устанавливается спектр для различных начальных данных, какие формы он может принимать?

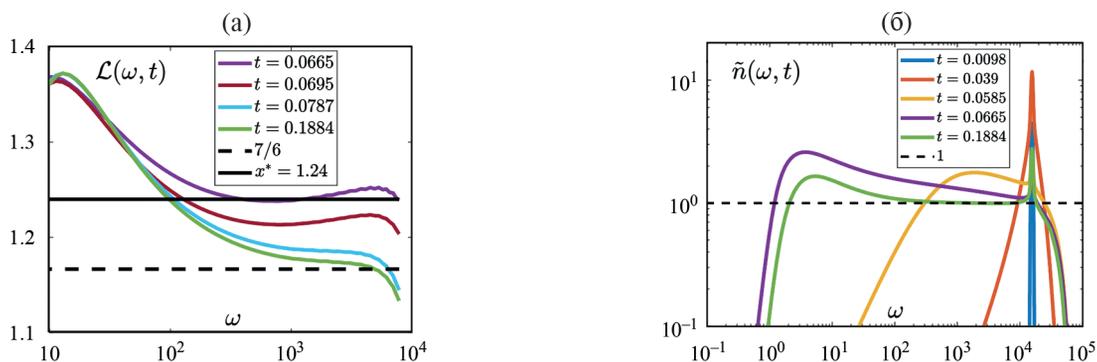
Алгоритм, описанный в этой работе, реализован в среде MATLAB и использован для решения уравнения (1.5) с ядром интеграла столкновений (1.3) и нулевым значением спектра в начальный момент времени. Численное решение строилось на отрезке $[\omega_{\min}, \omega_{\max}] = [10^{-1}, 10^5]$. Параметры накачки и диссипации заданы следующим образом: $c_f = 50$, $\sigma = 500$, $\omega_f = 125^2$, $\omega_L = 10$, $\omega_R = (1/3) \times 10^4$, $p = 4$, $q = 7$. Указанные параметры подобраны так, чтобы установление спектра происходило в широком инерционном диапазоне.

Результаты, приведенные ниже, получены с применением аппроксимации (4.2) со сгущением узлов (4.4) в окрестностях двух точек: $\delta_1 = \mathcal{L}(0)$, $\varepsilon_1 = 10^{-9.5}$, $\delta_2 = \mathcal{L}(\omega_f)$, $\varepsilon_2 = 10^{-1}$. Для построения кубатурных формул в областях $\Delta_0 - \Delta_4$, Π_1 , Π_2 (см. фиг. 1 а) использован метод Гаусса G. Параметры численного метода: $n = 181$, $M = 256$. В процессе установления использовано переключение между операторами B_{ϑ}^1 , B_{ϑ}^2 при $k_1 = k_2 = 1$. Для обеспечения сходимости метода установления с невязкой $\varepsilon_R = 10^{-6}$ длина отрезка по времени T и шаг метода установления Δ менялись в диапазонах $T \in [10^{-4}, 10^{-2}]$, $\Delta \in [10^{-2}, 1]$.

На фиг. 9 показан процесс установления спектра в обратном каскаде и соответствующие значения потока $Q(\omega)$. Видно, что в инерционном диапазоне $\omega \in [10, 10^4]$ поток устанавливается к константе $Q(\omega) = Q_0 \approx -27202$.

Из фиг. 9 видно, что процесс установления носит нетривиальный характер: сначала в области низких частот наблюдается резкий рост спектра, при этом спектр поднимается выше стационарного степенного решения (5.1) (оно изображено пунктиром на фиг. 9, а). Затем спектр плавно спускается и стабилизируется. Нетривиальная динамика потока частиц (5.2) в этом эксперименте видна на фиг. 9, б. Заметим, что при достаточно больших t ($t > 0.15$) решение устанавливается к спектру (5.1) с высокой точностью (относительная разность решений на предыдущем и следующем шагах по времени меньше 0.1%) и с высокой точностью совпадает со спектром Колмогорова–Захарова (максимальное отклонение при $\omega = 10$ приблизительно равно 0.0668).

Отметим, что в проведенном эксперименте на ранних этапах эволюции формируется неклассический спектр $n_{\omega} \sim \omega^{-x^*}$, $x^* \approx 1.24$ — линия на фиг. 9а, соответствующая $t = 0.0665$. Угол наклона этой линии больше, чем у спектра Колмогорова–Захарова: $x^* > 7/6$. Аналогичные спектры обсуждались в [10, 18, 19] при решении задач Коши для уравнения (1.2) с ядром (1.3) без членов, задающих накачку и диссипацию. Известно, что при определенных условиях на начальные данные в таких задачах возможен неограниченный рост значений спектра в окрестности нулевой моды [45], приводящий к разрушению решения в конечный момент времени $t = t^*$. Кроме того, в малой окрестности t^* при $t < t^*$ имеет место автомодельная эволюция второго рода со степенным убыванием решения $n_{\omega} \sim \omega^{-x}$ при достаточно больших ω [17].



