

Том 63, Номер 5

ISSN 0044-4669

Май 2023



ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ

www.sciencejournals.ru



СОДЕРЖАНИЕ

Том 63, номер 5, 2023 год

ОБЩИЕ ЧИСЛЕННЫЕ МЕТОДЫ

- Вычисление условия сильного резонанса в системе Гамильтона
А. Б. Батхин, З. Х. Хайдаров 697
- A Generalized Simplified Hermitian and Skew-Hermitian Splitting Preconditioner
for Double Saddle Point Problems
L. Meng, Y. W. He, J. Li 715
- On the Radial Basis Function Interpolation I: Spectral Analysis of the Interpolation
Matrix and the Related Operators
Jianping Xiao 716
- Аппроксимация таблично заданных функций: многокритериальный подход
А. П. Нелюбин, В. В. Подиновский 717
-

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ

- Синтез оптимальной системы с устойчивыми режимами скольжения
Л. Т. Ащепков 731
- Оптимизация множества достижимости линейной системы по отношению
к другому множеству
М. В. Балашов, Р. А. Камалов 739
-

ОБЫКНОВЕННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

- On Ranks of Matrices Over Noncommutative Domains
S. A. Abramov, M. Petkovšek, A. A. Ryabenko 760
-

УРАВНЕНИЯ В ЧАСТНЫХ ПРОИЗВОДНЫХ

- Alternative Direction Implicit Method for Solving First Order 2D Hyperbolic
Delay Differential Equations
S. Karthick, V. Subburayan 763
- Multiwave Interaction Solutions for a New Extended Equation in $(4 + 1)$ -Dimension
Yan Yang, Yinping Liu 764
-

МАТЕМАТИЧЕСКАЯ ФИЗИКА

- Восстановление двух функций в модели колебаний струны,
один конец которой помещен в подвижную среду
О. А. Андреева, А. Ю. Щеглов 765
- Метод моделирования турбулентного пограничного слоя на основе
аналитических законов стенки в формулировке метода
характеристических штрафных функций
О. В. Васильев, Н. С. Жданова 778
- Приближенное решение обратной задачи для интегродифференциального
уравнения теплопроводности с сингулярным возмущением
А. М. Денисов 795

Усвоение данных для двумерного уравнения амбиполярной диффузии в модели ионосферы земли <i>В. П. Дымников, Д. В. Кулямин, П. А. Останин, В. П. Шутяев</i>	803
Волноводная модель развитого турбулентного пограничного слоя <i>В. А. Жаров, И. И. Липатов, Р. С. Селим</i>	827
Метод квазирешений и проблема глобальной минимизации функционала невязки условно корректных обратных задач <i>М. Ю. Кокурин</i>	840
Задача сложного теплообмена с условиями типа Коши на части границы <i>П. Р. Месенев, А. Ю. Чеботарев</i>	856
О схеме Русанова третьего порядка точности для моделирования плазменных колебаний <i>Е. В. Чижонков</i>	864

ОБЩИЕ ЧИСЛЕННЫЕ МЕТОДЫ

УДК 517.5+004.421.6

ВЫЧИСЛЕНИЕ УСЛОВИЯ СИЛЬНОГО РЕЗОНАНСА В СИСТЕМЕ ГАМИЛЬТОНА

© 2023 г. А. Б. Батхин^{1,2,*}, З. Х. Хайдаров^{3,**}

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

² 141701, Долгопрудный, М.о., Институтский переулок, 9, МФТИ, Россия

³ 140104 Самарканд, Университетский бул., 15, СамГУ им. Ш. Рашидова, Узбекистан

*e-mail: batkhin@gmail.com

**e-mail: zafarxx@gmail.com

Поступила в редакцию 20.11.2022 г.
Переработанный вариант 05.01.2023 г.
Принята к публикации 02.02.2023 г.

Для исследования областей формальной устойчивости положения равновесия многопараметрической системы Гамильтона с тремя степенями свободы в случае общего положения предложен способ символьного вычисления условия существования резонанса третьего и четвертого порядков. Это условие формулируется в виде нулей квазиоднородного полинома от коэффициентов характеристического многочлена линейной части системы Гамильтона. Методами компьютерной алгебры (базисы Грёбнера исключающих идеалов) и степенной геометрии (степенные преобразования) для различных резонансных векторов это условие представляется в виде рациональных алгебраических кривых, с помощью которых получено описание разбиения области устойчивости по линейному приближению в пространстве коэффициентов характеристического многочлена на такие части, где отсутствуют сильные резонансы. Приведен пример описания резонансных множеств для двупараметрической системы маятникового типа. Все вычисления выполнены в системе компьютерной алгебры Maple. Библ. 20. Фиг. 3.

Ключевые слова: система Гамильтона, положение равновесия, нормальная форма, формальная устойчивость, резонансное условие, исключающий идеал.

DOI: 10.31857/S0044466923050071, **EDN:** PJRPGA

1. ВВЕДЕНИЕ

В колебательных системах резонансы играют существенную роль. Их присутствие, с одной стороны, приводит к появлению сложной динамики, когда энергия колебаний “перекачивается” между теми степенями свободы, чьи соответствующие частоты находятся в резонансе. С другой стороны, наличие нетривиальных решений резонансного уравнения позволяет получить дополнительные формальные первые интегралы и, как следствие, позволяет провести анализ устойчивости положения равновесия или асимптотически проинтегрировать систему уравнений движения, приведенную к нормальной форме.

Условно можно указать три основных типа устойчивости для теоретико-механических задач (см. [1]):

- строгая устойчивость (устойчивость по Ляпунову);
- формальная устойчивость (устойчивость по Мозеру);
- практическая устойчивость.

Устойчивость по Ляпунову является наиболее строгой и гарантирует равномерную ограниченность решений на бесконечном интервале времени относительно множества возмущений по начальным условиям и параметрам. Устойчивость по Мозеру слабее, но гарантирует скорость разбегания траекторий медленнее, чем любая степенная функция с произвольным положительным показателем. Практическая устойчивость означает только ограниченность решения на конечном интервале времени относительно набора малых возмущений по начальным условиям или параметрам системы.

Далее рассматриваем автономную гамильтонову систему с аналитической функцией $H(\mathbf{z}; \mathbf{P})$, положение равновесия (ПР) которой совпадает с началом координат. Тогда гамильтониан $H(\mathbf{z}; \mathbf{P})$ раскладывается в сходящийся ряд однородных полиномов H_k степени k от своих фазовых переменных $\mathbf{z} = (\mathbf{x}, \mathbf{y})$:

$$H(\mathbf{z}; \mathbf{P}) = \sum_{k=2}^{\infty} H_k(\mathbf{z}; \mathbf{P}), \quad (1)$$

где \mathbf{P} – вектор параметров.

Известно, что устойчивость ПР в первом приближении можно определить только для случая, когда квадратичная форма $H_2(\mathbf{z})$ знакоопределенная (теорема Лагранжа–Дирихле, см. [2]).

Если число степеней свободы не более двух, то

– устойчивость определяется теоремой Арнольда–Мозера (см. [3], гл. 4, п. 1) при отсутствии резонансов порядка четыре и меньше, что требует вычисления нормальной формы гамильтониана (1) до четвертого порядка;

– для резонансов порядка менее четырех условия устойчивости были получены в работах Маркеева и Сокольского (см. [3], гл. 4).

Когда число степеней свободы больше двух, устойчивость для большинства начальных условий определяется теоремой Арнольда (см., например, [3], гл. 5, п. 1).

С практической точки зрения вполне достаточна более слабая, чем устойчивость по Ляпунову, формальная устойчивость, предложенная Мозером (см. [4]).

Цель работы состоит в описании схемы исследования формальной устойчивости ПР системы Гамильтона с тремя степенями свободы, а также в получении явного представления условий существования резонансов кратности 1 в терминах коэффициентов характеристического многочлена линейной части гамильтоновой системы. Предварительные результаты работы были опубликованы в препринте авторов (см. [5]). В настоящей статье предъявлено в явном виде условие существования резонанса с резонансным вектором $\mathbf{p}^* = (q, 1, 1)$, $q \in \mathbb{Z}$, а также исправлены замеченные опечатки и устранены неточности.

Будем использовать следующие обозначения:

– полужирные символы типа \mathbf{x} , \mathbf{y} , \mathbf{u} , \mathbf{v} – столбцы-векторы в n -мерных вещественных \mathbb{R}^n или комплексных \mathbb{C}^n пространствах;

– полужирные символы типа \mathbf{p} , \mathbf{q} – векторы в n -мерной целочисленной решетке \mathbb{Z}^n ;

– $|\mathbf{p}| = \sum_{j=1}^n |p_j|$ – норма вектора;

– для $\mathbf{x} = (x_1, \dots, x_n)^\top$ и $\mathbf{p} = (p_1, \dots, p_n)^\top$ обозначим через $\mathbf{x}^{\mathbf{p}} \equiv \prod_{j=1}^n x_j^{p_j}$ моном и через $\langle \mathbf{p}, \mathbf{x} \rangle \equiv \sum_{j=1}^n p_j x_j$ скалярное произведение пары векторов.

2. МНОЖЕСТВО УСТОЙЧИВОСТИ ЛИНЕЙНОЙ СИСТЕМЫ ГАМИЛЬТОНА

В случае общего положения ряд (1) начинается с квадратичного гамильтониана $H_2(\mathbf{z}; \mathbf{P})$, определяющего локальную динамику вблизи ПР. Поведение фазового потока в первом приближении описывается линейной гамильтоновой системой

$$\dot{\mathbf{z}} = B(\mathbf{P})\mathbf{z}, \quad B(\mathbf{P}) = \frac{1}{2} J \frac{\partial^2 H_2(\mathbf{P})}{\partial \mathbf{z}^2}. \quad (2)$$

Напомним здесь основные свойства линейной гамильтоновой системы.

1. Если λ_j есть собственное число (СЧ) матрицы B , то $-\lambda_j$ также является ее СЧ (см. [2]). Все СЧ λ_j , $j = 1, \dots, 2n$, матрицы B могут быть упорядочены таким образом, что $\lambda_{j+n} = -\lambda_j$, $j = 1, \dots, n$. Обозначим через вектор $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n)$ множество базисных собственных значений.

2. Характеристический многочлен $\check{f}(\lambda)$ матрицы B содержит только четные степени λ , поэтому он является многочленом от $\mu = \lambda^2$. Такой многочлен назван в [6] *полухарактеристическим*:

$$f(\mu) = \sum_{k=0}^n a_{n-k}(\mathbf{P})\mu^k, \quad f_0 \equiv 1. \quad (3)$$

3. Если $\text{Re } \lambda_j \neq 0$ для некоторого j , то ПР неустойчиво.

4. Если все $\text{Re } \lambda_j = 0$, то поведение фазового потока в его окрестности может быть получено только при учете нелинейных членов.

Существует два типа задач об устойчивости многопараметрических систем:

- для определенных значений вектора параметров \mathbf{P} выяснить устойчивость ПР (частная задача);
- найти в пространстве параметров Π все значения \mathbf{P} , для которых ПР $\mathbf{z} = 0$ системы (1) устойчива, т.е. вычислить так называемое множество устойчивости Σ системы (1) (общая задача).

Здесь рассматривается схема решения общей задачи.

Определение 1. Множество устойчивости Σ линейной системы (2) – это множество всех значений параметров $\mathbf{P} \in \Pi$, для которых ПР $\mathbf{z} = 0$ устойчиво по Ляпунову.

В терминах корней многочлена (3) условие устойчивости ПР дается следующей теоремой.

Теорема 1 (см. [6]). *Положение равновесия $\mathbf{z} = 0$ линейной гамильтоновой системы (2) устойчиво по Ляпунову тогда и только тогда, когда*

- а) все корни μ_k многочлена (3) вещественны и неположительны;
- б) все элементарные делители матрицы B просты.

Невыполнение условия а) приводит к экспоненциальному разбеганию решений, которое не перекрывается нелинейными добавками, привносящими только степенной эффект в поведение решений. Невыполнение условия б) приводит к степенной неустойчивости, которая может быть перекрыта нелинейными добавками.

Условие вещественности и неположительности корней многочлена $f(\mu)$ определяется следующей теоремой.

Теорема 2. *Для того чтобы все корни многочлена $f(\mu)$ степени n были вещественны и неположительны, необходимо и достаточно выполнение условий*

$$a_j(\mathbf{P}) \geq 0, \quad j = 1, \dots, n, \quad D^{(k)}(f) \geq 0, \quad k = 0, \dots, n - 2,$$

где $D^{(k)}(f)$ есть k -й субдискриминант многочлена $f(\mu)$ (см. [7], [8]).

3. НОРМАЛЬНАЯ ФОРМА СИСТЕМЫ ГАМИЛЬТОНА

В дальнейшем считаем, что выполнено условие а) устойчивости теоремы 1. Если не выполнено условие б) теоремы 1, то устойчивость определяется по нормальной форме (НФ) общими методами.

Согласно [9, теорема 12] существует каноническое формальное преобразование в виде степенного ряда, которое приводит исходную систему Гамильтона к ее *нормальной форме*

$$\dot{\mathbf{u}} = \frac{\partial h}{\partial \mathbf{v}}, \quad \dot{\mathbf{v}} = -\frac{\partial h}{\partial \mathbf{u}},$$

задаваемой нормализованным гамильтонианом $h(\mathbf{u}, \mathbf{v})$:

$$h(\mathbf{u}, \mathbf{v}) = \sum_{j=1}^n \sigma_j \lambda_j u_j v_j + \sum h_{pq} \mathbf{u}^p \mathbf{v}^q, \quad \sigma_j = \pm 1, \quad (4)$$

который содержит только резонансные члены $h_{pq} \mathbf{u}^p \mathbf{v}^q$, удовлетворяющие условию

$$\langle \mathbf{p} - \mathbf{q}, \lambda \rangle = 0. \quad (5)$$

Здесь $0 \leq \mathbf{p}, \mathbf{q} \in \mathbb{Z}^n$, $|\mathbf{p}| + |\mathbf{q}| \geq 2$ и $h_{\mathbf{p}\mathbf{q}}$ – постоянные коэффициенты.

Резонансное уравнение (5) имеет два вида решений, которым соответствуют два вида резонансных членов в НФ (4):

- *вековые члены* вида $h_{\mathbf{p}\mathbf{p}} \mathbf{u}^{\mathbf{p}} \mathbf{v}^{\mathbf{p}}$, которые всегда присутствуют в гамильтоновой нормальной форме из-за особой структуры матрицы B линейной части системы (2); вековые члены являются модами только четных степеней от фазовых переменных и входят в соответствующие однородные формы;
- *строго резонансные члены*, которые соответствуют нетривиальным целочисленным решениям уравнения

$$\langle \mathbf{p}, \lambda \rangle = 0. \quad (6)$$

Сама процедура нормализации обычно выполняется в комплексных переменных. Для перехода к комплексным переменным используется линейное преобразование $\Phi : (\mathbf{x}, \mathbf{y}) \rightarrow (\zeta, \bar{\zeta})$, которое преобразует исходный гамильтониан к виду

$$h(\zeta, \bar{\zeta}) = \sum h_{\mathbf{p}\mathbf{q}} \zeta^{\mathbf{p}} \bar{\zeta}^{\mathbf{q}}, \quad (7)$$

где $\mathbf{p}, \mathbf{q} \in \mathbb{Z}^n$ и $|\mathbf{p}| + |\mathbf{q}| \geq 2$. Значение $|\mathbf{p}| + |\mathbf{q}|$ называется *порядком* соответствующего члена разложения.

Определение 2. Функция Гамильтона $h(\zeta, \bar{\zeta})$ называется *комплексной нормальной формой* вещественной системы Гамильтона для случая полупростых СЧ, если 1) его квадратичная часть h_2 имеет вид

$$h_2 = \sum_{j=1}^n \sigma_j \lambda_j \zeta_j \bar{\zeta}_j, \quad \sigma_j = \pm 1;$$

2) разложение (7) содержит только члены $h_{\mathbf{p}\mathbf{q}} \zeta^{\mathbf{p}} \bar{\zeta}^{\mathbf{q}}$, которые удовлетворяют резонансному уравнению (5). Константы σ_j являются инвариантами НФ.

Если резонансы отсутствуют, то имеется НФ Биркгофа (см. [10])

$$h(\zeta, \bar{\zeta}) = \sum_{k=1}^{\infty} h_{2k}(\zeta, \bar{\zeta}),$$

состоящая из однородных форм четных степеней $2k$, зависящих только от переменных вида $\zeta_j \bar{\zeta}_j$, $j = 1, \dots, n$, при этом каждая из величин $\zeta_j \bar{\zeta}_j$ является формальным первым интегралом. В переменных действие–угол (φ, ρ) НФ Биркгофа может быть записана в виде $h(\rho) = \sum_{k=1}^{\infty} h_k(\rho)$, где $\rho_j = \zeta_j \bar{\zeta}_j$, $j = 1, \dots, n$.

В этом случае система интегрируема, но преобразование Биркгофа обычно расходящееся, а в случае гладкой зависимости от параметров \mathbf{P} резонансные значения располагаются всюду плотно в пространстве параметров. Таким образом, сколь угодно малые изменения параметров приводят к появлению резонансных членов в НФ.

Дальнейшие результаты связаны с существованием резонансов в системе Гамильтона, поэтому напомним их определение и укажем условие, с помощью которого последующие результаты проще формулируются.

Определение 3 (см. [11], гл. I, п. 3). *Кратность резонанса* \mathfrak{k} – это число линейно независимых решений $\mathbf{p} \in \mathbb{Z}^n$ резонансного уравнения $\langle \mathbf{p}, \lambda \rangle = 0$. *Порядок резонанса* равен $q = \min |\mathbf{p}|$ по $\mathbf{p} \in \mathbb{Z}^n$, $\mathbf{p} \neq 0$, $\langle \mathbf{p}, \lambda \rangle = 0$. Если решение резонансного уравнения содержит только два собственных значения, то такой резонанс называется *двухчастотным резонансом*, если более двух – то *многочастотным резонансом*. Резонансы порядков 2, 3 и 4 назовем *сильными*, бóльших порядков – *слабыми* резонансами.

Условие A_k^n (см. [12]). Будем говорить, что для нормализованной системы Гамильтона выполняется условие A_k^n , если резонансное уравнение (6) не имеет целочисленных решений \mathbf{p} с $|\mathbf{p}| \leq k$.

4. СХЕМА ИССЛЕДОВАНИЯ ФОРМАЛЬНОЙ УСТОЙЧИВОСТИ

Определение 4. Положение равновесия $\mathbf{z} = 0$ системы с функцией Гамильтона $H(\mathbf{z})$ формально устойчиво, если существует возможно расходящийся степенной ряд $G(\mathbf{z})$, который является формальным положительно-определенным первым интегралом $\{G, H\} = 0$.

В [12] было дано схематическое описание метода изучения формальной устойчивости ПР. Этот метод основан на следующих ключевых результатах:

- вычисляется НФ системы Гамильтона в окрестности ПР,
- применяется теорема Брюно (см. [13]) о формальной устойчивости,
- используются q -аналоги объектов классической теории исключений.

При этом были сделаны следующие предположения:

- число степеней свободы системы больше двух,
- квадратичная форма $H_2(\mathbf{z})$ в разложении (1) невырожденна и не является знакоопределенной,
- функция Гамильтона $H(\mathbf{z}; \mathbf{P})$ гладко зависит от вектора параметров \mathbf{P} .

Пусть имеет место условие A_4^n , т.е. $\langle \mathbf{L}, \boldsymbol{\lambda} \rangle \neq 0$ для $\mathbf{L} \in \mathbb{Z}^n$, $0 < |\mathbf{L}| \leq 4$, тогда существует аналитическое каноническое преобразование $(\mathbf{x}, \mathbf{y}) \rightarrow (\boldsymbol{\rho}, \boldsymbol{\varphi})$ такое, что новый гамильтониан g имеет вид

$$g(\boldsymbol{\rho}, \boldsymbol{\varphi}) = g_1(\boldsymbol{\rho}) + g_2(\boldsymbol{\rho}) + r(\boldsymbol{\rho}, \boldsymbol{\varphi}),$$

где $g_1(\boldsymbol{\rho}) = \langle \boldsymbol{\lambda}, \boldsymbol{\rho} \rangle$, $g_2(\boldsymbol{\rho}) = \langle C\boldsymbol{\rho}, \boldsymbol{\rho} \rangle$, $C = [c_{ij}]_{i,j=1}^n$, и $r(\boldsymbol{\rho}, \boldsymbol{\varphi})$ является сходящимся степенным рядом переменных $(\boldsymbol{\rho}, \boldsymbol{\varphi})$ степени три или выше в $\boldsymbol{\rho}$.

При отсутствии сильных резонансов между собственными значениями линейной части гамильтоновой системы в окрестности ПР условие ее формальной устойчивости определяется следующей теоремой.

Теорема 3 (см. [13]). *Если условие A_4^n выполнено и для любых ненулевых целых векторов \mathbf{L} , которые являются решением уравнения*

$$\langle \mathbf{L}, \boldsymbol{\lambda} \rangle = 0,$$

квадратичная форма $\langle C\mathbf{L}, \mathbf{L} \rangle \neq 0$ при $\boldsymbol{\lambda} \neq 0$, то ПР $\mathbf{z} = 0$ гамильтоновой системы формально устойчиво.

Таким образом, для применения теоремы 3 о формальной устойчивости необходимо найти границы областей в пространстве параметров \mathbf{P} , определяемых резонансными многообразиями (понятие резонансного многообразия дано ниже), соответствующих сильным резонансам.

Определение 5. Резонансным многообразием $\mathcal{R}_n^{\mathbf{P}}$ в пространстве \mathbf{K} коэффициентов a_1, \dots, a_n полухарактеристического многочлена $f_n(\mu)$ степени n назовем такое алгебраическое многообразие, на котором вектор базовых собственных значений $\boldsymbol{\lambda}$ соответствующего характеристического многочлена $\check{f}(\boldsymbol{\lambda})$ является нетривиальным решением резонансного уравнения (6) для фиксированного целочисленного вектора \mathbf{p} . Аналитическое представление многообразия $\mathcal{R}_n^{\mathbf{P}}$ в неявной или параметрической формах далее обозначим через $R_n^{\mathbf{P}}$.

Постановка задачи. Для исследования формальной устойчивости ПР гамильтоновой системы необходимо в пространстве параметров \mathbf{P} найти множество устойчивости Σ линеаризованной системы, в ней определить области, в которых квадратичная форма $H_2(\mathbf{z})$ не является знакоопределенной; в найденных областях выделить их части S_k , в которых отсутствуют сильные резонансы, а затем в каждой из таких частей S_k выполнить процедуру нормализации гамильтониана до четвертого порядка включительно и применить теорему 3. Для этого будет достаточно выбрать какую-либо точку в каждой из частей S_k в пространстве параметров и воспользоваться одним из алгоритмов нормализации функции Гамильтона. Поскольку в каждой внутренней точке части S_k все собственные числа λ_k , $k = 1, \dots, n$, простые, то в этом случае легко применим алгоритм инвариантной нормализации (см. [14]).

В настоящей работе рассматривается описание резонансных многообразий порядков 2, 3 и 4 в пространстве коэффициентов \mathbf{K} полухарактеристического многочлена $f(\mu)$ для системы Гамильтона с тремя степенями свободы.

Основная задача — для многопараметрической системы Гамильтона с тремя степенями свободы дать описание областей в пространстве параметров системы, в которых отсутствуют сильные резонансы порядков 2, 3 и 4.

Рассмотрим более подробно, при каких условиях реализуются резонансы указанных выше порядков:

- порядок $q = 2$: $p = (1, 1, 0)$ — это случай кратных корней, который описывается дискриминантным множеством $R_3^{(1,1,0)} \equiv D(f) = 0$;
- порядок $q = 3$: для двухчастотного случая $p = (2, 1, 0)$, описывается q -дискриминантом $R_3^{(2,1,0)} \equiv D_4(f) = 0$;
- порядок $q = 4$: для двухчастотного случая $p = (3, 1, 0)$, описывается q -дискриминантом $R_3^{(3,1,0)} \equiv D_9(f) = 0$;
- в трехчастотном случае: для порядка 3 описывается условием $R_3^{(1,1,1)} = 0$, а для порядка 4 — условием $R_3^{(2,1,1)} = 0$.

Для решения этой задачи следует получить описание границ областей, свободных от сильных резонансов. Эти границы состоят из участков алгебраических многообразий, на которых резонансное уравнение (6) имеет нетривиальное решение.

Основную задачу разобьем на две вспомогательных: 1) получить аналитическое представление R_3^p в пространстве коэффициентов $\mathbf{K} = (a_1, a_2, a_3)$ кубического многочлена резонансных многообразий R_3^p для всех векторов \mathbf{p} порядков 2, 3 и 4;

2) выяснить взаимное расположение всех найденных выше резонансных многообразий, т.е. определить, каким образом указанные выше резонансные многообразия касаются или пересекаются в пространстве \mathbf{K} .

5. УСЛОВИЯ СУЩЕСТВОВАНИЯ СИЛЬНЫХ РЕЗОНАНСОВ

Общее описание процедуры получения условия существования двухчастотного и многочастотного резонансов выглядит следующим образом.

1. Для некоторого вектора $\mathbf{p}^* \in \mathbb{Z}_n$, удовлетворяющего резонансному уравнению $\langle \mathbf{p}^*, \boldsymbol{\lambda} \rangle = 0$, составляется полиномиальный идеал

$$\mathcal{I} = \{ \langle \mathbf{p}^*, \boldsymbol{\lambda} \rangle, \lambda_j^2 - \mu_j, j = 1, \dots, n \}.$$

2. Вычисляется базис Грёбнера \mathcal{G} этого идеала с подходящим мономиальным порядком переменных $\lambda_j, \mu_j, j = 1, \dots, n$ (подробнее см. [15], гл. 2, 3), так, чтобы первый полином этого базиса содержал только переменные μ_j . Этот полином является квазиоднородным полиномом в переменных $\mu_j, j = 1, \dots, n$. Он определяет условие существования резонанса для заданного вектора \mathbf{p}^* .

3. Для получения соответствующего резонансного условия для коэффициентов $a_j, j = 1, \dots, n$, многочлена $f(\mu)$ строится новый исключаящий базис Грёбнера \mathcal{F} идеала \mathcal{I} , содержащего полученное в п. 2 условие для μ_j и связи коэффициентов исходного полухарактеристического полинома с его корнями в виде элементарных симметрических полиномов третьей степени. При этом указывается порядок исключения переменных последовательно μ_j и $a_j, j = 1, \dots, n$. Первый полином вычисленного базиса \mathcal{F} , зависящий только от коэффициентов a_j , является условием существования резонанса, записанный в коэффициентах многочлена (3).

Описанная выше последовательность действий для нахождения условия существования резонансных соотношений реализована с помощью прикладной программы, а именно, использовалась система компьютерной алгебры (СКА) Maple. Для ее реализации применяются следующие пакеты и процедуры:

- Для вычисления и исследования идеалов \mathcal{G}, \mathcal{F} (см. выше пункты 1 и 3 соответственно) используется процедура Basis построения базисов Грёбнера для различных лексикографических порядков, а также некоторые дополнительные процедуры пакета Groebner.

- Для определения рода кривой использовалась команда genus, а для нахождения параметризации – parametrization из пакета algcurves.

Эта реализация апробирована на модельном примере, описанном в разд. 6.

5.1. Вычисление условия двухчастотного резонанса

Рассмотрим случай двухчастотного резонанса $\mathbf{p}^* = (q, 1, 0)$, где $q \in \mathbb{N}$. Здесь имеется пара соизмеримых СЧ, а третья СЧ несоизмеримо ни с каким из них. В пространстве параметров такой резонанс описывается в терминах резонансного множества $\mathcal{R}_{q^2}(f)$ полухарактеристического многочлена $f(\mu)$ (см. [16]).

Вычислим по выше указанной последовательности условие двухчастотного резонанса для $\mathbf{p}^* = (q, 1, 0)$. Для начала составим идеал \mathcal{J} , содержащий соотношения зависимости между корнями характеристического и полухарактеристического многочленов вида $\lambda_j^2 - \mu_j$, $j = 1, 2, 3$, а также резонансное соотношение $q\lambda_1 + \lambda_2$. Первый многочлен исключаящего базиса Грёбнера \mathcal{G} этого идеала с последовательным порядком исключения переменных λ_j, μ_j , $j = 1, 2, 3$, есть многочлен

$$q^2\mu_1 - \mu_2. \tag{8}$$

Равенство нулю данного многочлена дает условие на корни полухарактеристического полинома. Согласно п. 3, чтобы получить условие на коэффициенты, составляем новый идеал \mathcal{J} , включающий в себя полученное условие (8) и их связь с коэффициентами a_j многочлена (3) через элементарные симметрические многочлены

$$\mu_1 + \mu_2 + \mu_3 + a_1, \quad \mu_1\mu_2 + \mu_1\mu_3 + \mu_2\mu_3 - a_2, \quad \mu_1\mu_2\mu_3 + a_3. \tag{9}$$

Для идеала \mathcal{J} вычисляем исключаящий базис Грёбнера \mathcal{F} с соответствующим порядком исключения переменных μ_j, a_j , $j = 1, 2, 3$. Равенство нулю первого его многочлена:

$$\mathcal{R}_3^{(q,1,0)} \stackrel{\text{def}}{=} \left\{ (q^4 + q^2 + 1)^3 a_3^2 + q^4 (q^2 + 1)^2 a_2^3 + q^4 (q^2 + 1)^2 a_1^3 a_3 - q^6 a_1^2 a_2^2 - \right. \\ \left. - q^2 (q^4 + q^2 + 1)(q^4 + 4q^2 + 1) a_1 a_2 a_3 = 0 \right\}, \tag{10}$$

зависящего только от a_j , является условием существования двухчастотного резонанса в общем виде для некоторого натурального значения q .

Проверим полученный результат, сравнивая с полученными ранее условиями для случая, когда $q = 1, 2, 3$. Если в полученных в [17] формулах обобщенных субдискриминантов для f_3 положить $\omega = 0$, а $q = q^2$, то получим совпадающие с точностью до знака выражения для соответствующих резонансных многообразий:

- при $q = 1$ условие принимает вид

$$R_3^{(1,1,0)} \stackrel{\text{def}}{=} 4a_1^3 a_3 - a_1^2 a_2^2 - 18a_1 a_2 a_3 + 4a_2^3 + 27a_3^2 = 0,$$

- при $q = 2$ –

$$R_3^{(2,1,0)} \stackrel{\text{def}}{=} 400a_1^3 a_3 - 64a_1^2 a_2^2 - 2772a_1 a_2 a_3 + 400a_2^3 + 9261a_3^2 = 0,$$

- а при $q = 3$ –

$$R_3^{(3,1,0)} \stackrel{\text{def}}{=} 8100a_1^3 a_3 - 729a_1^2 a_2^2 - 96642a_1 a_2 a_3 + 8100a_2^3 + 753571a_3^2 = 0. \tag{11}$$

5.2. Вычисление условия трехчастотного резонанса

Рассмотрим для начала случай трехчастотного резонанса, для которого алгебраическая сумма трех собственных значений равна нулю, т.е. $\sum_{j=1}^3 p_j \lambda_j = 0$. Такой резонанс может иметь крат-

ность 1 или 2. Если кратность $\mathfrak{f} = 2$, то это означает, что имеет место попарная соизмеримость между базисными частотами λ_j характеристического многочлена. Такая ситуация исследуется с помощью условий существования двухчастотного резонанса. Далее рассматриваем только случай кратности 1.

Например, в случае базисных частот $\lambda_1 = 1 + \sqrt{2}$, $\lambda_2 = 1 - \sqrt{2}$, $\lambda_3 = -2$, СЧ попарно несоизмеримы, но их сумма равна нулю. Следовательно, резонансным вектором \mathbf{p}^* служит вектор $(1, 1, 1)$, что означает наличие трехчастотного резонанса кратности $\mathfrak{f} = 1$ порядка $q = 3$. Таким образом, будем рассматривать случай, когда взвешенная алгебраическая сумма всех трех базисных собственных чисел λ_j , $j = 1, 2, 3$, равна нулю для некоторого вектора $\mathbf{p}^* \in \mathbb{Z}^3$, т.е. $\langle \mathbf{p}^*, \lambda \rangle = 0$.

Для случая трехчастотного резонанса $\mathbf{p}^* = (1, 1, 1)$ аналогично, по выше указанной последовательности действий, вычислим условие существования резонанса. Составляем идеал, содержащий соотношения зависимости между корнями характеристического и полухарактеристического многочленов вида $\lambda_j^2 - \mu_j$, $j = 1, 2, 3$, а также резонансное соотношение $\lambda_1 + \lambda_2 + \lambda_3$. Первый многочлен базиса Грёбнера этого идеала, с последовательным порядком исключения переменных λ_j , μ_j , $j = 1, 2, 3$, есть многочлен

$$\mu_1^2 - 2\mu_1\mu_2 - 2\mu_3\mu_1 + \mu_2^2 - 2\mu_3\mu_2 + \mu_3^2.$$

Равенство нулю данного многочлена дает условие на корни полухарактеристического полинома. Для того чтобы получить условие на коэффициенты, составляем новый идеал, включающий в себя полученное условие на корни и их взаимосвязи с коэффициентами исходного полинома вида (9). Вычислив исключаящий базис Грёбнера для этого идеала, получим искомое условие, т.е. резонансное многообразие $\mathcal{R}_3^{(1,1,1)}$, которое задается первым полиномом этого базиса вида

$$\mathcal{R}_3^{(1,1,1)} \stackrel{\text{def}}{=} a_1^2 - 4a_2 = 0. \quad (12)$$

Рассмотрим трехчастотный резонанс порядка 4, задаваемый вектором $\mathbf{p}^* = (2, 1, 1)$. Так же и для предыдущего случая составляем идеал из резонансного соотношения $2\lambda_1 + \lambda_2 + \lambda_3$ и зависимостей между λ_j и μ_j . Вычисляем его базис Грёбнера и, приравнявая его первый многочлен к нулю, получаем условия на корни полинома f_3 :

$$16\mu_1^2 - 8\mu_1\mu_2 - 8\mu_1\mu_3 + \mu_2^2 - 2\mu_2\mu_3 + \mu_3^2 = 0.$$

Далее составляем еще один идеал с этим многочленом и полиномами (9). Также первый полином исключаящего базиса Грёбнера этого идеала задает условие на коэффициенты вида

$$\mathcal{R}_3^{(2,1,1)} \stackrel{\text{def}}{=} 16a_1^6 - 264a_1^4a_2 + 36a_1^3a_3 + 1425a_1^2a_2^2 - 630a_1a_2a_3 - 2500a_2^3 + 9261a_3^2 = 0. \quad (13)$$

Проделанные выше вычисления обобщаются на более общий случай трехчастотного резонанса кратности 1, где $\mathbf{p}^* = (q, 1, 1)$. Здесь резонансное соотношение выглядит как $q\lambda_1 + \lambda_2 + \lambda_3 = 0$, где $q \in \mathbb{Z}$. Составляем идеал \mathcal{J} из этого полинома и соотношений зависимости λ_j и μ_j и вычисляем его исключаящий базис Грёбнера \mathcal{G} . Первый полином полученного идеала представляет собой квадратичную форму относительно корней μ_j , $j = 1, 2, 3$, полухарактеристического многочлена f_3 :

$$q^4\mu_1^2 - 2q^2\mu_1\mu_2 - 2q^2\mu_1\mu_3 + \mu_2^2 - 2\mu_3\mu_2 + \mu_3^2.$$

Равенство нулю этой квадратичной формы дает условие существования трехчастотного резонанса. Теперь составляем новый идеал \mathcal{F} , состоящий из указанной выше квадратичной формы и многочленов из (9), и вычисляем его исключаящий базис Грёбнера \mathcal{F} . Его первый полином от коэффициентов a_j , $j = 1, 2, 3$, и целого числа q задает условие существования резонанса вида

$$\begin{aligned} \mathcal{R}_3^{(q,1,1)} \stackrel{\text{def}}{=} & q^4a_1^6 - 2q^2(q^4 + 4q^2 + 1)a_1^4a_2 + 2(q^4 - 3q^2 - 2)(q^2 - 1)^2a_1^3a_3 + (q^4 + 10q^2 + 1) \times \\ & \times (q^2 + 1)^2a_1^2a_2^2 - 2(q^4 - 9)(q^4 - 1)(q^2 - 1)a_1a_2a_3 - 4(q^2 + 1)^4a_2^3 + (q^2 - 1)^3(q^2 + 3)^3a_3^2 = 0. \end{aligned} \quad (14)$$

Если в (14) положить $q = 1$ или $q = 2$, то можно получить резонансные условия $\mathcal{R}_3^{(1,1,1)}$ или $\mathcal{R}_3^{(2,1,1)}$ соответственно.

Правые части полученных в формулах (10), (12)–(14) условий существования двух и трехчастотных резонансов в коэффициентах $\mathbf{a} = (a_1, a_2, a_3)$ многочлена $f_3(\mu)$ представляют собой квазиоднородные многочлены. Это означает, что носитель каждого из многочленов (т.е. множество векторных показателей степеней его мономов в пространстве \mathbb{R}^3) лежит в плоскости, нормаль которой – это вектор $\mathbf{N} = (1, 2, 3)$. Вектор \mathbf{N} принадлежит пространству \mathbb{R}_k^3 , сопряженному пространству показателей степеней \mathbb{R}^3 . Если выполнить такое линейное преобразование, задаваемое унимодулярной матрицей $M = [m_{ij}]_{i,j=1}^3$, $m_{ij} \in \mathbb{Z}$, $\det M = \pm 1$, которая вектор \mathbf{N} переведет внутрь одномерного координатного подпространства, то соответствующее преобразование в \mathbb{R}^3 с матрицей M^{-1} переведет каждый из носителей многочленов на плоскость, параллельную координатной плоскости (см. [18]). Следовательно, соответствующее этой матрице M степенное преобразование $\mathbf{a} \rightarrow \mathbf{v} = (v_1, v_2, v_3)$:

$$\ln \mathbf{a} = M^{-1} \ln \mathbf{v}, \tag{15}$$

преобразует каждый из квазиоднородных многочленов R_3^p от трех переменных a_1, a_2, a_3 к многочленам вида $v_3^k \tilde{R}_3^p(v_1, v_2)$. Таким образом, каждое из условий существования резонанса можно представить в виде плоской алгебраической кривой.

Унимодулярная матрица легко вычисляется с помощью известного алгоритма Эйлера, одна из реализаций которого приведена в [19].

Для вектора $\mathbf{N} = (1, 2, 3)$ соответствующая унимодулярная матрица M может быть выбрана вида

$$M = \begin{pmatrix} -2 & 1 & 0 \\ -3 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \text{тогда} \quad M \cdot \mathbf{N} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Следовательно, степенное преобразование (15) определяется матрицей

$$M^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 3 \end{pmatrix}$$

и имеет вид

$$a_1 = \alpha_1 v_3, \quad a_2 = \alpha_2 v_1 v_3^2, \quad a_3 = \alpha_3 v_2 v_3^3, \tag{16}$$

где ненулевые множители $\alpha_j, j = 1, 2, 3$, можно подбирать для дальнейшего упрощения коэффициентов многочленов R_3^p . Здесь α_j выбраны следующим образом:

$$\alpha_1 = 3, \quad \alpha_2 = 3, \quad \alpha_3 = 1. \tag{17}$$

Поскольку по условиям теоремы 2 коэффициенты $a_k, k = 1, 2, 3$, многочлена $f(\mu)$ должны быть неотрицательны, то из степенного преобразования (16) следует, что параметры $v_k, k = 1, 2, 3$, также должны быть неотрицательны. Следовательно, нас будет интересовать взаимное расположение кривых, соответствующих резонансным многообразиям, в первом квадранте координатной плоскости (v_1, v_2) .

5.3. Параметризация и упрощение резонансных условий

Используя степенные преобразования, упростим вычисленные резонансные условия. В эти выражения для резонансов подставим выражения в переменных $v_k, k = 1, 2, 3$, а именно, $a_1 = 3v_3, a_2 = 3v_1 v_3^2, a_3 = v_2 v_3^3$. Аналитическое представление резонансных многообразий R_3^p в новых переменных далее обозначается \tilde{R}_3^p .

Для общего двухчастотного резонансного условия (10) подстановка (16), (17) после сокращения на множитель v_1^6 дает алгебраическую кривую

$$\begin{aligned} \tilde{R}_3^{(q,1,0)} = 27q^4(q^2+1)^2v_1^3 - 81q^6v_1^2 - 9q^2(q^4+q^2+1)(q^4+4q^2+1)v_1v_2 + \\ + (q^4+q^2+1)^3v_2^2 + 27q^4(q^2+1)^2v_2 = 0. \end{aligned} \quad (18)$$

Из уравнения (18) можно получить условия $\tilde{R}_3^{(1,1,0)}$, $\tilde{R}_3^{(2,1,0)}$ и $\tilde{R}_3^{(3,1,0)}$ соответственно для $q = 1, 2, 3$.

Поступая аналогично с резонансным условием (14) подстановка (16), (17) после сокращения на множитель v_1^6 дает алгебраическую кривую

$$\begin{aligned} \tilde{R}_3^{(q,1,1)} = -108(q^2+1)^4v_1^3 + 81(q^4+10q^2+1)(q^2+1)^2v_1^2 - 18(q^4-9)(q^4-1)(q^2-1)v_1v_2 + \\ + (q^2-1)^3(q^2+3)^3v_2^2 - 486q^2(q^4+4q^2+1)v_1 + 54(q^4-3q^2-2)(q^2-1)^2v_2 + 729q^4 = 0. \end{aligned} \quad (19)$$

Из уравнения (19) для $q = 1$ и $q = 2$ получаем условия для многообразий $\mathcal{R}_3^{(1,1,1)}$ и $\mathcal{R}_3^{(2,1,1)}$ соответственно:

$$\tilde{R}_3^{(1,1,1)} = -4v_1 + 3, \quad \tilde{R}_3^{(2,1,1)} = -2500v_1^3 + 4275v_1^2 - 210v_2v_1 + 343v_2^2 - 2376v_1 + 36v_2 + 432.$$

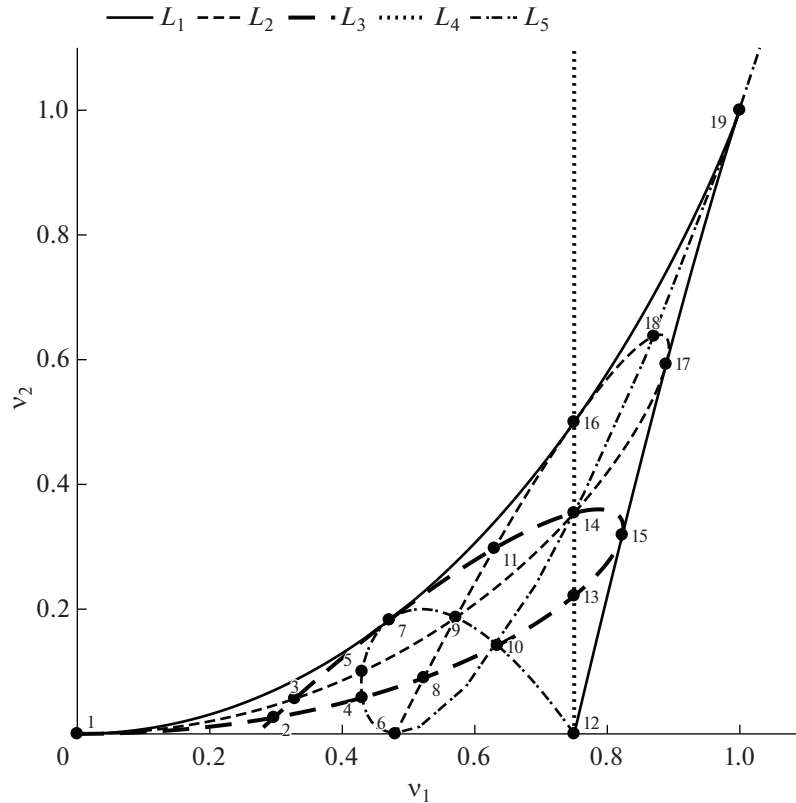
Для каждой из полученных пяти алгебраических кривых был вычислен род, который получился равным нулю во всех случаях. Это говорит о том, что все они рациональные кривые, допускающие рациональную параметризацию. Эти параметризации имеют следующий вид:

$$\begin{aligned} \text{для } \tilde{R}_3^{(1,1,0)}: \left\{ v_1 = -\frac{(9t_1-5)(t_1-1)}{4(2t_1-1)^2}, v_2 = \frac{(t_1-1)^2(7t_1-4)}{4(2t_1-1)^3} \right\}; \\ \text{для } \tilde{R}_3^{(2,1,0)}: \left\{ v_1 = -\frac{(49t_1-16)(743t_1-812)}{400(4t_1-7)^2}, v_2 = \frac{(49t_1-16)^2(97t_1-100)}{400(4t_1-7)^3} \right\}; \\ \text{для } \tilde{R}_3^{(3,1,0)}: \left\{ v_1 = -\frac{(8281t_1-729)(972271t_1-803439)}{218700(27t_1-91)^2}, v_2 = \frac{(8281t_1-729)^2(2617t_1-2025)}{54675(27t_1-91)^3} \right\}; \\ \text{для } \tilde{R}_3^{(1,1,1)}: \left\{ v_1 = \frac{3}{4}, v_2 = t_1 \right\}; \\ \text{для } \tilde{R}_3^{(2,1,1)}: \left\{ v_1 = \frac{39739t_1^2 - 77430t_1 + 54075}{4(111t_1-175)^2}, v_2 = -\frac{(t_1+15)(-225+353t_1)^2}{(111t_1-175)^3} \right\}. \end{aligned}$$

Пользуясь параметрическим представлением, можно построить плоские алгебраические кривые на координатной плоскости (v_1, v_2) .

Для полного представления о взаимном расположении многообразий, соответствующих сильным резонансам, дадим описание фиг. 1. Кривые обозначим символами L_k , $k = 1, \dots, 5$. Их особые точки, а также точки их взаимного пересечения обозначим символами P_j , где нумерация индексов j точек выбрана в соответствии с увеличением расстояния от начала координат.