Фиг. 10. Эволюция модуля логарифмической производной решения $n_\omega(t)$ в инерционном диапазоне: переход от неклассического спектра $n_\omega \sim \omega^{-x^*}$ к стационарному решению $n_\omega \sim \omega^{-7/6}$ (а), компенсированные спектры (б).

В рассмотренной нами задаче после выхода на автомодельное степенное решение $n_\omega \sim \omega^{-x^*}$ спектр под действием низкочастотной диссипации плавно изменяет свой наклон, и решение сходится к стационарному спектру Колмогорова–Захарова. Переход между указанными решениями происходит на временном интервале $t \in [0.065, 0.188]$, его удобно также наблюдать на графике модуля логарифмической производной $\mathcal{L}(\omega, t) = \left| \frac{\partial \ln n(\omega, t)}{\partial \omega} \right|$, фиг. 10а. Отметим также, что процесс перехода от автомодельного к стационарному спектру при отсутствии диссипации описан в [10, 19]. Однако он существенно отличается от наблюдаемого нами перехода: при отсутствии диссипации решение КУ неограниченно растет в малых ω (формируется конденсат). Эволюция конденсата описывается отдельным уравнением. После формирования конденсата при $t > t^*$ переход от автомодельного к стационарному спектру происходит вдоль волны, распространяющейся обратно из области низких в область высоких частот. Похожий сценарий перехода между спектрами возникает и в других моделях [46, 47]. Поскольку он выходит за рамки известной классификации Зельдовича, в указанных работах введен термин “автомодельное решение III рода”.

Существенной сложностью для численного анализа эволюции спектра в обратном каскаде является возникновение крутого фронта у левой границы частотного диапазона (в работе [16] он назван “bottleneck”). Этот фронт более отчетливо виден при построении так называемых компенсированных спектров, то есть графиков функции $\tilde{n}(\omega, t) = n_\omega(\omega, t) / (C_i |Q_0|^{1/3} \omega^{-7/6})$, которые в инерционном диапазоне сходятся к единице, фиг. 10б. За счет вариации параметров накачки и диссипации этот фронт можно уменьшить, но в любом случае его влияние на процесс вычислений является ощутимым: при формировании фронта длину временного интервала T в разработанном методе приходилось варьировать и уменьшать в несколько раз. Отметим однако, что предложенный алгоритм справляется с этой сложной проблемой и позволяет получить стационарные решения с высокой точностью и относительно небольшими затратами памяти и машинного времени. Приведенные результаты получены на персональной ЭВМ AMD Ryzen 5950X, DDR4 32Gb 3200MHz. Расчеты проводились в системе MATLAB на 16 ядрах с распараллеливанием. Расчет до установления с относительной разностью решений на предыдущем и следующем шагах по времени порядка 10^{-4} длился порядка суток. Отметим, что здесь для расчета интеграла столкновений использован самый простой вариант — формулы Гаусса, G, без адаптации к особенностям. Результаты разд. 2, 3 позволяют утверждать, что при использовании других методов временные показатели, названные выше, могут быть существенно улучшены. Подчеркнем также, что расчет интеграла столкновения для каждого $\omega = \omega_m, m = 1, \dots, M$, происходит независимо, поэтому метод допускает эффективное распараллеливание. Распараллеливание на 16 ядер показало практически кратное уменьшение времени вычислений.

Полученные результаты служат верификацией теории волновой турбулентности для нелинейного уравнения Шредингера (НУШ). Они также свидетельствуют о корректности и высокой точности проведенных расчетов. Отметим, что похожие результаты (только менее точные и на более коротких частотных диапазонах) обнаруживаются как в физических экспериментах, так и при прямом численном моделировании трехмерного НУШ. Подробнее по поводу последнего обстоятельства см. [16].

6. ЗАКЛЮЧЕНИЕ

В этой работе создан алгоритм поиска турбулентных спектров, являющихся решениями четырехволновых КУ, в случаях изотропного и однородного взаимодействия волн в нелинейных физических системах. Принципиальная сложность, возникающая при решении таких задач, состоит в существенной разномасштабности исследуемых процессов (масштабы частот накачки и диссипации, как правило, отличаются на несколько порядков) и в наличии особенностей, как у решений КУ, так и у ядер интегралов столкновения. Разработка вычислительных алгоритмов для решения задачи потребовала 1) построения высокоточных кубатурных формул с возможностью адаптации к указанным особенностям и 2) создания специальных аппроксимаций по частотной и временной переменным. Отметим в заключении важные нюансы, оставшиеся за рамками описанных работ.