Кривая L_1 играет особую роль, она задает границу области устойчивости ПР по линейному приближению. Эта кривая является образом дискриминантного множества $\mathcal{D}(f_3)$, которое делит пространство коэффициентов \mathbf{K} многочлена третьей степени на две части. В одной части все корни многочлена вещественные, а в другой части имеется пара комплексно-сопряженных корней и один вещественный корень. Согласно теореме 2, криволинейный треугольник $P_1P_9P_2$ является границей области Σ . Остальные резонансные кривые полностью или частично располагаются внутри этой области. Резонансные кривые L_2 и L_3 , соответствующие двухчастотному резонансу, полностью располагаются в ней, касаясь кривой L_1 . Это связано с тем, что если есть две частоты, участвующие в резонансе, то они как корни характеристического уравнения должны быть одной природы — либо одновременно вещественные, либо одновременно комплексные. Но пара комплексно-сопряженных корней не может находиться в резонансе, а третий корень всегда должен быть вещественным. Отметим, что ранее кривые L_1 , L_2 и L_3 были изображены в [16], но их параметризации были получены другим способом.



Фиг. 1. Резонансные многообразия в переменных v_1, v_2 .

Еще два резонансных многообразия, образами которых являются кривые L_4 и L_5 , соответствуют трехчастотным резонансам. В них, в отличие от двух частотных резонансов, могут участвовать корни различной природы. Это говорит о том, что трехчастотные резонансные многообразия будут находиться и в области вещественности корней, и в области, где существуют комплексные корни. Во всех изображенных точках на кривых кратность резонанса \mathfrak{k} становится равной 2. Легенда фиг. 1 следующая:

1. Кривая L_1 определяется уравнением $\tilde{R}_3^{(1,1,0)} = 0$. Она изображена сплошной линией. На ней имеется одна особая точка $P_9 = (1, 1)$ – точка возврата. Этой точке соответствует резонанс $(1, 1, 1)$ кратности 2, потому что имеются два набора независимых решений. В качестве примера можно привести резонансы вида $(1, 1, 0)$ и $(1, 0, 1)$, для которых алгебраическая сумма трех собственных значений равна нулю.

2. Кривая L_2 , изображенная тонкой штриховой линией, определяется уравнением $\tilde{R}_3^{(2,1,0)} = 0$ и является самопересекающейся кривой с точкой самопересечения $P_9 = (4/7, 64/343)$. Структура резонанса, соответствующая этой точке, будет описана ниже в п. 5, поскольку она одновременно принадлежит и кривой L_5 . Кривая L_2 также касается дискриминантной кривой L_1 в точках $P_{16} = (3/4, 1/2)$, $P_{17} = (8/9, 16/27)$.

3. Полу жирной штриховой линией изображена кривая L_3 . Она определяется уравнением $\tilde{R}_3^{(3,1,0)} = 0$ и является кривой самопересечения с особой точкой $P_2 = (27/91, 19683/753571)$. Она также касается дискриминантной кривой L_1 в точках $P_7 = (57/121, 243/1331)$, $P_{15} = (297/361, 2187/6859)$. Кривая L_3 пересекается с кривой L_2 в четырех точках с координатами $P_3 = (552/1681, 3888/68921)$, $P_8 = (1107/2116, 2187/24334)$, $P_{11} = (216/343, 34992/117649)$, $P_{14} = (3/4, 243/686)$.

4. Прямая пунктирная линия, обозначенная L_4 , определяется уравнением $\tilde{K}_3^{(1,1,1)} = 0$. Она не имеет никаких особенностей и пересекается с указанными выше кривыми в точках P_{12}, P_{13}, P_{14} и P_{16} , где $P_{12} = (3/4, 0)$, $P_{13} = (3/4, 86/2197)$.

5. Последняя кривая L_5 , изображенная штрихпунктирной линией, определяется уравнением $\tilde{K}_3^{(2,1,1)} = 0$. Она является самопересекающейся кривой с особой точкой $P_{10} = (111/175, 243/1715)$, которая также является точкой пересечения с кривой L_3 . С кривой L_1 она касается в точке P_7 и пересекается в точках P_{12} и P_{19} . С кривой L_2 она пересекается в точках P_5, P_6, P_9, P_{14} и P_{18} , где $P_5 = (43/100, 1/10)$, $P_6 = (12/25, 0)$, $P_{18} = (732/841, 15552/24389)$. С кривой L_3 она также касается в точке P_7 , а пересекается с ней в точках $P_4 = (1497/3481, 11907/205379)$ и P_{14} , описанной выше в п. 3.

Для проверки выполнения резонансных условий в вышевычисленных точках рассмотрим наглядно некоторые из них. Выполнение условий резонанса проверим для точки P_9 , которая является особенной для двухчастотной резонансной кривой L_2 и точкой пересечения с трехчастотной резонансной кривой L_5 . Подставив в выражения степенного преобразования (16) значения переменных v_1, v_2 , являющихся координатами точки P_9 , и выполнив замену $v_3 = -t$ при $t > 0$, получим однопараметрические выражения коэффициентов многочлена f_3 . Сам многочлен примет вид

$$f_3 = (7\mu + t)(7\mu + 16t)(7\mu + 4t)/343.$$

Однопараметрическим семейством корней этого полухарактеристического многочлена являются

$$\mu_1 = -t/7, \quad \mu_2 = -4t/7, \quad \mu_3 = -16t/7.$$

Учитывая, что в каждом из них имеется общий множитель вида $t/7$, то все семейство корней будет пропорционально числам $(-1, -4, -16)$. Пользуясь равенством зависимости корней характеристического и полухарактеристического полиномов, вычислим все значения λ_j , $j = 1, \dots, 6$, они будут пропорциональны значениям

$$\lambda_{1,4} = \pm i, \quad \lambda_{2,5} = \mp 2i, \quad \lambda_{3,6} = \pm 4i.$$

Зная значения корней, можно убедиться в выполнении резонансных соотношений. Так как точка P_9 принадлежит двум резонансным многообразиям, проверим эти условия для каждого из них.

1. Поскольку данная точка для L_2 является точкой самопересечения, то имеются две пары соизмеримых корней с отношением $2 : 1$, т.е. резонансное соотношение $\mathbf{p}^* = (2, 1, 0)$ выполняется для корней в виде равенства $2\lambda_1 + \lambda_2 = 2\lambda_2 + \lambda_3 = 2\lambda_4 + \lambda_5 = 2\lambda_5 + \lambda_6 = 0$.

2. В принадлежащей кривой L_5 точке должен выполняться трехчастотный резонанс $\mathbf{p}^* = (2, 1, 1)$. Так, соотношение $2 : 1 : 1$ выполняется в виде $2\lambda_1 + \lambda_5 + \lambda_6 = \lambda_2 + \lambda_3 + 2\lambda_4 = 0$.

Такую же проверку выполним для точки P_{10} . Она является особой для резонансной кривой L_5 и точкой пересечения с кривой L_3 . Выполнив подобные вышеприведенным подстановки, получим выражение для f_3 :

$$f_3 = (35\mu + 27t)(35\mu + 3t)(7\mu + 15t)/8575.$$

А соответствующее ему однопараметрическое семейство корней имеет вид

$$\mu_1 = -3t/35, \quad \mu_2 = -15t/7, \quad \mu_3 = -27t/35.$$

Если учесть, что в каждом из них имеется общий множитель вида $3t/35$, то все семейство корней будет пропорционально числам $(-1, -25, -9)$. А λ_j в данном случае будут пропорциональны

$$\lambda_{1,4} = \pm i, \quad \lambda_{2,5} = \mp 5i, \quad \lambda_{3,6} = \pm 3i.$$

Зная значения корней, можно аналогично убедиться в выполнении резонансных соотношений. Так как точка P_{10} принадлежит двум резонансным многообразиям, проверим эти условия для каждого из них.

1. Поскольку данная точка для L_5 является точкой самопересечения, то имеются две тройки соизмеримых корней с отношением $2 : 1 : 1$, т.е. резонансные соотношения $\mathbf{p}^* = (2, 1, 1)$ выполняются для корней в виде равенств $2\lambda_1 + \lambda_2 + \lambda_3 = 2\lambda_4 + \lambda_5 + \lambda_6 = 0$.

2. В принадлежащей кривой L_3 точке должен выполняться двухчастотный резонанс $\mathbf{p}^* = (3, 1, 0)$. Так, соотношение $3 : 1$ выполняется для корней в виде равенств $3\lambda_1 + \lambda_6 = \lambda_3 + 3\lambda_4 = 0$.

В аналогичной последовательности рассмотрим выполнение условия резонанса в точке P_{14} . Она, в свою очередь, не является особой для какой-либо кривой, но является точкой пересечения сразу четырех алгебраических кривых: L_2, L_3, L_4 и L_5 . Следовательно, в ней должно выполняться сразу четыре резонансных условия. Чтобы проверить это, также найдем корни многочлена f_3 . Выразим коэффициенты этого многочлена через параметр t и разложим его на множители:

$$f_3 = (14\mu + 3t)(7\mu + 6t)(14\mu + 27t)/1372,$$

а соответствующее однопараметрическое семейство корней примет вид

$$\mu_1 = -3t/14, \quad \mu_2 = -6t/7, \quad \mu_3 = -27t/14.$$

Если учесть, что в них имеется общий множитель вида $3t / 14$, то все семейство корней будет пропорционально числам $(-1, -4, -9)$. А λ_j в данном случае будут пропорциональны значениям

$$\lambda_{1,4} = \pm i, \quad \lambda_{2,5} = \mp 2i, \quad \lambda_{3,6} = \pm 3i.$$

Зная значения корней, можно аналогично убедиться в выполнении резонансных соотношений. Так как точка P_{14} принадлежит четырем резонансным многообразиям, проверим эти условия для каждого из них.

1. Точка P_{14} принадлежит L_2 , тогда должно выполняться резонансное соотношение $\mathbf{p}^* = (2, 1, 0)$, а для корней оно примет вид $2\lambda_1 + \lambda_2 = 2\lambda_4 + \lambda_5 = 0$.

2. В принадлежащей кривой L_3 точке должен выполняться двухчастотный резонанс $\mathbf{p}^* = (3, 1, 0)$. Так, соотношение $3 : 1$ выполняется для корней в виде $3\lambda_1 + \lambda_6 = \lambda_3 + 3\lambda_4 = 0$.

3. Для точки, принадлежащей кривой L_4 , должен выполняться трехчастотный резонанс $\mathbf{p}^* = (1, 1, 1)$. Так, соотношение $1 : 1 : 1$ выполняется следующим образом: $\lambda_1 + \lambda_5 + \lambda_6 = \lambda_2 + \lambda_3 + \lambda_4 = 0$.

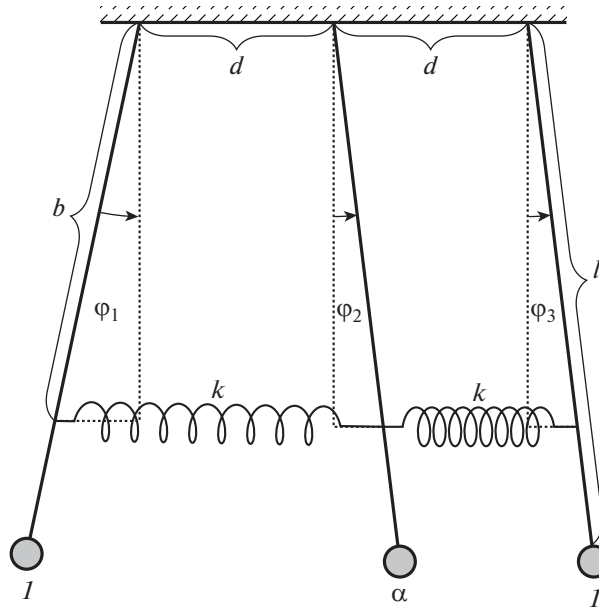
4. Для точки, принадлежащей кривой L_5 , должен выполняться трехчастотный резонанс $\mathbf{p}^* = (2, 1, 1)$. Так, соотношение $2 : 1 : 1$ выполняется для корней в виде равенства $\lambda_1 + 2\lambda_2 + \lambda_3 = \lambda_4 + 2\lambda_5 + \lambda_6 = 0$.

6. ПРИМЕР

Рассмотрим три математических маятника одинаковой длины l , точки подвеса которых расположены на равных расстояниях d на горизонтальной прямой, а массы маятников выбраны равными $\mathbf{m} = (1, \alpha, 1)$ соответственно. Пусть маятники соединены между собой невесомыми линейно-упругими пружинами жесткости k длиной d в недеформированном состоянии. Точки прикрепления пружин расположены на расстоянии $b \leq dl$ от точек подвеса маятников (см. фиг. 2).

Выбрав в качестве обобщенных координат ϕ углы $\phi_i, i = 1, 2, 3$, отклонения маятников от вертикали, запишем функцию Лагранжа этой системы:

$$L(\phi, \dot{\phi}) = \frac{1}{2} \langle T \dot{\phi}, \dot{\phi} \rangle + \Pi. \tag{20}$$



Фиг. 2. Три плоских математических маятника, соединенные невесомыми пружинами.

Матрица T диагональна: $T = \text{diag}(l^2, \alpha l^2, l^2)$. Потенциальная энергия Π есть сумма потенциальной энергии упругой деформации пружин Π_1 и энергии Π_2 математических маятников в однородном поле силы тяжести:

$$\Pi_1 = \frac{k}{2} \sum_{j=1}^2 \left\{ \sqrt{(\Delta_j^{(1)})^2 + (\Delta_j^{(2)})^2} - d \right\}, \quad \Pi_2 = -gl \sum_{j=1}^3 \cos \varphi_j - gl(\alpha - 1) \cos \varphi_2,$$

где $\Delta_j^{(1)} = b(\cos \varphi_{j+1} - \cos \varphi_j)$, $\Delta_j^{(2)} = b(\sin \varphi_{j+1} - \sin \varphi_j) + d$ суть проекции величин деформаций пружин между j -м и $(j + 1)$ -м маятниками на оси абсцисс и ординат соответственно.

Следуя [20], введем безразмерные переменные $\tau = \sqrt{g/l}t$, $\beta = b^2k/(gl)$.

Перейдем к гамильтоновой форме уравнений движения, но при этом будем использовать новые канонические переменные, определяемые нормальными модами колебаний. Для этого раскладываем функцию Лагранжа (20) в ряд Тейлора в окрестности положения равновесия $\varphi = 0$, получим линейную систему уравнений Лагранжа с СЧ

$$\lambda = (1, \sqrt{1 + \beta}, \sqrt{\beta\alpha + 2\beta + 1}).$$

Этим частотам соответствуют амплитудные векторы

$$\mathbf{u}_1 = (1, 1, 1)^\top, \quad \mathbf{u}_2 = (-1, 0, 1)^\top, \quad \mathbf{u}_3 = (1, -2/\alpha, 1)^\top,$$

которые задают матрицу перехода $U = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ к новым переменным \mathbf{Q} :

$$\varphi = U\mathbf{Q}.$$

Переход к канонически сопряженным импульсам \mathbf{P} осуществляется с помощью матрицы $(U^*)^{-1}$:

$$\mathbf{p} = (U^*)^{-1} \mathbf{P}, \quad \text{где} \quad \mathbf{p} = gl \langle \mathbf{m}, \dot{\varphi} \rangle.$$

Здесь звездочка означает транспонирование.

Тогда в новых переменных разложение функции Гамильтона вблизи начала координат имеет вид

$$H(\mathbf{Q}, \mathbf{P}) = H_2(\mathbf{Q}, \mathbf{P}) + H_4(\mathbf{Q}, \mathbf{P}) + \dots, \tag{21}$$

где первые два члена разложения следующие:

$$\begin{aligned}
 H_2 &= \frac{2P_1^2}{2+\alpha} + P_2^2 + \frac{\alpha P_3^2}{2+\alpha} + \left(1 + \frac{\alpha}{2}\right) Q_1^2 + (\beta + 1) Q_2^2 + \left(\frac{2}{\alpha} + 1\right) \left(\beta + 1 + \frac{2\beta}{\alpha}\right) Q_3^2, \\
 H_4 &= -\frac{2+\alpha}{24} Q_1^4 - \left(\beta + \frac{1}{2}\right) Q_1^2 Q_2^2 - \left(\frac{1}{\alpha} + \frac{1}{2}\right) \left(2\beta + 1 + \frac{4\beta}{\alpha}\right) Q_1^2 Q_3^2 - \\
 &\quad - \left(3\beta + 1 + \frac{2\beta}{\alpha}\right) Q_1 Q_2^2 Q_3 - \left(1 - \frac{4}{\alpha^2}\right) \left(\beta + \frac{1}{3} + \frac{2\beta}{\alpha}\right) Q_1 Q_3^3 - \\
 &\quad - \frac{4\beta + 1}{12} Q_2^4 - \left(2\beta + \frac{1}{2} + \frac{2\beta}{\alpha}\right) Q_2^2 Q_3^2 - \frac{(2+\alpha)}{3\alpha^3} (\alpha^2 - 2\alpha + 4) \left(\beta + \frac{1}{4} + \frac{2\beta}{\alpha}\right) Q_3^4.
 \end{aligned} \tag{22}$$

Все дальнейшие вычисления выполнялись в СКА Maple согласно схемы разд. 5.

Поскольку в (21) $H_3(\mathbf{Q}, \mathbf{P}) \equiv 0$, то резонансы порядка $q = 3$ проявят себя при приведении к НФ только в формах степени 6 и выше. Следовательно, в данном случае можно ограничиться исследованием резонансов порядка $q = 4$, а именно,

- в случае двухчастотного резонанса исследуем многообразии $\mathcal{R}_3^{(3,1,0)}$,
- в случае трехчастотного резонанса исследуем многообразии $\mathcal{R}_3^{(2,1,1)}$.

Обозначим через $A(H_2)$ матрицу квадратичной формы (22), тогда полухарактеристический многочлен матрицы $JA(H_2)$, соответствующей линейной системы Гамильтона, есть

$$f(\mu) = \mu^3 + \left(2\beta + 3 + \frac{2\beta}{\alpha}\right) \mu^2 + \left(\beta^2 + 4\beta + \frac{2\beta^2}{\alpha} + 3 + \frac{4\beta}{\alpha}\right) \mu + \left(\beta + 1 + \frac{2\beta}{\alpha}\right) (\beta + 1). \tag{23}$$

Таким образом, пространство параметров $\Pi \equiv (\alpha, \beta)$ задачи двумерно и значения параметров должны быть неотрицательны: $\{\alpha, \beta \geq 0\}$. Квадратичная форма (22) в указанной области пространства параметров является знакоопределенной, следовательно, нижнее равновесное положение маятников является устойчивым по Ляпунову. Поэтому далее будет приведено описание резонансных множеств 4-го порядка.

Подставляем коэффициенты a_1, a_2, a_3 многочлена (23) в уравнение (11) резонансного многообразия $\mathcal{R}_3(3,1,0)$, раскладываем его на множители и отбираем среди них только те, нули которых располагаются в первом квадранте плоскости (α, β) . Аналогично поступаем и с уравнением (13) резонансного многообразия $\mathcal{R}_3(2,1,1)$. В результате получается пять резонансных кривых – три соответствуют двухчастотному резонансу, а две – трехчастотному:

$$\mathcal{R}_a^{(3,1,0)} \stackrel{\text{def}}{=} \beta = 8\alpha / (2 + \alpha), \tag{24}$$

$$\mathcal{R}_b^{(3,1,0)} \stackrel{\text{def}}{=} \beta = 8, \tag{25}$$

$$\mathcal{R}_c^{(3,1,0)} \stackrel{\text{def}}{=} \beta = 4\alpha / (1 - 4\alpha), \tag{26}$$

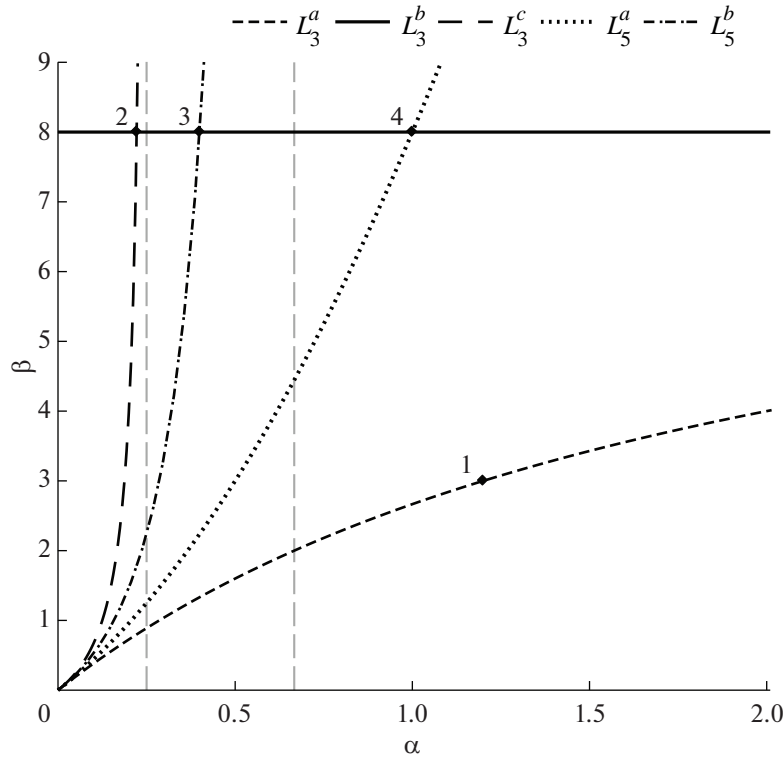
$$\mathcal{R}_a^{(2,1,1)} \stackrel{\text{def}}{=} \beta = 4\alpha^2 + 4\alpha, \tag{27}$$

$$\mathcal{R}_b^{(2,1,1)} \stackrel{\text{def}}{=} \beta = 8\alpha(2 - \alpha) / (3\alpha - 2)^2. \tag{28}$$

Эти кривые показаны на фиг. 3: двухчастотные резонансные кривые изображены сплошными и штриховыми линиями, а трехчастотные – штрихпунктирными.

Кривые обозначим символами $L_k^i, k = 3, 5, i = a, b, c$, где нижний индекс соответствует индексу кривых L_k , описанных в пунктах 3 и 5 подраздела 5.3. Их точки взаимного пересечения обозначим символами P_j , где нумерация индексов j точек выбрана в соответствии с увеличением расстояния от начала координат. Дадим краткое описание структуры СЧ на соответствующих резонансных кривых.

1. На кривой L_3^a , которая изображена длинной штриховой линией, с уравнением (24) СЧ следующие: $\lambda_{1,4} = \pm i, \lambda_{2,5} = \pm i \sqrt{(9\alpha + 2) / (\alpha + 2)}, \lambda_{3,6} = \pm 3i$. Таким образом, для всех значений пара-



Фиг. 3. Резонансные многообразия примера в переменных α, β .

метра α имеет место двухчастотный резонанс, кроме значения $\alpha = 6/5$, соответствующего точке P_1 (см. ниже).

2. На прямой L_3^b , показанной в виде сплошной линии, с уравнением (25) СЧ следующие: $\lambda_{1,4} = \pm i, \lambda_{2,5} = \pm 3i, \lambda_{3,6} = \pm i\sqrt{(9\alpha + 16)/\alpha}$. Эта прямая пересекается с другими кривыми в трех точках: P_2, P_3, P_4 .

3. На кривой L_3^c , которая изображена короткой штриховой линией, с уравнением (26) СЧ следующие:

$$\lambda_{1,4} = \pm i, \quad \lambda_{2,5} = \pm \frac{i}{\sqrt{1 - 4\alpha}}, \quad \lambda_{3,6} = \pm \frac{3i}{\sqrt{1 - 4\alpha}}.$$

4. На кривой L_5^a , изображенной пунктирной линией, с уравнением (27) СЧ следующие:

$$\lambda_{1,4} = \pm i, \quad \lambda_{2,5} = \pm i(2\alpha + 1), \quad \lambda_{3,6} = \pm i(2\alpha + 3).$$

На этой кривой для всех значений α имеет место трехчастотный резонанс, кроме значения $\alpha = 1$, соответствующего точке P_4 .

5. На кривой L_5^b , изображенной штрихпунктирной линией, с уравнением (28) СЧ следующие:

$$\text{при } 0 \leq \alpha < 2/3: \quad \lambda_{1,4} = \pm i, \quad \lambda_{2,5} = \pm i \frac{\alpha + 2}{2 - 3\alpha}, \quad \lambda_{3,6} = \pm i \frac{6 - \alpha}{2 - 3\alpha};$$

$$\text{при } \alpha > 2/3: \quad \lambda_{1,4} = \pm i, \quad \lambda_{2,5} = \pm i \frac{\alpha + 2}{3\alpha - 2}, \quad \lambda_{3,6} = \pm i \frac{6 - \alpha}{3\alpha - 2}.$$

На этой кривой для всех значений α имеет место трехчастотный резонанс, кроме точек P_2, P_3, P_4 .

Кривая L_3^a пересекается только с одной ветвью кривой L_5^b в точке $P_1 = (6/5, 3)$. Подставив в выражения (23) значения переменных α, β , являющихся координатами точки P_1 , получим

$$f_3 = (\mu + 1)(\mu + 4)(\mu + 9).$$

Корнями этого полухарактеристического многочлена являются значения

$$\mu_1 = -1, \quad \mu_2 = -4, \quad \mu_3 = -9.$$

А в корнях характеристического уравнения они запишутся в виде

$$\lambda_{1,4} = \pm i, \quad \lambda_{2,5} = \pm 2i, \quad \lambda_{3,6} = \pm 3i.$$

Зная значения корней, проверим выполнение резонансных соотношений. Так как точка P_1 принадлежит двум резонансным многообразиям, соответствующим двухчастотному и трехчастотному резонансам, проверим эти условия для каждого из них:

1. Для $\mathcal{R}_a^{(3,1,0)}$ имеются две пары соизмеримых корней с отношением 3:1, т.е. $3\lambda_1 + \lambda_6 = 3\lambda_4 + \lambda_3 = 0$.

2. Для $\mathcal{R}_b^{(2,1,1)}$ также должно выполняться соотношение 2:1:1, которое выглядит как $\lambda_1 + 2\lambda_5 + \lambda_3 = \lambda_4 + 2\lambda_2 + \lambda_6 = 0$.

Кривая L_3^b пересекается с кривой L_3^c в точке $P_2 = (2/9, 8)$, с кривой L_5^b – в точках $P_3 = (2/5, 8)$ и $P_4 = (1, 8)$. В точке P_4 она также пересекается с кривой L_5^a . Подставив в выражения (23) значения переменных α, β , являющихся координатами этих точек, получим соответствующие выражения для f_3 :

$$P_2: f_3 = (\mu + 1)(\mu + 9)(\mu + 81),$$

$$P_3: f_3 = (\mu + 1)(\mu + 9)(\mu + 49),$$

$$P_4: f_3 = (\mu + 1)(\mu + 9)(\mu + 25).$$

Аналогично, как и в вышеуказанном случае, здесь можно убедиться в выполнении резонансных соотношений.

7. ЗАКЛЮЧЕНИЕ

Для системы Гамильтона с тремя степенями свободы получено описание разбиения области устойчивости в пространстве коэффициентов \mathbf{K} полухарактеристического многочлена $f(\mu)$ на такие части, в которых гарантированно отсутствие сильных резонансов между СЧ линейной системы (2). Это позволяет, с одной стороны, выполнить исследование формальной устойчивости ПР для каждой из таких частей с помощью НФ четвертого порядка и теоремы 3, а с другой стороны, в случае наличия резонанса кратности 1 получить дополнительные формальные интегралы и выполнить асимптотическое интегрирование уравнений нормализованной системы Гамильтона.

Авторы выражают благодарность А.Д. Брюно за указанные замечания и полезное обсуждение работы.

СПИСОК ЛИТЕРАТУРЫ

1. Брюно А. О типах устойчивости в системах Гамильтона // Препринты ИПМ им. М.В. Келдыша. 2020. № 21. С. 1–24.
2. Зигель К., Мозер Ю. Лекции по небесной механике. Ижевск: НИЦ “Регулярная и хаотическая динамика”, 2001. 384 с.
3. Маркеев А.П. Точки либрации в небесной механике и космодинамике. М.: Наука, 1978. 352 с.
4. Moser J.K. New aspects in the theory of stability of Hamiltonian Systems // Comm. Pure Appl. Math. 1958. V. 11. № 1. P. 81–114.
5. Батхин А.Б., Хайдаров З.Х. Сильные резонансы в нелинейной системе Гамильтона // Препринты ИПМ им. М.В. Келдыша. 2022. № 59. С. 1–28.
6. Батхин А.Б., Брюно А.Д., Варин В.П. Множества устойчивости многопараметрических гамильтоновых систем // Приклад. матем. и мех. 2012. Т. 76. № 1. С. 80–133.

7. *Калинина Е.А., Утешев А.Ю.* Теория исключения: Учеб. пособие. СПб: Изд-во НИИ химии СПбГУ, 2002. 72 с.
8. *Basu S., Pollack R., Roy M.-F.* Algorithms in Real Algebraic Geometry. Algorithms and Computations in Mathematics 10. Berlin–Heidelberg–New York: Springer-Verlag, 2006. ix p.
9. *Брюно А.Д.* Аналитическая форма дифференциальных уравнений (II) // Тр. ММО. 1972. Т. 26. С. 199–239.
10. *Биркгоф Д.Д.* Динамические системы. Ижевск: Изд. дом “Удмуртский университет”, 1999. 408 с.
11. *Брюно А.Д.* Ограниченная задача трех тел: Плоские периодические орбиты. М.: Наука, 1990. 296 с.
12. *Bruno A.D., Batkhin A.B.* Survey of eight modern methods of Hamiltonian mechanics // Axioms. 2021. V. 10. № 4. <https://www.mdpi.com/2075-1680/10/4/293>.
13. *Брюно А.Д.* О формальной устойчивости систем Гамильтона // Матем. заметки. 1967. Т. 1. № 3. С. 325–330.
14. *Журавлев В.Ф., Петров А.Г., Шундерюк М.М.* Избранные задачи гамильтоновой механики. М.: ЛЕНАНД, 2015. 304 с.
15. *Кокс Д., Литтл Д., О’Ши Д.* Идеалы, многообразия и алгоритмы. Введение в вычислительные аспекты алгебраической геометрии и коммутативной алгебры. М.: Мир, 2000. 687 с.
16. *Батхин А.Б.* Резонансное множество многочлена и проблема формальной устойчивости // Вестн. ВолГУ. Сер. 1. Матем. Физ. 2016. № 4 (35). С. 5–23.
17. *Батхин А.Б.* Параметризация множества, определяемого обобщенным дискриминантом многочлена // Программирование. 2018. № 2. С. 5–17.
18. *Брюно А.Д., Солеев А.* Локальная униформизация пространственной кривой и многогранника Ньютона // Алгебра и анализ. 1991. Т. 3. № 1. С. 67–102.
19. *Брюно А.Д., Азимов А.А.* Вычисление унимодулярных матриц степенных преобразований // Программирование. 2023. № 1. С. 38–47.
20. *Маркеев А.П.* О движении связанных маятников // Нелинейная динамика. 2013. Т. 9. № 1. С. 27–38.

ОБЩИЕ ЧИСЛЕННЫЕ
МЕТОДЫ

УДК 519.612

**A GENERALIZED SIMPLIFIED HERMITIAN AND SKEW-HERMITIAN
SPLITTING PRECONDITIONER FOR DOUBLE SADDLE
POINT PROBLEMS¹⁾**

© 2023 г. L. Meng¹, Y. W. He^{2,1,*}, J. Li¹

¹ College of Mathematics and Statistics, Northwest Normal University, Lanzhou 730070, P. R. China

² Institute for Advanced Study in History of Science, Northwest University, Xi'an 710127, P. R. China

*e-mail: heyuwen18793436582@163.com

Поступила в редакцию 25.11.2021 г.

Переработанный вариант 05.12.2022 г.

Принята к публикации 02.02.2023 г.

Обобщенный упрощенный эрмитов и косоэрмитов предобусловливатель для задач с двойной седловой точкой. Рассмотрен обобщенный упрощенный эрмитов и косоэрмитов расщепляющий предобусловливатель для решения задач с двойной седловой точкой, а также распределение собственных значений этого предобусловливателя. Исследованы распределение собственных векторов и степень минимального многочлена предварительно обусловленной матрицы. Представлены численные эксперименты, показывающие эффективность такого предобусловливателя.

Ключевые слова: эрмитов и косоэрмитов предобусловливатель, двойная седловая точка, распределение собственных значений, распределение собственных векторов.

DOI: 10.31857/S0044466923050149, **EDN:** DXFIAI

¹⁾ Полный текст статьи печатается в английской версии журнала.

ОБЩИЕ ЧИСЛЕННЫЕ
МЕТОДЫ

УДК 519.63

**ON THE RADIAL BASIS FUNCTION INTERPOLATION I:
SPECTRAL ANALYSIS OF THE INTERPOLATION MATRIX
AND THE RELATED OPERATORS¹⁾**

© 2023 г. Jianping Xiao^{1,2,3,*}

¹ China Resources Networks Co., Ltd., Shenzhen, China

² National University of Singapore, SITY, Singapore

³ University of Michigan, Ann Arbor, Michigan, USA

*e-mail: jpxiao0325@gmail.com

Поступила в редакцию 11.10.2022 г.

Переработанный вариант 11.10.2022 г.

Принята к публикации 02.02.2023 г.

Использование радиальных базисных функций I: Спектральный анализ матрицы интерполяции и связанных с ней операторов. Исследованы спектральные свойства периодической матрицы интерполяции для радиальных базисных функций (RBF), а также связанных с ними гармонических операторов, полученных дискретизацией с использованием таких функций. Для гауссовских RBF эта процедура распространена на пространство произвольной размерности на сетке тензорного произведения. Экспериментальный результат числа обусловленности Бойда хорошо согласуется с аналитической оценкой для периодических RBF.

Ключевые слова: радиальные базисные функции, спектральные методы, нейронная сеть, тензорная декомпозиция.

DOI: 10.31857/S0044466923050204, **EDN:** DXJOLI

¹⁾ Полный текст статьи печатается в английской версии журнала.

ОБЩИЕ ЧИСЛЕННЫЕ МЕТОДЫ

УДК 519.651

АППРОКСИМАЦИЯ ТАБЛИЧНО ЗАДАННЫХ ФУНКЦИЙ: МНОГОКРИТЕРИАЛЬНЫЙ ПОДХОД¹⁾

© 2023 г. А. П. Нелюбин^{1,*}, В. В. Подиновский^{2,**}

¹ 101990 Москва, М. Харитоньевский пер., 4, ИМАШ РАН, Россия

² 101000 Москва, ул. Мясницкая, 20, НИУ ВШЭ, Россия

*e-mail: nelubin@gmail.com;

**e-mail: podinovski@mail.ru

Поступила в редакцию 08.09.2022 г.
Переработанный вариант 29.09.2022 г.
Принята к публикации 02.02.2023 г.

Развит новый подход к оцениванию параметров аппроксимации, при котором удаленность аппроксимирующей функции от заданного конечного множества точек оценивается векторным критерием, компонентами которого являются модули невязок во всех точках. При помощи этого критерия задается отношение предпочтения в удаленности и лучшей считается аппроксимирующая функция, недоминируемая по такому отношению. Изучена аппроксимация для нескольких отношений предпочтения, в том числе для отношения Парето и отношения, порожденного информацией о равноважности критериев. Рассмотрены вычислительные вопросы и исследованы взаимоотношения введенных аппроксимирующих функций с классическими (получаемыми методами наименьших квадратов, наименьших модулей и наименьшего максимального модуля отклонений). Библ. 15. Фиг. 8.

Ключевые слова: аппроксимация функций, регрессионный анализ, многокритериальный анализ, теория важности критериев.

DOI: 10.31857/S0044466923050174, **EDN:** GFZWFZ

ВВЕДЕНИЕ

Аппроксимация (приближение), т.е. замена одних математических объектов другими, в том или ином смысле близкими к исходным, широко применяется в математике и ее приложениях. В частности, она является основой для аналитического представления табличных данных функцией заданного вида с неизвестными параметрами, значения которых определяются так, чтобы график функций был по возможности максимально близок к заданным точкам согласно выбранной числовой мере близости. Задачи такого оценивания являются базовыми в теории аппроксимации функций и регрессионном анализе [1], [2].

Меры близости в регрессионном анализе достаточно обоснованно (в частности, с привлечением принципа максимума правдоподобия) выбираются в тех случаях, когда распределение вероятностей ошибок измерения подчинено некоторому известному закону, чаще всего нормальному закону, или закону Гаусса (при этом мерой служит сумма квадратов отклонений рассчитанных значений аппроксимирующей функции от табличных), реже — закону Лапласа (и тогда мерой служит сумма модулей таких отклонений).

В общем же случае при решении задач аппроксимации проблема обоснованного выбора меры близости остается нерешенной, на практике ей обычно уделяется недостаточно внимания (нередко о ней вообще не вспоминают) и в качестве меры берут сумму квадратов отклонений, так как это — “общепринятый подход”, для него имеются давно и хорошо разработанный математический аппарат аппроксимации и созданы компьютерные программы. А ведь выбор меры близости самым существенным образом влияет на результат аппроксимации.

В настоящей статье проблема выбора меры близости решается за счет того, что близость оценивается при помощи векторного критерия, состоящего из модулей отклонений значений ап-

¹⁾Работа выполнена при поддержке Международного центра анализа и выбора решений.

проксимирующей функции от заданных для всех табличных значений аргумента. Ценой общности такого подхода является многозначность получающейся аппроксимирующей функции.

1. МНОГОКРИТЕРИАЛЬНАЯ ЗАДАЧА АППРОКСИМАЦИИ

Исходными данными в задаче аппроксимации являются пары значений независимой переменной x и зависимой переменной y :

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n). \quad (1.1)$$

В регрессионном анализе независимая переменная x именуется фактором, а зависимая переменная y – откликом (наблюдённым значением).

Далее будем полагать, что аппроксимирующая функция (функция регрессии) является линейной:

$$y = ax + b, \quad (1.2)$$

где a и b – коэффициенты аппроксимации, которые нужно оценить с использованием исходных данных (1.1). Для такого оценивания вводятся в рассмотрение модули невязок δ_i , или уклонений теоретических значений зависимой переменной – рассчитанных по формуле (1.2) – от заданных:

$$\delta_1 = ax_1 + b - y_1, \quad \delta_2 = ax_2 + b - y_2, \quad \dots, \quad \delta_n = ax_n + b - y_n. \quad (1.3)$$

Общепринятым при аппроксимации является путь, основанный на введении числовой меры близости аппроксимирующей функции к исходным данным, такой, как сумма квадратов невязок, сумма модулей невязок или максимальный модуль невязок. Наиболее распространённым методом оценивания является метод наименьших квадратов (МНК) (the least squares method – LS), согласно которому значения коэффициентов регрессии определяются из условия:

$$\sum_{i=1}^n \delta_i^2 \rightarrow \min_{a,b}. \quad (1.4)$$

Для расчета значений коэффициентов согласно (1.4) имеются простые формулы [1], [2].

Достаточно распространённым, особенно в последние годы, стал метод наименьших модулей (МНМ) (the least modulus method – LM), согласно которому значения коэффициентов аппроксимации определяются из условия:

$$\sum_{i=1}^n |\delta_i| \rightarrow \min_{a,b}. \quad (1.5)$$

Разработаны специальные алгоритмы нахождения значений коэффициентов согласно (1.5), см. [3].

Метод, основанный на использовании наибольшего модуля невязок, называемый далее *методом наименьшего максимального модуля* (МНММ) (the least maximum modulus method – LMM), по сути, использует меру Чебышёва и предполагает получение значений коэффициентов регрессии из условия

$$\max_{i \in N} |\delta_i| \rightarrow \min_{a,b}, \quad (1.6)$$

где $N = \{1, 2, \dots, n\}$. Отметим, что нелинейные оптимизационные задачи (1.5) и (1.6) за счет введения дополнительных переменных можно свести к эквивалентным задачам линейного программирования [4].

Далее рассматривается альтернативный метод оценивания, учитывающий, что по своей сути задача оценивания параметров регрессии изначально является многокритериальной: критериями f_i , подлежащими минимизации ($i \in N$), служат модули невязок δ_i из (1.3), составляющие векторный критерий $f(a, b) = (f_1(a, b), f_2(a, b), \dots, f_n(a, b))$, где $f_i(a, b) = z_i = |\delta_i|$. Значение векторного критерия для пары (a, b) называется ее векторной оценкой: $z = f(a, b)$.

Пусть на множестве векторных оценок $Z = \text{Re}_+^n = [0, +\infty)^n$ задан строгий частичный порядок P^Γ , где Γ – информация, касающаяся оценивания удаленности: если верно $z' P^\Gamma z''$, то при модулях невязок z'_1, z'_2, \dots, z'_n аппроксимирующая функция считается ближе к заданным точкам (1.1), чем

при модулях невязок $z_1'', z_2'', \dots, z_n''$. Отношение P^Γ индуцирует на плоскости Re^2 строгий частичный порядок P_Γ :

$$(a', b')P_\Gamma(a'', b'') \Leftrightarrow f(a', b')P^\Gamma f(a'', b''),$$

имеющий аналогичный смысл: если верно $(a', b')P_\Gamma(a'', b'')$, то при значениях параметров a', b' аппроксимирующая функция считается ближе к заданным точкам (1.1), чем при значениях параметров a'', b'' . В качестве решения задачи оценивания параметров аппроксимации рассматривается множество точек (векторов) $(a, b) \in \text{Re}^2$, которые недоминируемы по отношению P_Γ . (Точка (a^*, b^*) называется недоминируемой по P_Γ , если не существует точки (a, b) такой, что верно $(a, b)P_\Gamma(a^*, b^*)$.)

При указанном многокритериальном подходе в качестве решения задачи аппроксимации выделяется не одно оптимальное значение коэффициентов (a^*, b^*) , а множество недоминируемых значений $U^\Gamma \subset \text{Re}^2$. Каждой паре $(a, b) \in U^\Gamma$ соответствует “своя” функция $y = f(x, a, b) = ax + b$ и ее график – прямая (1.2), а множеству U^Γ – множеству функций $F^\Gamma = \{f(x|a, b)|(a, b) \in U^\Gamma\}$ и множество их графиков – связка (совокупность) прямых. При этом фиксированному значению аргумента x ставится в соответствие множество чисел $Y^\Gamma(x) = \{y \in \text{Re} | y = ax + b, (a, b) \in U^\Gamma\}$. Иными словами, полученная указанным путем аппроксимирующая функция F^Γ оказывается многозначной. Такие аппроксимирующие функции будем называть *недоминируемыми*, или, более конкретно, *Γ -недоминируемыми*.

Многозначность аппроксимации (множественность выделяемых значений параметров) – плата за то, что при многокритериальной постановке задачи их оценивания не делается никаких дополнительных ограничительных допущений о характере (свойствах) исходных данных (1.1) типа фиксирования вида закона распределения вероятностей ошибок измерений и т.п.

Для снятия неопределенности из-за многозначности аппроксимации (за счет снижения информативности) можно множество $Y^\Gamma(x)$ представлять его центром тяжести (центроидом) $c^\Gamma(x)$. В частности, если $Y^\Gamma(x)$ – отрезок $[\alpha(x), \beta(x)]$, то $c^\Gamma(x)$ – его середина: $c^\Gamma(x) = 1/2(\alpha(x) + \beta(x))$. Если функция $y = c^\Gamma(x)$ оказывается линейной (на интересующем интервале оси Ox), то ее можно принять за однозначное упрощенное решение задачи линейной аппроксимации.

2. О ПОСТРОЕНИИ МНОЖЕСТВ U^Γ И $Y^\Gamma(x)$

Для приближенного построения (конечной аппроксимации) множества U^Γ можно построить “достаточно мелкую” сетку, покрывающую “подходящую” область в Re^2 (она определяется с учетом специфики задачи, т.е. расположения точек из (1.1) на плоскости), рассматривать эту сетку вместо множества Re^2 и выделять те из ее узлов, которые являются недоминируемыми по P_Γ . Подобный подход широко используется в многокритериальной оптимизации, когда множество допустимых вариантов решения является областью в Re^n (см., например, [5]). Он применялся также при построении множества ПН-средних, определяемого в рамках многокритериального подхода как совокупности точек прямой, каждая из которых недоминируема по отношению предпочтения на множестве векторов, компонентами которых служат расстояния от нее до заданных точек [6], [7].

В качестве ориентира при первоначальном определении области поиска в Re^2 можно использовать точки (a^\bullet, b^\bullet) , соответствующие прямым, проходящим через 2 или более точек множества (1.1). Авторы работы [8] называют такие точки “узловыми” и доказывают, что линию, определяемую по МНМ, следует искать среди $y = a^\bullet x + b^\bullet$. Также полезными могут оказаться точки $(a^\blacksquare, b^\blacksquare)$, соответствующие прямым, проходящим через 2 или более середин отрезков, соединяющих пары точек множества (1.1). В последующих примерах линии, определяемые по МНММ, оказываются среди $y = a^\blacksquare x + b^\blacksquare$.

Для сокращения области поиска множества U^Γ можно воспользоваться следующим замечанием. Зафиксируем значение параметра $a = a^0$ и рассмотрим локальную задачу поиска точек $(a^0, b) \in U^\Gamma(a^0)$, недоминируемых по P_Γ на множестве $\{(a, b) | a = a^0, b \in \text{Re}\}$. Обозначим $b_i = y_i - a^0 x_i$ для всех $i \in N$ и преобразуем выражение для уклонений (1.3):

$$\delta_i = a^0 x_i + b - y_i = b - b_i.$$

В результате получаем задачу определения множества ПН-средних $G^\Gamma(B(a^0))$ среди множества вещественных чисел $B(a^0) = \{b_1, b_2, \dots, b_n\}$ (см. [6], [7]). Если точка b не входит в множество средних $G^\Gamma(B(a^0))$, то это означает, что точка (a^0, b) оказывается доминируемой по P_Γ на рассматриваемом множестве $\{(a, b) | a = a^0, b \in \text{Re}\}$, а значит, и на всем множестве Re^2 . Следовательно, должно выполняться включение

$$\{(a, b) \in U^\Gamma | a = a^0\} \subseteq \{(a, b) | a = a^0, b \in G^\Gamma(B(a^0))\}.$$

В ходе перебора точек методом построения сетки, это замечание можно использовать следующим образом. Для каждого перебираемого значения a сначала можно определить ПН-средние $G^\Gamma(B(a))$, а затем перебирать только значения b из этого множества. Это может быть удобно, так как для некоторых типов информации Γ существуют быстрые аналитические методы определения ПН-средних.