В теории волновой турбулентности вывод КУ основывается на нескольких существенных предположениях (см. [4], гл. 6), и доказать, что они будут выполняться при долговременной эволюции спектра, до сих пор не удалось. Строгие оценки отклонения решения КУ от спектра исходного уравнения, описывающего динамику гамильтоновой системы, получены только на временных интервалах порядка ϵ -малое от характерного времени эволюции спектра. Такие оценки, использующие идею теоремы Ланфорда для кинетического уравнения Больцмана [50], представлены, например, в [51]. Для выполнения базовых предположений теории волновой турбулентности на больших временных интервалах в ходе вычислений требуется тонкая настройка начальных данных и членов, описывающих накачку и диссипацию. Кроме того, необходима верификация полученных численных решений КУ при сравнении с решениями исходных динамических уравнений. Сложности, которые возникают при таком сравнении, и его результаты для случая НУШ обсуждаются в [8]. Отметим, что в той работе показано высокоточное совпадение спектров решений динамического и кинетического уравнений на временных интервалах, где значения спектров изменяются приблизительно в четыре раза, и качественное соответствие спектров в существенно больших временных диапазонах.

Важным является вопрос о существовании и единственности решений четырехволновых КУ, который подробно изучался в ряде работ. Одна из основных, статья [48], посвящена исследованию обобщенных (в смысле распределений Шварца) решений задачи Коши для КУ, удовлетворяющих определенным условиям (см. п. 2.1 в [48]). В зависимости от начальных данных формулируются теоремы о локальных и глобальных существовании и единственности таких решений (см. теоремы 2.16, 2.18 в [48]). Для других типов КУ получено меньше строгих математических результатов. Доказаны, например, теоремы существования и единственности решений трехволновых КУ с ограниченным ядром интеграла столкновений в классе достаточно быстро убывающих на бесконечности функций [49].

Для КУ имеют место аналоги законов сохранения исходных дифференциальных моделей (например, для КУ (1.2), (1.3) — сохранение частиц и энергии, см. [8]). Эти инварианты представляются интегралами от спектра с множителем ω в определенных степенях по интервалу $\omega \in [0, \infty)$. Таким образом, важным является вопрос о консервативности предложенных вычислительных схем. В расчетах с применением описанного алгоритма показано, что указанные величины сохраняются с высокой точностью для спектров, локализованных внутри расчетного отрезка $[0, \omega_{\max}]$. Если в точке $\omega = \omega_{\max}$ спектр становится ненулевым, то возникает “утечка” сохраняющихся величин за границу ω_{\max} , обусловленная взаимодействием волн из областей $\omega < \omega_{\max}$ и $\omega > \omega_{\max}$. Эти взаимодействия задаются интегралами по подобластям $\Delta_1, \Delta_2, \Delta_4$ на иг. 1а, поэтому исключение этих подобластей из области Δ_ω при расчете интеграла столкновений приводит к запрету таких взаимодействий. Применение этого искусственного приема обеспечивает консервативность расчетов, однако является спорным с точки зрения физики. С одной стороны, запрет расширения спектра в область высоких частот позволяет избежать так называемой ультрафиолетовой катастрофы, с другой стороны, свободная эволюция спектра является важным аспектом адекватного учета свойств исходных динамических моделей. Здесь еще более проявляется актуальность численного анализа разномасштабных задач, поскольку для разрешения указанной проблемы необходимо существенно увеличивать ω_{\max} .

СПИСОК ЛИТЕРАТУРЫ

1. *Richardson L. F.* Weather Prediction by Numerical Processes. Boston: Cambridge University Press. 1922.
2. *Kolmogorov A. N.* The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers // Dokl. Akad. Nauk. SSSR. 1941. V. 30. N 9. P. 301–304.
3. *Obukhov A. M.* On the distribution of energy in the spectrum of turbulent flow // Bull. Acad. Sci. USSR, Geog. Geophys. 1941. V. 5. N 4. P. 453–466.