Для произвольного фиксированного значения x^0 множество $Y^\Gamma(x^0)$ графически представляется сечением вертикальной прямой $x = x^0$ связки прямых F^Γ . Для построения множеств $Y^\Gamma(x)$ можно использовать также графоаналитический метод, не требующий построения связки прямых и основанный на следующих соображениях. Произвольной точке (x^0, y^0) пространства Re^2 данных (x, y) соответствует прямая $b = -x^0a + y^0$ в пространстве Re^2 параметров (a, b) . Если эта прямая содержит хотя бы одну точку $(a^*, b^*) \in U^\Gamma$, то точке (x^0, y^0) соответствует, по крайней мере, одна прямая из связки аппроксимирующих прямых F^Γ . Фиксируем $x = x^0$ и будем менять y . Тогда соответствующая прямая $b = -x^0a + y^0$ будет сдвигаться параллельно самой себе: $b = -x^0a + y$, т.е. вертикальной прямой $x = x^0$ в пространстве Re^2 данных (x, y) соответствует множество прямых $b = -x^0a + y$ в пространстве Re^2 параметров (a, b) с одинаковым угловым коэффициентом $-x^0$.

Следовательно, для построения $Y^\Gamma(x^0)$ можно действовать так. Вначале следует построить множество U^Γ или его конечную аппроксимацию в системе координат параметров aOb . Далее нужно взять (начертить) прямую с наклоном $-x^0$. Параллельно сдвигая эту прямую (т.е. сохраняя угол наклона), нужно определить все те положения, в которых прямая $b = -x^0a + y$ будет пересекать множество U^Γ . Соответствующие этим положениям значения b на оси ординат Ob составят множество $Y^\Gamma(x^0)$. В частности, если U^Γ — компактное (т.е. замкнутое и ограниченное) связное множество, то $Y^\Gamma(x^0)$ — отрезок с концами y_{\min} и y_{\max} , соответствующими крайним (верхнему и нижнему) положениям передвигаемой прямой: $y_{\min} = b_{\min}$ и $y_{\max} = b_{\max}$, ибо линейное отображение $b = -x^0a + y^0$ является непрерывным и поэтому, как известно из курса анализа, переводит компактное связное множество в Re^2 в компактное и связное множество в Re .

Примеры построения множеств U^Γ и $Y^\Gamma(x)$ приводятся ниже. В них для конечной аппроксимации множеств U^Γ и $Y^\Gamma(x)$ использовалась специально разработанная компьютерная программа.

3. АППРОКСИМАЦИЯ ПРИ ОТСУТСТВИИ ИНФОРМАЦИИ Γ

Рассмотрим вначале случай, когда никакой информации Γ нет: $\Gamma = \emptyset$. Здесь в роли частично-го порядка P^Γ выступает отношение Парето P^\emptyset , которое определяется (с учетом желательности минимизации критериев) так [9]:

$$z' P^\emptyset z'' \Leftrightarrow z'_i \leq z''_i, \quad i = 1, 2, \dots, n, \quad z' \neq z''. \quad (3.1)$$

Структура множества U^\emptyset исследовалась в [10]. Было установлено, что оно компактно (замкнуто и ограничено), связно, возможно, не выпукло и является объединением конечного числа политопов (выпуклых оболочек конечных наборов двумерных векторов). Из этого свойства, как указывалось выше, следует, что множества $Y^\emptyset(x)$ — отрезки.

Изложенные теоретические положения иллюстрирует следующий числовой пример.

Пример 1. Пусть исходные данные (1.1) заданы в табл. 1.

Таблица 1. Исходные данные для задач 1–3

x_i	1	2	3	4	5	7
y_i	3	2	5	5	4	7

Здесь уравнения аппроксимирующих функций, определяемые МНК, МНМ и МНММ, соответственно, таковы:

$$f^{LS}(x) = \frac{47}{70}x + 1\frac{61}{70} \approx 0.671x + 1.871, \tag{3.2}$$

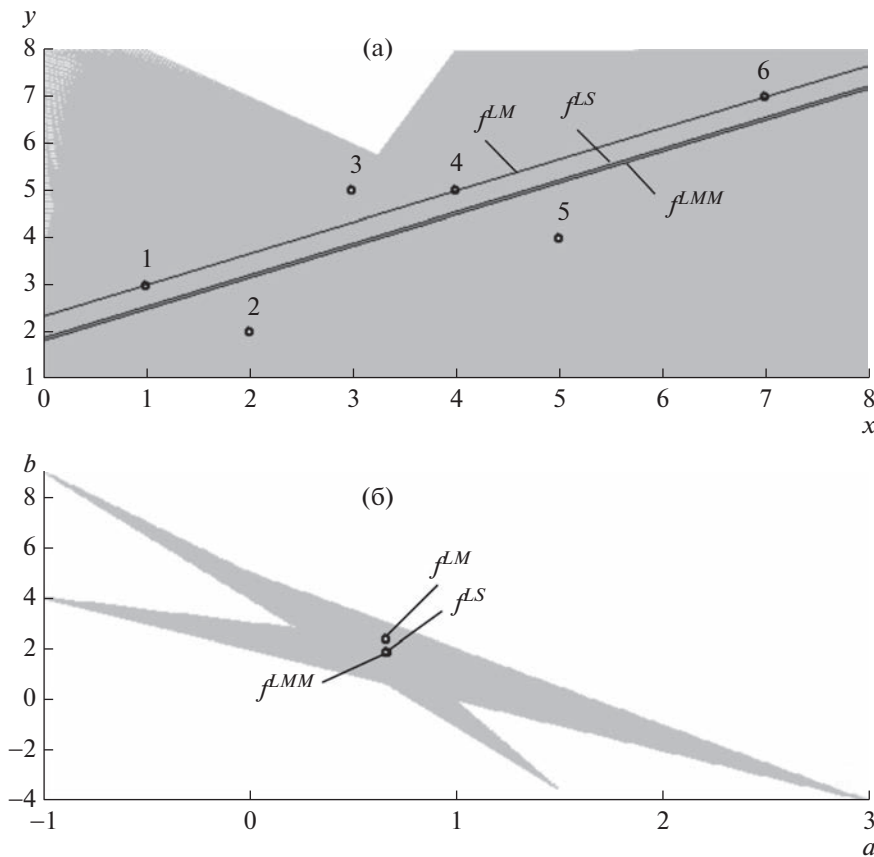
$$f^{LM}(x) = \frac{2}{3}x + 2\frac{1}{3} \approx 0.667x + 2.333, \tag{3.3}$$

$$f^{LMM}(x) = \frac{2}{3}x + 1\frac{5}{6} \approx 0.667x + 1.833. \tag{3.4}$$

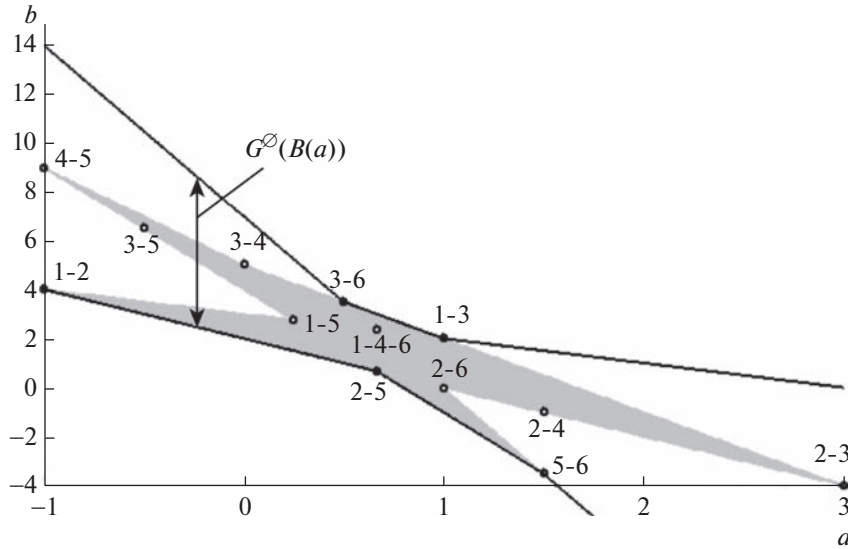
В этом примере при расчетах для аппроксимации множества U^∞ использовалась сетка с шагом 0.005 по параметру a и с шагом 0.05 по параметру b , покрывающая область $[-1, 3] \times [-4, 9]$ для параметров (a, b) . Чтобы избежать вычислительных ошибок при сравнении чисел с плавающей точкой, при вычислениях использовались только рациональные числа, представленные целыми числителем и знаменателем. Например, $a = \frac{\hat{a}}{\check{a}}$, где числитель \hat{a} и знаменатель $\check{a} > 0$ – целые числа. В таком представлении числитель и знаменатель уклонений (1.3) рассчитываются по следующим формулам:

$$\hat{\delta}_i = \hat{a}\check{x}_i\check{b}\check{y}_i + \hat{b}\check{a}\check{x}_i\check{y}_i - \check{y}_i\check{a}\check{x}_i\check{b}, \quad \check{\delta}_i = \check{a}\check{b}\check{x}_i\check{y}_i.$$

Связка аппроксимирующих функций представлена в области $[0, 8] \times [1, 8]$ для (x, y) на фиг. 1 сверху в виде затемненной области, а множество точек U^∞ представлено в области $[-1, 3] \times [-4, 9]$ для (a, b) на фиг. 1б.



Фиг. 1. Связка прямых и множество U^∞ .



Фиг. 2. Множество U^\emptyset , графики границ отрезков $G^\emptyset(B(a))$, а также узловые точки.

Для информации $\Gamma = \emptyset$ множество ПН-средних $G^\emptyset(B(a))$ (множество значений функций, реализующих среднее по Коши) определяется легко: это отрезок $[\min_{i \in N} b_i, \max_{i \in N} b_i]$. Графики нижней и верхней границы этого отрезка в зависимости от параметра a представлены на фиг. 2. Также на фиг. 2 изображены “узловые” точки (a^*, b^*) , соответствующие прямым, проходящим через 2 или более точек множества (1.1) (номера $i \in N$ этих точек указаны рядом через дефис, например, точке 1–4–6 соответствует линия, определяемая по МНМ). Очевидно, что эти точки (a^*, b^*) недоминируемы по Парето, так как для каждой такой точки (и только для нее) все соответствующие уклонения $\delta_i = 0$.

4. АППРОКСИМАЦИЯ ПРИ РАВНОВАЖНЫХ КРИТЕРИЯХ

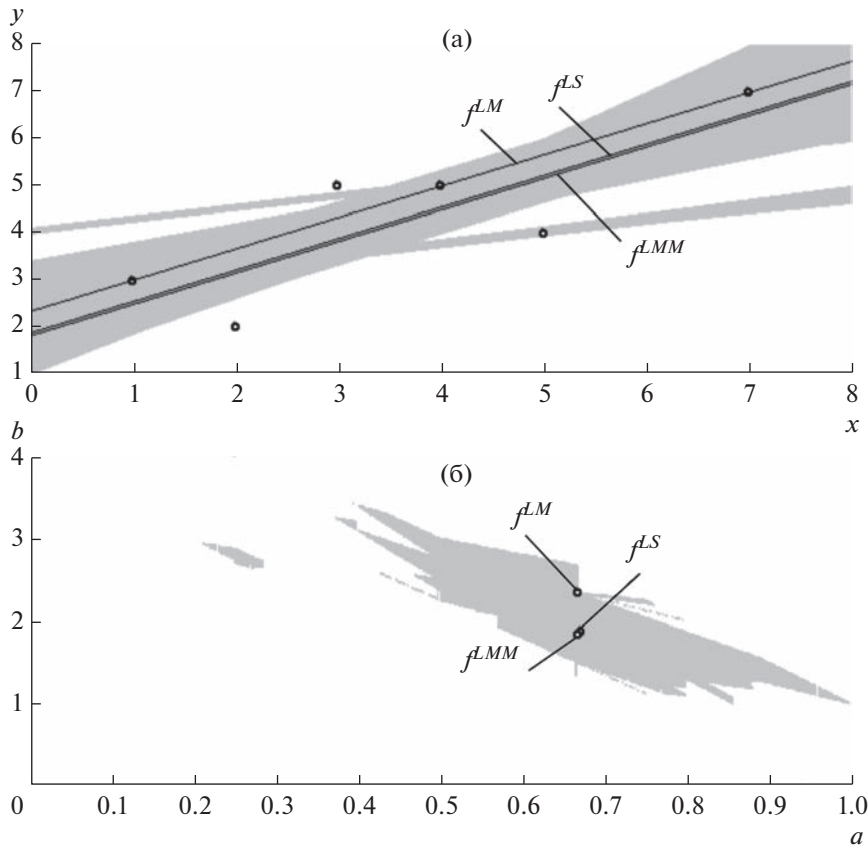
Пусть теперь все критерии имеют равную важность (см. [11], [12]): информация $\Gamma = E$. Содержательно равноважность критериев означает, что после произвольной перестановки компонент векторной оценки z получится векторная оценка, эквивалентная (равная по предпочтительности) исходной. В рассматриваемой постановке задачи аппроксимации все точки (1.1) одинаково существенны (“равноправны”) и поэтому критерии f_i равноважны.

Обозначим через $z_{\setminus} = (z_{[1]}, z_{[2]}, \dots, z_{[n]})$ вектор, полученный упорядочением по невозрастанию компонент вектора z : $z_{[1]} \geq z_{[2]} \geq \dots \geq z_{[n]}$. Отношение P^E определяется следующим аналитическим решающим правилом(см. [11]):

$$z' P^E z'' \Leftrightarrow z'_{[1]} \leq z''_{[1]}, \quad z'_{[2]} \leq z''_{[2]}, \quad \dots, \quad z'_{[n]} \leq z''_{[n]}, \quad z'_{\setminus} \neq z''_{\setminus}.$$

Поскольку $P^E \supset P^\emptyset$, то имеет место включение $U^E \subseteq U^\emptyset$. Оказывается, что здесь множество U^E может быть несвязным, и тогда связка аппроксимирующих прямых расщепляется на несколько частей. (Несвязность U^E – аналог несвязности средних по P_E [6].) Эту возможность иллюстрирует следующий пример.

Пример 2. Пусть исходные данные (1.1) заданы в табл. 1. При расчетах для аппроксимации множества U^E использовалась сетка с шагом 0.001 по параметру a и с шагом 0.01 по параметру b , покрывающая область $[0, 1.1] \times [0, 5]$ для параметров (a, b) . На фиг. 3 представлены связка графиков аппроксимирующих функций и множество U^E , а также графики функций (3.2)–(3.4) и исходные точки. Фигура 3 показывает, что множество U^E не связно, и из-за этого связка линий аппроксимации расщепилась на несколько частей, а множества $Y^E(x)$ не являются числовыми промежутками.



Фиг. 3. Связка прямых и множество U^E .

5. АППРОКСИМАЦИЯ ПРИ РАВНОВАЖНЫХ КРИТЕРИЯХ И ОТСУТСТВИИ КОМПЕНСАЦИИ

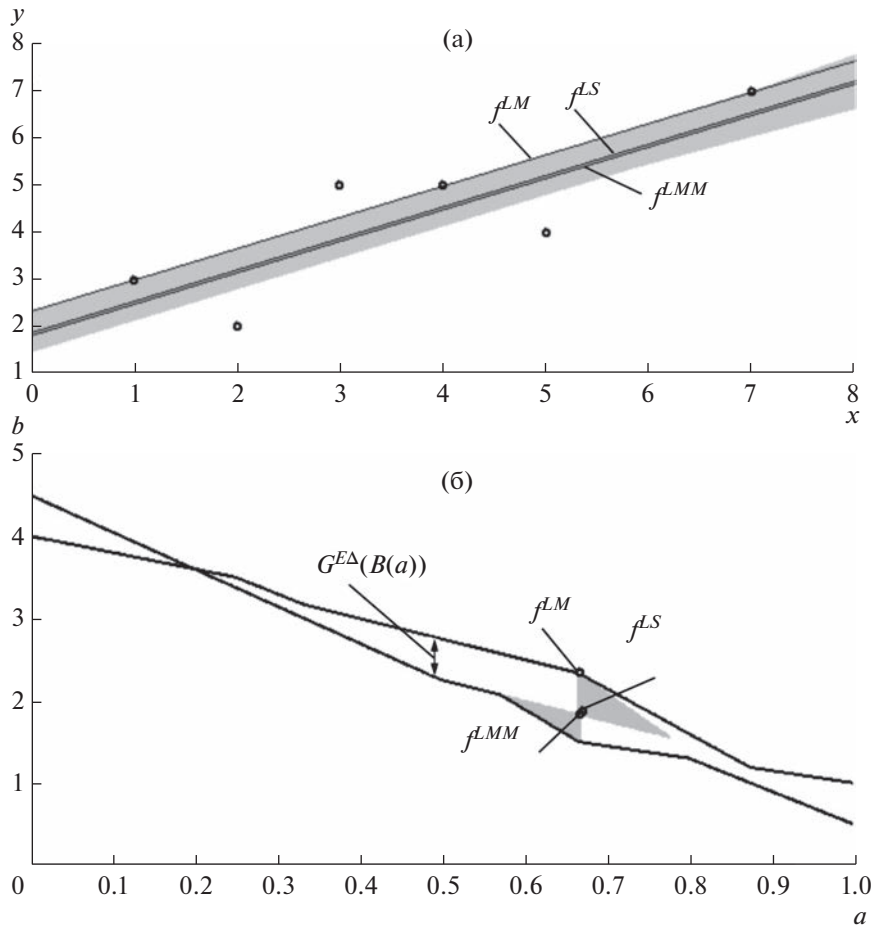
Рассмотрим случай, когда увеличение большего уклонения аппроксимирующей функции в одной из точек x_i не компенсируется уменьшением меньшего уклонения в другой такой точке. Такой подход можно рассматривать как уточнение МНММ (см. (1.6)). Для такой информации EL лексикографического типа (относительно упорядоченных по невозрастанию компонент векторных оценок) отношение P^{EL} определяется следующим решающим правилом:

$$z' P^{EL} z'' \Leftrightarrow [(z'_{i1} < z''_{i1}) \vee (z'_{i1} = z''_{i1}, z'_{i2} < z''_{i2}) \vee \dots \\ \dots \vee (z'_{in} = z''_{in}, i = 1, 2, \dots, n-1, z'_n < z''_n)].$$

Поскольку $P^{EL} \supset P^E$, то $U^{EL} \subseteq U^E$.

6. АППРОКСИМАЦИЯ ПРИ РАВНОВАЖНЫХ КРИТЕРИЯХ СО ШКАЛОЙ ПЕРВОЙ ПОРЯДКОВОЙ МЕТРИКИ

Пусть, наконец, все критерии имеют равную важность и шкала критериев является шкалой первой порядковой метрики (см. [12], [13]): информация $\Gamma = E\Delta$. Такая шкала обладает следующим свойством: если в произвольной векторной оценке z , в которой $z_i > z_j$, заменить z_i на $z_i - q$, а z_j на $z_j + q$, но при этом обеспечить выполнение неравенства $z_i - q \geq z_j + q$, где q — положительное число, то полученная таким образом векторная оценка будет более предпочтительна, чем исходная. Это означает, что при указанном переходе большее отклонение от одной из точек выборки уменьшается за счет равного по величине увеличения меньшего отклонения от другой точки выборки, так что набор отклонений становится более близким к идеальному набору из минимально возможных равных уклонений. Такой подход является аналогом известного в теории обществен-



Фиг. 4. Связка прямых, множество $U^{E\Delta}$ и графики границ отрезков $G^{E\Delta}(B(a))$.

ного благосостояния принципа Пигу-Дальтона: “Общество станет лучше, если единица дохода будет передана от более богатого индивида более бедному” [14].

В рассматриваемом случае имеем частичный порядок $P^{E\Delta}$, задаваемый следующим решающим правилом [12], [15]:

$$z' P^{E\Delta} z'' \Leftrightarrow z'_{[1]} \leq z''_{[1]}, \quad z'_{[1]} + z'_{[2]} \leq z''_{[1]} + z''_{[2]}, \quad \dots$$

$$\dots, \quad z'_{[1]} + z'_{[2]} + \dots + z'_{[n]} \leq z''_{[1]} + z''_{[2]} + \dots + z''_{[n]}, \quad z'_n \neq z''_n.$$

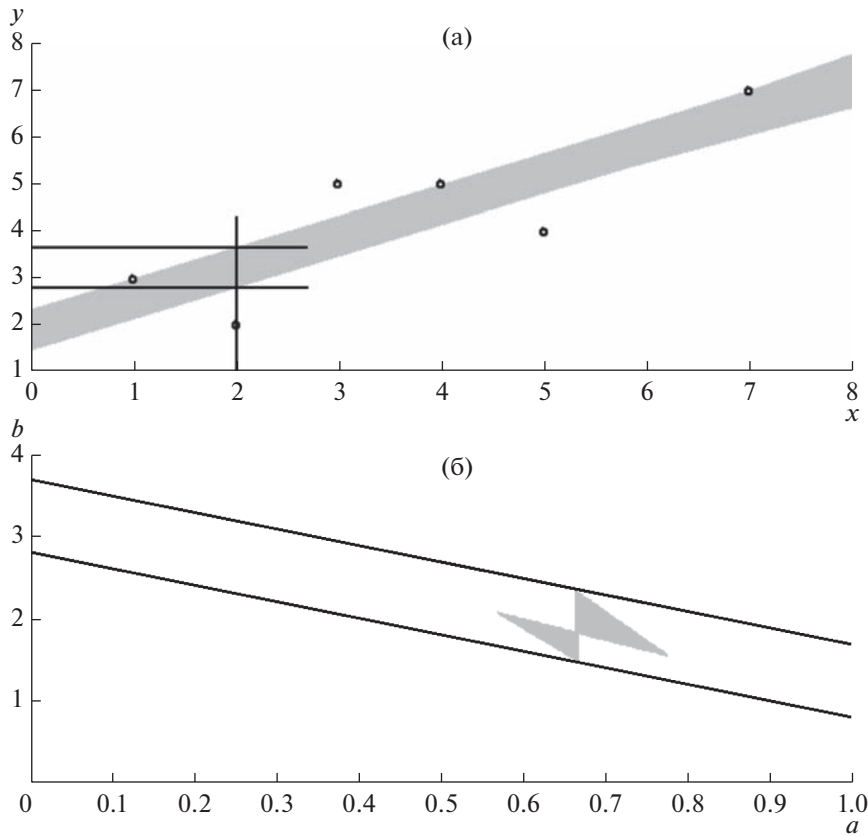
Поскольку $P^{E\Delta} \supset P^E$, имеет место включение $U^{E\Delta} \subseteq U^E$.

Пример 3. Пусть исходные данные (1.1) заданы в табл. 1. На фиг. 4 представлены связки линий аппроксимации и множество $U^{E\Delta}$, а также линии регрессии (3.2)–(3.4) и исходные точки.

ПН-средние $G^{E\Delta}(B(a))$ представляют собой отрезок $[\alpha, \beta]$, границы которого легко вычисляются по формулам (см. [7]):

$$\alpha = \frac{1}{2} \min_{p \in H} (b_{(p)} + b_{(n+1-p)}), \quad \beta = \frac{1}{2} \max_{p \in H} (b_{(p)} + b_{(n+1-p)}),$$

где $H = \{1, 2, \dots, h\}$, а h – целая часть числа $(n+1)/2$. Еще одно полезное свойство ПН-средних при информации $E\Delta$ – устойчивость к изменению положения точек множества $B(a)$. Поскольку при небольшом изменении параметра a на $a + \varepsilon$, положение точек b_i множества $B(a + \varepsilon)$ меняется относительно точек $B(a)$ на величину εx_i , то, согласно [7], максимальное изменение границ множества $G^{E\Delta}(B(a + \varepsilon))$ по сравнению с $G^{E\Delta}(B(a))$ не превышает $\varepsilon \max_{i \in N} |x_i|$. Из этого следует, что



Фиг. 5. Графоаналитический метод определения $Y^{E\Delta}(x)$ при $x = 2$.

функции границ $\alpha(a)$ и $\beta(a)$ должны быть непрерывными. Графики этих функций приведены на фиг. 4.

На фиг. 4 множество $U^{E\Delta}$ связно, и поэтому множества $Y^{E\Delta}(x)$ являются числовыми промежутками.

Рассмотрим графоаналитический метод определения $Y^{E\Delta}(x)$. На фиг. 5б представлено множество $U^{E\Delta}$ и графики прямой $b = -x^0a + y$ для точки $x^0 = 2$ в двух крайних положениях, которые определяют приближенные значения $b_{\min} = 2.85$ и $b_{\max} = 3.7$, которым соответствует $Y^{E\Delta}(2) = [2.85, 3.7]$. Это согласуется с результатом, который можно получить при помощи фиг. 5а, если построить вертикальную прямую $x = 2$ и рассечь ею связку прямых.

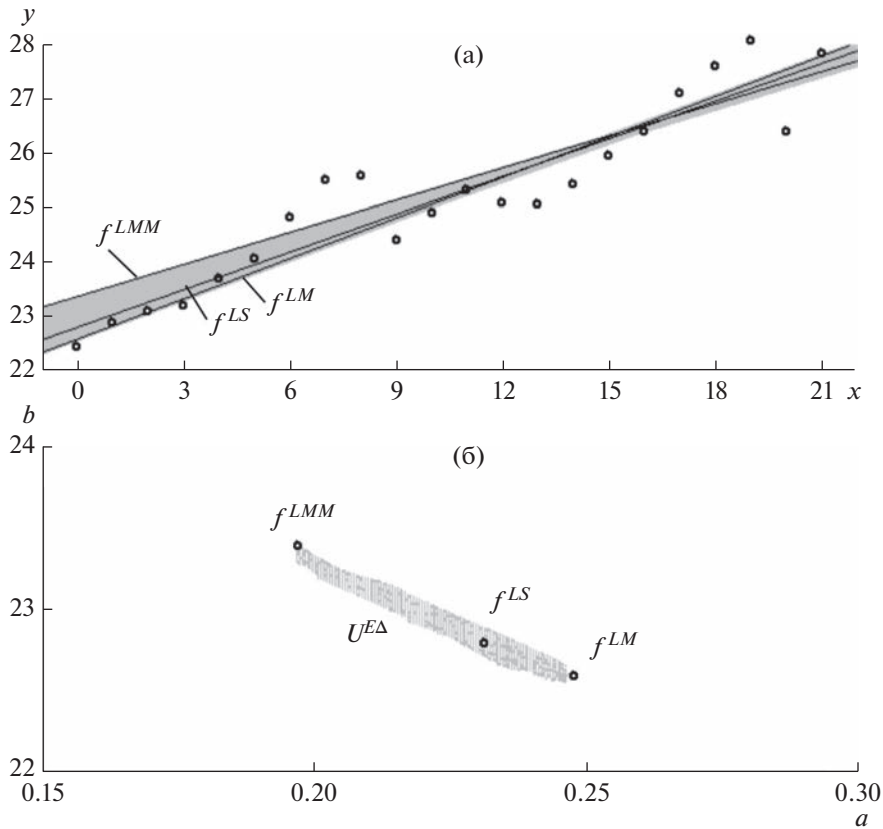
Рассмотрим более сложный пример с реальными данными.

Пример 4. В табл. 2 представлены исходные данные (1.1) по ВВП на душу населения в странах Евросоюза за 2000–2021 гг. [16]. Нумерация годов x_i для удобства начинается с 0, т.е. $x_1 = 0$ соответствует 2000 г. Величины y_i заданы в тысячах евро на человека.

Здесь уравнения аппроксимирующих функций, определяемые МНК, МНМ и МНММ, соответственно, таковы:

Таблица 2. Исходные данные по ВВП на душу населения в Евросоюзе

x_i	0	1	2	3	4	5	6	7	8	9	10
y_i	22.45	22.90	23.09	23.21	23.71	24.06	24.82	25.51	25.58	24.41	24.90
x_i	11	12	13	14	15	16	17	18	19	20	21
y_i	25.32	25.10	25.06	25.42	25.95	26.41	27.11	27.62	28.07	26.39	27.83



Фиг. 6. Связка прямых и множество $U^{E\Delta}$.

$$f^{LS}(x) \approx 0.2313x + 22.7953;$$

$$f^{LM}(x) \approx 0.2478x + 22.5944;$$

$$f^{LMM}(x) \approx 0.197x + 23.3885.$$

При расчетах для аппроксимации множества $U^{E\Delta}$ использовалась сетка с шагом 0.0005 по параметру a и с шагом 0.005 по параметру b , покрывающая область $[0.15, 0.3] \times [22, 24]$ для параметров (a, b) . Связка аппроксимирующих функций представлена в области $[-1, 22] \times [22, 28]$ для (x, y) на фиг. 6а, а множество точек $U^{E\Delta}$ представлено в области $[0.15, 0.3] \times [22, 24]$ для (a, b) на фиг. 6б.

По разбросу значений параметра a примерно от 0.2 до 0.25 можно судить о ежегодном среднем приращении показателя ВВП на данном периоде времени.

7. О ВЗАИМОСВЯЗИ НЕДОМИНИРУЕМЫХ И КЛАССИЧЕСКИХ АППРОКСИМИРУЮЩИХ ФУНКЦИЙ

Поскольку информация $E\Delta$ уточняет информацию E , то, как уже указывалось выше, верно $P^\emptyset \subset P^E \subset P^{E\Delta}$, и поэтому имеют место включения $U^\emptyset \supseteq U^E \supseteq U^{E\Delta}$, а соответствующие связи аппроксимирующих функций последовательно входят друг в друга. Это хорошо видно при сравнении фиг. 1, 3 и 4.

Интересен также вопрос о том, всегда ли графики функций f^{LS} , f^{LM} и f^{LMM} , полученные классическими методами – МНК, МНМ и МНММ – входят в связи прямых для $\Gamma = \emptyset$, $\Gamma = E$ и $\Gamma = E\Delta$, т.е. когда справедливы включения $f^{LS} \in F^\emptyset$, $f^{LM} \in F^E$ и $f^{LMM} \in F^{E\Delta}$.

Ответ на этот вопрос дает

Теорема. Справедливы соотношения $f^{LS} \in F^{E\Delta} \subseteq F^E \subseteq F^\emptyset, f^{LM} \in F^E \subseteq F^\emptyset$ и $f^{LMM} \in F^\emptyset$. Если функция f^{LM} (соответственно функция f^{LMM}) единственна, то $f^{LM} \in F^{E\Delta}$ (соответственно $f^{LMM} \in F^E$).

Доказательство. Вначале заметим, что функция $g(z)$ называется убывающей (невозрастающей) по P^Γ (на Z), если из $z' P^\Gamma z''$ следует $f(z') < f(z'')$ (соответственно $f(z') \leq f(z'')$). Известно следующее утверждение о монотонных функциях (см., например, [9, лемма 1.2.2]), которое применительно к рассматриваемому случаю можно сформулировать в следующем виде.

Утверждение. Пусть $Y \subseteq Z$ и невозрастающая по P^Γ функция $g(z)$ достигает минимума на Y в точке z^0 . Для того чтобы точка z^0 была недоминируемой по P^Γ , достаточно выполнения одного из следующих условий: $g(z)$ возрастает по P^Γ ; z^0 – единственная точка минимума $g(z)$ на Y .

Рассмотрим функцию $g(z|s) = \sum_{i=1}^n z_i^s$, где $s > 0, z \in \text{Re}_+^n = [0, +\infty)^n$. Очевидно, что из $z' P^\emptyset z''$, как и из $z' P^E z''$, следует $g(z'|s) < g(z''|s)$. Из свойств функций, выпуклых по Шуру (с учетом того, что меньшие значения компонент z_i предпочтительнее больших) [15], следует, что при $s > 1$ из $z' P^E z''$, следует $g(z'|s) < g(z''|s)$, а при $s = 1$ лишь $g(z|s) \leq g(z''|s)$. Следовательно, поскольку минимизация $g(z|1)$ при $z_i = |\delta_i|$ и $g(z|2)$ при $z_i = \delta_i$ лежит в основе МНК и МНМ, то $f^{LS} \in F^{E\Delta}$ и $f^{LM} \in F^E$; а если f^{LM} единственна, то $f^{LM} \in F^{E\Delta}$. Функция $g^-(z) = \max_{i \in N} z_i$ при $z_i = |\delta_i|$ не возрастает по P^\emptyset , а потому и по P^E , и по $P^{E\Delta}$. Следовательно, если она единственна, то $f^{LMM} \in F^{E\Delta}$. Теорема доказана.

Теорему иллюстрируют примеры 5–7 с “нетипичными” исходными данными.

Пример 5. Исходные данные (1.1) заданы табл. 3.

Здесь уравнения аппроксимирующих функций, определяемые МНК, МНМ и МНММ, соответственно, таковы:

$$f^{LS}(x) = 1.8x - 2.6; \quad f^{LM}(x) \equiv 1; \quad f^{LMM}(x) = 2.25x - 4.625.$$

Связка прямых для $F^{E\Delta}$, а также три прямые – графики всех аппроксимирующих функций – представлены на фиг. 7. Все функции f^{LS}, f^{LMM} и f^{LM} единственны. Все три соответствующие прямые входят в связку прямых, порождаемых недоминируемыми аппроксимирующими функциями, и соответствуют вершинам множества $U^{E\Delta}$.

Пример 6. Исходные данные (1.1) заданы табл. 4.

Здесь множество $F^{E\Delta}$ одноэлементно: $F^{E\Delta} = \{y = 2\}$, и

$$f^{LM}(x) \equiv l, \text{ где } l \in [1, 3], \text{ а также } f^{LM}(x) = x \text{ и } f^{LM}(x) = 5 - x; f^{LS}(x) \equiv f^{LMM}(x) \equiv 2.$$

Функции f^{LS} и f^{LMM} единственны и равны, а f^{LM} не единственна. Справедливы соотношения $f^{LS} \in F^{E\Delta}, f^{LMM} \in F^{E\Delta}$, и в $F^{E\Delta}$ входит только $f^{LM}(x) = l$ при $l = 2$.

Пример 7. Исходные данные (1.1) заданы табл. 5.

Множество $F^{E\Delta}$ одноэлементно: $F^{E\Delta} = \{y = 2\}$, и

$$f^{LS}(x) \equiv 2; f^{LM}(x) = kx + 2(1 - k), \text{ где } k \in [-1, 1]; f^{LMM}(x) = kx + 2 - k, \text{ где } k \in [-1, 1].$$

Таблица 3. Исходные данные для задачи 5

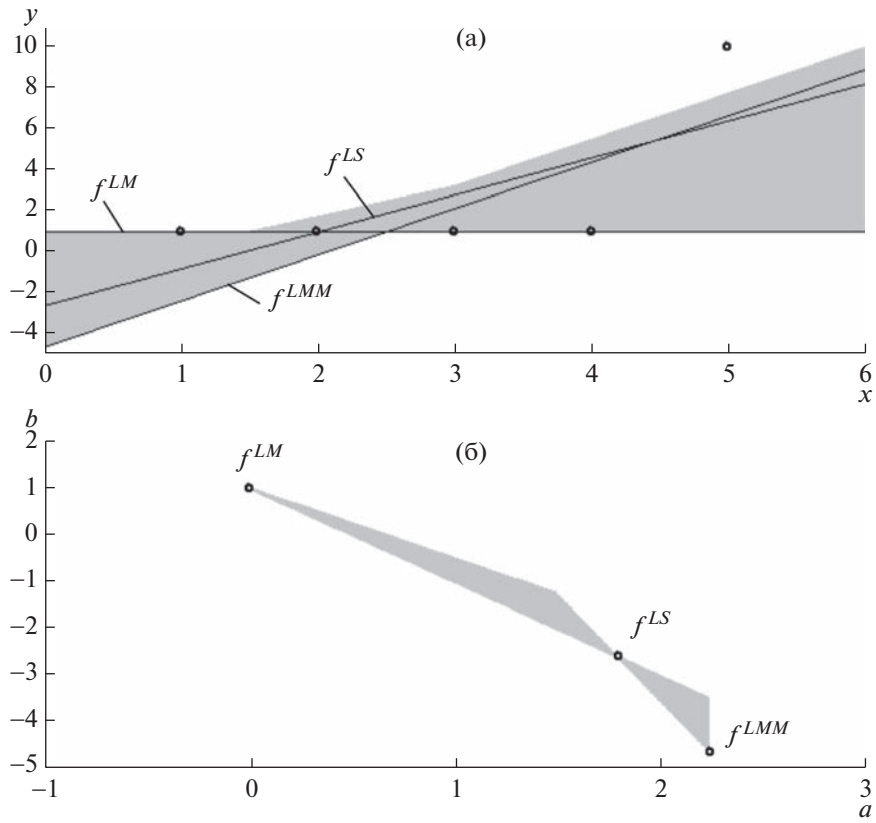
x_i	1	2	3	4	5
y_i	1	1	1	1	10

Таблица 4. Исходные данные для задачи 6

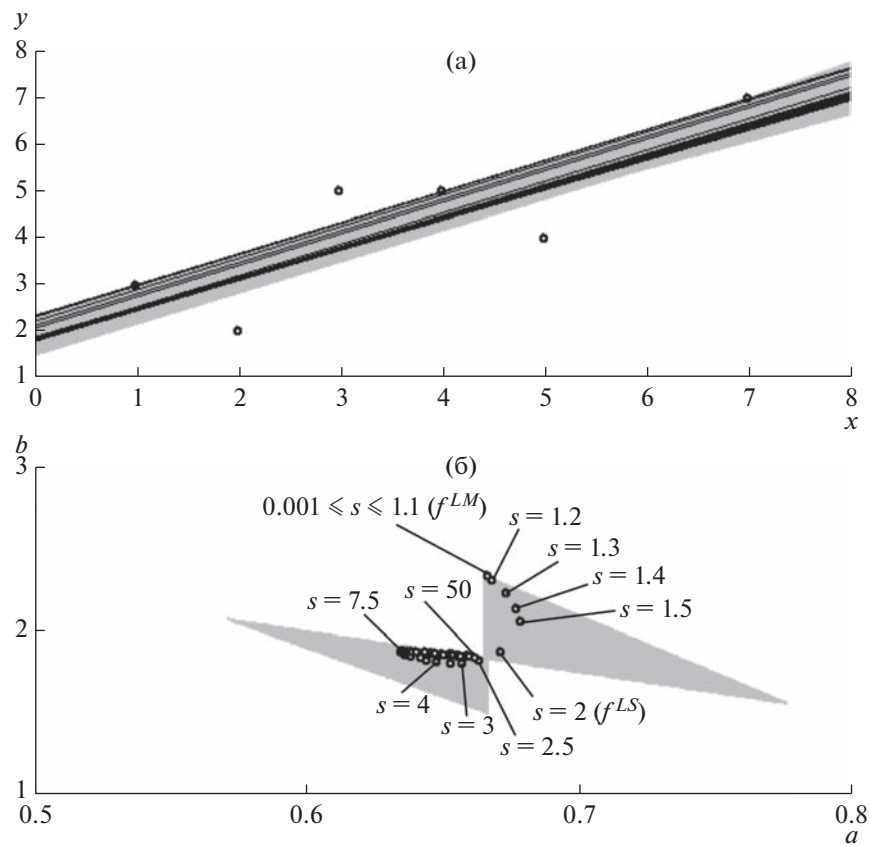
x_i	1	2	3	4
y_i	1	3	3	1

Таблица 5. Исходные данные для задачи 7

x_i	1	2	1
y_i	1	2	3



Фиг. 7. Связка прямых и множество $U^{E\Delta}$.



Фиг. 8. Связка прямых, множество $U^{E\Delta}$ и решения, определяемые (7.1).

Функции f^{LM} и f^{LMM} не единственны. Здесь $f^{LS} \in F^{E\Delta}$ и лишь $f^{LM}(x) \equiv 2$ и $f^{LMM}(x) \equiv 2$ входят в $F^{E\Delta}$ (т.е. при $k = 0$).

Рассмотрим еще один пример, связанный с теоремой и ее доказательством.

Пример 8. Рассмотрим, как расположены решения (a, b) , определяемые условиями

$$\sum_{i=1}^n |\delta_i^s| \rightarrow \min_{a,b} \quad (7.1)$$

при различных значениях показателя $s > 0$ для исходных данных (1.1) из табл. 1. Используем сетку с шагом 0.0005 по параметру a и с шагом 0.005 по параметру b , покрывая область $[0.5, 0.8] \times [1, 3]$. Получившиеся точки оказываются внутри множества $U^{E\Delta}$, как показано на фиг. 8 (ср. с фиг. 4). Это – иллюстрация соответствующего утверждения из доказательства теоремы.

8. ЗАКЛЮЧЕНИЕ

Множественный подход к определению параметров аппроксимации выделяет не одну прямую линию аппроксимации, а связку таких линий. Множественность выделяемых значений параметров аппроксимации – плата за то, что при многокритериальной постановке задачи их оценивания не делается никаких дополнительных ограничительных допущений о характере (свойствах) исходных данных (например, о законе распределения их вероятностей). Получающуюся при предлагаемом подходе многозначную аппроксимирующую функцию можно интерпретировать двояко: либо как совокупность возможных однозначных функций аппроксимации, либо как множество функций, характеризующих аппроксимацию в целом.

За единственную линию аппроксимации, особенно в случае равноважных критериев со шкалой первой порядковой метрики, можно взять геометрическое место точек – центров тяжести (центроидов) множеств $Y^\Gamma(x)$.

Актуальной остается проблема разработки эффективных вычислительных методов построения множеств недоминируемых значений параметров аппроксимации U^Γ , а также множеств $Y^\Gamma(x)$. Такие методы должны учитывать специфические особенности отношений P_Γ для различных видов информации Γ .

Предложенный подход очевидным образом обобщается на случай аппроксимации функций нескольких переменных, но вычислительные проблемы из-за повышения размерности при этом существенно усложняются.

СПИСОК ЛИТЕРАТУРЫ

1. Демидович Б.П., Марон И.А., Шувалова Э.З. Численные методы анализа. Приближение функций, дифференциальные и интегральные уравнения. М.: Наука, 1967.
2. Малов С.В. Регрессионный анализ: теоретические основы и практические рекомендации. СПб.: Изд-во С.-Пб. Ун-та, 2013.
3. Мудров В.И., Кушко В.Л. Метод наименьших модулей. М.: URSS, 2013.
4. Charnes A., Cooper W.W. Management models and industrial applications of linear programming. N.Y.: Wiley, 1961.
5. Статников Р.Б., Матусов И.Б. О недопустимых, допустимых и оптимальных решениях в задачах проектирования // Проблемы машиностр. и надежности машин. 2012. № 4. С. 10–19.
6. Подиновский В.В., Нелюбин А.П. Средние величины: многокритериальный подход // Проблемы управления. 2020. № 5. С. 3–16.
7. Подиновский В.В., Нелюбин А.П. Средние величины: многокритериальный подход. II // Проблемы управления. 2021. № 2. С. 33–41.
8. Тырсин А.Н., Азарян А.А. Точные алгоритмы реализации метода наименьших модулей на основе спуска по узловым прямым // Вычисл. матем. 2017. № 4. С. 21–32.
9. Подиновский В.В., Ногин В.Д. Парето-оптимальные решения многокритериальных задач. М.: Физматлит, 1982. Изд. второе, испр. и доп., 2007.

10. Зоркальцев В.И. Октаэдрические и евклидовы проекции точки на линейное многообразие // Тр. ин-та матем. и механ. УрО РАН. 2012. № 3. С. 106–118.
11. Подиновский В.В. Многокритериальные задачи с однородными равноценными критериями // Ж. вычисл. матем. и матем. физ. 1975. № 2. С. 330–344.
12. Подиновский В.В. Идеи и методы теории важности критериев в многокритериальных задачах принятия решений. М.: Наука, 2019.
13. Нелюбин А.П., Подиновский В.В. Алгоритмическое решающее правило, использующее ординальные коэффициенты важности критериев со шкалой первой порядковой метрики // Ж. вычисл. матем. и матем. физ. 2012. № 1. С. 43–59.
14. Fishburn P.C., Willig R.D. Transfer principles in income redistribution // J. of Public Economics. 1984. V. 25. P. 323–328.
15. Marshall A.W., Olkin I. Inequalities: Theory of majorization and its applications. N.Y.: Academic press, 1979.
16. Eurostat. Real GDP per capita, 2022. https://ec.europa.eu/eurostat/databrowser/view/sdg_08_10/default/table

**ОПТИМАЛЬНОЕ
УПРАВЛЕНИЕ**

УДК 519.612

**СИНТЕЗ ОПТИМАЛЬНОЙ СИСТЕМЫ
С УСТОЙЧИВЫМИ РЕЖИМАМИ СКОЛЬЖЕНИЯ**

© 2023 г. Л. Т. Ащепков^{1,*}

¹ 690922 Владивосток, о. Русский, п. Аякс, 10, ДВФУ, Россия

*e-mail: ashchepkov@yahoo.com

Поступила в редакцию 09.09.2022 г.
Переработанный вариант 09.09.2022 г.
Принята к публикации 15.12.2022 г.

Предложен метод синтеза оптимального управления, обеспечивающего существование и устойчивость режимов скольжения системы нелинейных обыкновенных дифференциальных уравнений. В методе используется вспомогательная задача оптимального управления. Решение задачи дает искомое управление в аналитической форме. Установлена устойчивость по Ляпунову тривиального решения замкнутой и доопределенной систем. Показано применение метода к линейным и квазилинейным системам уравнений. Приведен иллюстративный пример. Библ. 11. Фиг. 1.

Ключевые слова: синтез системы, оптимальное управление, достаточные условия оптимальности, скользящий режим, устойчивость.

DOI: 10.31857/S0044466923050058, **EDN:** GFFXOS

1. ВВЕДЕНИЕ

Проблема устойчивости систем автоматического регулирования возникла в XVIII веке с развитием промышленной революции и не утратила актуальности в настоящее время. Многочисленные исследования привели к созданию ряда новых научных методов и направлений. Не останавливаясь на них подробно, отметим близкую теме статьи теорию систем с переменной структурой. В ее основе лежит идея использования порождающих устойчивые режимы скольжения разрывных обратных связей (далее — *специальных управлений*). Историю развития, результаты, методологию и библиографию можно найти в публикациях [1–3]. Отдавая должное достигнутому в теории успехам, отметим, что вопрос об общем формализованном способе построения специальных управлений остается пока открытым.

В данной статье предложен метод синтеза специального оптимального управления для нелинейной системы определенного типа, использующий аналитическое решение оригинальной вспомогательной задачи оптимального управления. Построенное этим методом управление определено и непрерывно на всем фазовом пространстве системы, за исключением некоторого многообразия. Благодаря этому процедура доопределения системы на многообразии [4] упрощается и позволяет выписать уравнения скольжения в явном виде [1]. Итоговый фазовый портрет синтезированной системы состоит из участков оптимальных траекторий, попавших из начальных точек на подмножество многообразия (*область скольжения*), и участков траекторий, лежащих в области скольжения и притягивающихся к началу координат. За счет появления в решении задачи режимов скольжения достигается еще и важная в техническом отношении малая чувствительность траекторий синтезированной системы к возмущениям [1].

Приведенные в статье результаты легко обобщаются на случай эллипсоидальной области управления. Вместе с тем их распространение на области управления с негладкими границами, например, многогранники, наталкивается на серьезные технические трудности, связанные с возможным появлением режимов скольжения на пересечениях поверхностей разрыва управления. Заметим также, что вспомогательная задача оптимального управления является обратной по отношению к “единичному” управлению [5], веденному в качестве векторного аналога релейного скалярного управления.

2. ПОСТАНОВКА ЗАДАЧИ

Рассмотрим систему обыкновенных дифференциальных уравнений в координатной форме

$$\dot{x}_i = f_i(x_1, \dots, x_n) + \sum_{k=1}^m b_{ik}(x_1, \dots, x_n)u_k, \quad i = 1, \dots, n,$$

и векторно-матричном представлении

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

где x, u – векторы фазового состояния и управляющих воздействий с координатами $x_1, \dots, x_n, u_1, \dots, u_m$ соответственно и $\dot{x} = dx/dt$. Предположим, что векторная функция $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ со свойством $f(0) = 0$ и матричная функция $B: \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$ ранга $m \leq n$ в областях определения *гладкие* (достаточное число раз дифференцируемые). Кроме того, для заданной гладкой *положительно-определенной* функции $c(x)$ ($c(x) > 0, x \neq 0, c(0) = 0$) существуют гладкая скалярная функция $w(x)$ и положительные постоянные α, β , удовлетворяющие условиям

$$\begin{aligned} f(x)'w_x(x) + c(x) &= 0, \\ c(x) &\geq \alpha w(x), \quad w(x) \geq \beta \|x\|^2, \quad x \in \mathbb{R}^n. \end{aligned} \quad (2)$$

Здесь $f'w_x = f_1'w_{x_1} + \dots + f_n'w_{x_n}$ – скалярное произведение вектора f на градиент w_x функции w , штрих – знак транспонирования и $\|x\| = (x'x)^{1/2}$ – евклидова норма вектора x .

В предположениях (2) по второй теореме Ляпунова (см., например, [6, с. 17]) тривиальное решение уравнения $\dot{x} = f(x)$ равномерно асимптотически устойчиво. Непротиворечивость условий (2) в частном случае показана в разд. 5.

Требуется обеспечить устойчивость тривиального решения системы (1) с помощью специальных ограниченных управлений.

3. ВСПОМОГАТЕЛЬНАЯ ЗАДАЧА

В связи со сказанным сформулируем задачу оптимального управления

$$\begin{aligned} J = \int_0^{\infty} (c(x(t)) + \|s(x(t))\|) dt &\rightarrow \min, \\ \dot{x} = f(x) + B(x)u, \quad x(0) = x_0, \quad x(\infty) = 0, \quad \|u\| \leq 1, \quad t \geq 0, \end{aligned} \quad (3)$$

где x_0 – произвольный фиксированный вектор, $x(\infty)$ – предел функции $x(t)$ при $t \rightarrow \infty$ и

$$s(x) = B(x)'w_x(x), \quad x \in \mathbb{R}^n, \quad (4)$$

– гладкая векторная функция со значениями в \mathbb{R}^m . Ранг матрицы $(s_x B): \mathbb{R}^n \rightarrow \mathbb{R}^{m \times m}$ на множестве R решений уравнения $s(x) = 0$ примем равным m .

Непрерывную кусочно-гладкую траекторию $x(t), t \geq 0$, и кусочно-непрерывное управление $u(t), t \geq 0$, назовем *процессом*, если пара $x(t), u(t)$ отвечает условиям (3), за возможным исключением целевого требования, и соответствующее значение функционала J конечно. Вспомогательная задача состоит в нахождении оптимального процесса с наименьшим значением целевого функционала.

Как видно из постановки задачи, близость траекторий $x(t)$ к началу координат и множеству R оценивается суммой неотрицательных несобственных интегралов

$$\int_0^{\infty} c(x(t)) dt, \quad \int_0^{\infty} \|s(x(t))\| dt.$$

Минимизация функционала J предполагает уменьшение интегралов, т.е. влечет асимптотическое приближение траекторий к началу координат и множеству R (устойчивость) или приближение к началу координат по множеству R (скольжение), что отвечает требуемой цели управления системой (1).

Обсудим корректность постановки задачи (3). Если $x_0 = 0$, то решение $x = 0, u = 0$ задачи очевидно. Следовательно, множество оптимальных процессов, отвечающих начальным точкам $x_0 \in \mathbb{R}^n$, не пусто.

Принятые предположения гарантируют также существование и единственность множества R . Действительно, если $m = n$, то R по определению состоит из единственной точки $x = 0$. При $m < n$ из свойств функций w, s следует $w_x(0) = 0, s(0) = 0$. Поскольку ранг матрицы $s_x(0)$ равен m , то без потери общности можно считать последние m столбцов этой матрицы линейно независимыми. Положим $y = (x_1, \dots, x_{n-m}), z = (x_{n-m+1}, \dots, x_n)$ и представим уравнение $s(x) = 0$ в виде $s(y, z) = 0$. По предположению, функция $s(y, z)$ гладкая по совокупности аргументов и при $y = 0, z = 0$ удовлетворяет условиям $s(0, 0) = 0, \text{rank } s_z(0, 0) = m$. Согласно теореме о неявной функции (см., например, [7, с. 299]) последние условия гарантируют существование в некоторой окрестности точки $(0, 0)$ единственной гладкой функции $z = \varphi(y)$ со свойствами $0 = \varphi(0), s(y, \varphi(y)) = 0$. В силу единственности график функции $z = \varphi(y)$ и множество R в окрестности точки $(0, 0)$ совпадают. Следовательно, при $m < n$ в окрестности начала координат существует единственное множество R , представляющее собой гладкое многообразие.

4. РЕШЕНИЕ ВСПОМОГАТЕЛЬНОЙ ЗАДАЧИ

Решение задачи (3) проведем в четыре этапа.

4.1. Нижняя оценка функционала

Следуя [8], представим целевой функционал на произвольном фиксированном процессе $x(t), u(t)$ в виде

$$J = \int_0^\infty (c(x(t)) + \|s(x(t))\|) dt + \int_0^\infty (w(x(t)))' dt - w(x(t)) \Big|_{t=0}^{t=\infty}.$$

Выполняя во втором интеграле дифференцирование по t и учитывая (3), (4), получим

$$\begin{aligned} J &= w(x_0) + \int_0^\infty (c(x(t)) + \|s(x(t))\|) dt + \int_0^\infty w_x(x(t))' (f(x(t)) + B(x(t))u(t)) dt = \\ &= w(x_0) + \int_0^\infty (\|s(x(t))\| + s(x(t))'u(t)) dt. \end{aligned}$$

Найдем нижнюю оценку функционала J . Имеем

$$J = w(x_0) + \int_0^\infty (\|s(x(t))\| + s(x(t))'u(t)) dt \geq w(x_0) + \int_0^\infty \min_{x \in \mathbb{R}^n} \min_{\|u\| \leq 1} (\|s(x)\| + s(x)'u) dt.$$

Минимизируемая на шаре $\|u\| \leq 1$ подынтегральная функция в силу неравенства Коши–Буняковского имеет точную нижнюю оценку

$$\|s(x)\| + s(x)'u \geq \|s(x)\| - \|s(x)\| \|u\| \geq 0,$$

достижимую на управлении

$$u(x) = -\frac{s(x)}{\|s(x)\|}, \quad s(x) \neq 0. \tag{5}$$

В результате нижняя оценка функционала J примет вид

$$J = w(x_0) + \int_0^\infty (\|s(x(t))\| + s(x(t))'u(t)) dt \geq w(x_0). \tag{6}$$

4.2. Структура многообразия R

Формула (5) задает непрерывное управление $u(x)$ на всем пространстве \mathbb{R}^n , кроме многообразия R . Следуя [4], доопределим на R уравнения замкнутой системы

$$\dot{x} = f(x) + B(x)u(x). \quad (7)$$

Введем обозначения

$$\dot{s}(x, u) = s_x(x)(f(x) + B(x)u) \quad (8)$$

для производной функции $s(x)$ в силу системы (1), U — для шара $\|u\| \leq 1$ и $\dot{s}(x, U)$ — для области значений производной $\dot{s}(x, u)$ на шаре U при фиксированном x . Выделим, далее, на R множества S, P , полагая

$$x \in S \Leftrightarrow (0 \in \dot{s}(x, U), s(x) = 0), \quad (9)$$

$$x \in P \Leftrightarrow (0 \notin \dot{s}(x, U), s(x) = 0). \quad (10)$$

Согласно (9) включение $x \in S$ равносильно существованию решения $v = v(x)$ уравнения $s_x(x)(f(x) + B(x)v) = 0$ и выполнению условий $v(x) \in U$, $x \in R$. Учитывая обратимость матрицы $s_x(x)B(x)$, представим эти условия в эквивалентной форме:

$$v(x) = -(s_x(x)B(x))^{-1} s_x(x)f(x), \quad \|v(x)\| \leq 1, \quad s(x) = 0. \quad (11)$$

На множестве S функция $v(x)$ непрерывна и допустима по амплитудному ограничению (3). Рассматривая $v(x)$ как продолжение управления $u(x)$ на многообразии R и следуя [4], определим на S замкнутую систему дифференциальных уравнений

$$\dot{x} = f(x) + B(x)v(x), \quad x \in S. \quad (12)$$

Из (8), (11) вытекает тождественное на S равенство $\dot{s}(x, v(x)) = 0$, поэтому исходящие из точек множества S траектории системы (12) лежат в S . Множество S будем называть *областью скольжения* и уравнения (12) — *уравнениями скольжения*.

Установим свойства множества P . В произвольной фиксированной точке $\xi \in P$ множество $\dot{s}(\xi, U) = s_x(\xi)(f(\xi) + B(\xi)U)$ выпуклое, замкнутое, ограниченное в \mathbb{R}^m и по определению (10) не содержит 0. Согласно теореме отделимости (см., например, [9, с. 25]) для множеств $\dot{s}(\xi, U)$ и 0 найдется такой зависящий от ξ ненулевой вектор $a \in \mathbb{R}^m$, что справедливо строгое неравенство $a's(\xi, U) > a'0 = 0$. В подробной записи в обозначении (8) имеем

$$a's_x(\xi)(f(\xi) + B(\xi)u) > 0, \quad u \in U. \quad (13)$$

По предположению, строки матрицы $s_x(\xi)$ линейно независимы, поэтому $a's_x(\xi) \neq 0$. Обозначим через M поверхность $a's(x) = 0$. Вектор a в неравенстве (13) определен с точностью до положительного множителя и $a's_x(x)$ есть градиент функции $a's(x)$, поэтому без потери общности считаем $a's_x(\xi)$ нормалью поверхности M в точке ξ . В этих терминах неравенство (13) означает положительность проекции любого вектора скорости $\dot{x} = f(\xi) + B(\xi)u$, $u \in U$, на нормаль $a's_x(\xi)$. В частности, если траектория $x(t)$ замкнутой системы (7) имеет в момент $t = \tau$ общую с поверхностью M точку $x(\tau) = \xi$, то проекции односторонних векторов скорости $\dot{x}(\tau \pm 0) = f(x(\tau)) + B(x(\tau))u(x(\tau \pm 0))$ в точке $x(\tau)$ на нормаль $a's_x(\xi)$ тоже положительны. Последнее возможно тогда и только тогда, когда траектория $x(t)$ пересекает поверхность M в точке ξ без односторонних касаний. Это значит, что односторонние векторы скорости $\dot{x}(\tau \pm 0)$ не лежат в касательной плоскости к поверхности M в точке ξ . Поскольку $P \subset M$ и ξ — произвольная точка P , то каждая траектория замкнутой системы (7), имеющая с P общую точку, пересекает P . В связи с этим множество P будем называть *областью прошивания*.

Итак, многообразие R состоит из двух взаимно дополняющих областей скольжения S и прошивания P , заданных соответственно условиями

$$\begin{aligned} \|(s_x(x)B(x))^{-1} s_x(x)f(x)\| \leq 1, \quad s(x) = 0; \\ \|(s_x(x)B(x))^{-1} s_x(x)f(x)\| > 1, \quad s(x) = 0. \end{aligned} \tag{14}$$

4.3. Оптимальный процесс

Вернемся к вопросу о точности нижней оценки (6). Обозначим через $x(t)$ определенную на полуоси $t \geq 0$ непрерывную траекторию замкнутой системы (7), (12) с начальным условием $x(0) = x_0$. Выделим на полуоси $t \geq 0$ множества T_1, T_2, T_3 , полагая

$$t \in T_1 \Leftrightarrow x(t) \notin R; \quad t \in T_2 \Leftrightarrow x(t) \in S; \quad t \in T_3 \Leftrightarrow x(t) \in P.$$

Если $t \in T_1$, то с использованием (2), (5) и (7) находим

$$(w(x(t)))^\bullet = w_x(x(t))'(f(x(t)) + B(x(t))u(x(t))) = -c(x(t)) - \|s(x(t))\| \leq -c(x(t)) \leq -\alpha w(x(t)).$$

При $t \in T_2$ в силу (2), (4) и (12) так же получим

$$(w(x(t)))^\bullet = w_x(x(t))'(f(x(t)) + B(x(t))v(x(t))) = -c(x(t)) + s(x(t))'v(x(t)) = -c(x(t)) \leq -\alpha w(x(t)).$$

Следовательно, неравенство

$$(w(x(\tau)))^\bullet \leq -\alpha w(x(\tau)), \quad \tau \geq 0, \tag{15}$$

верно на всей полуоси времени, за возможным исключением моментов $\tau \in T_3$ и момента первого попадания траектории $x(t)$ на S . Интегрируя неравенство (15) по τ в пределах от нуля до $t \geq 0$, приходим к оценке

$$w(x(t)) \leq w(x_0)e^{-\alpha t}, \quad t \geq 0. \tag{16}$$

Из (16) следует соотношение $x(\infty) = 0$, значит, пара $x(t), u(t)$ с управлением

$$u(t) = u(x(t)), \quad t \in T_1; \quad u(t) = v(x(t)), \quad t \in T_2, \tag{17}$$

является процессом.

Покажем, что на этом процессе нижняя оценка (6) целевого функционала достигается. Действительно, подынтегральная функция в неравенстве (6) равна нулю, во-первых, в силу определений (17) и (5) управления $u(t)$ при $t \in T_1$ и, во-вторых, в силу равенства $s(x(t)) = 0$ при $t \in T_2 \cup T_3$. По построению, множества T_1, T_2, T_3 не пересекаются и в объединении составляют всю полуось $t \geq 0$, поэтому несобственный интеграл в неравенстве (6) равен нулю. В результате соотношение (6) выполняется как точное равенство и тогда процесс $x(t), u(t)$ – оптимальный.

4.4. Устойчивость синтезированной системы

Как показано в п. 4.3, каждая определенная при $t \geq 0$ оптимальная траектория $x(t)$ замкнутой системы (7), (12), исходящая в момент $t = 0$ из произвольной точки $x_0 \in \mathbb{R}^n$, удовлетворяет оценке (16). При заданном $\beta > 0$ в силу свойств функции w (неотрицательности на \mathbb{R}^n , непрерывности и нулевом значении в начале координат) по любому $\varepsilon > 0$ найдется такое $\delta > 0$, что выполняется неравенство $0 \leq w(x_0) < \beta\varepsilon^2$, если только $\|x_0\| < \delta$. Тогда с использованием (2) и (16) получим

$$\beta \|x(t)\|^2 \leq w(x(t)) \leq w(x_0)e^{-\alpha t} < \beta\varepsilon^2, \quad t \geq 0.$$

Отсюда вытекает справедливость неравенства $\|x(t)\| < \varepsilon$ при любых $t \geq 0, \|x_0\| < \delta$, что равносильно устойчивости по Ляпунову тривиального решения системы (7), (12).

Сформулируем общие выводы.

Теорема. Если условия задачи (3) отвечают принятым в разд. 2, 3 предположениям, то при любом x_0 решение задачи существует. Оптимальное управление (5) имеет разрыв на многообразии R , состоящем из областей скольжения S и прошивания P вида (14). Залегающие на S части оптимальных траекторий описываются уравнениями скольжения (12). Управление (11) в области скольжения непрерывное и допустимое по амплитудному ограничению (3). Тривиальное решение замкнутой системы (7), (12) устойчиво по Ляпунову. Минимум целевого функционала равен $w(x_0)$.

Теорема полностью доказана приведенными выше рассуждениями.

5. ЧАСТНЫЕ СЛУЧАИ ЗАДАЧИ

5.1. Линейные стационарные системы

Рассмотрим задачу (3) с функциями

$$f(x) = Ax, \quad c(x) = x'Cx, \quad w(x) = x'Wx, \quad x \in \mathbb{R}^n, \quad (18)$$

в следующих предположениях. Квадратные матрицы A , C , W порядка n постоянные, матрица A устойчивая – действительные части всех ее собственных значений отрицательные, матрица C симметричная и положительно-определенная, матрица W удовлетворяет матричному алгебраическому уравнению Ляпунова

$$A'W + WA + C = 0. \quad (19)$$

Согласно [10] уравнение (19) задает симметричную положительно-определенную матрицу W однозначно.

Для функций (18) с указанными матрицами условия (2) выполняются. Действительно, в силу (19) квадратичная форма $w(x)$ обращает уравнение (2) в тождество на всем пространстве \mathbb{R}^n . Неравенства (2) верны при $\alpha = \min_{\|x\|=1} c(x)/w(x)$, $\beta = \min_{\|x\|=1} w(x)$.

Считая дополнительно матрицу $B(x) = B$ ранга m постоянной, приходим к стационарной задаче оптимального управления

$$J = \int_0^{\infty} (x(t)'Cx(t) + 2\|B'Wx(t)\|) dt \rightarrow \min, \quad (20)$$

$$\dot{x} = Ax + Bu, \quad x(0) = x_0, \quad x(\infty) = 0, \quad \|u\| \leq 1, \quad t \geq 0,$$

отвечающей всем условиям теоремы из разд. 4. На основании теоремы задача (20) имеет решение при любых x_0 , оптимальное управление $u(x)$ задано формулой (5) с функцией $s(x) = B'Wx$, управление (11) имеет вид

$$v(x) = -(B'WB)^{-1} B'WAx, \quad s(x) = 0,$$

тривиальное решение замкнутой системы

$$\dot{x} = Ax + Bu(x), \quad s(x) \neq 0; \quad \dot{x} = Ax + Bv(x), \quad s(x) = 0$$

экспоненциально устойчиво, минимум целевого функционала равен $x_0'Wx_0$.

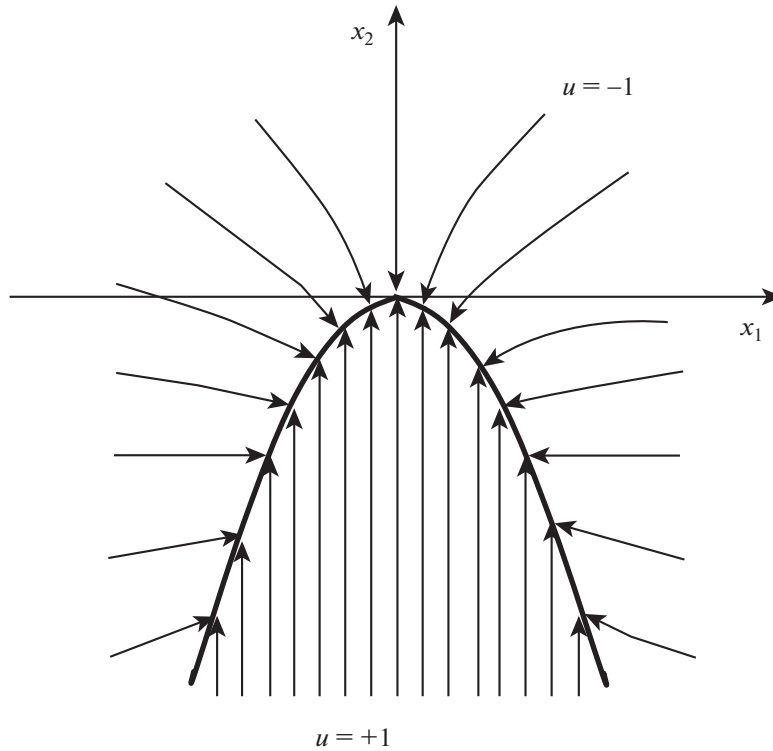
Приведенные результаты полностью совпадают с выводами [11]. Экспоненциальная устойчивость вытекает из оценки (16) и неравенства $x'Wx \leq \mu \|x\|^2$, где μ – максимум квадратичной формы $x'Wx$ на сфере $\|x\| = 1$.

5.2. Квазилинейные системы

Пусть в задаче (20) матрицы A , C , W по-прежнему отвечают предположениям п. 5.1 и переменная матрица $s_x(x)B(x)$ имеет ранг m на многообразии $s(x) = 0$ при $s(x) = B(x)'Wx$. Тогда утверждения теоремы остаются верными.

Проиллюстрируем сказанное на примере билинейной системы. Положим

$$A = \begin{pmatrix} -1 & 0 \\ 0 & -3 \end{pmatrix}, \quad B(x) = \begin{pmatrix} x_1 \\ 1 \end{pmatrix}, \quad C = \begin{pmatrix} 2 & 0 \\ 0 & 6 \end{pmatrix}, \quad W = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$



Фиг. 1. Фазовый портрет оптимальной синтезированной системы в примере. Движение фазовой точки к началу координат происходит по оси ординат или по параболе в режиме скольжения.

В соответствии с выводами теоремы находим оптимальное релейное управление $u(x_1, x_2) = -\text{sign}(x_1^2 + x_2)$, область скольжения $x_1^2 + x_2 = 0$ и уравнения скольжения

$$\begin{aligned} \dot{x}_1 &= -x_1(1 - v(x_2)), & \dot{x}_2 &= -2x_2(1 - v(x_2)), \\ v(x_2) &= x_2/(1 - 2x_2), & x_1^2 + x_2 &= 0. \end{aligned} \tag{21}$$

Область прошивания отсутствует. Картина расположения оптимальных траекторий показана на фигуре.

Выберем малое $\varepsilon > 0$ и зафиксируем произвольную точку $\xi(\tau) = (|\tau|, -\tau^2)$, $|\tau| > \varepsilon$, области скольжения. Рассмотрим две траектории, исходящие в момент $t = 0$ из точки $\xi(\tau)$: пробную траекторию $y(t) = (y_1(t), y_2(t))$, $t \geq 0$, исходной системы уравнений при $u = 0$ и оптимальную траекторию $x(t) = (x_1(t), x_2(t))$, $t \geq 0$, описываемую уравнениями скольжения (21).

Сравним траектории $x(t), y(t)$ по переходным периодам t_1, t_2 – соответствующим моментам первого попадания на границу квадрата $[-\varepsilon, \varepsilon] \times [-\varepsilon, \varepsilon]$, $\varepsilon < |\tau|$. Для пробной траектории верны тождества

$$\dot{y}_1(t)/y_1(t) \equiv -1, \quad \dot{y}_2(t)/y_2(t) \equiv -3, \quad t \geq 0, \tag{22}$$

и уравнения

$$|y_1(\theta_1)| = \varepsilon, \quad y_2(\theta_2) = -\varepsilon \tag{23}$$

имеют единственные положительные корни θ_1, θ_2 . Проинтегрируем тождества (22) по t на отрезках $[0, \theta_1]$ и $[0, \theta_2]$ соответственно с использованием начальных значений и равенств (23). Получим

$$\theta_1 = \ln(|\tau|/\varepsilon) > \theta_2 = \ln(\tau^2/\varepsilon)^{1/3}.$$

Следовательно, $t_2 = \theta_1 = \ln(|\tau|/\varepsilon)$. Таким же способом находим переходный период t_1 оптимальной траектории

$$t_1 = t_2 - \frac{1}{6} \ln \frac{1 + 3\tau^2}{1 + 3\varepsilon^2}, \quad |\tau| > \varepsilon.$$

Из полученной формулы вытекает, что для любых τ , $|\tau| > \varepsilon$, оптимальная траектория имеет меньший переходный период, чем пробная траектория, и разность $t_2 - t_1$ неограниченно возрастает с увеличением $|\tau|$. Кроме того, в силу точной нижней оценки (6) оптимальная траектория оказывается лучше пробной еще и по целевому функционалу.

В общем случае пробная траектория системы (1) (при управлении $u = 0$) оптимальна во вспомогательной задаче (3) в том и только в том случае, если полностью лежит в области скольжения. Этот вывод непосредственно следует из точной нижней оценки (6) целевого функционала.

СПИСОК ЛИТЕРАТУРЫ

1. *Уткин В.И.* Системы с переменной структурой: состояние, проблемы, перспективы // Автоматика и телемехан. 1983. № 9. С. 5–25. *Utkin V. I.* Variable structure systems: present and future // Automat. Remote Control. 1983. V. 44. № 9. P. 1105–1120.
2. *Ferrara A., Incremona G.P., Cucuzzella M.* Advanced and optimization based sliding mode control. Philadelphia: SIAM, 2019.
3. *Steinberger M., Horn M., Fridman L.M.* Variable-structure systems and sliding-mode control. Berlin: Springer, 2020.
4. *Филиппов А.Ф.* Дифференциальные уравнения с разрывной правой частью // Матем. сб. 1960. Т. 51 (93). № 1. С. 99–128.
5. *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. № 9. P. 1671–1680.
6. *Афанасьев В.Н., Колмановский В.Б., Носов В.Р.* Математическая теория конструирования систем управления. М.: Высш. шк., 2003.
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. *Карманов В.Г.* Математическое программирование. М.: Наука, 1986.
10. *Параев Ю.И.* Уравнения Ляпунова и Риккати. Томск: Томский ун-т, 1989.
11. *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.

ОПТИМАЛЬНОЕ
УПРАВЛЕНИЕ

УДК 517.98

ОПТИМИЗАЦИЯ МНОЖЕСТВА ДОСТИЖИМОСТИ ЛИНЕЙНОЙ СИСТЕМЫ ПО ОТНОШЕНИЮ К ДРУГОМУ МНОЖЕСТВУ¹⁾

© 2023 г. М. В. Балашов^{1,*}, Р. А. Камалов¹

¹ 117997 Москва, Профсоюзная ул., 65, Институт проблем управления РАН им. В.А. Трапезникова, Россия

*e-mail: balashov73@mail.ru

Поступила в редакцию 02.11.2022 г.
Переработанный вариант 21.11.2022 г.
Принята к публикации 02.02.2023 г.

Рассматриваются задача максимально быстрого по времени выполнения включения во множество достижимости линейной управляемой автономной системы некоторого выпуклого компакта, а также задача поиска максимального времени, при котором выполнено включение множества достижимости в некоторый выпуклый компакт. При этом ищутся начальная точка и время, для которых экстремальное время в соответственной задаче реализуется. Рассмотрена дискретизация задачи на сетке единичных векторов и с помощью сведения к задаче линейного программирования получены приближенное решение задачи, а также оценки погрешности решения. Задача объединяет общая идеология, восходящая к задаче поиска чебышёвского центра. Библ. 24. Фиг. 4. Табл. 4.

Ключевые слова: множество достижимости, равномерная выпуклость, условие непустой внутренности, многозначный интеграл, линейное программирование, аппроксимация в метрике Хаусдорфа.

DOI: 10.31857/S004446692305006X, **EDN:** PJPZUE

1. ВВЕДЕНИЕ И ОСНОВНЫЕ ОБОЗНАЧЕНИЯ

1.1. Введение

Рассмотрим линейную управляемую систему

$$x' \in Ax + U, \tag{1}$$

где $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$, $0 \in U \subset \mathbb{R}^n$ – выпуклый компакт. Пусть $M(t)$ – множество достижимости этой системы с начальным условием $x(0) = 0$. Напомним, что для всякого $t \geq 0$ множество достижимости $M(t)$ для системы (1) есть

$$M(t) = \{x(t) \in \mathbb{R}^n : x'(s) = Ax(s) + u(s) \text{ п.в. } s \in [0, t], u(s) \in U - \text{измеримый селектор}\}. \tag{2}$$

Хорошо известно в силу результата Ляпунова о векторных мерах (см. [1], а также [2], теорема 1, гл. 2, § 2.2, что множество достижимости выпукло и замкнуто. Легко получить представление множества достижимости для случая $x(0) = 0$ в виде интеграла

$$M(t) = \int_0^t e^{As} U ds. \tag{3}$$

Если в системе (1) $x(0) = x_0 \in \mathbb{R}^n$, то множество достижимости системы (1) с начальным условием $x(0) = x_0$ задается выражением

$$M(x(0), t) = e^{At} x_0 + \int_0^t e^{As} U ds = e^{At} x_0 + M(t). \tag{4}$$

¹⁾Результаты разделов 2, 3 получены М.В. Балашовым за счет гранта Российского научного фонда № 22-11-00042, <https://rscf.ru/project/22-11-00042/> в ИПУ РАН.

Многозначный интеграл здесь и далее мы будем понимать в смысле интеграла Аумана (см. [3])

$$\int_0^t e^{As} U ds = \left\{ \int_0^t e^{As} u(s) ds : u(s) \in U - \text{измеримый селектор} \right\}.$$

Будем понимать $M(+\infty)$ как предел в метрике Хаусдорфа интегралов

$$\lim_{t \rightarrow +\infty} \int_0^t e^{As} U ds = \lim_{t \rightarrow +\infty} M(t),$$

если таковой существует.

Включение $0 \in U$ обеспечивает монотонность множеств достижимости: $M(t_1) \subset M(t_2)$ при $t_1 \leq t_2$. Также заметим, что в силу аддитивности интеграла Аумана получаем

$$M(t_2) = M(t_1) + \int_{t_1}^{t_2} e^{As} U ds, \quad t_1 \leq t_2.$$

Отметим, что включение $0 \in U$ является техническим. Действительно, если $0 \notin U$, то для произвольной точки $u_0 \in U$, определив $U_0 = U - u_0$, имеем

$$\int_0^t e^{As} U ds = \int_0^t e^{As} U_0 ds + \int_0^t e^{As} u_0 ds.$$

Сформулируем рассматриваемые задачи.

Задача 1. Пусть $M_0 \subset \mathbb{R}^n$ – выпуклый компакт. Мы хотим найти минимальное время $t \geq 0$ и такую точку $x \in \mathbb{R}^n$, что множество $x + M(t)$ содержит M_0 . Иными словами, ищется решение задачи

$$t \rightarrow \min_{t \geq 0, x \in \mathbb{R}^n} x + M(t) \supset M_0. \quad (5)$$

Пусть $t = T \geq 0$ и $x = x_0$ – решение задачи (5). Легко видеть, что если рассмотреть траектории системы $x' \in Ax + U$ с начальным условием $x(0) = e^{-AT} x_0$, то T есть минимальное время, за которое в силу (4), множество $M(e^{-AT} x_0, T) = x_0 + M(T)$ будет содержать множество M_0 . Иными словами, если известно решение (T, x_0) задачи 1, то гарантируется следующее утверждение: при начальном условии $x(0) = e^{-AT} x_0$ нам требуется время не более T , чтобы попасть траекториями системы (1) в любую точку множества M_0 . При этом для любого другого начального условия $x(0) \in \mathbb{R}^n$ найдется точка множества M_0 , для достижения которой при любом выборе управления $u(t) \in U$ в системе (1) потребуется время не меньше T .

Действительно, предположим, что есть вектор начальных условий $x(0)$ такой, что для времени $t_0 < T$ выполнено включение $M(x(0), t_0) = e^{At_0} x(0) + M(t_0) \supset M_0$. Тогда для $x = e^{At_0} x(0)$ выполнено включение $x + M(t_0) \supset M_0$, что противоречит минимальности T в (5).

Таким образом, задача 1 имеет смысл задачи быстрогодействия. По ее решению (T, x_0) строится начальное условие $x(0) = e^{-AT} x_0$ для системы (1) и определяется минимальное время $T > 0$, в течение которого решения системы могут достичь любой точки множества M_0 . Для любого другого начального условия результат по времени будет не меньше.

Задача 2. Другая близкая постановка к задаче (5) – задача

$$t \rightarrow \max_{t \geq 0, x \in \mathbb{R}^n} x + M(t) \subset M_0. \quad (6)$$

Пусть $t = T \geq 0$ и $x = x_0$ – решение задачи (6). Легко видеть, что концы траекторий системы $x' \in Ax + U$ с начальным условием $x(0) = e^{-AT} x_0$ будут в момент времени T в силу (4) содержаться

во множестве M_0 , т.е. $x(T) \in M_0$ для всех решений $x(\cdot)$. При этом указанное время максимально возможное по всем другим начальным условиям $x(0) \in \mathbb{R}^n$ в системе (1).

Действительно, предположим, что есть вектор начальных условий $x(0)$ такой, что для времени $t_0 > T$ выполнено включение $\mathcal{M}(x(0), t_0) = e^{At_0}x(0) + M(t_0) \subset M_0$. Тогда для $x = e^{At_0}x(0)$ выполнено включение $x + M(t_0) \subset M_0$, что противоречит максимальнойности T в (6).

Если (T, x_0) – решение задачи 2, то вопрос о включении $\mathcal{M}(e^{-AT}x_0, s) = e^{A(s-T)}x_0 + M(s) \subset M_0$ для промежуточных значений $s \in (0, T)$ должен решаться отдельно. В общем случае для некоторых $s < T$ предыдущее включение может быть неверно (см. [4]). Приведем простейший пример системы без управления. Пусть $e^{At} = e^{-at} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$, $a > 0$ – малое число. Пусть $(x, y)^T \in \mathbb{R}^2$ – вектор начальных условий, $x, y > 0$ и M_0 есть круг радиуса $r > 0$ с центром в нуле, содержащий (x, y) . Тогда $\|e^{At}(x, y)^T\|^2 = e^{-2at}((x+ty)^2 + y^2)$, и при $t = \frac{1}{a}$ значение $\|e^{At}(x, y)^T\|^2$ имеет порядок $e^{-2} \frac{y^2}{a^2}$ при $a \rightarrow +0$. Число $e^{-2} \frac{y^2}{a^2}$ больше r^2 при малых $a > 0$, однако решение асимптотически устойчиво и тем самым при больших t попадает в любую окрестность нуля, в частности, и в M_0 .

Заметим, что для произвольного начального условия $x(0) \in M_0$ существует решение $x(\cdot)$ системы (1) со свойством $x(t) \in M_0$ при всех $t \in [0, T]$ тогда и только тогда, когда $(Ax + U) \cap T_B(M_0, x) \neq \emptyset$ для всех $x \in M_0$ (см. [5], гл. 4, теорема 1, утверждение 1, с. 180, где $T_B(M_0, x)$ – верхний касательный конус ко множеству M_0 в точке $x \in M_0$). Результаты о локальной по времени выживаемости всех траекторий (1) во множестве M_0 можно найти в [6], Theorem 1.12.

Задача 2 имеет связь с частным случаем хорошо известной линейной задачи быстрогодействия (см. [7–10], [11, гл. 2, § 7, теорема 2.20]). Пусть требуется найти оптимальное управление и минимальное время, при котором некоторая траектория системы (1) с начальным условием $x(0) = 0$ попадет на заданное терминальное множество N_0 . То есть решается задача

$$t \rightarrow \min_{t \geq 0} M(t) \cap N_0 \neq \emptyset. \quad (7)$$

Пусть также множество N_0 является каверной, т.е. N_0 есть замыкание множества $\mathbb{R}^n \setminus M_0$, где M_0 – телесный выпуклый компакт. С учетом эквивалентности условий $M(t) \cap \text{int } N_0 = \emptyset$ и $M(t) \subset M_0$, приходим к постановке $t \rightarrow \max$ при условии $M(t) \subset M_0$. В случае $M(t_1) \subset \text{int } M(t_2)$ при всех $t_1 < t_2$ последняя задача эквивалентна (7). Добавив переменную x , мы получаем задачу 2. Отметим, что если в задаче 2 зафиксировать $x = 0$, то получится задача, эквивалентная (7) для множества N_0 – каверны.

Если U в задаче (1) есть множеством помех, то решение задачи 2 дает нам максимальное время T и начальное условие $x(0) = e^{-AT}x_0$, гарантирующее нахождение концов всех траекторий системы при реализации любой помехи в допустимом подмножестве M_0 фазового пространства в момент времени T . При этом T максимально по всем начальным условиям $x(0) \in \mathbb{R}^n$, т.е. для любого другого начального условия $x(0) \in \mathbb{R}^n$ выполнение включения $\mathcal{M}(x(0), t) \subset M_0$ влечет $t \leq T$. Кроме того, для всякого $t \in (0, T)$ начальное условие $x(0) = e^{-At}x_0$ обеспечивает включение $\mathcal{M}(e^{-At}x_0, t) \subset M_0$. Действительно, с учетом (4)

$$\mathcal{M}(e^{-At}x_0, t) = e^{At}e^{-At}x_0 + M(t) \subset x_0 + M(T) \subset M_0.$$

Эти свойства имеют важное значение для различных приложений, когда все траектории системы (1) должны попасть в заданное подмножество M_0 фазового пространства в заданный момент времени t .

Отличительной чертой задач 1 и 2 является оптимизация не только по времени, но и по пространственной переменной x . При этом в постановке задач 1 и 2 время и фазовая переменная разведены в разные слагаемые, что позволяет эффективно применять простые алгоритмы с вы-

пуклыми множествами для решения указанных задач. Возможны и другие постановки задач. Ниже мы будем ссылаться на задачу 1 по ссылке (5), а на задачу 2 по ссылке (6).

Заметим, что задача (5) может не иметь решения. Например, решения нет в случае, когда спектр $\sigma(A) < 0$ и $M(+\infty) \stackrel{*}{=} M_0 = \{x \in \mathbb{R}^n : x + M_0 \subset M(+\infty)\} = \emptyset$. В свою очередь задача (6) может быть разрешима при любом $T \geq 0$ в том смысле, что для всякого $T \geq 0$ найдется $x \in \mathbb{R}^n$ такой, что $x + M(T) \subset M_0$. Такая ситуация возможна при условии $\sigma(A) < 0$ и $M_0 \stackrel{*}{=} M(+\infty) \neq \emptyset$. Ниже мы будем предполагать, что решения задач (5) и (6) существуют, строго положительны по времени и конечны.

Идея предлагаемого алгоритма состоит в дискретизации множеств $M(t)$, M_0 , заданных через опорные функции, на сетке единичных векторов. Кроме того, мы также будем рассматривать дискретизацию по времени $0 = t_0 < t_1 < \dots < t_k < \dots < t_K = T_0$, где параметр $T_0 > 0$ задает максимальное время, до которого мы ищем решение.

В настоящей работе будет применяться аппроксимация множеств $M(t)$, M_0 с помощью опорных функций (см. детали в [12], [13]). Хорошо известно (см. [14], табл. 1), что разумная аппроксимация выпуклых компактов на сетке единичных векторов с помощью неравенств с опорной функцией реализуется в небольшой размерности: $n \leq 5$ на современном персональном компьютере. Поэтому рассмотренные в работе алгоритмы имеют практический смысл в пространстве небольшой размерности.

Существуют также подходы к исследованию множеств достижимости со специальными множествами, например, множествами-зонотопами (см. [15]) и эллипсоидальной техникой (см. [16]). Однако аппроксимация с помощью опорных функций до сих пор является одной из наиболее распространенных.

Задача (5) очень похожа на задачу поиска чебышёвского центра множества или на более общую задачу о накрытии выпуклого компакта гомотетичным образом другого выпуклого компакта (см. [17]): для выпуклых компактов $A, B \subset \mathbb{R}^n$, $0 \in B$, требуется решить задачу

$$t \rightarrow \min_{t \geq 0, x \in \mathbb{R}^n} \text{ при условии } x + tB \supset A. \quad (8)$$

В [17] получены достаточные условия для единственного решения этой задачи и оценки погрешности между точным и приближенным решениями в метрике Хаусдорфа. В задаче (5) появляется еще один непрерывный параметр — время t . Поэтому здесь мы рассмотрим дискретизацию по пространственной переменной, описывая множества на языке опорных функций, и по времени. При этом мы используем идеи оценок погрешности по пространству из [17]. С помощью аналогичного подхода решается задача (6).

1.2. Основные обозначения и вспомогательные факты

Через \mathbb{R}^n будем обозначать вещественное евклидово пространство n измерений со скалярным произведением (x, y) для всех $x, y \in \mathbb{R}^n$. Обозначим через $B_R(x)$ замкнутый шар с центром x радиуса $R > 0$.

Напомним, что суммой Минковского—Понтрягина множеств $A, B \subset \mathbb{R}^n$ называется множество $A + B = \{a + b : a \in A, b \in B\}$, а геометрической разностью, или разностью Минковского—Понтрягина, множеств $A, B \subset \mathbb{R}^n$ называется множество

$$A \stackrel{*}{=} B = \{x \in \mathbb{R}^n : x + B \subset A\} = \bigcap_{x \in B} (A - x).$$

Расстояние от точки $x \in \mathbb{R}^n$ до множества $A \subset \mathbb{R}^n$ обозначим через $\varrho_A(x) = \inf_{a \in A} \|x - a\|$. Расстоянием в метрике Хаусдорфа между компактными множествами $A, B \subset \mathbb{R}^n$ называется величина

$$h(A, B) = \max \left\{ \max_{x \in A} \varrho_B(x), \max_{x \in B} \varrho_A(x) \right\}.$$

Полунормой компакта $A \subset \mathbb{R}^n$ называется число $\|A\| = h(A, \{0\}) = \max_{a \in A} \|a\|$.

Опорной функцией множества $A \subset \mathbb{R}^n$ в точке $p \in \mathbb{R}^n$ называется $s(p, A) = \sup_{a \in A} (p, a)$. Применяя теорему об отделимости, получаем, что для любого выпуклого компакта A выполняется равенство

$$A = \{x \in \mathbb{R}^n : (p, x) \leq s(p, A), \|p\| = 1\}.$$

Отметим, что для множества $M(t) = \int_0^t e^{As} U ds$ легко вычислить опорную функцию при условии, что опорная функция компактного множества управлений U может быть точно вычислена аналитически, либо с помощью простого алгоритма, например, симплекс-метода, для всякого $p \in \mathbb{R}^n$. Тогда в силу линейности многозначного интеграла

$$s(p, M(t)) = \int_0^t s(p, e^{As} U) ds = \int_0^t s(e^{A^T s} p, U) ds. \tag{9}$$

Для выпуклого замкнутого множества A нормальным конусом в точке $x \in A$ называется множество $N(A, x) = \{p \in \mathbb{R}^n | (p, x - a) \geq 0 \forall a \in A\}$. Определим также $N_1(A, x) = N(A, x) \cap \partial B_1(0)$. Через $\text{co } A$, $\text{cone } A$ будем обозначать соответственно выпуклую и коническую (иногда называют выпуклую коническую) оболочку множества $A \subset \mathbb{R}^n$. С учетом теоремы Каратеодори

$$\text{co } A = \left\{ \sum_{k=1}^{n+1} \lambda_k a_k : \lambda_k \in [0, 1], \sum_{k=1}^{n+1} \lambda_k = 1, a_k \in A \right\}, \quad \text{cone } A = \left\{ \sum_{k=1}^n \lambda_k a_k : \lambda_k \geq 0, a_k \in A \right\}.$$

Обозначим через $P_A x$ метрическую проекцию точки $x \in \mathbb{R}^n$ на выпуклое замкнутое множество $A \subset \mathbb{R}^n$. Для $x, y \in \mathbb{R}^n$ определим $(x, y) = \{(1 - \lambda)x + \lambda y : \lambda \in (0, 1)\}$.

Будем говорить, что выпуклое компактное множество $A \subset \mathbb{R}^n$ равномерно выпукло, если существует строго возрастающая функция $\delta_A : [0, \text{diam } A] \rightarrow [0, +\infty)$, $\delta_A(0) = 0$, модуль равномерной выпуклости, такая, что для всех $x, y \in A$ выполнено включение $\frac{1}{2}(x + y) + B_{\delta_A(\|x-y\|)}(0) \subset A$ (см. [18], гл. 3). Рассмотрим равномерно выпуклые множества A, B и их сумму $C = A + B$. Если $c_1, c_2 \in C$, $\|c_1 - c_2\| = t > 0$, то $c_i = a_i + b_i$, $a_i \in A, b_i \in B, i = 1, 2$. При этом либо $\|a_1 - a_2\| \geq \frac{1}{2}t$, либо $\|b_1 - b_2\| \geq \frac{1}{2}t$. Отсюда следует равномерная выпуклость суммы C с модулем

$$\delta_C(t) \geq \min \left\{ \delta_A \left(\frac{1}{2}t \right), \delta_B \left(\frac{1}{2}t \right) \right\}. \tag{10}$$

В силу теоремы Дзя–Нордлендера (см. [18, гл. 3, § 3]) максимальный модуль равномерной выпуклости среди центрально-симметричных тел диаметра $2R$ имеет евклидов шар, для него $\delta_{B_R(0)}(t) = R - \sqrt{R^2 - \frac{t^2}{4}} \geq \frac{t^2}{8R}$. Отметим наконец, что в \mathbb{R}^n класс строго и равномерно выпуклых компактов совпадает. Это легко вытекает из определения модуля выпуклости и соображений компактности.

Множество A называется R -сильно выпуклым (или сильно выпуклым с радиусом R) (см. [12], [19]), если оно может быть представлено в виде пересечения замкнутых шаров радиуса R . Легко видеть, что для сильно выпуклого множества $A \subset \mathbb{R}^n$ с радиусом R модуль выпуклости не меньше, чем для шара радиуса R при всех допустимых значениях аргумента t , т.е. $\delta_A(t) \geq \frac{t^2}{8R}$. Известно (см. [12], следствие 2), что для R_i -сильно выпуклых множеств $A_i, i = 1, 2$, множество $A_1 + A_2$ является $(R_1 + R_2)$ -сильно выпуклым.