4. *Nazarenko S.* Wave Turbulence. Heidelberg, Germany: Springer, 2012.
5. *Zakharov V. E., L'vov V.S., Falkovich G.* Kolmogorov Spectra of Turbulence I: Wave Turbulence. Germany: Springer, 1992.
6. *Zakharov V. E., Musher S. L., Rubenchik A. M.* Hamiltonian approach to the description of non-linear plasma phenomena // *Physics Reports*. 1985. V. 129. N 5. P. 285–366.
7. *Dyachenko S., Newell A. C., Pushkarev A., Zakharov V. E.* Optical turbulence: weak turbulence, condensates and collapsing filaments in the nonlinear Schrödinger equation // *Physica D: Nonlinear Phenomena*. 1992. V. 57. N 1–2. P. 96–160.
8. *Zhu Y., Semisalov B. V., Krstulovic G., Nazarenko S. V.* Testing wave turbulence theory for the Gross-Pitaevskii system // *Phys. Rev. E* 2022. V. 106. Art #014205.
9. *Krasitskii V. P.* On the canonical transformation of the theory of weakly nonlinear waves with nondecay dispersion law // *Sov. Phys. JETP*. 1990. V. 98. P. 1644–1655.
10. *Semikoz D. V., Tkachev I. I.* Kinetics of Bose Condensation // *Phys. Rev. Lett.* 1995. V. 74. N 16. P. 3093–3097.
11. *Pitaevskii L. P., Stringari S.* Bose–Einstein Condensation and Superfluidity. Oxford University Press, 2016. 554 p.
12. *Maloney J. V., Newell A. C.* Nonlinear Optics. 1st Ed. Boca Raton: CRC Press. 2019.
13. *Zurek W. H.* Cosmological experiments in superfluid helium? // *Nature*. 1985. V. 317. P. 505–508.
14. *Zakharov V. E.* Statistical theory of gravity and capillary waves on the surface of a finite-depth fluid // *European Journal of Mechanics — B/Fluids*. 1999. V. 18. N 3. P. 327–344.
15. *Pushkarev A., Resio D., Zakharov V.* Weak turbulent approach to the wind-generated gravity sea waves // *Physica D*. 2003. V. 184. P. 29–63.
16. *Zhu Y., Semisalov B. V., Krstulovic G., Nazarenko S. V.* Direct and Inverse Cascades in Turbulent Bose-Einstein Condensates // *Phys. Rev. Lett.* 2023. V. 130. N 13. Art # 133001.
17. *Semisalov B. V., Grebenev V. N., Medvedev S. B., Nazarenko S. V.* Numerical analysis of a self-similar turbulent flow in Bose–Einstein condensates // *Communications in Nonlinear Science and Numerical Simulation*. 2021. V. 102. Art. #105903.
18. *Lacaze R., Lallemand P., Pomeau Y., Rica S.* Dynamical formation of a Bose–Einstein condensate // *Physica D. Nonlinear Phenomena*. 2001. V. 152. P. 779–786.
19. *Connaughton C., Pomeau Y.* Kinetic theory and Bose–Einstein condensation // *Comptes Rendus Physique*. 2004. V. 5. N. 1. P. 91–106.
20. *Hossain M. A., Islam Md. S.* Generalized Composite Numerical Integration Rule Over a Polygon Using Gaussian Quadrature // *Dhaka Univ. J. Sci.* 2014. V. 62. N 1. P. 25–29.
21. *Суетин П. К.* Классические ортогональные многочлены. 3-е изд., перераб. и доп. М.: Физматлит, 2005. 480 с.
22. *Szegö G.* Orthogonal polynomials. N.Y.: Amer. Math. Soc., 1959.
23. *Davis P. J., Rabinowitz P.* Methods of Numerical Integration. 2nd Ed. California: Academic Press, INC, 1984. 626 p.
24. *Trefethen L. N.* Spectral methods in MATLAB. Philadelphia: SIAM, 2000, 165 p.
25. *Бабенко К. И.* Основы численного анализа. М.: Физматлит, 1986. 744 с.
26. *Trefethen L. N.* Approximation Theory and Approximation Practice. Philadelphia: SIAM, 2018. 371 p.
27. *Гавриков М. Б.* Методы без насыщения в вычислительной математике // *Препринты ИПМ им. М. В. Келдыша*. 2019. № 075. 40 с.
28. *Lubinsky D. S., Rabinowitz P.* Rates of Convergence of Gaussian Quadrature for Singular Integrands // *Mathematics of Computation*. 1984. V. 43, N 167. P. 219–242.
29. *Mori M.* An IMT-Type Double Exponential Formula for Numerical Integration // *Publ. RIMS, Kyoto Univ.* 1978. V. 14. P. 713–729.
30. *Tee T. W., Trefethen L. N.* A rational spectral collocation method with adaptively transformed Chebyshev grid points // *SIAM J. Sci. Comput.* 2006. V. 28. N 5. P. 1798–1811.
31. *Семисалов Б. В., Кузьмин Г. А.* К вопросу о приближении гладких функций с погранслойными составляющими // *Труды УрО РАН*. 2021. Т. 27. С. 111–124.
32. *Iri M., Moriguti S., Takasawa Y.* On a certain quadrature formula // *RIMS Kokyuroku Kyoto Univ.* 1970. V. 91. P. 82–118 (in Japanese).
33. *Zakharov V. E., Filonenko N.* The energy spectrum for stochastic oscillation of a fluid's surface // *Dokl. Akad. Nauk*. 1966. V. 170. P. 1292–1295.