Сеткой $\mathbb{G} = \{p_i\}_{i=1}^I$ мы будем называть набор единичных векторов p_i . Будем говорить, что сетка \mathbb{G} имеет *мелкость* $\Delta \in \left(0, \frac{1}{2}\right)$ (см. [12, определение 6], а также [20, § 2.6], [22]), если для всякого единичного вектора $p \in \mathbb{R}^n$ существуют множество $I_p \subset \{1, \dots, I\}$ и числа $\alpha_i > 0, i \in I_p$, такие, что

$$p = \sum_{i \in I_p} \alpha_i p_i, \quad \|p_i - p_j\| < \Delta \quad \forall i, j \in I_p.$$

Мы будем рассматривать внешнюю *многогранную аппроксимацию* выпуклого компактного множества $A \subset \mathbb{R}^n$ на сетке \mathbb{G} вида

$$\hat{A} = \{x \in \mathbb{R}^n \mid (p_i, x) \leq s(p_i, A), p_i \in \mathbb{G}\}. \quad (11)$$

Предложение 1 (см. [23], теорема 2.2). Пусть выпуклый компакт $A \subset \mathbb{R}^n$ является равномерно выпуклым с модулем $\delta_A(\cdot)$. Пусть \mathbb{G} – сетка мелкости $\Delta \in \left(0, \frac{1}{2}\right)$ и $\frac{\Delta}{4 - \Delta^2} < \frac{\delta_A(\text{diam} A)}{\text{diam} A}$. Тогда

$$h(A, \hat{A}) \leq \frac{8}{7} \varepsilon(\Delta) \Delta,$$

где $\varepsilon(\cdot)$ – решение уравнения $\frac{\delta_A(\varepsilon)}{\varepsilon} = \frac{\Delta}{4 - \Delta^2}$.

Если A – произвольный выпуклый компакт, то в качестве $\varepsilon(\Delta)$ в предложении 1 нужно взять диаметр множества A .

Для сильно выпуклого множества с радиусом $R > 0$ погрешность аппроксимации уточняется в следующем предложении.

Предложение 2 (см. [22], замечание 1). Пусть $R > 0$, \mathbb{G} – сетка мелкости $\Delta \in \left(0, \frac{1}{2}\right)$ и $M = \{x \in \mathbb{R}^n \mid (p_i, x) \leq 1, p_i \in \mathbb{G}\}$. Тогда $B_1(0) \subset M$ и

$$h(B_R(0), RM) \leq R \frac{\Delta^2}{1 - \frac{1}{2} \Delta^2}. \quad (12)$$

Если множество $B \subset \mathbb{R}^n$ является R -сильно выпуклым, то погрешность аппроксимации задается такой же формулой: $h(B, \hat{B}) \leq R \frac{\Delta^2}{1 - \frac{1}{2} \Delta^2}$.

В силу теоремы Дзя–Нордлендера $\varepsilon(\Delta)$ в типичных ситуациях имеет порядок $\Delta^\alpha, \alpha \in (0, 1]$.

Нам также понадобится один комбинаторный результат, доказанный в [17].

Предложение 3 (см. [17], лемма 1). Пусть $A, B \subset \mathbb{R}^n$ – такие выпуклые компакты, что $B \stackrel{*}{=} A = \{x_0\}$. Определим $D = (A + x_0) \cap \partial B$ и

$$P = \bigcup_{x \in D} N_1(B, x).$$

Тогда $0 \in \text{co} P$. Заметим, что по теореме Каратеодори $0 \in \text{co}\{p_i\}_{i=1}^m, m \leq n + 1, p_i \in P$.

2. ЗАДАЧА (5), АЛГОРИТМ И ПОГРЕШНОСТЬ

Дискретный вариант задачи (5) имеет вид

$$t_i \rightarrow \min_{i, x \in \mathbb{R}^n} x + \hat{M}(t_i) \supset \hat{M}_0, \quad (13)$$

где $0 = t_0 < t_1 < \dots < t_K = T_0$ – некоторое разбиение отрезка $[0, T_0]$. Для выпуклого компактного множества M_0 через \hat{M}_0 , как и выше, будем обозначать его сеточную аппроксимацию на сетке \mathbb{G} . Договоримся решать задачу на отрезке времени $[0, T_0]$. Если при выбранном T_0 решение не найдено, то либо увеличиваем T_0 , либо возвращаем null.

Легко видеть, что если система неравенств

$$(p_j, x) + s(p_j, M(t_i)) \geq s(p_j, M_0) \quad \forall p_j \in \mathbb{G} \tag{14}$$

несовместна, то истинное решение задачи (5) $T > t_i$.

Пусть t_k – минимальное время из дискретного набора, для которого система (14) совместна, а \hat{x}_0 – некоторое решение системы (14).

Лемма 1. Пусть $\varepsilon > 0$ выбрано так, что $\varepsilon \geq h(\hat{M}(t_k), M(t_k))$ и $\int_{t_{k-1}}^{t_k} e^{As} U ds \subset B_\varepsilon(0)$, где t_k – минимальное время, для которого система (14) разрешима. Тогда для любого решения \hat{x}_0 задачи (14) для $i = k$ истинное решение T задачи (5) удовлетворяет включению

$$-\hat{x}_0 \in (M(T) + B_{2\varepsilon}(0)) \overset{*}{\subset} M_0. \tag{15}$$

Доказательство. Пусть \hat{x}_0 – решение (14) для $i = k$. Тогда

$$-\hat{x}_0 \in \hat{M}(t_k) \overset{*}{\subset} \hat{M}_0 \subset (M(t_k) + B_\varepsilon(0)) \overset{*}{\subset} M_0. \tag{16}$$

Как было показано выше, $T > t_{k-1}$. Если $T \in (t_{k-1}, t_k]$, то из включения (16) получаем

$$-\hat{x}_0 \in \left(M(T) + \int_T^{t_k} e^{As} U ds + B_\varepsilon(0) \right) \overset{*}{\subset} M_0 \subset (M(T) + B_{2\varepsilon}(0)) \overset{*}{\subset} M_0. \tag{17}$$

Если $T > t_k$, то $M(t_k) \subset M(T)$ и

$$-\hat{x}_0 \in (M(t_k) + B_\varepsilon(0)) \overset{*}{\subset} M_0 \subset (M(T) + B_\varepsilon(0)) \overset{*}{\subset} M_0. \tag{18}$$

Объединяя формулы (17) и (18), получаем, что в любом случае выполнена формула (15). Лемма 1 доказана.

Включение $\int_{t_{k-1}}^{t_k} e^{As} U ds \subset B_\varepsilon(0)$ легко проверить. Для его справедливости надо, чтобы шаг по времени $\Delta t_k = t_k - t_{k-1}$ удовлетворял условию $\Delta t_k \leq \frac{\varepsilon}{\mu \|U\|}$, где $\mu = \max_{s \in [0, t]} \|e^{As}\|$, а $\|U\| = \max_{u \in U} \|u\|$.

Теорема 1. Пусть (T, x_0) – решение задачи (5) и $T < +\infty$. Пусть выполнены условия $\varepsilon \geq h(\hat{M}(t_k), M(t_k))$ и $\int_{t_{k-1}}^{t_k} e^{As} U ds \subset B_\varepsilon(0)$, где t_k – наименьшее время из дискретного набора, при котором верно (14).

1. Пусть множество $M(T)$ равномерно выпукло с модулем $\delta_{M(T)}$. Тогда

$$\delta_{M(T)+B_{2\varepsilon}(0)}(\|x_0 - \hat{x}_0\|) \leq 2\varepsilon.$$

Заметим, что множество $M(T) + B_{2\varepsilon}(0)$ равномерно выпукло (см. (10)).

2. Определим $D = (M_0 - x_0) \cap \partial(M(T))$ и $P = \bigcup_{x \in D} N_1(M(T), x)$. Пусть $\delta > 0$ и $B_\delta(0) \subset \text{co } P$. Тогда $\|x_0 - \hat{x}_0\| \leq \frac{2\varepsilon}{\delta}$.

Доказательство. Доказательство повторяет доказательство теорем 1 и 2 из [17].

Прежде всего покажем, что решение (T, x_0) единственно в условиях пунктов 1 и 2 теоремы.

Пусть множество $M(T)$ равномерно строго выпукло. Если допустить, что есть две точки x_1 и x_2 , для которых $x_i + M(T) \supset M_0$, $i = 1, 2$, то в силу строгой выпуклости $M(T)$ выполнено включение

ние $\frac{x_1 + x_2}{2} + \text{int } M(T) \ni x$ для всякой точки $x \in M_0$. В силу компактности M_0 имеем $\frac{x_1 + x_2}{2} + \text{int } M(T) \supset M_0$. Это противоречит минимальности T .

Пункт 2 теоремы также гарантирует единственность решения (T, x_0) . Действительно, если вектор $v \neq 0$, то найдется вектор $p \in P$ такой, что $(p, v) > 0$. От противного, если для всех $p \in P$ верно $(p, v) \leq 0$, то для любой выпуклой комбинации $\sum_{i=1}^{n+1} \lambda_i p_i$ для любых $p_i \in P$ выполнено $(\sum_{i=1}^{n+1} \lambda_i p_i, v) \leq 0$. Отсюда $\sup_{p \in \text{co } P} (p, v) \leq 0$, т.е. 0 – граничная точка $\text{co } P$. Получили противоречие с условием $B_\delta(0) \subset \text{co } P$. Итак, пусть $p \in P$ и $(p, v) > 0$. Если $x \in D$ и $p \in N(M(T), x)$, то $(p, x + v) = (p, x) + (p, v) > s(p, M(T))$. Таким образом, сдвиг множества $M - x_0$ на любой вектор $v \neq 0$ выходит из $M(T)$, так как $x \in M - x_0$ и $x + v \notin M(T)$.

1. Без ограничения общности будем считать, что $x_0 = 0$. Фиксируем произвольное решение $\hat{x}_0 \in \mathbb{R}^n$ задачи (13), (14). По лемме 1 выполнено включение (15). Это означает, что для любого $x \in D = M_0 \cap \partial(M(T))$ выполнено включение $x - \hat{x}_0 \in M(T) + B_{2\varepsilon}(0)$.

В силу предложения 3 найдутся точки $x_i \in D$, векторы $p_i \in N_1(M(T), x_i)$ и числа $\lambda_i > 0$, $1 \leq i \leq m \leq n + 1$ такие, что $\sum_{i=1}^m \lambda_i = 1$ и $\sum_{i=1}^m \lambda_i p_i = 0$. Отсюда $\sum_{i=1}^m \lambda_i (-\hat{x}_0, p_i) = 0$. Следовательно, найдется индекс $i \in \{1, \dots, m\}$ такой, что $(p_i, -\hat{x}_0) \geq 0$, $p_i \in N_1(M(T), x_i)$, $x_i \in D$. Зафиксируем данное i и соответствующие x_i, p_i .

Пусть $H^+ = \{x \in \mathbb{R}^n : (p_i, x) \geq (p_i, x_i)\}$. Тогда $x_i - \hat{x}_0 \in H^+ \cap (M(T) + B_{2\varepsilon}(0)) =: S$, причем гиперплоскость $\partial H^+ + 2\varepsilon p_i$ опорная ко множеству $M(T) + B_{2\varepsilon}(0)$ в точке $x_i + 2\varepsilon p_i$.

Итак, вписанный шар максимального радиуса во множестве $M(T) + B_{2\varepsilon}(0)$ с центром в сегменте S имеет радиус не более 2ε (2ε – расстояние между параллельными гиперплоскостями ∂H^+ и $\partial H^+ + 2\varepsilon p_i$). Поэтому расстояние от x_i до $x_i - \hat{x}_0$, т.е. $\|\hat{x}_0\|$, оценивается из условия $2\varepsilon \geq \delta_{M(T)+B_{2\varepsilon}(0)}(\|\hat{x}_0 - 0\|)$. Пункт 1 доказан.

2. Снова, не ограничивая общности, можно считать $x_0 = 0$. Фиксируем $\tau \in (0, 1)$. Пусть $P_1 \subset P$ – такое конечное множество, что $h(P, P_1) < (1 - \tau)\delta$. Тогда $B_{\tau\delta}(0) \subset \text{co } P_1$.

Пусть L – грань множества $\text{co } P_1$, $L = \text{co}\{p_j\}_{j=1}^J$, $p_j \in P_1$. Заметим, что $\text{co}(L \cup \{0\})$ есть n -мерная пирамида с боковыми ребрами длины 1, вершиной 0 и основанием L . Определим $e = P_{\text{aff } L} 0$, $\|e\| \geq \tau\delta$ ($\text{aff } L$ – аффинная оболочка L).

Угол φ между векторами e и p_j для всякого $j = 1, \dots, J$ можно оценить как $\cos \varphi = \frac{\|e\|}{\|p_j\|} = \|e\| \geq \tau\delta$. Через $\angle(a, b)$ будем обозначать угол между ненулевыми векторами $a, b \in \mathbb{R}^n$.

Если решение задачи (13), (14) $\hat{x}_0 \in \mathbb{R}^n$ удовлетворяет включению $-\hat{x}_0 \in \text{cone}\{p_j\}_{j=1}^J = \text{cone } L$, то найдется такой номер j , $1 \leq j \leq J$, что $\angle(-\hat{x}_0, p_j) \leq \varphi$. Покажем это. Пусть $w = \{-t\hat{x}_0 \mid t \geq 0\} \cap \text{aff } L$ и $R = \sqrt{1 - \|e\|^2} = \|p_j - e\|$ для любого j . Пусть точка w – относительно внутренняя точка симплекса $\text{co}\{p_j\}_{j=1}^m$, $m \leq n$, и $a_j = p_j - w$, $1 \leq j \leq m$. Так как w относительно внутренняя точка, то найдутся $\lambda_j > 0$, $\sum_{j=1}^m \lambda_j = 1$, для которых $\sum_{j=1}^m \lambda_j p_j = w$ или $\sum_{j=1}^m \lambda_j a_j = 0$. Отсюда $\sum_{j=1}^m \lambda_j (a_j, w) = 0$ и, значит, найдется номер j , для которого $(a_j, w) \geq 0$.

Для этого номера j в треугольнике $p_i 0w$ $\angle(p_i - w, -w) \geq \frac{1}{2}\pi$ и тем самым $\|p_i - w\|^2 + \|e\|^2 \leq \|p_i - w\|^2 + \|w\|^2 \leq 1$, т.е. $r = \|p_i - w\| \leq R$.

По теореме косинусов из треугольника $p_i 0w$ косинус угла $\gamma = \angle(-\hat{x}_0, p_i) = \angle(w, p_i)$ равен

$$\cos \gamma = \frac{1 + \|w\|^2 - r^2}{2\|w\|} \geq \frac{1 + \|w\|^2 - R^2}{2\|w\|} = \frac{\|w\|^2 + \|e\|^2}{2\|w\|} \geq \|e\| = \cos \varphi.$$

Итак, $\angle(-\hat{x}_0, p_i) \leq \varphi$.

Рассмотрим соответствующую точку $x_i \in D = M_0 \cap \partial(M(T))$, для которой $p_i \in N_1(M(T), x_i)$. Для краткости обозначим $z_0 = x_i, p_0 = p_i$.

Пусть $H^+ = \{x \in \mathbb{R}^n \mid (p_0, x - z_0) \geq 0\}$. В силу формулы (15) точка $z_0 - \hat{x}_0$ лежит между плоскостями ∂H^+ и $\partial H^+ + 2\epsilon p_0$, и угол между $-\hat{x}_0$ и p_0 не более φ . Отсюда

$$\|\hat{x}_0\| \leq \frac{2\epsilon}{\cos \angle(-\hat{x}_0, p_0)} \leq \frac{2\epsilon}{\cos \varphi} \leq \frac{2\epsilon}{\tau\delta}.$$

Устремляя $\tau \rightarrow 1 - 0$, получаем утверждение теоремы. В заключение отметим, что условие непустой внутренности $B_\delta(0) \subset \text{co } P$ в п. 2 теоремы является ключевым. Именно оно гарантирует выполнение линейной оценки по ϵ в п. 2. Теорема доказана.

Опорную функцию множества достижимости $M(t)$ для любого единичного вектора $p \in \mathbb{R}^n$ легко вычислить по формуле (9). Если U – любое компактное множество, для которого точно вычисляется опорная функция $s(p, U)$, то функция $s(p, e^{As}U) = s(e^{A^T s} p, U)$ точно вычисляется для каждого s и является липшицевой по s . Поэтому проблем с ее интегрированием по отрезку $[0, t]$ не возникает. В частности, в рассмотренных ниже примерах погрешность вычисления соответствующего интеграла может быть сделана сколь угодно малой.

Замечание 1. Напомним, что во введении мы предположили, что решение задачи (5) существует и реализуется за конечное время. Достаточно, чтобы нашлось $t > 0$, для которого $M(t) \stackrel{*}{=} M_0 \neq \emptyset$.

Для строгой=равномерной выпуклости $M(T)$ достаточно, чтобы множество U удовлетворяло условию: для любого единичного вектора $p \in \mathbb{R}^n$ опорное подмножество $(e^{As}U)(p)$ одноточечно для п.в. $s \in [0, T]$. При этом опорное множество $M(T)(p) = \int_0^T (e^{As}U)(p) ds$ одноточечно для любого $\|p\| = 1$, и, значит, множество $M(T)$ строго выпукло. При этом множество $M(T)$ не обязательно сильно выпукло. Например, для системы с нулевой динамикой $x' \in 0 \cdot x + U = U, x(0) = 0, U$ – строго, но не сильно выпуклый компакт (например, U – шар ℓ_α^n -нормы при $\alpha > 2$).

Вопрос о модуле равномерной выпуклости $M(T)$ для конкретных матриц A и множеств U должен решаться отдельно и в настоящей работе не рассматривается. Отметим, что если $M(T)$ сильно выпукло с радиусом R , оценка $\|x_0 - \hat{x}_0\|$ в п. 1 имеет порядок $\sqrt{(R + \epsilon)\epsilon} \asymp R\Delta$ по радиусу и мелкости сетки.

Заметим, что и для нестрого выпуклого множества U (например, отрезка) множество $M(T)$ может быть строго-равномерно, но не сильно выпуклым. Соответствующий пример будет рассмотрен в приложении 2.

В п. 2 теоремы 1 требуется выполнение включения $B_\delta(0) \subset \text{co } P$, которое априори бывает трудно проверить. Тем не менее, если из каких-то соображений это включение установлено, то, как показано в примерах, возможно оценить погрешность алгоритма. Иногда проверить указанное условие можно после решения на сетке задачи-аппроксимации. Ниже приводится такой пример.

Пример 1. Рассмотрим пример, в котором обсудим ряд условий, позволяющих установить включение $0 \in \text{int co } P$ в теореме 1 по решению задачи (13). Пусть $D = \{d_i\}_{i=1}^{n+1} \subset \mathbb{R}^n$ и $M_0 = \text{co } D$ –

полноразмерный симплекс. Допустим, что сетка \mathbb{G} выбирается так, что $\hat{M}_0 = M_0$. Предположим, что некоторое решение \hat{x}_0, t_k задачи (13) удовлетворяет условиям $\hat{x}_0 = 0, M_0 \cap \partial \hat{M}(t_k) = D$. Предположим, что внутренность $M(t_k)$ содержит шар радиуса $r > 0$ и $r > 10\epsilon$. Потребуем строгую выпуклость $M(T)$.

Определим грани $S_i = \text{co}(D \setminus \{d_i\})$ для всех $1 \leq i \leq n + 1$.

Тогда в обозначениях теоремы 1 по лемме 1

$$\hat{M}(t_k) \subset M(t_k) + B_\epsilon(0) \subset M(T) + B_{2\epsilon}(0), \quad M'_k := \hat{M}(t_k) \overset{*}{\subset} B_{2\epsilon}(0) \subset M(T). \quad (19)$$

Предположим, что для каждого i найдется вектор $v_i \neq 0$ такой, что $S_i + v_i \subset M'_k \overset{*}{\subset} B_{2\epsilon}(0)$. Последнее условие означает, что сетка \mathbb{G} имеет достаточно малую мелкость, а симплекс M_0 имеет ребра меньшие, чем характерный размер M'_k . Отметим, что проверка существования векторов v_i сводится к нахождению ненулевых векторов $v = v_i$, удовлетворяющих системе линейных неравенств $s(p, S_i) + (p, v) + 4\epsilon \|p\| \leq s(p, \hat{M}(t_k))$ для всех $p \in \mathbb{G}$ и может быть решена с помощью линейного программирования. Покажем, что если требуемые векторы v_i существуют, то $(x + M_0) \cap \partial M(T) = D$ для некоторого $x \in \mathbb{R}^n$.

Действительно, найдется вектор $x \in \mathbb{R}^n$ такой, что выполнено включение $x + M_0 \subset M(T)$, а значит, для всякого i имеют место включения $x + S_i \subset M(T), S_i + v_i \subset M'_k \overset{*}{\subset} B_{2\epsilon}(0) \subset M(T) \overset{*}{\subset} B_{2\epsilon}(0) \subset \text{int } M(T)$. При этом, в силу непрерывности операции сложения векторов, можно выбрать $v_i \neq x$ и сохранить включение $S_i + v_i \subset M'_k \overset{*}{\subset} B_\epsilon(0) \subset \text{int } M(T)$. Следовательно, $S_i + (x, v_i] \subset \text{int } M(T)$ для всех $1 \leq i \leq n + 1$. Если для некоторого i точка $x + d_i$ является внутренней для множества $M(T)$, то найдется $\lambda \in (0, 1)$ такое, что $(S_i + \lambda v_i + (1 - \lambda)x) \cup (d_i + \lambda v_i + (1 - \lambda)x) \subset \text{int } M(T)$. Это противоречит минимальности T . Утверждение $(x + M_0) \cap \partial M(T) = D$ доказано. Также из строгой выпуклости $M(T)$ и замечания 1 вытекает, что $M(T) \overset{*}{\subset} M_0 = \{x\}$. Далее можно считать $x = 0$, заменив M_0 на $M_0 + x$.

Проанализируем множество $P = \bigcup_{1 \leq i \leq n+1} N_1(M(T), d_i)$ и установим выполнение условия $0 \in \text{int } \text{co } P$.

В силу предложения 3 $0 \in \text{co } P$. Предположим, что $0 \in \partial \text{co } P$. Тогда по теореме Каратеодори найдутся n векторов $p_i \in N_1(M(T), d_i)$ (без ограничения общности с номерами $1 \leq i \leq n$) таких, что $0 \in \text{co}\{p_i\}_{i=1}^n$. Рассмотрим S_{n+1} . Как показано выше, это множество можно сдвинуть на вектор $v = v_{n+1} \neq 0: S_{n+1} + v \subset \text{int } M(T)$. Значит, $S_{n+1} + (0, v] \subset \text{int } M(T)$. Так как для всякого $1 \leq i \leq n$ точка $d_i + v \in \text{int } M(T)$, то $(p_i, v) < 0$. Тем самым для любой нетривиальной выпуклой комбинации $\sum_{i=1}^n \lambda_i p_i$ векторов $\{p_i\}_{i=1}^n$ (где $\lambda_i \geq 0, \sum_{i=1}^n \lambda_i > 0$) имеем $(\sum_{i=1}^n \lambda_i p_i, v) < 0$. Значит, выпуклая комбинация $\sum_{i=1}^n \lambda_i p_i$ не может равняться нулю. Противоречие показывает, что $0 \in \text{int } \text{co } P$.

Оценим теперь $|T - t_k|$. С одной стороны, $T > t_{k-1}$. Для оценки T сверху предположим, что $T > t_k$. Тогда

$$\emptyset \neq M(T) \overset{*}{\subset} M_0 = \left(M(t_k) + \int_{t_k}^T e^{As} U ds \right) \overset{*}{\subset} M_0,$$

причем T минимально возможное. По формуле (16) $(M(t_k) + B_\epsilon(0)) \overset{*}{\subset} M_0 \neq \emptyset$, откуда $T \leq \tau$, где $\tau > t_k$ – минимально возможное число, для которого

$$\int_{t_k}^{\tau} e^{As} U ds \supset B_\epsilon(0). \quad (20)$$

Условие (20) достаточно трудно для проверки и, в частности, оно может не выполняться. В реальных задачах это условие может проверяться численно. В случае $0 \in \text{int } U$ можно дать оценку сверху на T через ε .

Лемма 2. Пусть $\mu = \min_{\|h\|=1, t \in [0, T_0]} \|e^{At} h\|$ и $B_d(0) \subset U$ для некоторого $d > 0$. Тогда если $t_k + \frac{\varepsilon}{\mu d} \leq T_0$, то

$$t_{k-1} < T \leq t_k + \frac{\varepsilon}{\mu d}.$$

Доказательство. Из определения μ получаем

$$\int_{t_k}^{\tau} e^{As} U ds \supset \int_{t_k}^{\tau} \mu B_d(0) ds = (\tau - t_k) \mu d \cdot B_1(0).$$

Поэтому $(\tau - t_k) \mu d \leq \varepsilon$ и $T \leq \tau \leq t_k + \frac{\varepsilon}{\mu d}$.

Некоторые оценки типа (20) для плоского случая содержатся в приложении 1.

3. ЗАДАЧА (6), АЛГОРИТМ И ПОГРЕШНОСТЬ. ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Рассмотрим теперь задачу (6). Напомним, что мы ищем решение t на отрезке $[0, T_0]$. Дискретный вариант задачи (6) имеет вид

$$t_i \rightarrow \max_{x \in \mathbb{R}^n} x + \hat{M}(t_i) \subset \hat{M}_0, \tag{21}$$

все обозначения имеют тот же смысл, что и в разд. 2. Включение $x + \hat{M}(t_i) \subset \hat{M}_0$ имеет вид системы

$$(p_j, x) + s(p_j, M(t_i)) \leq s(p_j, M_0) \quad \forall p_j \in \mathbb{G}. \tag{22}$$

Лемма 3. Пусть t_k – решение (21), и \hat{x}_0 – некоторое решение системы (22). Определим $\varepsilon > 0$ так, что $\varepsilon \geq h(M_0, \hat{M}_0)$ и $\int_{t_k}^{t_{k+1}} e^{As} U ds \subset B_\varepsilon(0)$. Тогда для $i = k$ и для истинного решения T задачи (6) выполнено включение

$$\hat{x}_0 \in (M_0 + B_{2\varepsilon}(0)) \overset{*}{\subset} M(T). \tag{23}$$

Доказательство. Пусть \hat{x}_0 – решение (21) для $i = k$. С учетом равенства $M(t_{k+1}) = M(t_k) + \int_{t_k}^{t_{k+1}} e^{As} U ds \subset M(t_k) + B_\varepsilon(0)$ имеем

$$\begin{aligned} \hat{x}_0 \in \hat{M}_0 \overset{*}{\subset} \hat{M}(t_k) &\subset (M_0 + B_\varepsilon(0)) \overset{*}{\subset} M(t_k) = \\ &= (M_0 + B_{2\varepsilon}(0)) \overset{*}{\subset} (M(t_k) + B_\varepsilon(0)) \subset (M_0 + B_{2\varepsilon}(0)) \overset{*}{\subset} M(t_{k+1}). \end{aligned}$$

Легко видеть, что если $\hat{M}_0 \overset{*}{\subset} \hat{M}(t_i) = \emptyset$, то и $M_0 \overset{*}{\subset} M(t_i) = \emptyset$. Отсюда $M_0 \overset{*}{\subset} M(t_{k+1}) = \emptyset$ и тем самым $T < t_{k+1}$. Следовательно, $M(T) \subset M(t_{k+1})$ и выполнено включение (23).

Используя формулу (23), получаем аналогично разд. 2 следующий результат относительно погрешности $\|x_0 - \hat{x}_0\|$ при условии, что t_k – решение (21).

Теорема 2. Пусть (T, x_0) – решение задачи (6) и $T < +\infty$. Пусть $\varepsilon \geq h(M_0, \hat{M}_0)$ и $\int_{t_k}^{t_{k+1}} e^{As} U ds \subset B_\varepsilon(0)$, где t_k – наибольшее время из дискретного набора, при котором верно (22).

1. Пусть множество M_0 равномерно выпуклое с модулем δ_{M_0} . Тогда

$$\delta_{M_0 + B_{2\varepsilon}(0)}(\|x_0 - \hat{x}_0\|) \leq 2\varepsilon.$$

2. Определим $D = (M(T) + x_0) \cap \partial M_0$ и $P = \bigcup_{x \in D} N_1(M_0, x)$. Пусть $\delta > 0$ и $B_\delta(0) \subset \text{co } P$. Тогда $\|x_0 - \hat{x}_0\| \leq 2\varepsilon/\delta$.

Замечание 2. Как и в теореме 1, решение (T, x_0) единственно. Замечание, аналогичное замечанию 1, имеет место и для теоремы 2.

Предположим, $T < t_k$. Тогда из включения

$$\hat{x}_0 \in (M_0 + B_\varepsilon(0)) \overset{*}{\subset} M(t_k) = (M_0 + B_\varepsilon(0)) \overset{*}{\subset} M(T) \overset{*}{\subset} \int_T^{t_k} e^{As} U ds$$

получаем, что T не менее максимального числа τ , $\tau < t_k$, которое определяется из условия

$$\int_\tau^{t_k} e^{As} U ds \supset B_\varepsilon(0). \quad (24)$$

Действительно, из доказательства леммы 3 и выполнения включения (24) имеем

$$\emptyset \neq (M_0 + B_\varepsilon(0)) \overset{*}{\subset} M(t_k) = \left((M_0 + B_\varepsilon(0)) \overset{*}{\subset} \int_\tau^{t_k} e^{As} U ds \right) \overset{*}{\subset} M(\tau) \subset M_0 \overset{*}{\subset} M(\tau),$$

и тем самым $\tau \leq T$ (T – максимально число, для которого $M_0 \overset{*}{\subset} M(T) \neq \emptyset$).

Лемма 4. Пусть $\mu = \min_{\|h\|=1, t \in [0, T_0]} \|e^{At} h\|$ и $B_d(0) \subset U$ для некоторого $d > 0$. Тогда $t_k - \frac{\varepsilon}{\mu d} \leq T < t_{k+1}$.

Доказательство. Повторяет доказательство леммы 2.

Заметим, что если $t_k - \frac{\varepsilon}{\mu d} \leq 0$, то оценка леммы 4 не информативна, так как утверждает, что $0 \leq T < t_{k+1}$.

Доказательство лемм 2 и 4 основано на включении (20) или (24), они доказываются при условии, что 0 – внутренняя точка множества U . Однако может выполняться включение $0 \in \partial U$. В этом случае удастся доказать включение (20) или (24) для некоторых конкретных видов матрицы A и множества U . Этим вопросам посвящено приложение 1, где рассматривается ряд примеров в плоском случае.

В заключение рассмотрим, как меняется множество достижимости при изменении начальной точки и времени на их приближенные значения. Пусть (T, x_0) – истинное решение задачи (5) или (6), а (\hat{T}, \hat{x}_0) – некоторое их приближение. Без ограничения общности считаем, что $T \leq \hat{T}$. Тогда отличие множеств достижимости системы (1) с начальными условиями $x(0) = e^{-AT} x_0$ и $x(0) = e^{-A\hat{T}} \hat{x}_0$ и временами T и \hat{T} соответственно есть (см. (4))

$$\begin{aligned} & \left\| e^{AT} e^{-AT} x_0 + \int_0^T e^{As} U ds, e^{A\hat{T}} e^{-A\hat{T}} \hat{x}_0 + \int_0^{\hat{T}} e^{As} U ds \right\| \leq \|x_0 - \hat{x}_0\| + \\ & + \left\| \int_0^T e^{As} U ds, \int_0^{\hat{T}} e^{As} U ds + \int_T^{\hat{T}} e^{As} U ds \right\| \leq \|x_0 - \hat{x}_0\| + \left\| \int_T^{\hat{T}} e^{As} U ds \right\| \leq \|x_0 - \hat{x}_0\| + C \|U\| \cdot |T - \hat{T}|, \end{aligned}$$

где $C = \max_{s \in [T, \hat{T}]} \|e^{As}\|$.

Таким образом, основной вклад в погрешность вносит погрешность по времени $|T - \hat{T}|$.

Для сильно выпуклого множества достижимости с радиусом R погрешность $\|x_0 - \hat{x}_0\|$ решения задач (13) и (21) имеет порядок $\sqrt{(R + \varepsilon)\varepsilon}$. В случае условия непустой внутренней (пункт 2 в теоремах 1 и 2) указанная погрешность имеет порядок ε .

Величина $|T - \hat{T}|$ в силу лемм 2 и 4 оценивается $\varepsilon/(\mu d)$ в обозначениях этих лемм. В случае, если точка 0 – граничная точка множества $U \subset \mathbb{R}^2$ и спектр матрицы $A \in \mathbb{R}^{2 \times 2}$ не вещественный, получаем, что $|T - \hat{T}|$ имеет порядок $\varepsilon^{1/3}$ в случае гладкой границы U , когда возможно касание 0 шаром, целиком содержащимся в U . В случае, когда 0 лежит на отрезке, являющемся частью гра-

ницы ∂U , $U \subset \mathbb{R}^2$, получаем, что $|T - \hat{T}|$ имеет порядок $\varepsilon^{1/2}$. Эти плоские случаи рассмотрены в приложении 1.

4. ПРИМЕРЫ

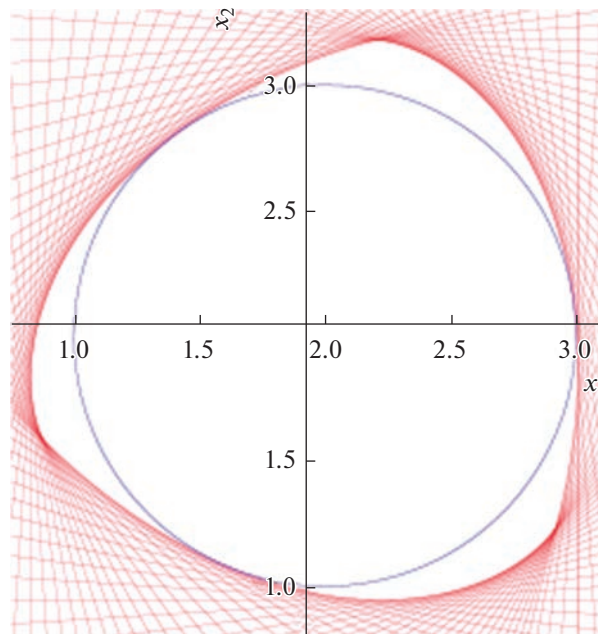
В примерах используется равномерная сетка в \mathbb{R}^2 вида $\left\{ \left(\cos \frac{2\pi k}{I}, \sin \frac{2\pi k}{I} \right) \right\}_{k=0}^{I-1}$, где I – число элементов сетки. Мелкость сетки имеет порядок $1/I$. Все картинки нарисованы для $I = 100$. Во всех примерах для всех i шаг по времени $t_{i+1} - t_i = 10^{-7}$ (избыточно малый). Мы старались подобрать условия примеров так, чтобы читатель мог их максимально просто интерпретировать.

Пример 2. Рассмотрим матрицу $A \in \mathbb{R}^{2 \times 2}$,

$$A = \begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix}, \quad e^{As} = e^{-s} \begin{pmatrix} \cos s & -\sin s \\ \sin s & \cos s \end{pmatrix}.$$

Пусть множество $U \subset \mathbb{R}^2$ является треугольником $\text{co}\{(-1,1), (1,-1), (\sqrt{3}, \sqrt{3})\}$, а множество $M_0 \subset \mathbb{R}^2$ является кругом $M_0 = \{(x_1, x_2)^T \in \mathbb{R}^2 \mid (x_1 - 2)^2 + (x_2 - 2)^2 \leq 1\}$.

Решается задача (5) (см. фиг. 1). В табл. 1 приведены результаты численного моделирования для данного примера. Все обозначения соответствуют ранее описанным, и их значения были получены с помощью алгоритмов, описанных выше. Радиус сильной выпуклости множеств $M(t_k)$ равен примерно $R = 2.81$ и вычисляется численно, его можно также оценить чисто теоретически (см. следствие теоремы 1 из [24]). В условиях п. 2 теоремы 2 значение $\delta = 0.5$. Истинное время оценивается из условия $t_k - 10^{-7} \leq T \leq \tau$. Последняя строка данной таблицы подтверждает оценку $\|x_0 - \hat{x}_0\|$, полученную в п. 2 теоремы 2. Скорость сходимости алгоритма по пространственной переменной является квадратичной по мелкости сетки и соответственно линейной по $\varepsilon = h(\hat{M}(t_k), M(t_k))$ в силу неравенства $\delta > 0$.



Фиг. 1. Пример 1.

Таблица 1

I	200	400	600	2000	10 000
t_k	1.70047	1.70056	1.70058	1.70059	1.70060
\hat{x}_0	(1.89550, 1.40899)	(1.89548, 1.40902)	(1.89547, 1.40901)	(1.89548, 1.40901)	(1.89548, 1.40901)
ε	0.00275	0.00069	0.00031	2.77699×10^{-5}	1.11079×10^{-6}
τ (20)	1.91705	1.80585	1.77063	1.72128	1.70475
τ по лемме 6	1.93234	1.80997	1.77225	1.72155	1.70475
$\frac{1}{I^2} \ \hat{x}_0 - \hat{x}_0(10\,000)\ $	1.07092	2.93539	0.710737	1.96163	0

Пример 3. Снова матрица $A \in \mathbb{R}^{2 \times 2}$ имеет вид

$$A = \begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix}, \quad e^{As} = e^{-s} \begin{pmatrix} \cos s & -\sin s \\ \sin s & \cos s \end{pmatrix},$$

$U = \text{co}\{(-1, 1), (1, -1), (\sqrt{3}, \sqrt{3})\}$, а множество $M_0 \subset \mathbb{R}^2$ является эллипсом с центром симметрии $x_0 = (-1, 1)^T$ и матрицей $Q = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix}$.

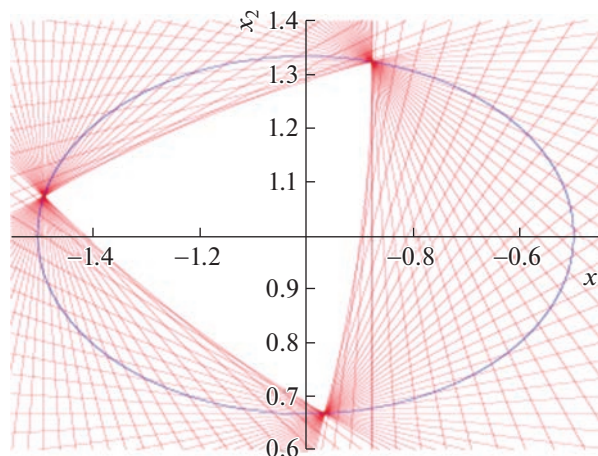
По теореме 3 (см. [12]) множество M_0 является сильно выпуклым с радиусом $R = 3/4$.

Решается задача (6) (см. фиг. 2). В условиях п. 2 теоремы 2 значение $\delta = 0.11$. Истинное время T оценивается из условия $\tau \leq T \leq t_k + 10^{-7}$. В табл. 2 приводятся результаты численного моделирования для данного примера.

Пример 4. Пусть матрица $A \in \mathbb{R}^{2 \times 2}$ задана следующим образом:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad e^{As} = \begin{pmatrix} e^s & 0 \\ 0 & e^{2s} \end{pmatrix}.$$

Зададим $U = \text{co}\{(0.5, 0), (-0.5, 0), (0, 1), (0, -1)\}$. Множество $M_0 \subset \mathbb{R}^2$ является эллипсом с центром симметрии $x_0 = (2, 2)^T$ и матрицей $Q = \begin{pmatrix} \frac{1}{9} & 0 \\ 0 & \frac{1}{25} \end{pmatrix}$. Тогда по теореме 3 (см. [12]) множество M_0 является сильно выпуклым с радиусом $R = \frac{25}{3}$.



Фиг. 2. Пример 2.

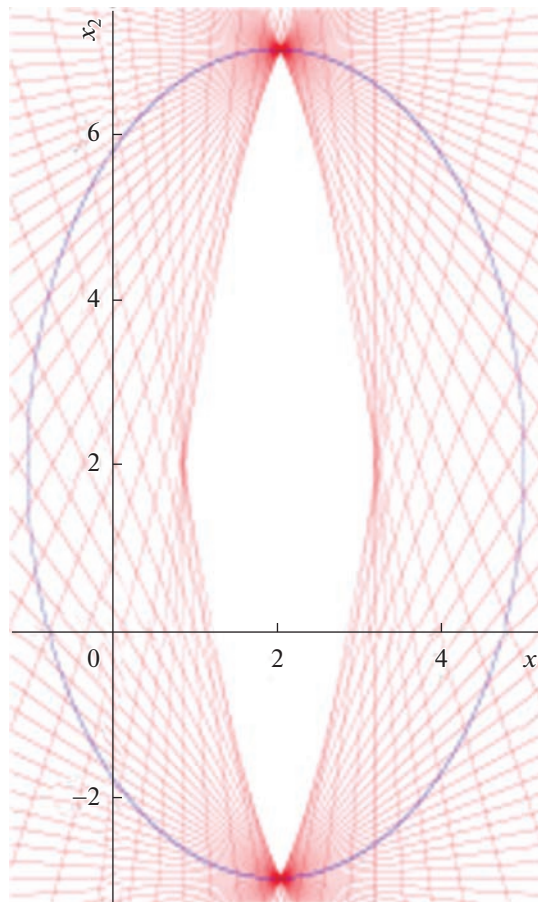
Таблица 2

I	200	400	600	2000	10 000
t_k	0.26734	0.26729	0.26729	0.26729	0.26728
\hat{x}_0	(-1.22649, 0.86960)	(-1.22653, 0.86966)	(-1.22653, 0.86965)	(-1.22653, 0.86966)	(-1.22653, 0.86966)
ε	0.00074	0.00019	8.22512×10^{-5}	7.40224×10^{-6}	2.96088×10^{-7}
τ (24)	0.21863	0.24216	0.25083	0.26243	0.26634
τ по лемме 6	0	0.07499	0.12055	0.20153	0.24479
$\frac{1}{I^2} \ \hat{x}_0 - \hat{x}_0(10\,000)\ $	2.79663	0.26175	2.65212	0.71218	0

На фиг. 3 изображено решение задачи (6). В отличие от предыдущего примера здесь не реализуется случай п. 2 теоремы 2 ($\delta = 0$). При этом 0 – внутренняя точка множества U , а именно, $B_{1/\sqrt{5}}(0) \subset U$. Истинное время T оценивается из условия $\tau \leq T \leq t_k + 10^{-7}$. В табл. 3 приведены результаты численного моделирования для данного примера.

Пример 5. Пусть матрица $A \in \mathbb{R}^{2 \times 2}$ задана следующим образом:

$$A = \begin{pmatrix} -0.5 & 0 \\ 0 & -3 \end{pmatrix}, \quad e^{As} = \begin{pmatrix} e^{-0.5s} & 0 \\ 0 & e^{-3s} \end{pmatrix}.$$



Фиг. 3. Пример 3.

Таблица 3

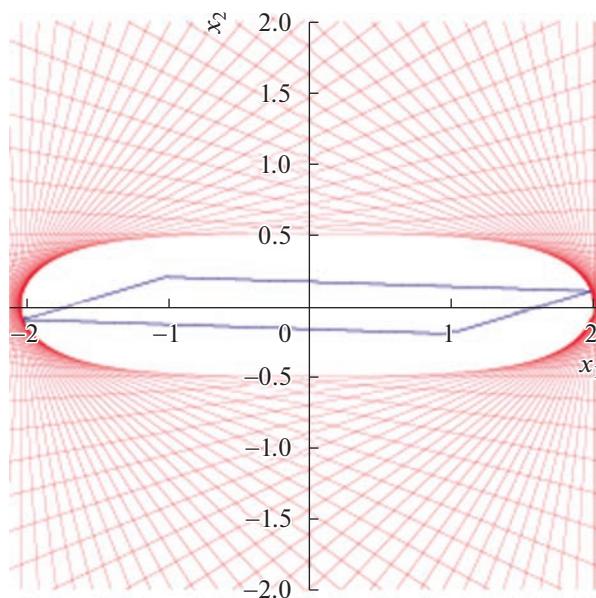
I	200	400	600	2000	10 000
t_k	1.19895	1.19895	1.19895	1.19895	1.19895
\hat{x}_0	(2.02830, 2)	(2.01418, 2)	(2.00949, 2)	(2.00303, 2)	(2.00158, 2)
ε	0.00823	0.00206	0.00091	8.22471×10^{-5}	3.28987×10^{-6}
τ (24)	1.19393	1.19769	1.19839	1.19890	1.19895
τ по лемме 4	0.99780	1.14868	1.17661	1.19694	1.19887
$\frac{1}{I^2} \ \hat{x}_0 - \hat{x}_0(10\,000)\ $	5.34503	5.04090	4.74569	2.90680	0

Таблица 4

I	200	600	2000	10 000
t_k	2.23484	2.23490	2.23491	2.23491
\hat{x}_0	(-0.00021, 0.00052)	(-0.00021, 0.00054)	(-3.90502×10^{-5} , 1.02372×10^{-5})	(-2.93699×10^{-5} , 7.72814×10^{-5})
ε	0.06393	0.00710	0.00064	2.55644×10^{-5}
$\frac{1}{I^2} \ \hat{x}_0 - \hat{x}_0(10\,000)\ $	0.09606	0.29784	0.05379	0

Зададим $U = B_{1.5}(0)$. Множество $M_0 \subset \mathbb{R}^2$ является параллелограммом с вершинами $(-2, -0.1)$, $(-1, 0.2)$, $(2, 0.1)$, $(1, -0.2)$.

Решается задача (5). В табл. 4 приведены результаты численного моделирования для данного примера. Радиус сильной выпуклости множества $M(t_k)$ примерно равен 64.76 и может быть вычислен по формуле $\int_0^{t_k} R(s) ds$ (см. [19]), где $R(s) = \lambda_1^2(s)/\lambda_2(s) = \frac{3}{2}e^{-2s}$ – радиус сильной выпуклости эллипса $e^{As}B_{3/2}(0)$ с максимальной полуосью $\lambda_1(s) = \frac{3}{2}e^{-0.5s}$ и минимальной полуосью



Фиг. 4. Пример 4.

$\lambda_2(s) = \frac{3}{2}e^{-3s}$. Для $I = 3000$ векторов в сетке время τ , полученное из включения (20), равно $\tau = 2.44260$. При меньшем I получается $\tau = 0$ в силу малости параметра $\mu = e^{-3t_k}$ из леммы 2. Скорость сходимости алгоритма по фазовой переменной является линейной по мелкости сетки, что обусловлено равенством $\delta = 0$.

Приложение 1

ВНУТРЕННОСТЬ МНОГОЗНАЧНОГО ИНТЕГРАЛА В ПЛОСКОМ СЛУЧАЕ

Лемма 5. Пусть собственные значения матрицы $A \in \mathbb{R}^{2 \times 2}$ – комплексно-сопряженные числа $\alpha \pm i\beta$, $\beta > 0$, и $e^{As} = e^{\alpha s} Q(\beta s)$, где $Q(s) = \begin{pmatrix} \cos s & -\sin s \\ \sin s & \cos s \end{pmatrix}$. Пусть $U \subset \mathbb{R}^2$ – выпуклый компакт, $r > 0$, $x \in \mathbb{R}^2$, $\|x\| = r$ и выполняются условия $B_r(x) \subset U$, $0 \in B_r(x) \cap \partial U$. Определим $C = \min\{e^{\alpha t_1}, e^{\alpha t_2}\}$. Тогда для любых $0 < t_1 < t_2$

$$\int_{t_1}^{t_2} e^{As} U ds \supset B_\delta(0),$$

где $\delta = Cr(t_2 - t_1) \left(1 - \frac{\sin(\beta(t_2 - t_1)/2)}{\beta(t_2 - t_1)/2} \right)$.

Отметим, что условия леммы означают гладкость границы множества U в точке 0. Более того, внутри множества U точку $0 \in \partial U$ можно коснуться шаром радиуса r .

Доказательство. Имеем цепочку включений

$$\int_{t_1}^{t_2} e^{As} U ds \supset \int_{t_1}^{t_2} e^{As} B_r(x) ds \supset C \int_{t_1}^{t_2} Q(\beta s) B_r(x) ds = C \int_{t_1}^{t_2} Q(\beta s) B_r(0) ds + C \int_{t_1}^{t_2} Q(\beta s) x ds.$$

Первое слагаемое справа с учетом равенства $Q(\beta s) B_r(0) = B_r(0)$ оценивается как

$$C \int_{t_1}^{t_2} Q(\beta s) B_r(0) ds \supset Cr(t_2 - t_1) B_1(0).$$

Пусть $x = r(\cos \theta, \sin \theta)^T$ для некоторого $\theta \in [0, 2\pi)$. Для второго слагаемого получаем

$$C \int_{t_1}^{t_2} Q(\beta s) x ds = \frac{2Cr \sin(\beta(t_2 - t_1)/2)}{\beta} \begin{pmatrix} \cos\left(\frac{\beta(t_2 + t_1)}{2} + \theta\right) \\ \sin\left(\frac{\beta(t_2 + t_1)}{2} + \theta\right) \end{pmatrix} =: z.$$

Для точки z имеем $\|z\| = \frac{2Cr \sin(\beta(t_2 - t_1)/2)}{\beta}$, откуда

$$\delta \geq Cr(t_2 - t_1) - \|z\| = Cr(t_2 - t_1) \left(1 - \frac{\sin(\beta(t_2 - t_1)/2)}{\beta(t_2 - t_1)/2} \right).$$

Лемма 6. Пусть собственные значения матрицы $A \in \mathbb{R}^{2 \times 2}$ – комплексно-сопряженные числа $\alpha \pm i\beta$, $\beta > 0$, и $e^{As} = e^{\alpha s} Q(\beta s)$, где $Q(s)$ из леммы 5. Пусть $U \subset \mathbb{R}^2$ – выпуклый компакт, $r > 0$, $\theta \in [0, 2\pi)$ и $I = \text{co}\{\pm r(\cos \theta, \sin \theta)^T\} \subset \partial U$. Определим $C = \min\{e^{\alpha t_1}, e^{\alpha t_2}\}$. Тогда для любых $0 < t_1 < t_2$, $t_2 - t_1 < \pi/\beta$,

$$\int_{t_1}^{t_2} e^{As} U ds \supset B_\delta(0),$$

где $\delta = \frac{2}{\beta} Cr \left(1 - \cos\left(\frac{1}{2}\beta(t_2 - t_1)\right)\right)$.

Доказательство. Имеем цепочку включений $\int_{t_1}^{t_2} e^{As} U ds \supset \int_{t_1}^{t_2} e^{As} I ds \supset C \int_{t_1}^{t_2} Q(\beta s) I ds$. Пусть $p = (\cos \varphi, \sin \varphi)^T$ – единичный вектор. Тогда

$$s \left(p, \int_{t_1}^{t_2} e^{As} U ds \right) \geq C \int_{t_1}^{t_2} s(p, Q(\beta s) I) ds = Cr \int_{t_1}^{t_2} |\sin(\beta s + \psi)| ds,$$

где $\psi = \theta - \varphi + \pi/2$. Делая замену переменной $\tau = \beta s + \psi$, для $T_i = \beta t_i + \psi$, $i = 1, 2$, получаем $s \left(p, \int_{t_1}^{t_2} e^{As} U ds \right) \geq \frac{1}{\beta} Cr \int_{T_1}^{T_2} |\sin \tau| d\tau$. Поскольку $T_2 - T_1 < \pi$, то

$$\int_{T_1}^{T_2} |\sin \tau| d\tau \geq 2 \int_0^{\frac{1}{2}(T_2 - T_1)} \sin \tau d\tau = 2 \left(1 - \cos\left(\frac{1}{2}(T_2 - T_1)\right)\right).$$

Подставляя последнюю оценку в предыдущее неравенство, получаем утверждение леммы.

Заметим, что лемма 6 дает второй порядок радиуса шара по $t_2 - t_1$, а лемма 5 – третий.

Пусть $U \subset \mathbb{R}^2$ – выпуклый компакт. Допустим, что существует такое число $\rho > 0$, что $U \cap B_\rho(0) \supset H \cap B_\rho(0)$, где $H = \{x \in \mathbb{R}^2 | (p, x) \leq 0\}$ – полупространство для некоторого единичного вектора $p \in \mathbb{R}^2$. Тогда $\int_{t_1}^{t_2} e^{As} U ds \asymp (t_2 - t_1)^2$, $t_2 - t_1 \rightarrow +0$.

Действительно, пусть I – отрезок с центром в нуле такой, что $I \subset \partial H$. Рассмотрим круг $B_R(x)$, $\|x\| = R$, который касается ∂H в точке 0 и $I + B_R(x) \supset U$. Такие R и отрезок I можно подобрать в силу компактности U и того, что ∂H – опорная прямая к U в точке 0. Тогда

$$\int_{t_1}^{t_2} e^{As} U ds \subset \int_{t_1}^{t_2} e^{As} I ds + \int_{t_1}^{t_2} e^{As} B_R(x) ds.$$

По леммам 5, 6 слагаемые в правой части последней формулы имеют порядок $(t_2 - t_1)^2 + (t_2 - t_1)^3 \asymp (t_2 - t_1)^2$, $t_2 - t_1 \rightarrow +0$.

Таким образом, наличие отрезка во множестве U , $0 \in \partial U$, с центром в нуле важнее телесности U для максимальности вписанного в интеграл шара с центром в нуле.

Лемма 7. Пусть $e^{At} = \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix}$, где $\lambda_1 > \lambda_2$. Определим $\lambda = \lambda_1 - \lambda_2 > 0$ и $C = \min\{e^{\lambda_2 t_1}, e^{\lambda_2 t_2}\}$.

Пусть $U \subset \mathbb{R}^2$ – выпуклый компакт и $I \subset \partial U$, где $I = \text{co}(\pm r(\cos \theta, \sin \theta)^T)$, $r > 0$, $\theta \in (0, \pi)$, $\theta \neq \pi/2$. Тогда для любых $0 < t_1 < t_2$

$$\int_{t_1}^{t_2} e^{As} U ds \supset B_\delta(0),$$

где $\delta = Cr \frac{\lambda K^2 \sin \theta}{\pi^2} (t_2 - t_1)^2$. Здесь $K = K(t_1, t_2, \lambda, \theta) = \min_{i=1,2} \left\{ \frac{e^{\lambda t_i} |\text{ctg} \theta|}{1 + e^{2\lambda t_i} \text{ctg} \theta} \right\}$.

Доказательство. Пусть $M = \int_{t_1}^{t_2} e^{As} U ds$. Тогда

$$M \supset \int_{t_1}^{t_2} e^{\lambda_2 s} \begin{pmatrix} e^{\lambda s} & 0 \\ 0 & 1 \end{pmatrix} I ds \supset Cr \int_{t_1}^{t_2} \text{co}(\pm e^{\lambda s} \cos \theta, \sin \theta)^T ds.$$

Пусть $p = (\cos \varphi, \sin \varphi)^T$ – единичный вектор, $\varphi \in [0, 2\pi)$. Тогда для

$$J = \int_{t_1}^{t_2} |e^{\lambda s} \cos \varphi \cos \theta + \sin \varphi \sin \theta| ds$$

имеем $s(p, M) \geq CrJ$. Оценим J снизу.

1. Рассмотрим случай $\theta \in (0, \frac{1}{2}\pi)$. Здесь

$$J = \int_{t_1}^{t_2} \sqrt{\sin^2 \theta + \cos^2 \theta e^{2\lambda s}} |\sin(\alpha(s) + \varphi)| ds,$$

где $\alpha(s) = \operatorname{arctg}(e^{\lambda s} \theta)$. Введем новую переменную $\tau = \alpha(s) + \varphi$, $ds = \frac{d\tau}{\alpha'(s)} = \frac{1 + e^{2\lambda s} \operatorname{ctg}^2 \theta}{\lambda e^{\lambda s} \operatorname{ctg} \theta} d\tau$. Отсю-

да, с учетом включения $\alpha(s) \in (0, \frac{1}{2}\pi)$,

$$\begin{aligned} J &= \frac{\sin \theta}{\lambda} \int_{\alpha(t_1)+\varphi}^{\alpha(t_2)+\varphi} \frac{(1 + \operatorname{tg}^2 \alpha(s))^{3/2}}{\operatorname{tg} \alpha(s)} |\sin \tau| d\tau = \frac{\sin \theta}{\lambda} \int_{\alpha(t_1)+\varphi}^{\alpha(t_2)+\varphi} \frac{|\sin \tau|}{(\tau - \varphi) \cos^3(\tau - \varphi)} d\tau \geq \\ &\geq \frac{\sin \theta}{\lambda} \int_{\alpha(t_1)+\varphi}^{\alpha(t_2)+\varphi} |\sin \tau| d\tau \geq \frac{2 \sin \theta}{\lambda} \int_0^{\frac{1}{2}(\alpha(t_2)-\alpha(t_1))} \sin \tau d\tau = \frac{4 \sin \theta}{\lambda} \sin^2 \left(\frac{\alpha(t_2) - \alpha(t_1)}{4} \right) \geq \\ &\geq \frac{4 \sin \theta}{\lambda} \frac{4}{\pi^2} \frac{1}{4^2} (\alpha(t_2) - \alpha(t_1))^2 \frac{\sin \theta}{\lambda \pi^2} (\alpha(t_2) - \alpha(t_1))^2. \end{aligned}$$

По теореме о среднем найдется $\xi \in [t_1, t_2]$ такое, что

$$|\alpha(t_2) - \alpha(t_1)| = \frac{\lambda e^{\lambda \xi} |\operatorname{ctg} \theta|}{1 + e^{2\lambda \xi} \operatorname{ctg}^2 \theta} (t_2 - t_1) \geq \lambda K (t_2 - t_1).$$

Отсюда получаем утверждение леммы.

2. Пусть $\theta \in (\frac{1}{2}\pi, \pi)$. Заменой $\psi = \pi - \theta$, $\alpha(s) = \operatorname{arctg}(e^{\lambda s} \operatorname{ctg} \psi)$ и $\tau = \alpha(s) + \varphi$ приводим интеграл J к виду

$$\begin{aligned} J &= \frac{\sin \psi}{\lambda} \int_{\alpha(t_1)+\varphi}^{\alpha(t_2)+\varphi} \frac{(1 + \operatorname{tg}^2 \alpha(s))^{3/2}}{\operatorname{tg} \alpha(s)} |\cos(\alpha(s) + \varphi)| d\tau \geq \frac{\sin \theta}{\lambda} \int_{\alpha(t_1)+\varphi-\frac{1}{2}\pi}^{\alpha(t_2)+\varphi-\frac{1}{2}\pi} |\sin \tau| d\tau \geq \\ &\geq \frac{2 \sin \theta}{\lambda} \int_0^{\frac{1}{2}(\alpha(t_2)-\alpha(t_1))} |\sin \tau| d\tau. \end{aligned}$$

Далее доказательство повторяет пункт 1.

Лемма 8. Пусть $e^{At} = e^{\lambda t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$. Определим $C = \min\{e^{\lambda t_1}, e^{\lambda t_2}\}$. Пусть $U \subset \mathbb{R}^2$ – выпуклый компакт и $I \subset \partial U$, где $I = \operatorname{co}(\pm r(\cos \theta, \sin \theta)^T)$, $r > 0$, $\theta \in (0, \pi)$. Тогда для любых $0 < t_1 < t_2$

$$\int_{t_1}^{t_2} e^{As} U ds \supset B_\delta(0),$$

где $\delta = Cr \frac{K^2 \sin \theta}{\pi^2} (t_2 - t_1)^2$. Здесь $K = K(t_1, t_2, \theta) = \min_{i=1,2} \left\{ \frac{1}{1 + (t_i + \operatorname{ctg} \theta)^2} \right\}$.

Доказательство. Повторяя рассуждения леммы 7, получаем для $M = \int_{t_1}^{t_2} e^{As} U ds$, $p = (\cos \varphi, \sin \varphi)$ и $J = \int_{t_1}^{t_2} |(\cos \theta + s \sin \theta) \cos \varphi + \sin \theta \sin \varphi| ds$ имеем $s(p, M) \geq CrJ$. Для J получаем $J = \sin \theta \int_{t_1}^{t_2} \sqrt{1 + (s + \operatorname{ctg} \theta)^2} |\sin(\alpha(s) + \varphi)| ds$, где $\alpha(s) = \operatorname{arctg}(s + \operatorname{ctg} \theta)$. Заменой $\tau = \alpha(s) + \varphi$ находим

$$\begin{aligned} J &= \sin \theta \int_{\alpha(t_1)+\varphi}^{\alpha(t_2)+\varphi} (1 + (s + \operatorname{ctg} \theta)^2)^{3/2} |\sin \tau| d\tau = \sin \theta \int_{\alpha(t_1)+\varphi}^{\alpha(t_2)+\varphi} (1 + \operatorname{tg}^2 \alpha(s))^{3/2} |\sin \tau| d\tau = \\ &= \sin \theta \int_{\alpha(t_1)+\varphi}^{\alpha(t_2)+\varphi} \frac{|\sin \tau|}{|\cos^3(\tau - \varphi)|} d\tau \geq 2 \sin \theta \int_0^{\frac{1}{2}(\alpha(t_2)-\alpha(t_1))} \sin \tau d\tau = 4 \sin \theta \sin^2 \left(\frac{\alpha(t_2) - \alpha(t_1)}{4} \right). \end{aligned}$$

Отсюда $J \geq 4 \sin \theta \frac{4}{\pi^2} \frac{(\alpha(t_2) - \alpha(t_1))^2}{4^2}$. Применение теоремы о среднем завершает доказательство.

Приложение 2

ОДИН ПРИМЕР МНОЖЕСТВА ДОСТИЖИМОСТИ

Приведем пример множества достижимости нетривиальной линейной системы, которое строго выпукло, но не сильно выпукло. Напомним, что $f(s) \asymp g(s)$, $s \rightarrow 0$, если существует такое $\delta > 0$, что $C_1 \leq \frac{f(s)}{g(s)} \leq C_2$ для всех $s \in (-\delta, \delta)$ и некоторых $C_2 \geq C_1 > 0$. Пусть $U = \operatorname{co}\{(0, 0, \pm 1)\}$,

$$\lambda \in \mathbb{R}, \quad A = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}, \quad e^{As} = e^{\lambda s} \begin{bmatrix} 1 & s & \frac{1}{2}s^2 \\ 0 & 1 & s \\ 0 & 0 & 1 \end{bmatrix}. \quad \text{Для } \varepsilon \in (0, 1) \text{ рассмотрим векторы } p = \frac{1}{3}(2, -2, 1) \text{ и}$$

$q = q(\varepsilon) = \frac{(2, -2, 1 - \varepsilon)}{\sqrt{9 - 2\varepsilon + \varepsilon^2}}$. Легко проверить, что $\|p - q\| \asymp \varepsilon$, $\varepsilon \rightarrow 0$, и для $t > 1 + \sqrt{\varepsilon}$ ($s_{1,2} = 1 \pm \sqrt{\varepsilon}$ есть корни уравнения $(e^{A^T s} q(\varepsilon), (0, 0, 1)^T) = 0$)

$$M(t)(p) - M(t)(q) = \int_{1-\sqrt{\varepsilon}}^{1+\sqrt{\varepsilon}} e^{\lambda s} (s^2, 2s, 2)^T ds, \quad \|M(t)(p) - M(t)(q)\| \geq \int_{1-\sqrt{\varepsilon}}^{1+\sqrt{\varepsilon}} 2e^{\lambda s} ds \asymp \sqrt{\varepsilon}, \quad \varepsilon \rightarrow 0.$$

Для произвольного фиксированного $R > 0$ значение $\sqrt{\varepsilon}/(R\varepsilon)$ не ограничено сверху при малых $\varepsilon > 0$. Следовательно, $M(t)$ не сильно выпукло, так как для сильно выпуклого множества с радиусом $R > 0$ опорные элементы этого множества должны удовлетворять условию Липшица на единичной сфере с константой R (см. [19], [12], следствие 4). Заметим, что множество $M(t)$ строго выпукло, так как для всякого вектора $p \in \partial B_1(0)$ множество $(e^{As} U)(p)$ одноточечно при всех $s \in [0, t]$, за исключением не более чем двух значений.

СПИСОК ЛИТЕРАТУРЫ

1. *Liapounoff A.A.* Sur les fonctions-vecteurs completement additives // Izv. Akad. Nauk SSSR Ser. Mat. 1940. V. 4. № 6. P. 465–478.
2. *Lee E.B., Markus L.* Foundations of Optimal Control Theory, John Wiley; 1st Printing, ed. 1967, 588 p.
3. *Aumann R.* Integrals of set-valued functions // J. Math. Anal. Appl. 1965. V. 12. № 1. P. 1–12.
4. *Polyak B.T., Smirnov G.* Large deviations for non-zero initial conditions in linear systems // Automatica. 2016. V. 74. P. 297–307.
5. *Aubin J.-P., Cellina A.* Differential inclusions, Springer-Verlag, 1984.
6. *Aubin J.-P.* A Survey of viability theory // SIAM J. Control and Optimizat. 1990. V. 28. № 4. P. 749–788.

7. Kelley H.J. Gradient theory of optimal flight paths // ARSJ. 1960. V. 30. P. 947–953.
<https://doi.org/10.2514/8.5282>
8. Bryson A.E., Denham W.F. A steepest-ascent method for solving optimum programming problems // J. Appl. Mech. 1962. V. 29. P. 247–257; <https://www.gwern.net/docs/ai/1962-bryson.pdf>
9. Eichmeir Ph., Lauß Th., Oberpeilsteiner S., Nachbagauer K., Steiner W. The adjoint method for time-optimal control problems // J. Comput. Nonlinear Dynam. 2021. V. 16. № 2. P. 021003 (12 pages).
<https://doi.org/10.1115/1.4048808>
10. Cannarsa P., Sinestrari C. Convexity properties of the minimum time function // Calcul. Variat. Part. Different. Equat. 1995. V. 3. № 3. P. 273–298;
<https://doi.org/10.1007/bf01189393>
11. Boltyanskii V.G. Mathematical methods of optimal control, Holt, Rinehart and Winston (1st ed.), 1971.
12. Половинкин Е.С. Сильно выпуклый анализ // Матем. сб. 1996. Т. 187. № 2. С. 103–130.
13. Le Guernic C., Girard A. Reachability analysis of linear systems using support functions // Nonlinear Analysis: Hybrid Systems. 2010. V. 4. P. 250–262.
14. Althoff M., Frehse G., Girard A. Set propagation techniques for reachability analysis // Ann. Rev. Control, Robotics, and Autonomous Systems, Ann. Rev. 2021. V. 4. № 1. hal-03048155.
<https://doi.org/10.1146/annurev-control-071420-081941>
15. Serry M., Reissig G. Over-approximating reachable tubes of linear time-varying systems, arXiv:2102.04971v1.
16. Kurzhanski A.B., Varaiya P. Dynamics and control of trajectory tubes, ser. Systems and Control: Foundations and Applications. Birkhauser/Springer, 2014, theory and computation.
17. Балашов М.В. Покрытие множества выпуклым компактом: оценки погрешности и вычисление // Матем. заметки. 2022. Т. 112. № 3. С. 337–349.
18. Дистель Дж. Геометрия банаховых пространств, Киев, Вища школа, 1980.
19. Frankowska H., Olech C. R-convexity of the integral of the set-valued functions, Contributions to analysis and geometry, Johns Hopkins Univ. Press, Baltimore, MD, 1981. P. 117–129.
20. Половинкин Е.С., Балашов М.В. Элементы выпуклого и сильно выпуклого анализа, М., Физматлит, 2007. 2-е изд.
21. Балашов М.В., Половинкин Е.С. M-сильно выпуклые подмножества и их порождающие множества // Матем. сб. 2000. Т. 191. № 1. С. 27–64.
22. Balashov M.V. On polyhedral approximations in an n -dimensional space // Comput. Math. Math. Phys. 2016. V. 56. № 10. P. 1679–1685.
23. Balashov M.V., Repovš D. Polyhedral approximations of strictly convex compacta // J. Math. Anal. Appl. 2011. V. 374. P. 529–537.
24. Балашов М.В. Сильная выпуклость множеств достижимости линейных систем // Матем. сб. 2022. Т. 213. № 5. С. 30–49.

**ОБЫКНОВЕННЫЕ
ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ**

УДК 517.929

ON RANKS OF MATRICES OVER NONCOMMUTATIVE DOMAINS¹⁾

© 2023 г. S. A. Abramov^{1,*}, M. Petkovšek^{2,**}, A. A. Ryabenko^{1,***}

¹ 119333 Moscow, Vavilova str., 40, Federal Research Center “Computer Science and Control”
of the Russian Academy of Science, Russia

² University of Ljubljana, Faculty of Mathematics and Physics, Jadranska 19, SI-1000, Ljubljana, Slovenia

*e-mail: sergeyabramov@mail.ru

**e-mail: Marko.Petkovsek@fmf.uni-lj.si

***e-mail: anna.ryabenko@gmail.com

Поступила в редакцию 30.08.2022 г.
Переработанный вариант 30.09.2022 г.
Принята к публикации 02.02.2023 г.

Рассматриваются матрицы над некоторой областью целостности R , т.е. над кольцом, не обязательно коммутативным, без делителей нуля. Обсуждаются понятия рангов по строкам и столбцам. (Коэффициенты линейных зависимостей принадлежат R ; левые коэффициенты используются для строк, правые коэффициенты для столбцов.) Доказывается, что наличие ненулевых левых и правых общих кратных для произвольных ненулевых элементов R (условие Ore) является необходимым и достаточным условием равенства рангов по строкам и столбцам произвольной матрицы над R . Предлагается алгоритм вычисления ранга заданной матрицы. Наша реализация этого алгоритма в Maple охватывает области дифференциальных и (q -)разностных операторов как обычных, так и с частными производными и разностями. Библ. 8.

Ключевые слова: некоммутативная область, матрицы над областями, ранги по строкам и столбцам, компьютерная алгебра.

DOI: 10.31857/S0044466923050022, **EDN:** DXCTUA

О рангах матриц над некоммутативными областями. Ниже под областью всюду понимается кольцо, не обязательно коммутативное, не содержащее нетривиальных делителей нуля; везде далее R обозначает некоторую область.

Определение 1. Пусть A – матрица над R . Строки u_1, u_2, \dots, u_r матрицы A линейно зависимы над R , если существуют такие не равные одновременно нулю $f_1, f_2, \dots, f_r \in R$, что $f_1 u_1 + f_2 u_2 + \dots + f_r u_r = 0$, в противном случае эти строки линейно независимы над R . Столбцы v_1, v_2, \dots, v_s матрицы A линейно зависимы над R , если существуют такие не равные одновременно нулю $g_1, g_2, \dots, g_s \in R$, что $v_1 g_1 + v_2 g_2 + \dots + v_s g_s = 0$, в противном случае эти столбцы линейно независимы над R .

Наибольшее число линейно независимых строк (столбцов) матрицы A называется *рангом по строкам* или *левым рангом* (рангом по столбцам или *правым рангом*) матрицы A .

Определение ранга матрицы над некоторым полем как наибольшего числа ее линейно независимых строк и доказательство того, что это число равно ее наибольшему числу линейно независимых столбцов – это одно из начал классической линейной алгебры. Мы даем пример некоммутативной области, когда эти два ранга не совпадают, и естественно возникает вопрос о характеристике тех областей, над которыми совпадение имеет место. Мы показываем, что интересующие нас области суть те, которые удовлетворяют условиям Ore, т.е. области, в которых для любых ненулевых элементов существуют ненулевые левое и правое общие кратные.

В литературе можно найти различные подходы к определению ранга матрицы над областью. Выбор определения иногда диктуется удобством доказательства некоторой теоремы, что может вести к несовпадению с результатами, полученными при использовании других неэквивалентных определений. И даже для эквивалентных определений доказательство этой эквивалентности

¹⁾ Полный текст статьи печатается в английской версии журнала.

сти может оказаться довольно сложным. Например, в [1] ранг матрицы над кольцом полиномов Оре (см. [2] или [3]) определен как наибольшее число линейно независимых строк. Авторы статьи [1] отмечают, что их определение отличается от данного в [4], разд. 0.6, где ранг матрицы A над R определен как ранг левого модуля M , порожденного строками A над R . Теорема А.2 в [1] устанавливает, что для матриц над кольцом полиномов Оре одной переменной эти два числа совпадают, но доказательство этой теоремы выглядит очень непросто.

В книге Я.Б. Лопатинского [5] подчеркивается важность концепции ранга для исследования интегральных многообразий систем линейных дифференциальных уравнений с частными производными, и в этой книге дается доказательство равенства левого и правого рангов именно для дифференциальных операторов, при этом ранг понимается так же, как в определении 1.

В [1] дается доказательство равенства рангов для матриц над кольцом (некоммутативных) полиномов Оре, оснащенном автоморфизмом σ и отображением на себя δ , являющимся дифференцированием по отношению к σ (см., например, [2] или [3]). Абстракция таких полиномов одной переменной не покрывает, например, дифференциальных операторов с частными производными. Рассмотренные в [1] кольца таких полиномов одной переменной над коммутативными полями коэффициентов являются евклидовыми и соответствуют очень специальному случаю.

В нашем доказательстве мы исходим из более общего предположения – считаем, что область R удовлетворяет условиям Оре. Эти условия выполняются для полиномов Оре многих переменных (в частности, для дифференциальных операторов с частными производными), что доказано в [6].

Не очевидно, что предложенная в [7] теория левых и правых определителей позволяет получить короткое доказательство левого и правого рангов матрицы A над R в смысле определения 1, хотя и позволяет установить, что строки A линейно зависимы если и только если линейно зависимы ее столбцы.

Представляется, что в доступной литературе отсутствует полное доказательство равенства рангов (в смысле определения 1) по строкам и столбцам для матриц над удовлетворяющей условиям Оре областью R .

В нашей статье доказано, что рассмотрение вместе с областью R ее тела (т.е. поля, возможно, некоммутативного) частных, скажем, тела F “дробей” вида $p^{-1}q$, $p, q \in R$, $p \neq 0$, дает совпадение левого и правого рангов над F с левым и правым рангами над R . Естественно, что вычисление ранга над полем, пусть и некоммутативным, удобнее, чем над областью. Мы показываем, что если имеется алгоритм вычисления ненулевого левого общего кратного произвольных ненулевых элементов области, то с помощью гауссовых исключений возможно вычисление ранга заданной матрицы. Вычисление над телом F мы проводим без дробей. В компьютерно-алгебраической среде Maple выполнена реализация этого подхода, она ориентирована на матрицы с элементами, являющимися линейными дифференциальными, разностными и q -разностными операторами с частными производными, сдвигами или q -сдвигами. Предлагается команда

OreAlgebraGaussianElimination,

доступная в

<http://www.ccas.ru/ca/orealgebraussianelimination>.

Даются иллюстративные примеры.

Предварительная версия статьи опубликована в докладах конференции ISSAC'22 [8], где вместо условий Оре рассмотрены другие условия: для любого положительного целого n как строки произвольной матрицы из $R^{(n+1) \times n}$, так и столбцы произвольной матрицы из $R^{(n+1) \times n}$ линейно зависимы над R . В полной статье этот результат усилен: достаточно рассмотреть единственное значение $n = 1$.

Авторы признательны А.Э. Гутерману, А.И. Зобнину, Дж. Лабану, В. Левандовскому, А.А. Михалеву, А.В. Михалеву, Ф. Шизаку за полезные советы.

СПИСОК ЛИТЕРАТУРЫ

1. Beckermann B., Cheng H., Labahn G. Fraction-free row reduction of matrices of Ore polynomials // J. Symbolic Comput. 2006. V. 41. P. 513–543.

2. Ore O. Theory of non-commutative polynomials // Ann. of Math. (2). 1933. V. 34. P. 480–508.
3. Bronstein M., Petkovšek M. An introduction to pseudo-linear algebra // Theoret. Comput. Sci. 1996. V. 157. P. 3–33.
4. Кон П.М. Свободные кольца и их связи. М.: Мир, 1975.
5. Лопатинский Я.Б. Теория общих граничных задач. Киев: Наукова Думка, 1984.
6. Chyzak F., Salvy B. Non-commutative elimination in Ore algebras proves multivariate identities // J. Symbolic Comput. 1998. V. 26. P. 187–227.
7. Ore O. Linear equations in non-commutative fields // Ann. of Math. (2). 1931. V. 32. P. 463–477.
8. Abramov S., Petkovšek M., Ryabenko A., On linear dependence of rows and columns in matrices over non-commutative domains // Proceedings of the 2022 International Symposium on Symbolic and Algebraic Computation, ACM. 2022. P. 39–43.

**УРАВНЕНИЯ
В ЧАСТНЫХ ПРОИЗВОДНЫХ**

УДК 519.633

**ALTERNATIVE DIRECTION IMPLICIT METHOD FOR SOLVING
FIRST ORDER 2D HYPERBOLIC DELAY DIFFERENTIAL EQUATIONS¹⁾**

© 2023 г. S. Karthick¹, V. Subburayan^{1,*}

¹ *Department of Mathematics, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur-603 203, Tamilnadu, India*

**e-mail: suburayan123@gmail.com*

Поступила в редакцию 26.11.2022 г.
Переработанный вариант 26.11.2022 г.
Принята к публикации 02.02.2023 г.

Неявный численный метод переменных направлений для решения двумерной гиперболической задачи первого порядка с запаздыванием. Рассматриваются двумерные гиперболические дифференциальные уравнения первого порядка с запаздыванием. Показан характер распространения неоднородности решения по пространству. Предложен неявный разностный метод переменных направлений и метод против потока конечных разностей. Доказано, что метод обладает первым порядком сходимости. Дан ряд иллюстрирующих численных примеров, а также показано применение этого подхода к задаче с запаздыванием.

Ключевые слова: неявный метод переменных направлений, двумерное гиперболическое уравнение, схема против потока, принцип максимума.

DOI: 10.31857/S0044466923050113, **EDN:** DXETDR

¹⁾ Полный текст статьи печатается в английской версии журнала.

**УРАВНЕНИЯ
В ЧАСТНЫХ ПРОИЗВОДНЫХ**

**MULTIWAVE INTERACTION SOLUTIONS
FOR A NEW EXTENDED EQUATION IN $(4 + 1)$ -DIMENSION¹⁾**

© 2023 г. Yan Yang¹, Yinping Liu^{1,2,*}

¹ East China Normal University (ECNU), School of Mathematical Sciences, Shanghai, China

² BShanghai Key Laboratory of PMMP, Shanghai, China

*e-mail: ypliu@cs.ecnu.edu.cn

Поступила в редакцию 20.12.2022 г.

Переработанный вариант 20.12.2022 г.

Принята к публикации 02.02.2023 г.

Многоволновые решения для нового эволюционного уравнения в пространстве $(4 + 1)$ -измерений. Исследовано новое $(4 + 1)$ -мерное нелинейное эволюционное уравнение. Изучена его интегрируемость по Пенлеве методом Крушкаля, затем решение многоволнового взаимодействия для этого уравнения исследовано с помощью нескольких различных подходов. Показано, что это уравнение имеет “богатые” волновые решения со сложным характером взаимодействия.

Ключевые слова: интегрируемость по Пенлеве, многоволновое взаимодействие, алгоритм декомпозиции – солитона.

DOI: 10.31857/S0044466923050228, **EDN:** DXNYQA

¹⁾ Полный текст статьи печатается в английской версии журнала.

ВОССТАНОВЛЕНИЕ ДВУХ ФУНКЦИЙ В МОДЕЛИ КОЛЕБАНИЙ
СТРУНЫ, ОДИН КОНЕЦ КОТОРОЙ ПОМЕЩЕН
В ПОДВИЖНУЮ СРЕДУ¹⁾© 2023 г. О. А. Андреянова^{1,*}, А. Ю. Щеглов^{2,**}¹ 119991 Москва, Ленинские горы, 1, МГУ им. М.В. Ломоносова, Россия² 518172 Провинция Гуандун, Шэньчжэнь, район Лунган, Даюньсирьчэн, ул. Гоцзидасююань, 1,
Университет МГУ-ППИ в Шэньчжэне, Китай

*e-mail: oksashka@gmail.com

**e-mail: shcheg@cs.msu.ru

Поступила в редакцию 05.08.2022 г.
Переработанный вариант 23.10.2022 г.
Принята к публикации 02.02.2023 г.

Рассматривается обратная задача определения коэффициентов в модели малых поперечных колебаний однородной конечной струны, один конец которой помещен в подвижную среду, а другой свободен. Колебания моделируются уравнением гиперболического типа на отрезке. Одно краевое условие имеет неклассический вид. Дополнительной информацией для решения обратной задачи являются значения решения прямой задачи при известном фиксированном значении пространственного аргумента. В рамках обратной задачи определения требуются функция в неклассическом краевом условии и функциональный множитель в правой части уравнения. Доказаны теорема единственности и теорема существования решения обратной задачи. Для прямой задачи установлены условия однозначной разрешимости в виде, упрощающем исследование обратной задачи. Для численного решения обратной задачи предложен алгоритм поэтапного раздельного восстановления искомых функций с использованием метода последовательных приближений для решения интегральных уравнений. Библ. 23.

Ключевые слова: итерационный алгоритм, уравнение колебаний, обратная задача.**DOI:** 10.31857/S0044466923050046, **EDN:** PJPDA

1. ВВЕДЕНИЕ. МОДЕЛЬ И ПОСТАНОВКА ЗАДАЧ

Пусть при положительных значениях $a, \beta, T, l < aT$, прямая задача имеет вид:

$$u_{tt}(x, t) = a^2 u_{xx}(x, t) + f(x)g(t), \quad (x, t) \in Q_T, \quad (1)$$

$$-\beta u_x(x, t)|_{x=0} = \mu(t) - u_t(x, t)|_{x=0}, \quad 0 \leq t \leq T, \quad (2)$$

$$u_x(x, t)|_{x=l} = 0, \quad 0 \leq t \leq T, \quad (3)$$

$$u(x, t)|_{t=0} = \varphi(x), \quad u_t(x, t)|_{t=0} = \psi(x), \quad 0 \leq x \leq l, \quad (4)$$

где $Q_T = \{(x, t) : 0 < x < l, 0 < t \leq T\}$. Уравнение (1) и краевые условия (2), (3) описывают изменение амплитуды $u(x, t)$ малых поперечных колебаний струны, левый конец которой “испытывает сопротивление внешней среды, пропорциональное скорости ее движения” (см. [1, Гл. II, § 1, п. 7]), а правый конец струны не закреплен. Произведение функций $f(x)g(t)$ определяет линейную плотность силы, действующей на боковую поверхность струны, коэффициент a – скорость распространения колебаний вдоль струны, коэффициент β – натяжение (сила натяжения) струны по направлению ее оси, функция $\mu(t)$ – скорость внешней среды в плоскости поперечных колебаний струны при $x = 0$. Условия (4) являются начальными.

¹⁾Работа выполнена при частичной финансовой поддержке National Natural Science Foundation of China (No. 12171036) и Beijing Natural Science Foundation (Key Project No. Z210001).

Задача (1)–(4) может рассматриваться как модель неравномерного движения бура в глубоких нефтяных скважинах, что определяет интерес к исследованию задач в схожих постановках [2], [3]. Идентификация параметров в задачах для уравнений колебаний привлекает повышенное внимание при анализе геофизических процессов [4–7], в задачах оптимального управления [8–10], в постановках обратных задач [11–14]. Численное решение таких задач сопровождается использованием проработанных, ориентированных на приближенное решение некорректных задач, алгоритмов [15], [16] для компьютерной обработки экспериментальных данных.

Пусть в рамках обратной задачи при заданных положительных значениях $a, \beta, T, l < aT$, известны функции $\varphi(x), \psi(x), g(t)$ при $x \in [0, l]$ и $t \in [0, T]$, и дополнительно задана функция $h(t)$, $t \in [0, T]$, такая, что

$$h(t) = u(l, t), \quad t \in [0, T], \quad (5)$$

где $u(x, t)$ – решение задачи (1)–(4). В рамках обратной задачи пусть требуется восстановить функции $f(s)$ и $\mu(t)$ при $s \in [0, l]$, $t \in [0, \hat{T}]$, $\hat{T} = T - (l/a)$, и затем получить решение $u(x, t)$ прямой задачи (1)–(4) на множестве $\Lambda_{l, T} = \{(x, t) : 0 \leq t \leq T - (l - x)/a, 0 \leq x \leq l\}$, так, чтобы найденные функции $f(s)$, $\mu(t)$ и $u(x, t)$ удовлетворяли уравнению (1) на множестве $\Lambda_{l, T}$, условию (2) при $t \in [0, \hat{T}]$, условиям (3), (5) при $t \in [0, T]$ и начальным условиям (4). Множество $\Lambda_{l, T} \subset \bar{Q}_T$, где \bar{Q}_T – замыкание множества Q_T , имеет границу в форме прямоугольной трапеции с основаниями в виде отрезков, на одном из которых $x = 0$ и $t \in [0, \hat{T}]$, и на другом $x = l$ и $t \in [0, T]$. На перпендикулярной основаниям боковой стороне трапеции задаются начальные условия (4) при $t = 0 \forall x \in [0, l]$. Другая, наклонная, боковая сторона трапеции, ограничивающей множество $\Lambda_{l, T}$, является частью характеристики уравнения (1), соединяющей вершины трапеции с координатами $(x, t) = (l, T)$ и $(x, t) = (0, \hat{T})$.

Выделим достаточные условия существования и единственности решения прямой задачи в форме, позволяющей перейти к исследованию условий единственности и условий существования решения обратной задачи. На основе анализа решения обратной задачи построим итерационный алгоритм ее приближенного решения.

2. ПРЯМАЯ ЗАДАЧА

Теорема 1. Если заданы функции $\varphi(x), \psi(x), f(x), g(t), \mu(t)$ такие, что значение $\mu'(0)$ существует, и выполняются условия

$$\varphi(x) \in C^2[0, l], \quad \psi(x) \in C^1[0, l], \quad g(t) \in C^1[0, T], \quad \varphi'(l) = \psi'(l) = 0, \quad (6)$$

$$f(x) \in C[0, l], \quad \mu(t) \in C[0, T], \quad \mu(0) = \psi(0) - \beta\varphi'(0), \quad \mu'(0) = a^2\varphi''(0) - \beta\psi'(0) + f(0)g(0), \quad (7)$$

то задача (1)–(4) имеет единственное решение $u(x, t) \in C^2(\bar{Q}_T)$.

Доказательство. Установим, что решение задачи (1)–(4) совпадает на множестве \bar{Q}_T с решением $v(x, t)$ задачи Коши, получаемой продолжением известных функций $\varphi(x), \psi(x)$ и $f(x)$ в виде функций $\hat{\varphi}(x), \hat{\psi}(x)$ и $\hat{f}(x)$ на значения аргумента $x \in (-\infty, 0) \cup (l, +\infty)$. Для этого функция $\psi(x)$ с отрезка $[0, l]$ продолжается за точку $x = 0$ на луч $x \in (-\infty, 0)$ линейной функцией с получением непрерывно дифференцируемого продолжения $\psi(x)$ в виде функции $\hat{\psi}(x)$ для $x \in (-\infty, l]$. Функция $f(x)$ продолжается за точку $x = 0$ на луч $x \in (-\infty, 0)$ постоянным значением $f(0)$ с получением непрерывной функции $\hat{f}(x)$ для $x \in (-\infty, l]$. Функция $\varphi(x)$ доопределяется за точкой $x = 0$ на луче $x \in (-\infty, 0)$ значениями дополнительно вычисляемой функции $z(x)$ так, чтобы при $x = 0$ для решения $v(x, t)$ задачи Коши выполнялось краевое условие (2), и продолжение функции $\varphi(x)$ являлось дважды непрерывно дифференцируемым при $x \in (-\infty, l]$. Продолжение функций $\varphi(x), \psi(x)$ и $f(x)$ за точку $x = l$ на луч $[l, +\infty)$ осуществляется четным образом, т.е. осесимметрично относительно прямой $x = l$ для точек графиков функций $\varphi(x), \psi(x)$ и $f(x)$, т.е. так, чтобы решение $v(x, t)$ удовлетворяло однородному условию II рода (3) в точке $x = l$.

Рассмотрим для функции $v(x, t)$ представленную в предыдущем абзаце задачу Коши только в тех точках (x, t) полуплоскости $t \geq 0$, которые располагаются на характеристиках уравнения (1),

проходящих через точки множества \bar{Q}_T и, следовательно, задание исходных данных задачи (1)–(4) в которых непосредственно влияет на значения решения задачи (1)–(4) при $(x, t) \in \bar{Q}_T$. Пусть

$$v_{tt}(x, t) = a^2 v_{xx}(x, t) + \hat{f}(x)g(t), \quad (x, t) \in D_{l,T}, \tag{8}$$

$$v(x, t)|_{t=0} = \hat{\phi}(x), \quad v_t(x, t)|_{t=0} = \hat{\psi}(x), \quad -aT \leq x \leq l + aT, \tag{9}$$

где $D_{l,T} = \{(x, t) : -(T - t)a \leq x \leq l + (T - t)a, 0 \leq t \leq T\}$ – множество точек с координатами (x, t) на плоскости с границей в форме равнобедренной трапеции с основаниями в виде отрезка при $t = 0$ $\forall x \in [-aT, l + aT]$ и отрезка при $t = T$ $\forall x \in [0, l]$.

Доказательство проведем для значений T из полуинтервала $((l/a), (3l/a)]$, что вполне позволит продемонстрировать схему построения и исследования решений задачи (8), (9) и задачи (1)–(4) с возможностью перенесения такого анализа на бóльшие значения $T > (3l/a)$.

Пусть в задаче (8), (9)

$$\hat{\phi}(x) = \begin{cases} z(x), & x \in [-aT, 0), \\ \varphi(x), & x \in [0, l], \\ \varphi(2l - x), & x \in (l, 2l], \\ z(2l - x), & x \in (2l, l + aT], \end{cases} \tag{10}$$

$$\hat{\psi}(x) = \begin{cases} \psi(0) + \psi'(0)x, & x \in [-aT, 0), \\ \psi(x), & x \in [0, l], \\ \psi(2l - x), & x \in (l, 2l], \\ \psi(0) - \psi'(0)(x - 2l), & x \in (2l, l + aT], \end{cases} \quad \hat{f}(x) = \begin{cases} f(0), & x \in [-aT, 0), \\ f(x), & x \in [0, l], \\ f(2l - x), & x \in (l, 2l], \\ f(0), & x \in (2l, l + aT], \end{cases} \tag{11}$$

где функция $z(x)$ определяется так, чтобы $\hat{\phi}(x) \in C^2[-aT, l + aT] \forall T \in ((l/a), (3l/a)]$ при выполнении условий (2), (3).

Однозначная разрешимость задачи (8), (9) установлена [17] в рамках исследования решения задачи Коши при $x \in (-\infty, +\infty)$ для неоднородного уравнения $u_{tt} = a^2 u_{xx} + w(x, t)$ при условии непрерывной дифференцируемости функции $w(x, t)$ с использованием формулы Даламбера для неоднородного уравнения, известной также как формула Дюамеля. Докажем, что такая же однозначная разрешимость задачи (8), (9) имеет место и в случае функций $f(x) \in C[0, l]$ и $g(t) \in C^1[0, T]$.