34. *Salzer H. E.* Lagrangian interpolation at the Chebyshev points $x_{n,v} = \cos(v\pi / n)$, $v = O(1)n$; some unnoted advantages // *Computer J.* 1972. V. 15. P. 156–159.
35. *Baltensperger R., Berrut J.-P., Nl B.* Exponential convergence of a linear rational interpolant between transformed Chebyshev points // *Math. Comp.* 1999. V. 68. P. 1109–1120.
36. *Berrut J.-P., Mittelmann H. D.* Adaptive point shifts in rational approximation with optimized denominator // *J. of comp. and appl. math.* 2004. V. 164–165. P. 81–92.
37. *Jafari-Varzaneh H.A., Hosseini S. M.* A new map for the Chebyshev pseudospectral solution of differential equations with large gradients // *Numerical Algorithms.* 2015. V. 69. P. 95–108.
38. *Идимешев С. В.* Дробно-рациональная аппроксимация в начально-краевых задачах с фронтами // *Вычислительные технологии.* 2020. Т. 25. № 2. С. 63–79.
39. *Huang W., Sloan D. M.* The Pseudospectral Method for Solving Differential Eigenvalue Problems // *J. Comp. Phys.* 1994. V. 111. N 2. P. 399–409.
40. *Семисалов Б. В.* Об одном подходе к численному решению задач Дирихле произвольной размерности // *СибЖВМ.* 2022. Т. 25. № 1. С. 77–95.
41. *Rump S. M.* Verification Methods: Rigorous Results using Floating-point Arithmetic // *Acta Numerica.* 2010. V. 19. P. 287–449. Rump S. M. INTLAB
42. INTerval LABoratory. In *Developments in Reliable Computing.* Dordrecht: Kluwer Academic Publishers, 1999. P. 77–104.
43. *Бибердорф Э. А., Блинова М. А., Попова Н. И.* Модификации метода дихотомии матричного спектра и их применение к задачам устойчивости // *СибЖВМ.* 2022. Т. 21. № 2. С. 139–154.
44. *A. M. Blokhin, A. S. Ibragimova* Numerical Method for 2D Simulation of a Silicon MESFET with a Hydrodynamical Model Based on the Maximum Entropy Principle // *SIAM Journal on Scientific Computing.* 2009. V. 31. P. 2015–2046.
45. *Connaughton C., Josserand C., Picozzi A., Pomeau Y., Rica S.* Condensation of Classical N linear Waves // *Phys. Rev. Lett.* 2005. V. 95. Art. #263901.
46. *Bell N. K., Nazarenko S. V.* Reflected wave solution of Alfvén wave turbulence // *J. of Physics A: Mathematical and Theoretical.* 2018. V. 51. N 40. Art. # 405501.
47. *Nazarenko S. V., Grebenev V. N., Medvedev S. B., Galtier S.* The focusing problem for the Leith model of turbulence: a self-similar solution of the third kind // *Journal of Physics A: Mathematical and Theoretical.* 2019. V. 52. N 15. Art. # 155501.
48. *Escobedo M., Velázquez J. J. L.* On the theory of Weak Turbulence for the N linear Schrödinger Equation // *Memoirs of the American Mathematical Society.* 2015. V. 238. N 1124.
49. *Васкевич В. Л., Щербаков А. И.* Сходимость последовательных приближений в задаче Коши для интегродифференциального уравнения с квадратичной нелинейностью // *Матем. труды.* 2018. Т. 21. № 2. С. 136–149.
50. *Lanford O. E.* Time evolution of large classical systems. Chapter in book: Moser J. (eds) *Dynamical Systems, Theory and Applications.* Lecture Notes in Physics, vol 38. Berlin, Heidelberg: Springer. 1975.
51. *Deng Y., Hani Z.* Full derivation of the wave kinetic equation // *Inventiones mathematicae.* 2023. <https://doi.org/10.1007/s00222-023-01189-2>