Введенные по формулам (11) функции $\hat{\psi}(x)$ и $\hat{f}(x)$ сохраняют гладкость функций $\psi(x)$ и $f(x)$, имеющуюся в условиях (6), (7) теоремы: $\hat{\psi}(x) \in C^1[-aT, l + aT]$, $\hat{f}(x) \in C[-aT, l + aT]$. На основе формулы Даламбера для неоднородного уравнения $v_{tt} = a^2 v_{xx} + \hat{f}(x)g(t)$ при условиях $\hat{\phi}(x) \in C^2[-aT, l + aT]$, $\hat{\psi}(x) \in C^1[-aT, l + aT]$, $\hat{f}(x)g(t) \in C^1(D_{l,T})$ однозначное решение $v(x, t) \in C^2(D_{l,T})$ задачи (8), (9) существует (по теореме 7.6 в [17]) и имеет вид:

$$v(x, t) = \frac{\hat{\phi}(x - at) + \hat{\phi}(x + at)}{2} + \frac{1}{2a} \int_{x-at}^{x+at} \hat{\psi}(s) ds + \frac{1}{2a} \int_0^t g(\tau) \int_{x-(t-\tau)a}^{x+(t-\tau)a} \hat{f}(s) ds d\tau, \quad (x, t) \in D_{l,T}. \tag{12}$$

Дифференцируя формулу (12) для всех $(x, t) \in D_{l,T}$ и затем вводя в двух повторных интегралах подынтегральные аргументы $\xi = x + (t - \tau)a$ и $\eta = x - (t - \tau)a$, получаем последовательно формулы для производных решения $v(x, t)$:

$$v_x(x, t) = \frac{\hat{\phi}'(x - at) + \hat{\phi}'(x + at)}{2} + \frac{\hat{\psi}(x + at) - \hat{\psi}(x - at)}{2a} - \frac{1}{2a^2} \int_{x-at}^x g\left(t - \frac{x - \eta}{a}\right) \hat{f}(\eta) d\eta + \frac{1}{2a^2} \int_x^{x+at} g\left(t + \frac{x - \xi}{a}\right) \hat{f}(\xi) d\xi, \quad (x, t) \in D_{l,T}; \tag{13}$$

$$v_t(x, t) = a \frac{\hat{\phi}'(x + at) - \hat{\phi}'(x - at)}{2} + \frac{\hat{\psi}(x + at) + \hat{\psi}(x - at)}{2} + \frac{1}{2a} \int_{x-at}^x g\left(t - \frac{x-\eta}{a}\right) \hat{f}(\eta) d\eta + \frac{1}{2a} \int_x^{x+at} g\left(t + \frac{x-\xi}{a}\right) \hat{f}(\xi) d\xi, \quad (x, t) \in D_{l,T}; \tag{14}$$

$$v_{xx}(x, t) = \frac{\hat{\phi}''(x - at) + \hat{\phi}''(x + at)}{2} + \frac{\hat{\psi}'(x + at) - \hat{\psi}'(x - at)}{2a} + \frac{g(0)}{2a^2} (\hat{f}(x - at) + \hat{f}(x + at)) - \frac{1}{a^2} g(t) \hat{f}(x) + \frac{1}{2a^3} \int_{x-at}^x g'\left(t - \frac{x-\eta}{a}\right) \hat{f}(\eta) d\eta + \frac{1}{2a^3} \int_x^{x+at} g'\left(t + \frac{x-\xi}{a}\right) \hat{f}(\xi) d\xi, \quad (x, t) \in D_{l,T}; \tag{15}$$

$$v_{tt}(x, t) = a^2 \frac{\hat{\phi}''(x + at) + \hat{\phi}''(x - at)}{2} + a \frac{\hat{\psi}'(x + at) - \hat{\psi}'(x - at)}{2} + g(0) \frac{\hat{f}(x - at) + \hat{f}(x + at)}{2} + \frac{1}{2a} \int_{x-at}^x g'\left(t - \frac{x-\eta}{a}\right) \hat{f}(\eta) d\eta + \frac{1}{2a} \int_x^{x+at} g'\left(t + \frac{x-\xi}{a}\right) \hat{f}(\xi) d\xi, \quad (x, t) \in D_{l,T}. \tag{16}$$

Формулы (13)–(16) дают значения непрерывных ограниченных производных $v_x(x, t)$, $v_t(x, t)$, $v_{xx}(x, t)$, $v_{tt}(x, t)$ решения задачи (8), (9) при $\hat{\phi}(x) \in C^2[-aT, l + aT]$, $\hat{\psi}(x) \in C^1[-aT, l + aT]$, при $g(t) \in C^1[0, T]$ и при $\hat{f}(x) \in C[-aT, l + aT]$. Представленные формулами (15), (16) производные $v_{xx}(x, t)$, $v_{tt}(x, t)$ удовлетворяют уравнению (8), а функция $v(x, t)$ и производная $v_t(x, t)$ – условиям (9). Из этого следует, что задание в уравнении (8) неоднородности в виде произведения $\hat{f}(x)g(t)$ позволяет ослабить условия на функцию $f(x)$ до непрерывности $f(x)$ на отрезке $[0, l]$ по сравнению с условием непрерывной дифференцируемости из теоремы 7.6 в [17] с выражением однозначного решения задачи (8), (9) по той же формуле (12). Итак, задача (8), (9) имеет решение $v(x, t) \in C^2(D_{l,T})$ и при функциях $\hat{\phi}(x)$, $\hat{\psi}(x)$, $\hat{f}(x)$, сохраняющих на отрезке $[-aT, l + aT]$ гладкость функций $\phi(x)$, $\psi(x)$, $f(x)$, удовлетворяющих условиям теоремы.

Определим функцию $z(x)$. Подставляя правые части формул (13) и (14) в условие (2) при $x = 0$ и выполняя замену $s = -at$, с учетом формул (10), (11) получаем при $s \in [-l, 0)$ формулу

$$z'(s) = -\frac{2}{a + \beta} \mu\left(\frac{-s}{a}\right) + \frac{a - \beta}{a + \beta} \varphi'(-s) + \frac{1}{a} \psi(0) + \frac{s}{a} \psi'(0) + \frac{a - \beta}{(a + \beta)a} \psi(-s) + \frac{1}{a^2} f(0) \int_s^0 g\left(\frac{\eta - s}{a}\right) d\eta + \frac{a - \beta}{(a + \beta)a^2} \int_0^{-s} g\left(-\frac{s + \xi}{a}\right) f(\xi) d\xi, \quad s \in [-l, 0). \tag{17}$$

Интегрируя производную $z'(s)$ на полуинтервале $[x, 0) \subset [-l, 0)$, $\forall x \in [-l, 0)$ с учетом условия $\lim_{x \rightarrow 0-0} z(x) = \varphi(0)$ получаем

$$z(x) = \hat{z}(x) = \frac{2}{a + \beta} \int_x^0 \mu\left(\frac{-\tau}{a}\right) d\tau + \frac{2a\varphi(0)}{a + \beta} - \frac{a - \beta}{a + \beta} \varphi(-x) + \frac{\psi(0)}{a} x + \frac{\psi'(0)}{2a} x^2 - \frac{a - \beta}{(a + \beta)a} \int_x^0 \psi(-\tau) d\tau - \frac{1}{a^2} f(0) \int_x^0 \int_{x\tau}^0 g\left(\frac{\eta - \tau}{a}\right) d\eta d\tau - \frac{a - \beta}{(a + \beta)a^2} \int_x^0 \int_0^{-\tau} g\left(-\frac{\tau + \xi}{a}\right) f(\xi) d\xi d\tau, \quad x \in [-l, 0). \tag{18}$$

Из формулы (18) при $x = -l$ после замены подынтегральных аргументов $\tau = -\theta$, $v = -\eta$ в повторных интегралах получаем значение

$$\hat{z}(-l) = \frac{2}{a + \beta} \int_{-l}^0 \mu\left(\frac{-\tau}{a}\right) d\tau + \frac{2a\varphi(0)}{a + \beta} - \frac{a - \beta}{a + \beta} \varphi(l) - \frac{l}{a} \psi(0) + \frac{l^2}{2a} \psi'(0) - \frac{a - \beta}{(a + \beta)a} \int_0^l \psi(\tau) d\tau - \frac{1}{a^2} f(0) \int_0^{\theta} \int_0^{\theta} g\left(\frac{\theta - v}{a}\right) dv d\theta - \frac{a - \beta}{(a + \beta)a^2} \int_0^{\theta} \int_0^{\theta} g\left(\frac{\theta - \xi}{a}\right) f(\xi) d\xi d\theta. \tag{19}$$

Затем аналогично, подставляя правые части формул (13) и (14) в условие (2) при $x = 0$, получаем формулы для функций $z'(s)$ и $z(x)$ последовательно для $s, x \in [-\min\{aT, 2l\}, -l]$. После этого и в случае $aT \in (2l, 3l]$ получаем, подставляя правые части формул (13) и (14) в условие (2), формулы для функций $z'(s)$ и $z(x)$ при $s, x \in [-aT, -2l]$ с непрерывным продолжением функции $z(x)$ и ее производных $z'(x)$, $z''(x)$ в точках $x = -l$ и $x = -2l$. Итак, сначала при $s \in [-\min\{aT, 2l\}, -l]$ получаем

$$\begin{aligned} z'(s) = & \frac{-2}{a+\beta} \mu\left(\frac{-s}{a}\right) - \frac{a-\beta}{a+\beta} \varphi'(s+2l) + \frac{1}{a} \psi(0) + \frac{s}{a} \psi'(0) + \frac{a-\beta}{(a+\beta)a} \psi(s+2l) + \\ & + \frac{1}{a} f(0) \int_0^{-s/a} g(\tau) d\tau + \frac{a-\beta}{(a+\beta)a} \int_{-(s+l)/a}^{-s/a} g(\tau) f(-s-a\tau) d\tau + \\ & + \frac{a-\beta}{(a+\beta)a} \int_0^{-(s+l)/a} g(\tau) f(2l+s+a\tau) d\tau, \quad s \in [-\max\{aT, 2l\}, -l]. \end{aligned} \quad (20)$$

И с учетом полученного выше формулой (19) значения $\hat{z}(-l)$, интегрируя производную $z'(s)$ на полуинтервале $[x, -l] \subset [-\max\{aT, 2l\}, -l] \forall x \in [-\max\{aT, 2l\}, -l]$, получаем

$$\begin{aligned} z(x) = \hat{z}(x) = & \frac{2}{a+\beta} \int_x^{-l} \mu\left(\frac{-s}{a}\right) ds + \hat{z}(-l) + \frac{a-\beta}{a+\beta} (\varphi(l) - \varphi(0)) + \frac{(x+l)}{a} \psi(0) - \\ & - \frac{(x^2-l^2)}{a} \psi'(0) - \frac{f(0)}{a} \int_x^{-l} \int_0^{-s/a} g(\tau) d\tau ds - \frac{a-\beta}{(a+\beta)a} \left[\int_{x+2l}^l \psi(\xi) d\xi + \int_{x-(s+l)/a}^{-l} \int_{-s/a}^{-s/a} g(\tau) f(-s-a\tau) d\tau ds + \right. \\ & \left. + \int_x^{-l} \int_0^{-(s+l)/a} g(\tau) f(2l+s+a\tau) d\tau ds \right], \quad x \in [-\max\{aT, 2l\}, -l]. \end{aligned} \quad (21)$$

При этом, если $aT \in (2l, 3l]$, то при $x = -2l$ получаем значение

$$\begin{aligned} \hat{z}(-2l) = & \frac{2}{a+\beta} \int_{-2l}^{-l} \mu\left(\frac{-s}{a}\right) ds + \hat{z}(-l) + \frac{a-\beta}{a+\beta} (\varphi(l) - \varphi(0)) - \frac{l}{a} \psi(0) - \frac{3l^2}{a} \psi'(0) - \frac{1}{a} f(0) \int_{-2l}^{-l} \int_0^{-s/a} g(\tau) d\tau - \\ & - \frac{a-\beta}{(a+\beta)a} \left[\int_0^l \psi(\xi) d\xi + \int_{-2l}^{-l} \int_{-(s+l)/a}^{-s/a} g(\tau) f(-s-a\tau) d\tau ds + \int_{-2l}^{-l} \int_0^{-(s+l)/a} g(\tau) f(2l+s+a\tau) d\tau ds \right]. \end{aligned}$$

Затем, если $aT \in (2l, 3l]$, то получаем при $s \in [-aT, -2l]$ сначала

$$\begin{aligned} z'(s) = & \frac{-2}{a+\beta} \mu\left(\frac{-s}{a}\right) - \frac{a-\beta}{a+\beta} \hat{z}'(s+2l) + \frac{f(0)}{a} \int_0^{-s/a} g(\tau) d\tau + \\ & + \frac{a-\beta}{(a+\beta)a} \left[2l\psi'(0) + f(0) \int_0^{-(s+2l)/a} g(\tau) d\tau + \int_{-(s+l)/a}^{-s/a} g(\tau) f(-s-a\tau) d\tau + \right. \\ & \left. + \int_{-(s+2l)/a}^{-(s+l)/a} g(\tau) f(2l+s+a\tau) d\tau \right], \quad s \in [-aT, -2l], \end{aligned} \quad (22)$$

где значения используемой в правой части формулы (22) производной $\hat{z}'(s+2l)$, $s \in [-aT, -2l]$, уже определены выше в формуле (18). И с учетом полученного значения $\hat{z}(-2l)$, интегрируя производную $z'(s)$ на полуинтервале $[x, -2l] \subset [-aT, -2l]$, получаем $\forall x \in [-aT, -2l]$

$$\begin{aligned}
 z(x) = \hat{z}(x) = & \frac{2}{a + \beta} \int_x^{-2l} \mu\left(\frac{-s}{a}\right) ds + \hat{z}(-2l) + \frac{a - \beta}{a + \beta} (\varphi(0) - \hat{z}(x + 2l)) - \frac{f(0)}{a} \int_x^{-2l-s/a} \int_0^0 g(\tau) d\tau ds + \\
 & + 2 \frac{(a - \beta)l}{(a + \beta)a} \psi'(0)(x + 2l) - \frac{a - \beta}{(a + \beta)a} \int_x^{-2l} \left[\int_{-(s+l)/a}^{-s/a} g(\tau) f(-s - a\tau) d\tau + \right. \\
 & \left. + \int_{-(s+2l)/a}^{-(s+l)/a} g(\tau) f(2l + s + a\tau) d\tau + f(0) \int_0^{-(s+2l)/a} g(\tau) d\tau \right] ds, \quad x \in [-aT, -2l].
 \end{aligned} \tag{23}$$

При этом $\lim_{s \rightarrow 0-0} z(s) = \varphi(0)$. Из входящих в условие (7) равенств на значения $\mu(0), \mu'(0): \mu(0) = \psi(0) - \beta\varphi'(0), \mu'(0) = a^2\varphi''(0) - \beta\psi'(0) + f(0)g(0)$, следует, что

$$\begin{aligned}
 \lim_{s \rightarrow 0-0} z'(s) &= -\frac{2}{a + \beta} \mu(0) + \frac{a - \beta}{a + \beta} \varphi'(0) + \left(\frac{1}{a} + \frac{a - \beta}{(a + \beta)a} \right) \psi(0) = \varphi'(0), \\
 \lim_{s \rightarrow 0-0} z''(s) &= \frac{2\mu'(0)}{(a + \beta)a} - \frac{a - \beta}{a + \beta} \varphi''(0) + \left(1 - \frac{a - \beta}{a + \beta} \right) \frac{\psi'(0)}{a} - \left(1 + \frac{a - \beta}{a + \beta} \right) \frac{f(0)g(0)}{a^2} = \varphi''(0).
 \end{aligned}$$

Из непрерывной дифференцируемости правых частей формул (17), (20), (22) и из непрерывности построенной выше функции $z(x), x \in [-aT, 0)$ следует, что $z(x) \in C^2[-aT, 0)$. Из полученных равенств $\lim_{s \rightarrow 0-0} z(s) = \varphi(0), \lim_{s \rightarrow 0-0} z'(s) = \varphi'(0), \lim_{s \rightarrow 0-0} z''(s) = \varphi''(0)$ и из $z(x) \in C^2[-aT, 0)$ следует, что начальная функция $\hat{\varphi}(x)$ задачи (8), (9) имеет ограниченную непрерывную производную $\hat{\varphi}''(x)$ при $x \in [-aT, l + aT]$.

Из установленных свойств функции $\hat{\varphi}(x) \in C^2[-aT, l + aT]$, и из $\hat{\psi}(x) \in C^1[-aT, l + aT], g(t) \in C^1[0, T], \hat{f}(x) \in C[-aT, l + aT]$ следует, что решение $v(x, t) \in C^2(D_{l,T})$ задачи (8), (9) удовлетворяет условиям (2), (3) и однозначно определяется формулой (12) на множестве $D_{l,T}$. Таким образом, формула (12) с учетом условий теоремы дает и на множестве $\bar{Q}_T \subset D_{l,T}$ единственное на множестве $\bar{Q}_T \subset D_{l,T}$, ограниченное решение $u(x, t) = v(x, t) \in C^2(\bar{Q}_T)$ задачи (1)–(4). Теорема 1 доказана.

3. РЕШЕНИЕ ПРЯМОЙ ЗАДАЧИ В НЕКОТОРЫХ ПОДОБЛАСТЯХ

Конкретизируем формулы для решения $u(x, t)$ прямой задачи (1)–(4) в той подобласти, где решение $u(x, t)$ представлено непосредственно формулой Даламбера (12), и в тех подобластях изменения аргументов x и t , которым принадлежит отрезок прямой $x = l$ при $t \in [0, T]$, на котором задается в виде функции $h(t)$ дополнительная информация (5) при постановке обратной задачи (1)–(5).

С рассчитанными значениями $z(x)$, представленными формулами (18) и (21), функция $\hat{\varphi}(x)$ в начальном условии задачи (8), (9) на отрезке $[l - aT, l + aT]$ – части всего множества $[-aT, l + aT]$ определения функции $\hat{\varphi}(x)$, принимает при $T \in ((l/a), (3l/a)]$ вид

$$\hat{\varphi}(x) = \begin{cases} \frac{2}{a + \beta} \int_x^0 \mu\left(\frac{-\tau}{a}\right) d\tau + F(x), & x \in [l - aT, 0), \\ \varphi(x), & x \in [0, l], \\ \varphi(2l - x), & x \in (l, 2l], \\ \frac{2}{a + \beta} \int_{2l-x}^0 \mu\left(\frac{-\tau}{a}\right) d\tau + F(2l - x), & x \in (2l, l + aT], \end{cases} \tag{24}$$

где $\hat{\phi}(x) \in C^2[-aT, l + aT]$, и при этом $[l - aT, l + aT] \subset [-aT, l + aT]$. В формуле (24) из полученных при доказательстве теоремы 1 формул имеем

$$F(s) = 2 \frac{a\phi(0)}{a + \beta} - \frac{a - \beta}{a + \beta} \phi(-s) + \frac{\psi(0)}{a} s + \frac{\psi'(0)}{2a} s^2 - \frac{a - \beta}{(a + \beta)a} \int_s^0 \psi(-\tau) d\tau - \frac{f(0)}{a^2} \int_s^0 \int_\tau^0 g\left(\frac{\eta - \tau}{a}\right) d\eta d\tau - \frac{a - \beta}{(a + \beta)a^2} \int_s^0 \int_0^{-\tau} g\left(\frac{\tau + \xi}{-a}\right) f(\xi) d\xi d\tau, \quad s \in [-\min\{aT - l, l\}, 0]. \tag{25}$$

И в случае если значения a, T, l таковы, что $aT \in (2l, 3l]$, то на продолжении области определения функции $F(s)$ за точку $s = -l$ на отрезок $[-aT, -l]$ получаем

$$F(x) = \phi(0) + \frac{x}{a} \psi(0) + \frac{3l^2 - 2x^2}{2a} \psi'(0) - \frac{a - \beta}{(a + \beta)a} \left(\int_0^l \psi(\xi) d\xi + \int_{x+2l}^l \psi(\xi) d\xi \right) - \frac{f(0)}{a} \int_x^{0-s/a} \int_0^{\tau} g(\tau) d\tau - \frac{a - \beta}{(a + \beta)a^2} \left[\int_{-l}^{0-s/a} \int_0^{\tau} g(\tau) f(-s - a\tau) d\tau ds + \int_x^{-l} \int_{-(s+l)/a}^{-s/a} g(\tau) f(-s - a\tau) d\tau ds + \int_x^{-l-(s+l)/a} \int_0^{\tau} g(\tau) f(s + a\tau + 2l) d\tau ds \right], \quad x \in [-\max\{l - aT, l\}, -l]. \tag{26}$$

Непосредственно формула Даламбера (12) дает решение $u(x, t)$ задачи (1)–(4) в виде

$$u(x, t) = \frac{\phi(x - at) + \phi(x + at)}{2} + \frac{1}{2a} \int_{x-at}^{x+at} \psi(s) ds + \frac{1}{2a} \int_0^t g(\tau) \int_{x-(t-\tau)a}^{x+(t-\tau)a} f(s) ds d\tau, \quad (x, t) \in \Omega_0,$$

где $\Omega_0 = \{(x, t) : at \leq x \leq l - at, 0 \leq t \leq (l/(2a))\}$ – характеристический треугольник на плоскости с координатами (x, t) с вершиной в точке $((l/2), (l/(2a)))$ и основанием на отрезке прямой $t = 0$, расположенном между точками $(0, 0)$ и $(l, 0)$ с $x \in [0, l]$.

Далее с представленной формулой (24) функцией $\hat{\phi}(x)$ на подмножестве Ω_1 области определения \bar{Q}_T задачи (1)–(4) решение $u(x, t)$ принимает вид

$$u(x, t) = \frac{\phi(x - at) + \phi(2l - x - at)}{2} + \frac{1}{2a} \left[\int_{x-at}^{2l-at-x} \psi(\xi) d\xi + 2 \int_{2l-at-x}^l \psi(\xi) d\xi + \int_{x-at}^x \int_0^{t-(x-\xi)/a} g(\tau) d\tau f(\xi) d\xi + \int_x^l \int_0^{t+(x-\xi)/a} g(\tau) d\tau f(\xi) d\xi + \int_{2l-at-x}^l \int_0^{t+(x+\xi-2l)/a} g(\tau) d\tau f(\xi) d\xi \right], \quad (x, t) \in \Omega_1, \tag{27}$$

где $\Omega_1 = \{(x, t) : ((l - x)/a) \leq t \leq (x/a), (l/2) \leq x \leq l\}$ – равнобедренный треугольник на плоскости (x, t) , имеющий основание на прямой $x = l$ при $t \in [0, (l/a)]$, и противоположную основанию вершину в точке с координатами $((l/2), (l/(2a)))$.

И затем на множестве Ω_2 , ограниченном сторонами равнобедренного треугольника на плоскости (x, t) с основанием на прямой $x = l$ при $t \in [(l/a), T]$, и противоположной основанию вершиной в точке с координатами $((3l - aT)/2, (l + 2a)/(2a))$, имеем

$$u(x, t) = \frac{1}{a + \beta} \int_{x-at}^0 \mu\left(\frac{-\tau}{a}\right) d\tau + \frac{1}{2} F(x - at) + \frac{1}{a + \beta} \int_{2l-at-x}^0 \mu\left(\frac{-\tau}{a}\right) d\tau + \frac{1}{2} F(2l - at - x) + \frac{\psi(0)}{a} (at - l) + \frac{\psi'(0)}{2a} (2(at + x - l)l - a^2 t^2 - x^2) + \frac{1}{a} \int_0^l \psi(\xi) d\xi + \frac{1}{2a} \left[f(0) \int_0^{t-(x/a)} ((t - \tau)a - x) g(\tau) d\tau + \int_0^{x-t-(x-\xi)/a} \int_0^{\tau} g(\tau) d\tau f(\xi) d\xi + \int_x^l \int_0^{t+(x-\xi)/a} g(\tau) d\tau f(\xi) d\xi + \right] \tag{28}$$

$$+ \int_0^{l+(x+\xi-2l)/a} \int_0^{t-(2l-x)/a} g(\tau) d\tau f(\xi) d\xi + f(0) \int_0^{t-(2l-x)/a} ((t-\tau)a + x - 2l) g(\tau) d\tau, \quad (x, t) \in \Omega_2,$$

где

$$\Omega_2 = \left\{ (x, t) : \frac{2l-x}{a} \leq t \leq T - \frac{l-x}{a}, \frac{3l-aT}{2} \leq x \leq l \right\}.$$

Объединение множеств $\Omega_1 \cup \Omega_2$ содержит весь отрезок, на котором задается на плоскости (x, t) дополнительное условие (5), используемое в постановке обратной задачи (1)–(5).

4. ОБРАТНАЯ ЗАДАЧА И УСЛОВИЯ ЕЕ РАЗРЕШИМОСТИ

Определение 1. Три функции $f(s)$, $\mu(\tau)$, $u(x, t)$, где $s \in [0, l]$, $\tau \in [0, \hat{T}]$, $\hat{T} = T - (l/a)$, $(x, t) \in \Lambda_{l, T}$, называются решением обратной задачи (1)–(5), если при известных положительных значениях a , β , T и l , таких, что $l < aT$, при заданных функциях $\varphi(x)$, $\psi(x)$, $g(t)$, где $x \in [0, l]$ и $t \in [0, T]$, а также при дополнительно заданной функции $h(t)$, $t \in [0, T]$, таких, что функции $\varphi(x)$, $\psi(x)$, $g(t)$ удовлетворяют условию (6) и условию

$$g(0) \neq 0, \quad (29)$$

у функции $h(t)$ существует значение $h'(0)$, и для нее выполняются условия

$$h(t) \in C[0, T], \quad h(0) = \varphi(l), \quad h'(0) = \psi(l), \quad (30)$$

а сами функции $f(s)$, $\mu(\tau)$, $u(x, t)$ таковы, что $u(x, t) \in C^2(\Lambda_{l, T})$, существует значение $\mu'(0)$, для функций $f(s)$, $\mu(\tau)$ выполняются условия

$$\begin{aligned} f(x) \in C[0, l], \quad \mu(t) \in C[0, \hat{T}], \quad \mu(0) &= \psi(0) - \beta\varphi'(0), \\ \mu'(0) &= a^2\varphi''(0) - \beta\psi'(0) + f(0)g(0), \end{aligned} \quad (31)$$

а также выполняется уравнение (1) на множестве $\Lambda_{l, T}$, выполняется краевое условие (2) при $t \in [0, \hat{T}]$, и выполняются условия (3)–(5).

Теорема 2. Если заданы функции $\varphi(x)$, $\psi(x)$, $g(t)$, $h(t)$, удовлетворяющие условиям (6), (29), (30), то обратная задача (1)–(5) не может иметь более одного решения.

Доказательство. Предположим противное утверждению теоремы 2, т.е. существование двух различных решений обратной задачи (1)–(5): трех функций $f_1(s)$, $\mu_1(\tau)$, $u_1(x, t)$ и трех функций $f_2(s)$, $\mu_2(\tau)$, $u_2(x, t)$. Тогда при заданных функциях $\varphi(x)$, $\psi(x)$, $g(t)$, $h(t)$, с учетом условия (5), формула (27) при $x = l$ для значений $t \in [0, (l/a)]$ приобретает вид равенства

$$h(t) = \varphi(l - at) + \frac{1}{a} \int_{l-at}^l \psi(s) ds + \frac{1}{a} \int_0^t g(\tau) \int_{l-(t-\tau)a}^l f(s) ds d\tau, \quad t \in \left[0, \frac{l}{a}\right].$$

Изменяя порядок интегрирования в повторном интеграле, из этого равенства получаем уравнение относительно функции $f(s)$, $s \in [0, l]$:

$$\frac{1}{a} \int_{l-at}^l \left(\int_0^{t-(l-s)/a} g(\tau) d\tau \right) f(s) ds = H_1(t), \quad t \in \left[0, \frac{l}{a}\right], \quad (32)$$

где

$$H_1(t) = h(t) - \varphi(l - at) - \frac{1}{a} \int_{l-at}^l \psi(s) ds, \quad t \in \left[0, \frac{l}{a}\right].$$

Вычитая из уравнения (32) с функцией $f_1(s)$ на месте $f(s)$ это же уравнение (32) с функцией $f_2(s)$ на месте $f(s)$, относительно разности $\Delta f(s) = f_1(s) - f_2(s)$, $s \in [0, l]$, получаем уравнение

$$\frac{1}{a} \int_{l-at}^l \left(\int_0^{t-(l-s)/a} g(\tau) d\tau \right) \Delta f(s) ds = 0, \quad t \in \left[0, \frac{l}{a} \right]. \quad (33)$$

Дифференцируя уравнение (33) по переменной t дважды и затем производя замену $l - at = x$, получаем

$$\Delta f(x) + \int_x^l K_1(s, x) \Delta f(s) ds = 0, \quad x \in [0, l], \quad (34)$$

– линейное интегральное уравнение Вольтерра II рода относительно функции $\Delta f(x)$, где

$$K_1(s, x) = \frac{1}{ag(0)} g' \left(\frac{s-x}{a} \right), \quad 0 \leq x \leq s \leq l. \quad (35)$$

В силу следующей из условий (6), (29) теоремы 2 непрерывности ядра (35) уравнение (34) имеет [18] только решение $\Delta f(s) \equiv 0$, из чего следует, что $f_1(s) = f_2(s) = \bar{f}(s)$, $s \in [0, l]$.

Перейдем к доказательству единственности функции $\mu(t)$ для $t \in [0, \hat{T}]$. Доказательство проведем для возможных значений $T \in ((l/a), (3l/a)]$, что вполне позволит продемонстрировать схему исследования единственности решений обратной задачи с возможностью перенесения анализа на большие значения $T > (3l/a)$. С учетом условия (5) рассмотрим формулу (28) при $x = l$ для $t \in [(l/a), T]$ как уравнение относительно функции $\mu(\tau)$ для $\tau \in [0, \hat{T}]$:

$$h(t) = \frac{2}{a + \beta} \int_{l-at}^0 \mu \left(-\frac{\tau}{a} \right) d\tau + F(l - at) + \frac{at - l}{a} \psi(0) - \frac{(l - at)^2}{2a} \psi'(0) + \frac{1}{a} \int_0^l \psi(\xi) d\xi + \quad (36)$$

$$+ \frac{f(0)}{a} \int_0^{t-(l/a)} ((t - \tau)a - l) g(\tau) d\tau + \frac{1}{a} \int_0^l \left(\int_0^{t-(l-\xi)/a} g(\tau) d\tau \right) f(\xi) d\xi, \quad t \in \left[\frac{l}{a}, T \right].$$

Выполняя в интеграле, содержащем функцию $\mu(t)$, замену подынтегрального аргумента $\tau = -a\theta$, и замену $\omega = t - (l/a)$, из равенства (36) получаем уравнение

$$\int_0^\omega \mu(\theta) d\theta = H_2(\omega), \quad \omega \in [0, \hat{T}], \quad (37)$$

где

$$H_2(\omega) = \frac{a + \beta}{2a} h \left(\omega + \frac{l}{a} \right) - \frac{a + \beta}{2a} F(-a\omega) - \frac{a + \beta}{2a} \psi(0)\omega + \frac{a + \beta}{4} \psi'(0)\omega^2 - \frac{a + \beta}{2a^2} \int_0^l \psi(\xi) d\xi - \quad (38)$$

$$- \frac{a + \beta}{2a} f(0) \int_0^\omega (\omega - \tau) g(\tau) d\tau - \frac{a + \beta}{2a^2} \int_0^l \left(\int_0^{\omega+(\xi/a)} g(\tau) d\tau \right) f(\xi) d\xi, \quad \omega \in [0, \hat{T}].$$

Вычитая из уравнения (37) с функцией $f_1(s) = \bar{f}(s)$ и с функцией $\mu_1(\theta)$ на месте $\mu(\theta)$ это же уравнение (37) с функцией $f_2(s) = \bar{f}(s)$ с функцией $\mu_2(\theta)$ на месте $\mu(\theta)$, получаем относительно разности $\Delta\mu(\theta) = \mu_1(\theta) - \mu_2(\theta)$ уравнение

$$\int_0^\omega \Delta\mu(\theta) d\theta = 0, \quad \omega \in [0, \hat{T}]. \quad (39)$$

Дифференцируя уравнение (39) по аргументу ω , получаем равенство $\Delta\mu(\omega) \equiv 0 \forall \omega \in [0, \hat{T}]$, из чего следует, что $\mu_1(t) = \mu_2(t) = \bar{\mu}(t)$, $t \in [0, \hat{T}]$.

Итак, установлено попарное равенство функций $\mu_1(\tau) = \mu_2(\tau)$ и $f_1(s) = f_2(s)$ при $\tau \in [0, \hat{T}]$ и $s \in [0, l]$. Ввиду совпадения исходных коэффициентов на всей области их определения по теореме 1 совпадут и решения $u_1(x, t) = u_2(x, t)$, $(x, t) \in \Lambda_{l, T}$, прямой задачи (1)–(4). Таким образом, предположение о существовании двух различных решений $f_1(s), \mu_1(\tau), u_1(x, t)$ и $f_2(s), \mu_2(\tau), u_2(x, t)$ обратной задачи (1)–(5) опровергнуто за счет их попарного совпадения. Теорема 2 доказана.

Теорема 3. Пусть заданы функции $\varphi(x), \psi(x), g(t), h(t)$, удовлетворяющие условиям (6), (29) и условиям

$$h(t) \in C^2[0, T], \quad h(0) = \varphi(l), \quad h'(0) = \psi(l). \quad (40)$$

Тогда существует решение обратной задачи (1)–(5).

Доказательство. Доказательство проведем для значений $T \in ((l/a), (3l/a)]$, что продемонстрирует схему доказательства существования решений обратной задачи (1)–(5), с возможностью перенесения аналогичного анализа на большие значения $T > (3l/a)$.

Дифференцируя уравнение (32) дважды по t и производя замену $l - at = x$, имеем

$$f(x) + \int_x^l K_1(s, x) f(s) ds = \frac{1}{g(0)} H_1'' \left(\frac{l-x}{a} \right), \quad x \in [0, l]. \quad (41)$$

Уравнение (41) в силу следующей из условий (6), (29) теоремы 3 непрерывности его ядра $K_1(x, s)$, определяемого формулой (35), и, также следующей из условий (6), (29), (40), непрерывности правой части уравнения (41), задаваемой с функцией

$$H_1'' \left(\frac{l-x}{a} \right) = h'' \left(\frac{l-x}{a} \right) - a^2 \varphi''(x) + a \psi'(x), \quad x \in [0, l],$$

имеет [18] единственное решение $f(x) \in C[0, l]$, для которого также из уравнения (41) получаем значение

$$f(l) = \frac{1}{g(0)} h''(0) - \frac{a^2}{g(0)} \varphi''(l).$$

Таким образом, функция $f(x) \in C[0, l]$, как часть решения обратной задачи (1)–(5), существует и определяется из решения уравнения (41).

Дифференцируя уравнение (37) по аргументу ω , получаем для функции $\mu(\omega)$ формулу

$$\mu(\omega) = H_2'(\omega), \quad \omega \in [0, \hat{T}], \quad (42)$$

где, исходя из формулы (38), с учетом представлений (25) и (26) для функции $F(s)$ на разных частях отрезка $[0, \hat{T}]$, правая часть равенства (42) принимает вид

$$\begin{aligned} H_2'(\omega) = & \frac{a+\beta}{2a} h' \left(\omega + \frac{l}{a} \right) + \frac{a+\beta}{2} F'(-a\omega) - \frac{a+\beta}{2a} \psi(0) + \frac{a+\beta}{2} \psi'(0)\omega - \\ & - \frac{a+\beta}{2a} f(0) \int_0^\omega g(\tau) d\tau - \frac{a+\beta}{2a^2} \int_0^l g \left(\omega + \frac{\xi}{a} \right) f(\xi) d\xi, \quad \omega \in [0, \hat{T}], \end{aligned} \quad (43)$$

при

$$\begin{aligned} \frac{a+\beta}{2} F'(-a\omega) = & \frac{a-\beta}{2} \varphi'(a\omega) + \frac{a+\beta}{2a} \psi(0) - \frac{a+\beta}{2} \psi'(0)\omega + \frac{a-\beta}{2a} \psi(-a\omega) + \\ & + \frac{a+\beta}{2a^2} f(0) \int_0^{a\omega} g \left(\omega - \frac{\eta}{a} \right) d\eta + \frac{a-\beta}{2a^2} \int_0^{a\omega} g \left(\omega - \frac{\xi}{a} \right) f(\xi) d\xi, \quad \omega \in \left[0, \frac{l}{a} \right], \end{aligned} \quad (44)$$

и при

$$\begin{aligned} \frac{a+\beta}{2} F'(-a\omega) = \frac{a+\beta}{2} \left[\frac{\psi(0)}{a} - \psi'(0)\omega - \varphi'(2l-a\omega) + \frac{a-\beta}{(a+\beta)a} \psi(2l-a\omega) + \frac{f(0)}{a^2} \int_0^{a\omega} g\left(\omega - \frac{\eta}{a}\right) d\eta \right] + \\ + \frac{a-\beta}{2a^2} \int_0^l g\left(\omega - \frac{\xi}{a}\right) f(\xi) d\xi + \frac{a-\beta}{2a^2} \int_l^{a\omega} g\left(\omega - \frac{\xi}{a}\right) f(2l-\xi) d\xi, \quad \omega \in \left[\frac{l}{a}, \widehat{T}\right]. \end{aligned} \quad (45)$$

Из условий теоремы 3 следует непрерывность функции $H_2'(\omega)$, $\omega \in [0, \widehat{T}]$, представляемой формулой (43) с использованием равенств (44), (45). Следовательно, непрерывна и функция $\mu(t)$, $t \in [0, \widehat{T}]$, определяемая формулой (42). С учетом равенств

$$\begin{aligned} h'\left(\frac{l}{a}\right) &= -a\varphi'(0) + \psi(0) + \frac{1}{a} \int_0^l g\left(\frac{\xi}{a}\right) f(\xi) d\xi, \\ h''\left(\frac{l}{a}\right) &= a^2\varphi''(0) - a\psi'(0) + g(0)f(0) + \frac{1}{a} \int_0^l g'\left(\frac{\xi}{a}\right) f(\xi) d\xi, \end{aligned}$$

и формул (43), (44) имеем

$$\begin{aligned} \mu(0) &= H_2'(\omega)\Big|_{\omega=0} = \frac{a+\beta}{2a} h'\left(\frac{l}{a}\right) + \frac{a-\beta}{2} \varphi'(0) + \frac{a+\beta}{2a} \psi(0) + \frac{a-\beta}{2a} \psi(0) - \frac{a+\beta}{2a} \psi(0) - \\ &- \frac{a+\beta}{2a^2} \int_0^l g\left(\frac{\xi}{a}\right) f(\xi) d\xi = -\frac{a+\beta}{2} \varphi'(0) + \frac{a+\beta}{2a} \psi(0) + \frac{a+\beta}{2} \varphi'(0) + \frac{a-\beta}{2a} \psi(0) = \psi(0) - \beta\varphi'(0), \\ \mu'(0) &= H_2''(\omega)\Big|_{\omega=0} = \left[\frac{a+\beta}{2a} h''\left(\omega + \frac{l}{a}\right) + \frac{(a-\beta)a}{2} \varphi''(a\omega) + \frac{a-\beta}{2} \psi'(-a\omega) + \frac{a+\beta}{2a} f(0)g(0) + \right. \\ &\left. + \frac{a-\beta}{2a} g(0)f(a\omega) - \frac{a+\beta}{2a} f(0)g(\omega) - \frac{a+\beta}{2a^2} \int_0^l g'\left(\omega + \frac{\xi}{a}\right) f(\xi) d\xi \right] \Bigg|_{\omega=0} = a^2\varphi''(0) - \beta\psi'(0) + f(0)g(0). \end{aligned}$$

Полученные равенства соответствуют выполнению на функцию $\mu(t)$ условий на значение $\mu(0) = \psi(0) - \beta\varphi'(0)$ и на значение $\mu'(0) = a^2\varphi''(0) - \beta\psi'(0) + f(0)g(0)$ из ограничений (7) и (31), входящих в условия теоремы 1 и определения 1 решения обратной задачи. Следовательно, функция $\mu(t)$ удовлетворяет условиям определения 1, как часть решения обратной задачи (1)–(5), существует и определяется формулой (42).

В итоге решение $f(x)$ уравнения (41) и функция $\mu(t)$, определяемая формулой (42), удовлетворяют условиям теоремы 1, в соответствии с которой с учетом включения $\Lambda_{l,T} \subset \overline{Q}_T$ существует единственное решение $u(x,t) \in C^2(\Lambda_{l,T})$ прямой задачи (1)–(4). В результате три полученные функции $f(x)$, $\mu(t)$, $u(x,t)$ составляют решение обратной задачи (1)–(5). Теорема 3 доказана.

5. ИТЕРАЦИОННЫЙ АЛГОРИТМ РЕШЕНИЯ ОБРАТНОЙ ЗАДАЧИ

Для приближенного решения обратной задачи (1)–(5) при известных исходных функциях $\varphi(x)$, $\psi(x)$, $g(t)$ и при заданных вместо функции $h(t)$ значения уровня погрешности $\delta > 0$ и функции $h_\delta(t) \in C[0, T]$ такой, что $\|h_\delta(t) - h(t)\|_{C[0, T]} \leq \delta$, при выполнении условий (6), (29), (30) на заданные функции, могут быть использованы регуляризирующие алгоритмы [15], [16].

С учетом того, что получение решения обратной задачи (1)–(5) эквивалентно по существу последовательному решению интегральных уравнений Вольтерра I рода (32) и (37), для приближенного решения уравнений (32) и (37), в правых частях которых на местах значений функции $h(t)$ располагаются значения приближенно известной функции $h_\delta(t)$ с известным уровнем погрешности $\delta > 0$, можно использовать алгоритмы, ориентированные на решение именно интегральных уравнений Вольтерра I рода.

Уравнение (32) имеет ядро с тождественно нулевыми значениями на диагонали области определения его аргументов, т.е. при равенстве аргументов. При этом это ядро имеет на диагонали

строго знакоопределенную первую производную. Поэтому для решения уравнения (32) могут быть использованы методы, представленные в работах [19–22]. При этом предлагаемые в этих работах алгоритмы основаны на различных подходах к получению устойчивых приближенных решений интегральных уравнений Вольтерра I рода: сеточных и функциональных, с введением регуляризирующего параметра и точным определением условий сходимости приближенных методов.

Для решения уравнения Вольтерра I рода (37), имеющего ядро с положительными значениями на диагонали, помимо представленных в работах [19–22] алгоритмов, может быть использован предложенный в работе [23] переход к решению интегрального уравнения Вольтерра II рода с малым параметром $\alpha > 0$ при искомой функции в дополнительно вводимом в уравнение внеинтегральном слагаемом:

$$\alpha \mu(\omega) + \int_0^{\omega} \mu(\theta) d\theta = H_2(\omega), \quad \omega \in [0, \hat{T}]. \quad (46)$$

Сходимость приближенного решения уравнения (46) к точному решению уравнения (37) обусловливается в работе [23] согласованием погрешности $\delta > 0$ задания исходных данных для решения обратной задачи (1)–(5) с параметром $\alpha > 0$ уравнения (46).

Разнообразие алгоритмических схем, предложенных в работах [19–23], близость их конструкций к уравнениям (32) и (37) делают использование таких специфичных методов решения уравнений Вольтерра I рода достаточно привлекательным и не требующим алгоритмических усложнений.

Как альтернатива перечисленным методам может быть использован итерационный алгоритм, основанный на приближенном решении уравнения (41) методом последовательных приближений, с получением последовательности $\{f^{(n)}(x)\}$, приближающей функцию $f(x)$, по рекуррентной формуле

$$f^{(n+1)}(x) = \frac{1}{g(0)} H_1'' \left(\frac{l-x}{a} \right) - \int_x^l K_1(s, x) f^{(n)}(s) ds, \quad f^{(1)}(x) = \frac{1}{g(0)} H_1'' \left(\frac{l-x}{a} \right), \quad x \in [0, l].$$

После получения приближенного решения уравнения (41) приближенные значения функции $\mu(t)$ могут быть получены по формуле (42).

При использовании представленного итерационного алгоритма приближенного решения интегрального уравнения (41) и формулы (42) некорректность обратной задачи (1)–(5) связана лишь с необходимостью предварительного двукратного численного дифференцирования приближенно заданной на отрезке $[0, T]$ непрерывной функции $h_8(t)$.

СПИСОК ЛИТЕРАТУРЫ

1. Тихонов А.Н., Самарский А.А. Уравнения математической физики. М.: Гос. изд-во технико-теоретической литературы, 1953.
2. Moiseev E.I., Kholomeeva A.A. Solvability of the mixed problem for the wave equation with a dynamic boundary condition // Differential Equations. 2012. V. 48. № 10. P. 1392–1397.
3. Moiseev E.I., Kholomeeva A.A., Frolov A.A. Boundary displacement control for the oscillation process with boundary conditions of damping type for a time less than critical // J. of Math. Sci. 2021. V. 257. № 1. P. 74–84.
4. Yanovskaya T.V., Asbel I.J. The determination of velocities in the upper mantle from the observations on p-waves // Geophys. J. Royal Astronom. Soc. 1964. V. 8. № 3. P. 313–318.
5. Baev A.V., Glasko V.B. The solution of the inverse kinematic problem of seismology by means of a regularizing algorithm // J. Comput. Math. and Math. Phys. 1976. V. 16. № 4. P. 96–106.
6. Belishev M.I. Boundary control in reconstruction of manifolds and metrics (the BC-method) // Inverse Problems. 1997. V. 13. № 5. P. R1–R45.
7. Yilmaz O. Seismic data analysis. 1. Tulsa: SEG, 2001.
8. Vasil'ev F.P., Kurzhan'skij M.A., Razgulin A.V. On using Fourier method for solving a problem of string vibration control // Moscow University Comput. Math. and Cybernetics. 1993. № 2. P. 3–8.
9. Il'in V.A., Moiseev E.I. Optimization of boundary controls of string vibrations // Russian Math. Surveys. 2005. V. 60. № 6. P. 1093–1119.
10. Kapustin N.Yu., Kholomeeva A.A. Spectral solution of a boundary value problem for equation of mixed type // Lobachevskii Journal of Math. 2019. V. 40. № 7. P. 981–983.

11. *Gopinath B., Sondi M.* Determination of the shape of the human vocal tract from acoustical measurements // *Bell System Tech. J.* 1970. V. 49. P. 1195–1214.
12. *Cannon J.R., Du Chateau P.* An inverse problem for an unknown source term in a wave equation // *SIAM J. Appl. Math.* 1983. V. 43. № 3. P. 553–564.
13. *Zhang Guan Quan.* On an inverse problem for 1-dimensional wave equation // *Sci. China. Ser. A. Math., Phys., Astron. and Tech. Sci.* 1989. V. 32. № 3. P. 257–274.
14. *Denisov A.M.* Existence of a solution of the inverse coefficient problem for a quasilinear hyperbolic equation // *J. Comput. Math. and Math. Phys.* 2019. V. 59. № 4. P. 550–558.
15. *Тихонов А.Н., Гончарский А.В., Степанов В.В., Ягола А.Г.* Регуляризирующие алгоритмы и априорная информация. М.: Наука, 1983.
16. *Денисов А.М.* Введение в теорию обратных задач. М.: Изд-во Моск. ун-та, 1994.
17. *Свешников А.Г., Боголюбов А.Н., Кравцов В.В.* Лекции по математической физике. М.: Изд-во Моск. ун-та, 1993.
18. *Lovitt W.V.* Linear integral equations. New York: McGraw-Hill Book Company, 1924.
19. *Sergeev V.O.* Regularization of the Volterra equation of the first kind // *Dokl. Akad. Nauk SSSR.* 1971. V. 197. № 3. P. 531–534.
20. *Апарцин А.С., Бакушинский А.Б.* Приближенное решение интегральных уравнений Вольтерра I рода методом квадратурных сумм // В сб.: “Дифференц. и интегральные уравнения”. Вып. 1. Иркутск, Гос. ун-тет. 1972. С. 120–128.
21. *Magnitskii N.A.* A method of regularizing Volterra equations of the first kind // *Comput. Math. and Math. Phys.* 1975. V. 15. № 5. P. 221–228.
22. *Магницкий Н.А.* О приближенном решении некоторых интегральных уравнений Вольтерра I рода // *Вестн. Московского университета. Серия 15. Вычисл. матем. и кибернетика.* 1978. № 1. С. 91–98.
23. *Denisov A.M.* On approximate solution of Volterra equation of the first kind // *Comput. Math. and Math. Phys.* 1975. V. 15. № 4. P. 237–239.

**МЕТОД МОДЕЛИРОВАНИЯ ТУРБУЛЕНТНОГО ПОГРАНИЧНОГО
СЛОЯ НА ОСНОВЕ АНАЛИТИЧЕСКИХ ЗАКОНОВ СТЕНКИ
В ФОРМУЛИРОВКЕ МЕТОДА ХАРАКТЕРИСТИЧЕСКИХ
ШТРАФНЫХ ФУНКЦИЙ¹⁾**

© 2023 г. О. В. Васильев^{1,*}, Н. С. Жданова^{1,**}

¹ 125047 Москва, Миусская пл., 4, Институт прикладной математики им. М.В. Келдыша РАН, Россия

*e-mail: oleg.v.vasilyev@gmail.com

**e-mail: nat.zhdanova@gmail.com

Поступила в редакцию 12.09.2022 г.

Переработанный вариант 07.10.2022 г.

Принята к публикации 02.02.2023 г.

Предложен метод расчета пристеночных областей турбулентных течений для численного моделирования вязкого сжимаемого газа с применением уравнений Навье–Стокса, осредненных по Рейнольдсу. В основе метода – дифференциальное условие сшивки внешнего решения с пристеночной функцией, позволяющее использовать обобщение метода характеристических штрафных функций для переноса касательного напряжения из внешней области пограничного слоя на поверхность тела, при этом область сшивки задается неявно через локализованный источниковый член в уравнении пограничного слоя, записанного как функция расстояния от стенки, нормированного на масштаб вязкой длины. Касательное напряжение на стенке при этом определяется в процессе численного решения специального дифференциального уравнения, включающего в себя характеристические штрафные функции и аналитический закон стенки. Разработанный метод заметно снижает требования к пристеночному разрешению расчетной сетки без существенного усложнения вычислительного алгоритма и позволяет полностью устранить плохо-определенное условие точки сшивки решений. Численная реализация разработанного подхода проведена с применением вершинно-центрированного метода контрольных объемов и структурированных расчетных сеток. Его применимость продемонстрирована на примере решения двух тестовых задач: течение в двумерном канале и турбулентное обтекание бесконечно тонкой пластины. Библ. 38. Фиг. 8. Табл. 8.

Ключевые слова: характеристические штрафные функции, турбулентное течение, закон Рейхарда, метод пристеночных функций.

DOI: 10.31857/S0044466923050198, **EDN:** GGGNJE

ВВЕДЕНИЕ

В настоящее время в связи с большой вычислительной стоимостью прямого численного моделирования аэродинамических течений с большими числами Рейнольдса (см. [1]) активно развиваются вихреразрешающие методы на основе осредненных уравнений Навье–Стокса (RANS) (см. [2–5]) и гибридные подходы, в которых течение в пристеночной области моделируется на основе моделей класса RANS, а течение в областях, удаленных от твердых поверхностей, – Методом Крупных Вихрей (LES) (см. [6], [7]). К гибридному классу методов можно также отнести методы моделирования отсоединенных вихрей (DES) (см. [8–10]) с плавным переходом от RANS к LES решениям. Несмотря на сравнительно умеренные требования RANS, гибридных RANS-LES и DES подходов к пристеночному разрешению расчетных сеток, ограничения все равно остаются существенными и накладывают жесткие требования к объему вычислительных ресурсов, увеличивают время счета задачи и усложняют построение расчетной сетки.

Ограничения на размер ячеек сетки вблизи поверхности тела могут быть заметно снижены при использовании методов на основе пристеночных функций (см. [11–13]), что достигается за-

¹⁾Работа выполнена при финансовой поддержке РФФ (проект 20-41-09018).

меной граничных условий прилипания на поверхности тела условием сшивки пристеночной функции с внешней областью турбулентного пограничного слоя. Пристеночные функции также могут применяться и в слабой формулировке, через условие проскальзывания, допускающее перенос касательных напряжений из внешней области пограничного слоя на поверхность тела (см. [14], [15]). Такая формулировка предпочтительнее с вычислительной точки зрения в силу ее гибкости, но не гарантирует точного соответствия толщины вытеснения пограничного слоя, так как в пристеночной области вычисляется приближенное решение со скоростью проскальзывания и турбулентной вязкостью, экстраполированными из внешней области решения, что приводит к уменьшению толщины вытеснения пограничного слоя по сравнению с точным решением усредненных уравнений Навье–Стокса.

В традиционных подходах (см. [16]) граничные условия определяются путем решения нелинейных уравнений в точке сшивки, при этом сама точка сшивки заранее неизвестна, так как неявно задана расстоянием от стенки, нормированным на масштаб вязкой длины, который, в свою очередь, является функцией касательного напряжения на стенке. Для этой цели решение в точке сшивки интерполируется с ближайших узлов (см. [17]). Следует отметить возможность использования метода пристеночных функций совместно с дискретным методом погруженных границ, в котором в точке сшивки также используются дополнительные ограничительные связи, обеспечивающие выполнение граничных условий на поверхности тела (см. [18–20]).

Основная идея разработанного метода заключается в замене алгебраического условия сшивки внешнего решения с пристеночной функцией на дифференциальную формулировку, позволяющую использовать обобщение метода характеристических штрафных функций (см. [21], [22]) для переноса касательного напряжения из внешней области пограничного слоя на поверхность тела, при этом область сшивки задается неявно через локализованный источниковый член в уравнении пограничного слоя, записанного как функция расстояния от стенки, нормированного на масштаб вязкой длины. Такой подход, во-первых, позволяет полностью устранить плохо-определенное условие точки сшивки решений и, во-вторых, свести систему дифференциальных уравнений с нелинейными алгебраическими связями к системе уравнений с дифференциальными обратными связями, основанными на методе характеристических штрафных функций, обеспечивающих эту связь. Последнее обстоятельство определяет возможное развитие метода, связанное с применением дифференциальных стандартных (см. [23], [17]) и неравновесных (см. [24], [25]) динамических пристеночных функций (для задач с сильными отрывами). В целом новый метод позволяет заметно снизить требования к пристеночному разрешению расчетной сетки без существенного усложнения вычислительного алгоритма. Следует отметить отличие разработанного метода от гибридных подходов с использованием методов погруженных границ для задания граничных условий на поверхности обтекаемых тел (см. [18], [19]), так как в предложенном подходе используются характеристические штрафные функции для сшивки пристеночных функций с внешним решением. Разработанный метод может быть также обобщен для течений со сложной геометрией на основе уже разработанных методов штрафных функций (см. [21], [22], [26], [27]).

Статья организована следующим образом. В разд. 1 описан разработанный метод моделирования турбулентного пограничного слоя на основе аналитических законов стенки в формулировке метода характеристических штрафных функций. В разд. 2 кратко описана методика расчета, используемая для численной реализации разработанного подхода. Постановка и результаты численного моделирования тестовых задач приведены в разд. 3. В заключении сделаны выводы и обозначены направления дальнейшего развития темы.

1. ПОСТАНОВКА ЗАДАЧИ И МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

1.1. Осредненные уравнения движения вязкого сжимаемого газа

В качестве базовой математической модели для описания течений вязкой сжимаемой среды использована система осредненных по Фавру уравнений Навье–Стокса, принимающая следующий вид:

$$\frac{\partial p}{\partial t} + \frac{\partial(\rho u_j)}{\partial x_j} = 0, \quad (1)$$

$$\frac{\partial \rho u_i}{\partial t} + \frac{\partial}{\partial x_j} (\rho u_i u_j) = -\frac{\partial p}{\partial x_i} + \frac{\partial \hat{\tau}_{ij}}{\partial x_j}, \quad (2)$$

$$\frac{\partial \rho e}{\partial t} + \frac{\partial}{\partial x_j} [(\rho e + p) u_j] = \frac{\partial}{\partial x_j} [u_i \hat{\tau}_{ij} - q_j], \quad (3)$$

где

$$p = \rho RT, \quad (4)$$

$$e = \frac{1}{2} u_i u_i + \frac{p}{\rho(\gamma - 1)}, \quad (5)$$

$$q_j = -c_p \left(\frac{\mu}{Pr_L} + \frac{\mu_T}{Pr_T} \right) \frac{\partial T}{\partial x_j}, \quad (6)$$

$$\hat{\tau}_{ij} = 2\mu \tilde{S}_{ij} + \tau_{ij},$$

$$\tau_{ij} = 2\mu_T \tilde{S}_{ij}, \quad (7)$$

$$\tilde{S}_{ij} = \text{dev}(S_{ij}) = S_{ij} - \frac{1}{3} \frac{\partial u_k}{\partial x_k} \delta_{ij},$$

$$S_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right).$$

В приведенных выше уравнениях используются следующие обозначения: ρ – осредненная по Рейнольдсу плотность, p – осредненное по Рейнольдсу давление, ρu_j – массовый поток, u_j – осредненная по Фавру скорость, T – осредненная по Фавру температура, e – осредненная по Фавру полная энергия на единицу массы. Параметр R обозначает газовую постоянную, переменные c_v и c_p – соответственно удельные теплоемкости при постоянном объеме и давлении, значение коэффициента теплоемкости $\gamma = c_p/c_v$ без потери общности принято равным $\gamma = 1.4$, что соответствует двухатомному газу. Переменная q_j представляет собой сумму векторов ламинарного и моделируемого турбулентных тепловых потоков, где $Pr_L = 0.72$ и $Pr_T = 0.9$ – соответственно ламинарное и турбулентное числа Прандтля. Турбулентная вязкость, обозначенная как μ_T , неизвестна и для замыкания подразумевает использование модели турбулентности. Переменная $\hat{\tau}_{ij}$ представляет собой сумму тензоров вязкого (молекулярного) и турбулентного напряжений, τ_{ij} – тензор турбулентных напряжений, S_{ij} – тензор осредненных скоростей деформации, \tilde{S}_{ij} – дивергентная часть тензора S_{ij} . Для простоты в работе рассмотрены течения с малыми числами Маха и постоянной динамической молекулярной вязкостью μ .

1.2. Модель турбулентности

Для иллюстрации разработанного подхода в качестве модели турбулентности для замыкания уравнений (6) и (7) использована модель Спаларта–Аллмараса (см. [2]), широко применяемая для моделирования течений газа (см. [28]), включая высокоскоростные течения с существенным влиянием сжимаемости. Стандартная модель Спаларта–Аллмараса для переменной $\rho \tilde{\nu}$ может быть записана в следующем виде:

$$\begin{aligned} \frac{\partial \rho \tilde{\nu}}{\partial t} + \frac{\partial}{\partial x_j} (\rho \tilde{\nu} u_j) = c_{b1} (1 - f_{t2}) \tilde{S} \rho \tilde{\nu} - \left[c_{w1} f_w - \frac{c_{b1}}{\kappa^2} f_{t2} \right] \rho \left(\frac{\tilde{\nu}}{\delta} \right)^2 + \\ + \frac{\partial}{\partial x_j} \left[\left(\frac{\mu}{\sigma} + \frac{\rho \tilde{\nu}}{\sigma} \right) \frac{\partial \tilde{\nu}}{\partial x_j} \right] - \left(\frac{\mu}{\sigma \rho} + \frac{\tilde{\nu}}{\sigma} \right) \frac{\partial \rho}{\partial x_j} \frac{\partial \tilde{\nu}}{\partial x_j} + c_{b2} \frac{\rho}{\sigma} \frac{\partial \tilde{\nu}}{\partial x_j} \frac{\partial \tilde{\nu}}{\partial x_j}, \end{aligned} \quad (8)$$

где турбулентная вязкость вычисляется согласно

$$\mu_T = \rho \tilde{\nu} f_{v1}, \quad (9)$$

а вспомогательные переменные определяются соотношениями

$$f_{v1} = \frac{\chi^3}{\chi^3 + c_{v1}^3},$$

$$\chi = \tilde{v}/\nu,$$

$$\tilde{S} = \max \left[0.3\sqrt{2\Omega_{ij}\Omega_{ij}}, \sqrt{2\Omega_{ij}\Omega_{ij}} + \frac{\tilde{v}}{\kappa^2\delta^2} f_{v2} \right], \tag{10}$$

$$\Omega_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i} \right),$$

$$f_{v2} = 1 - \frac{\chi}{1 + \chi f_{v1}},$$

$$f_w = g \left[\frac{1 + c_{w3}^6}{g^6 + c_{w3}^6} \right]^{1/6},$$

$$g = r + c_{w2} (r^6 - r),$$

$$r = \min \left[\frac{\tilde{v}}{\tilde{S}\kappa^2\delta^2}, 10 \right], \tag{11}$$

$$f_{i2} = c_{i3} \exp(-c_{i4}\chi^2),$$

ν – кинематическая молекулярная вязкость, $\delta(\mathbf{x})$ – расстояние от точки поля \mathbf{x} до ближайшей поверхности, а для улучшения устойчивости вычислений переменная \tilde{S} ограничена снизу величиной $0.3\sqrt{2\Omega_{ij}\Omega_{ij}}$. Постоянные коэффициенты задаются как $c_{b1} = 0.1355$, $c_{b2} = 0.622$, $\sigma = 2/3$, $\kappa = 0.41$, $c_{w2} = 0.3$, $c_{w3} = 2$, $c_{v1} = 7.1$, $c_{w1} = \frac{c_{b1}}{\kappa} + \frac{1 + c_{b2}}{\sigma}$. В расчетах используется “стандартная” версия модели Спаларта–Аллмараса без дополнительного генерационного слагаемого f_{i1} , что подразумевает отсутствие необходимости использовать переменную f_{i2} , т.е. $c_{i3} = 0$.

В качестве граничного условия на поверхности стенки используется

$$\tilde{v} = 0. \tag{12}$$

При моделировании внешних течений в набегающем потоке используется константная турбулентная вязкость $\tilde{v} = 3\nu_\infty$.

1.3. Метод пристеночного моделирования

При моделировании турбулентных течений с большим числом Рейнольдса в рамках вихререшающих методов (см. [2–5]), несмотря на использование моделей турбулентности, прямое разрешение структур решения уравнений (1)–(3), (8) в пристеночной области требует значительных вычислительных затрат, связанных с существенным увеличением количества узлов сетки вблизи поверхности. Ограничения на размер расчетных ячеек вблизи поверхности тела могут быть заметно снижены при использовании методов на основе пристеночных функций (см. [11–13]), что достигается заменой граничных условий прилипания на поверхности тела условием сшивки пристеночной функции $f(\tilde{y}^+)$ с внешней областью турбулентного пограничного слоя:

$$u_{||}(\mathbf{x}) \Big|_{\delta(\mathbf{x}) = \frac{\nu}{u_\tau} \delta_{EL}^+} = u_\tau f(\delta_{EL}^+), \tag{13}$$

где условие точки сшивки определено как

$$\delta(\mathbf{x}) = \frac{\nu}{u_\tau} \delta_{EL}^+, \tag{14}$$

$u_{\parallel}(\mathbf{x})$ – компонента скорости, параллельная поверхности в точке \mathbf{x} , \tilde{x} и \tilde{y} – обобщенные координаты вдоль и по нормали к поверхности, $\tilde{y}^+ = u_{\tau}\tilde{y}/v$ – нормированная координата, $u_{\tau} = (\tau_w/\rho)^{1/2}$ – скорость трения, τ_w – напряжение трения на стенке, δ_{EL}^+ – нормированное расстояние сшивки решения, а $f(\tilde{y}^+)$ определена как функция нормированного расстояния от поверхности. Для простоты изложения рассмотрим формулировку метода для двухмерных течений, а обобщение метода для трехмерных течений будет приведено в конце раздела.

Для заданного внешнего поля скорости $u_{\parallel}(\mathbf{x})$ условие сшивки (13) является нелинейным уравнением для определения u_{τ} и точки сшивки для заданной точки \tilde{x} на поверхности тела. Линейное приближение уравнения (13) для коррекции решения δu_{τ} может быть записано как

$$u_{\parallel}(\mathbf{x})\Big|_{\delta(\mathbf{x})=\frac{v}{u_{\tau}}\delta_{EL}^+} + \frac{\partial u_{\parallel}(\mathbf{x})}{\partial \mathbf{n}}\Big|_{\delta(\mathbf{x})=\frac{v}{u_{\tau}}\delta_{EL}^+} \left(-\frac{v\delta_{EL}^+}{u_{\tau}^2} \delta u_{\tau} \right) \approx u_{\tau} f(\delta_{EL}^+) + \delta u_{\tau} f(\delta_{EL}^+), \quad (15)$$

где нормаль \mathbf{n} определена через функцию расстояния $\delta(\mathbf{x})$: $\mathbf{n} = \nabla \delta(\mathbf{x})$. Отметим, что второй член левой части уравнения (15) возникает из-за изменения расстояния $\delta(\mathbf{x})$ точки сшивки решения при изменении u_{τ} для фиксированного δ_{EL}^+ . Решение уравнения (15) для коррекции скорости трения δu_{τ} принимает следующий вид:

$$\delta u_{\tau} \approx u_{\tau} \frac{u_{\parallel}(\mathbf{x})\Big|_{\delta(\mathbf{x})=\frac{v}{u_{\tau}}\delta_{EL}^+} - u_{\tau} f(\delta_{EL}^+)}{u_{\tau} f(\delta_{EL}^+) + \frac{\partial u_{\parallel}(\mathbf{x})}{\partial \mathbf{n}}\Big|_{\delta(\mathbf{x})=\frac{v}{u_{\tau}}\delta_{EL}^+} \delta(\mathbf{x})}. \quad (16)$$

В случае сильной формулировки условия сшивки уравнение (16) может быть использовано для итерационной коррекции решения уравнения (13) методом Ньютона.

В слабой постановке задачи дискретная коррекция (16) может быть заменена на временную релаксацию решения $u_{\tau}(\tilde{x}, t)$ на масштабе времени η_f в каждой точке обобщенной координаты \tilde{x} , определенной вдоль поверхности тела:

$$\frac{\partial u_{\tau}}{\partial t} = \frac{u_{\tau}}{\eta_f} \frac{u_{\parallel}(\mathbf{x})\Big|_{\delta(\tilde{x})=\frac{v}{u_{\tau}}\delta_{EL}^+} - u_{\tau} f(\delta_{EL}^+)}{u_{\tau} f(\delta_{EL}^+) + \frac{\partial u_{\parallel}(\mathbf{x})}{\partial \mathbf{n}}\Big|_{\delta(\tilde{x})=\frac{v}{u_{\tau}}\delta_{EL}^+} \delta(\tilde{x})}, \quad (17)$$

где толщина пограничного слоя, неявно определенная уравнением (14), для большей наглядности записана в явном виде как $\delta(\tilde{x})$, при этом граничное условие прилипания для скорости \mathbf{u} на границе тела заменено на условие непротекания для нормальной компоненты скорости $u_{\perp}(\mathbf{x}) = \mathbf{u} \cdot \mathbf{n}$:

$$u_{\perp}\Big|_{\tilde{y}=0} = 0 \quad (18)$$

и условие переноса касательных напряжений на поверхность тела:

$$(v + v_T) \frac{\partial u_{\parallel}}{\partial \mathbf{n}}\Big|_{\tilde{y}=0} = u_{\tau}^2(\tilde{x}, t). \quad (19)$$

Для переноса турбулентной вязкости с внешней области течения на границу в модели Спаларта–Аллмараса в уравнениях (8), (10) и (11) функция расстояния $\delta(\mathbf{x})$ до ближайшей поверхности заменена на

$$\delta(\mathbf{x}) = \max\left(\delta(\mathbf{x}), \frac{v}{u_{\tau}} \delta_{EL}^+\right), \quad (20)$$

а граничное условие (12) для турбулентной вязкости заменено на

$$\left. \frac{\partial \tilde{\nu}}{\partial \mathbf{n}} \right|_{\tilde{y}=0} = 0. \tag{21}$$

Уравнение (16) с неявным определением расстояния точки сшивки $\delta(\tilde{x})$ усложняет решение задачи, так как определение величины тангенциальной компоненты скорости u_{\parallel} требует использования локального интерполянта с ближайших узлов (см., например, [17]). Задачу можно сильно упростить и полностью исключить применение локального интерполянта и точного определения точки сшивки введением дополнительного поля скорости трения $u_{\tau}(\mathbf{x}, t)$ и заменой уравнения (17) для $u_{\tau}(\tilde{x}, t)$ на следующее уравнение в частных производных для поля $u_{\tau}(\mathbf{x}, t)$ во всей области решения:

$$\frac{\partial u_{\tau}}{\partial t} - \underbrace{\mathcal{H}\left(\delta_{\text{EL}}^+ - \frac{u_{\tau}}{\nu} \delta(\mathbf{x})\right) \frac{L}{\eta_s} \frac{\partial u_{\tau}}{\partial \mathbf{n}}}_{\text{перенос } u_{\tau} \text{ на поверхность}} = \chi_{\delta} \underbrace{\left(\frac{\frac{u_{\tau}}{\nu} \delta(\mathbf{x}) - \delta_{\text{EL}}^+}{\sigma^+} \right)}_{\text{релаксация } u_{\tau} \text{ в локализованной области сшивки}} \frac{u_{\parallel}(\mathbf{x}) - u_{\tau} f(\delta_{\text{EL}}^+)}{\eta_f u_{\tau} f(\delta_{\text{EL}}^+) + \frac{\partial u_{\parallel}(\mathbf{x})}{\partial \mathbf{n}} \delta(\mathbf{x})} + \underbrace{\nu_n \Delta u_{\tau}}_{\text{сглаживание}}, \tag{22}$$

где L – характерный масштаб длины, η_s – характерный масштаб переноса решения с области сшивки на поверхность тела, $\mathcal{H}(\xi)$ – функция Хевисайда, отключающая перенос u_{τ} во внешней области решения, $\chi_{\delta}(\xi)$ – локализованная функция сшивки решений, например, функция Гаусса

$$\chi_{\delta}(\xi) = \exp(\xi^2/2), \tag{23}$$

σ^+ – нормированная толщина сшивки решений, $\nu_n \Delta u_{\tau}$ – численная диффузия, используемая для сглаживания вспомогательного поля u_{τ} . Отметим, что в уравнении (22) условие сшивки (13) обеспечивается условием релаксации скорости трения в локализованной области сшивки, из которой u_{τ} , в свою очередь, переносится на поверхность тела для последующего использования в граничном условии (19). Второй член в левой части уравнения (22) соответствует характеристической штрафной функции (см. [21], [22], [29]), обеспечивающей на масштабе времени η_s направленный перенос скорости трения на поверхность тела.

Уравнение (22) является основой разработанного метода, в дальнейшем, для краткости, именуемого *методом пенализированных пристеночных функций*. Обратим внимание, что несмотря на то что разработанный подход продемонстрирован только в контексте модели Спаларта–Аллмараса, метод пенализированных пристеночных функций может быть использован совместно с любой моделью турбулентности, позволяющей перенос касательного напряжения из внешней области пограничного слоя на поверхность тела.

В случае трехмерного течения задача может быть рассмотрена в локальных двухмерных координатах, где координата \tilde{x} вдоль поверхности тела соответствует направлению тангенциальной компоненты скорости в точке сшивки. В этом случае уравнение (22) можно использовать без изменений, а переменная $u_{\parallel}(\mathbf{x}) = \|\mathbf{u}_{\parallel}(\mathbf{x})\|$ соответствует величине параллельной компоненты скорости $\mathbf{u}_{\parallel}(\mathbf{x}) = \mathbf{u} - (\mathbf{u} \cdot \mathbf{n})\mathbf{n}$. Граничные условия (18) и (19) могут быть переписаны в следующем виде:

$$\mathbf{u} \cdot \mathbf{n} \Big|_{\tilde{y}=0} = 0, \tag{24}$$

$$(\nu + \nu_T) \frac{\partial \mathbf{u}_{\parallel}}{\partial \mathbf{n}} \Big|_{\tilde{y}=0} = \left(\frac{\mathbf{u}_{\parallel}}{u_{\parallel}} \Big|_{\tilde{y}=0} \right) u_{\tau}^2(\tilde{x}, t). \tag{25}$$

2. МЕТОД ЧИСЛЕННОГО РАСЧЕТА

Предложенный в работе метод реализован на базе программного комплекса NOISEtte, описанного в [30], детали вычислительной методики и некоторые особенности распараллеливания расчетов представлены в [31].

Пространственная дискретизация конвективной части системы уравнений (1)–(3), (8) основана на конечно-объемном подходе с определением искоемых переменных в узлах сетки, вокруг

которых построены расчетные ячейки (дуальная сетка). Для повышения порядка точности используется схема на основе квазиодномерной реконструкции переменных вдоль ребра сетки (EBR-схемы) (см. [32]). Для пространственной аппроксимации вязких членов применяется конечно-элементный метод Галеркина на основе линейных базисных функций.

Интегрирование по времени проводилось по неявной трехслойной схеме второго порядка аппроксимации с последующей линеаризацией по Ньютону разностной по пространству системы уравнений. На каждой ньютоновской итерации для решения системы линейных уравнений применяется стабилизированный метод бисопряженных градиентов (BiCGStab) (см. [33]).

Каждому шагу интегрирования по времени основной системы уравнений предшествует численное решение уравнения (22). При этом используются неявная схема первого порядка для аппроксимации производной по времени и метод направленных разностей – для производной по пространству. Последний обладает численной диффузией пропорциональной шагу сетки, поэтому члены с искусственной вязкостью $\nu_n \Delta u_\tau$ в уравнении (22) опускаются. Функция сшивки решения задается в виде функции Гаусса (23). Значения нормированной толщины сшивки решений σ^+ , масштабы релаксации η_f и η_s и начальное поле функции u_τ определяются постановкой задачи и параметрами расчетной сетки. Для представленных в работе задач они указаны в соответствующих разделах с описанием численных результатов.

3. РЕЗУЛЬТАТЫ РЕШЕНИЯ ТЕСТОВЫХ ЗАДАЧ

Представленный метод пристеночного моделирования применен для численного решения двух тестовых задач – задачи о течении в канале и задачи об обтекании пластины.

Задачи решались в безразмерных переменных, при этом использовались следующие характерные масштабы: ρ_∞ – плотность невозмущенного потока, U_∞ – скорость набегающего потока, μ_∞ – молекулярная вязкость невозмущенного потока, L – характерный линейный масштаб задачи. Связь размерных и безразмерных переменных, обозначенных знаком $\hat{(\cdot)}$, имеет вид

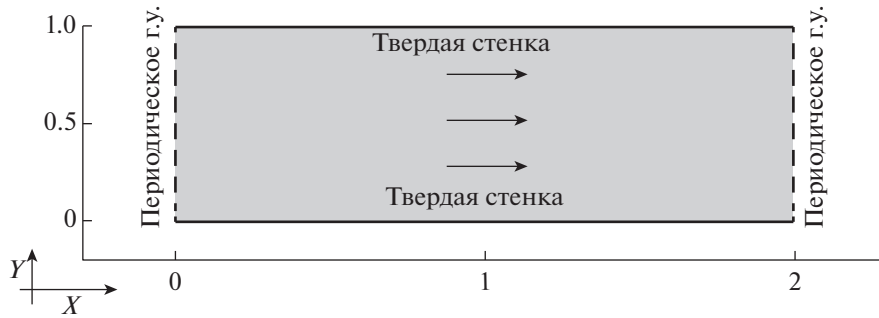
$$\rho = \rho_\infty \hat{\rho}, \quad u = U_\infty \hat{u}, \quad x = L \hat{x}, \quad \mu = \mu_\infty \hat{\mu}, \quad t = \frac{L}{U_\infty} \hat{t}, \quad p = \rho_\infty U_\infty^2 \hat{p}, \quad T = \frac{U_\infty^2}{R} \hat{T}. \quad (26)$$

Для удобства приведенные далее результаты вычислительных экспериментов представлены в безразмерном виде, при этом знак $\hat{(\cdot)}$ опускается. Отметим также, что задачи рассматривались при малых числах Маха набегающего потока, из-за чего влияние плотности и температуры на решение незначительно.

Все расчеты проводились до установления стационарного режима течения, при котором для значений относительной невязки продольной скорости и турбулентной вязкости справедливы следующие оценки: $\frac{\|\Delta u\|_2}{\|u\|_2} \leq 10^{-6}$ и $\frac{\|\Delta \nu_T\|_2}{\|\nu_T\|_2} \leq 10^{-8}$.

3.1. Турбулентное течение в плоском канале

Рассматривается течение между двумя плоскими пластинами, находящимися на расстоянии $H = 2h$ друг от друга, величина H выбрана характерным линейным масштабом. Задача решается в двумерной постановке, пластины расположены параллельно друг другу вдоль оси OX , а поток движется слева направо параллельно этой оси со скоростью, соответствующей числу Маха $M = 0.1$ (фиг. 1). Течение является квазиодномерным, поэтому на правой и левой границах расчетной области заданы периодические граничные условия для всех переменных. Верхняя и нижняя границы соответствуют адиабатическим твердым стенкам, на верхней границе задано условие прилипания $\mathbf{u} = (0, 0)$, на нижней – либо такое же условие для модели Спаларта–Аллмараса, либо условие непротекания для нормальной компоненты скорости (18) и условие переноса касательных напряжений на поверхность тела (19) для метода пенализированных пристеночных функций. Значение u_τ в (19) определялось из численного решения уравнения (22). Поле динамической скорости во всей расчетной области в начальный момент времени задавалось посто-



Фиг. 1. Постановка задачи о течении в плоском канале.

янным $-u_\tau = 7 \times 10^{-2}$, параметры релаксации $-\eta_f = \eta_s = 10^{-2}$, квадрат толщины шивки решений $-(\sigma^+)^2 = 80$. В качестве пристеночной функции использовался закон Рейхарда (см. [34]):

$$f_{\text{Rei}}(y^+) = \frac{1}{\kappa} \ln(1 + \kappa y^+) + 7.8 \left[1 - \exp\left(-\frac{y^+}{11}\right) - \frac{y^+}{11} \exp\left(-\frac{y^+}{3}\right) \right], \quad (27)$$

где $\kappa = 0.41$ – постоянная фон Кармана.

Течение характеризуется среднерасходной скоростью $u_b = \frac{1}{H} \int_0^H u(y) dy$, которая присутствует в математической модели в форме градиента давления, добавляемого в уравнение сохранения импульса как источниковый член. Значения u_b , кинематической вязкости ν и полуширины канала h полностью определяют физические свойства течения. При этом число Рейнольдса может быть посчитано по значениям среднерасходной скорости и скорости трения:

$$Re_b = \frac{u_b H}{\nu}, \quad Re_\tau = \frac{u_\tau h}{\nu}$$

соответственно. Для течений в канале имеет место следующее эмпирическое соотношение между Re_b и Re_τ (см. [35]), которое в выбранных единицах измерения принимает вид

$$Re_b \approx 14.64 Re_\tau^{7/8}.$$

Все расчеты проведены для значения $Re_\tau = 400$, что соответствует $Re_b \approx 14000$. Величина градиента давления, обеспечивающего течение, определяется заданным трением на стенке, т.е. $\int_V |\nabla p| dV = \int_S \tau_w dS$. Это соотношение задает величину источникового члена для уравнения импульса, в безразмерных переменных, соответствующих $\frac{dp}{dx} = -2u_\tau^2 = -6.53 \times 10^{-3}$ (при $Re_\tau = 400$).

Начальные условия соответствуют невозмущенному потоку со следующими безразмерными параметрами: $(u, v) = (0, 0)$, $\nu = Re_b^{-1}$, $\rho = 1.0$, $\nu_T = Re_b^{-1}$.

Во всех расчетах, представленных далее, использованы структурированные декартовы расчетные сетки, в которых N – число сеточных узлов, Δx и Δy – поперечный и продольный размеры ячеек соответственно, $q = \Delta y_{j+1} / \Delta y_j$ – коэффициент разбега поперечного размера ячейки в направлении по нормали от стенки, Δy_{bw}^+ – нормированный размер первой пристеночной ячейки на нижней границе канала, Δy_{tw}^+ – нормированный размер первой пристеночной ячейки на верхней границе (если его значение не указано, то $\Delta y_{bw}^+ = \Delta y_{tw}^+$).

Для оценки точности численного моделирования турбулентного пограничного слоя с применением модели Спаларта–Аллмараса и определения оптимального сеточного разрешения проведена серия расчетов на сетках, отличающихся размером пристеночной ячейки и коэффициентом разбега. В табл. 1 представлены параметры использованных сеток. Продольный размер расчетной ячейки не изменяется и равен $\Delta x = 0.05$. Значение Δy_{bw}^+ здесь посчитано относительно теоретического значения скорости трения u_τ , соответствующего заданному числу Рейнольдса Re_τ .

Таблица 1. Параметры сеток для расчетов с применением модели Спаларта–Аллмараса

Сетка	Δy	q	N	Δy_{bw}^+
1	6.25×10^{-4}	1.0	14409	0.5
2	1.25×10^{-3}	1.0	7209	1.0
3	5.00×10^{-3}	1.0	1809	4.0
4	1.25×10^{-2}	1.0	729	10.0
5	1.25×10^{-3}	1.1	765	1.0
6	1.25×10^{-3}	1.2	585	1.0
7	1.25×10^{-3}	1.25	549	1.0
8	1.25×10^{-3}	1.3	513	1.0

Таблица 2. Результаты расчетов с применением модели Спаларта–Аллмараса

Сетка	u_τ	u_c	$ v _{\max}$
1	5.718×10^{-2}	1.1454	2.68×10^{-9}
2	5.719×10^{-2}	1.1448	3.74×10^{-9}
3	5.846×10^{-2}	1.1267	7.69×10^{-7}
4	5.937×10^{-2}	1.0537	2.48×10^{-5}
5	5.720×10^{-2}	1.1438	3.43×10^{-8}
6	5.720×10^{-2}	1.1433	1.95×10^{-7}
7	5.720×10^{-2}	1.1428	4.52×10^{-7}
8	5.721×10^{-2}	1.1417	7.17×10^{-7}

Полученные численные результаты для всех сеток суммированы в табл. 2, где u_τ – скорость трения, u_c – максимум продольной скорости, $|v|_{\max}$ – максимум абсолютного значения поперечной скорости. По значению Re_τ можно получить теоретическую оценку скорости трения на стенке, которая при $Re_\tau = 400$ равна $u_\tau = 5.714 \times 10^{-2}$. Из табл. 2 следует, что наиболее близкими к этой оценке являются результаты, полученные на сетках 1 и 2. Это также подтверждается сопоставлением нормированных профилей скорости и закона Рейхарда (27) (см. фиг. 2а). Результаты, полученные на сетках 3 и 4, показывают неудовлетворительную точность расчетов, что связано с недостаточным сеточным разрешением в пристеночной области.

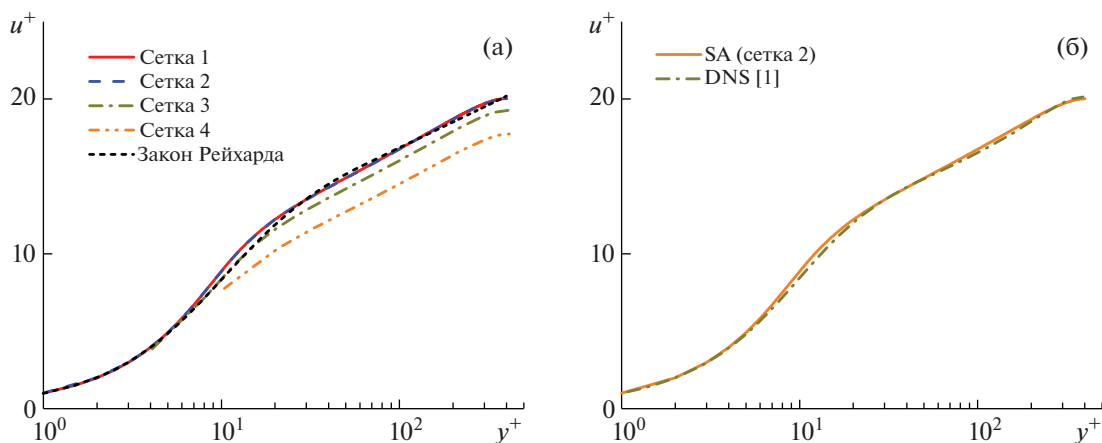
**Фиг. 2.** Результаты расчетов с применением модели Спаларта–Аллмараса (SA): (а) – сравнение с законом Рейхарда, (б) – сравнение с результатами прямого численного моделирования (DNS) (см. [36]).

Таблица 3. Параметры сеток для расчетов с применением метода пенализированных пристеночных функций

Сетка	$y < 0.5$		$y \geq 0.5$		q
	Δy_{bw}	Δy_{bw}^+	Δy_{tw}	Δy_{tw}^+	
1	1.94×10^{-2}	15.54	} 1.25×10^{-3}	1.0	} 1.2
2	1.12×10^{-2}	8.95			
3	4.49×10^{-3}	3.59			
4	2.47×10^{-3}	1.97			
5	1.25×10^{-3}	0.99	1.25×10^{-3}	1.0	} 1.0
6	0.63×10^{-3}	0.50	0.63×10^{-3}	0.5	
7	0.31×10^{-3}	0.25	0.31×10^{-3}	0.25	

Таблица 4. Зависимость результатов расчетов от сеточного разрешения

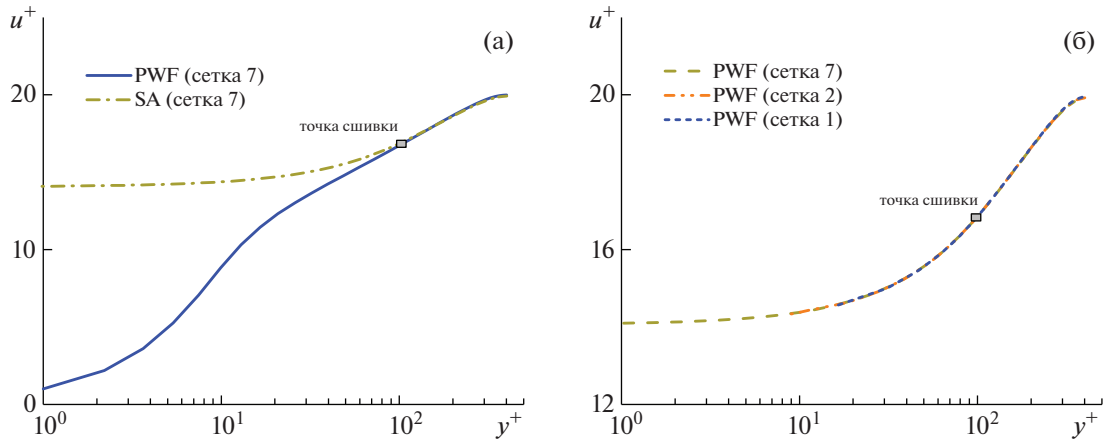
Сетка	u_τ	$\ E\ _2$
1	5.713×10^{-2}	6.75×10^{-3}
2	5.727×10^{-2}	5.06×10^{-3}
3	5.737×10^{-2}	3.98×10^{-3}
4	5.743×10^{-2}	2.95×10^{-3}
5	5.751×10^{-2}	9.70×10^{-4}
6	5.754×10^{-2}	3.48×10^{-4}
7	5.755×10^{-2}	3.45×10^{-3}

Решение, полученное для сетки 2, хорошо согласуется с результатами численного моделирования методом DNS (см. [36]). Это подтверждает фиг. 2б, на которой сопоставлены профили продольной скорости.

Результаты расчетов для сеток 5–8, приведенные в табл. 2, показывают, что приемлемое численное решение можно также получить и на расчетных сетках меньшего размера, в которых размер пристеночной ячейки увеличивается в направлении от границы с коэффициентом q . Результаты в виде профилей скорости не приводятся, так как их отличие от результатов, полученных на сетке 2, пренебрежимо мало.

Далее будут рассмотрены результаты численного решения данной задачи с применением метода пенализированных пристеночных функций. Предполагается, что точность моделирования пограничного слоя вдоль верхней границы канала обеспечивается сеточным разрешением, а вдоль нижней границы – пристеночными функциями. Для проведения расчетов построены сетки с параметрами, приведенными в табл. 3. Сетки 5–7 – равномерные, с одинаковым пристеночным разрешением на верхней и нижней границах $\Delta y_{bw}^+ = \Delta y_{tw}^+ \leq 1$, сетки 1–4 имеют достаточное разрешение пограничного слоя только на верхней границе ($\Delta y_{tw}^+ \leq 1$), на нижней – $\Delta y_{bw}^+ > 1$. В сетках 1–4 поперечный размер ячеек увеличивается в направлении нормали $(0, -1)$ от верхней границы канала с коэффициентом разбега $q = 1.2$ до достижения размера Δy_{bw} .

Рассмотрим численное решение, полученное на самой мелкой сетке – сетке 7. На фиг. 3а приведен нормированный профиль скорости, полученный при расчетах на этой сетке, сопоставленный с результатами расчетов на той же сетке моделью Спаларта–Аллмараса. В области $y^+ < \delta_{EL}^+ = 100$ можно видеть существенное отличие двух решений, что обусловлено принципиально разными граничными условиями на нижней границе канала. При $y^+ \geq \delta_{EL}^+$ они хорошо согласуются друг с другом, при этом норма отклонения решения с помощью метода пенализированных пристеночных функций от решения Спаларта–Аллмараса, посчитанная по полю продольной скорости, равна $\|E\|_2 = 3.45 \times 10^{-3}$ (см. табл. 4).



Фиг. 3. Результаты расчетов с применением метода пенализированных пристеночных функций (PWF): (а) — сравнение с моделью Спаларта–Аллмараса (SA) на разрешенной сетке, (б) — сравнение для сеток различного пристеночного разрешения.

Для расчетов на сетках 1–6 в табл. 4 приведено значение скорости трения u_τ на нижней стенке и норма отклонения от решения, полученного на сетке 7. Можно видеть, что в отсутствие сеточного разрешения пограничного слоя (сетки 1–3) ошибка численного решения остается приемлемой. Это подтверждается данными на фиг. 3б, где профили продольной скорости, полученные при расчетах на сетках 1 и 2, сопоставлены с результатами, полученными на сетке 7 ($\Delta y_{bw}^+ \sim 0.25$). Справа от точки сшивки решений ($y^+ \geq \delta_{EL}^+$) наблюдается хорошее согласование результатов.

Результаты исследования влияния параметров метода пенализированных пристеночных функций на точность численного решения суммированы в табл. 5–7, где приведены значения скорости трения на нижней стенке u_τ и отклонения поля продольной скорости от решения для сетки 7. Все расчеты с различными значениями параметров релаксации, толщины сшивки решений и положения точки сшивки проведены для расчетной сетки 2.

Таким образом, из данных табл. 5 следует слабая чувствительность решения к значениям параметров η_f и η_s , что обусловливается нечувствительностью установившегося стационарного решения к фазовым ошибкам метода пенализированных пристеночных функций. При решении нестационарных задач характеристические времена η_f и η_s должны быть существенно меньше характерного времени течения, чтобы минимизировать влияние фазовых ошибок, вызванных временем задерживания переноса касательного напряжения в уравнении (22) и времени релаксации условия сшивки (14).

В табл. 6 приведены результаты, полученные для различной толщины сшивки решений. Здесь N_f — количество точек расчетной сетки, находящихся внутри области сшивки. Можно видеть, что для обеспечения приемлемой точности численного решения достаточно нескольких расчетных точек, располагающихся внутри области сшивки, что демонстрирует существенное смягчение требований к пристеночному разрешению по сравнению с требованиями, накладываемыми моделями турбулентности, в частности, моделью Спаларта–Аллмараса.

Таблица 5. Зависимость результатов расчетов от масштаба релаксации (сетка 2, $\Delta y_{bw}^+ = 8.95$)

η_f/η_s	u_τ	$\ E\ _2$
$10^{-1}/10^{-1}$	5.692×10^{-2}	1.12×10^{-3}
$10^{-2}/10^{-2}$	5.727×10^{-2}	5.06×10^{-3}
$10^{-3}/10^{-3}$	5.727×10^{-2}	4.95×10^{-3}
$10^{-4}/10^{-4}$	5.727×10^{-2}	4.94×10^{-3}

Таблица 6. Зависимость результатов расчетов от толщины сшивки решений (сетка 2, $\Delta y_{bw}^+ = 8.95$)

$(\sigma^+)^2$	N_f	u_τ	$\ E\ _2$
80	4	5.727×10^{-2}	5.06×10^{-3}
400	11	5.727×10^{-2}	5.26×10^{-3}
800	16	5.731×10^{-2}	6.12×10^{-3}

Таблица 7. Зависимость результатов расчетов от положения точки сшивки решений (сетка 2, $\Delta y_{bw}^+ = 8.95$)

δ_{EL}^+	u_τ	$\ E\ _2$
50	5.841×10^{-2}	2.12×10^{-2}
75	5.742×10^{-2}	7.73×10^{-3}
100	5.727×10^{-2}	4.95×10^{-3}
120	5.727×10^{-2}	8.44×10^{-3}

В табл. 7 приведены результаты, демонстрирующие влияние положения точки сшивки решений, определяемого нормированным расстоянием δ_{EL}^+ . Так как точка сшивки является свободным параметром метода пенализированных пристеночных функций, то ее расположение может быть выбрано в любой области пограничного слоя, при этом точность численного решения определяется соответствием профиля скорости пристеночной функции с решением задачи (1)–(3) с используемой моделью турбулентности для замыкания. Кроме того, для ослабления требований к пристеночному разрешению точка сшивки должна быть достаточно удалена от вязкого подслоя. Так, из табл. 7 видно, что для данного числа Рейнольдса оптимальным положением является точка сшивки, соответствующая $\delta_{EL}^+ \approx 100$.

3.2. Обтекание пластины

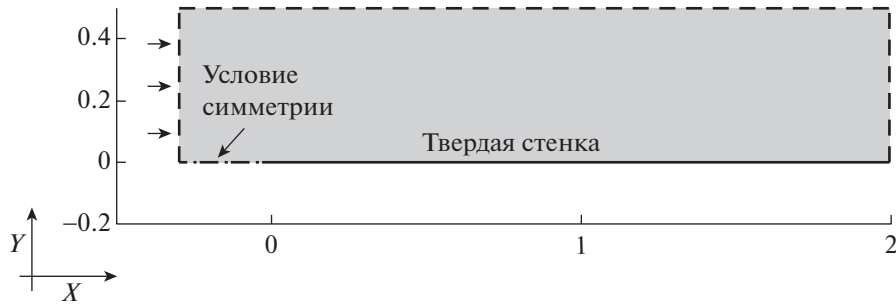
Для верификации метода пенализированных пристеночных функций в двумерной постановке решена задача численного моделирования турбулентного обтекания пластины. Параметры задачи соответствуют тесту NASA, представленному в [37]. При анализе результатов использовано численное решение из [37], полученное с помощью кода CFL3D, основанного на применении структурированных расчетных сеток, элементно-центрированного численного алгоритма и модели Спаларта–Аллмараса.

В качестве пристеночной функции использовался модифицированный закон Рейхарда (см. [38]), который более точно воспроизводит логарифмическую подобласть пограничного слоя:

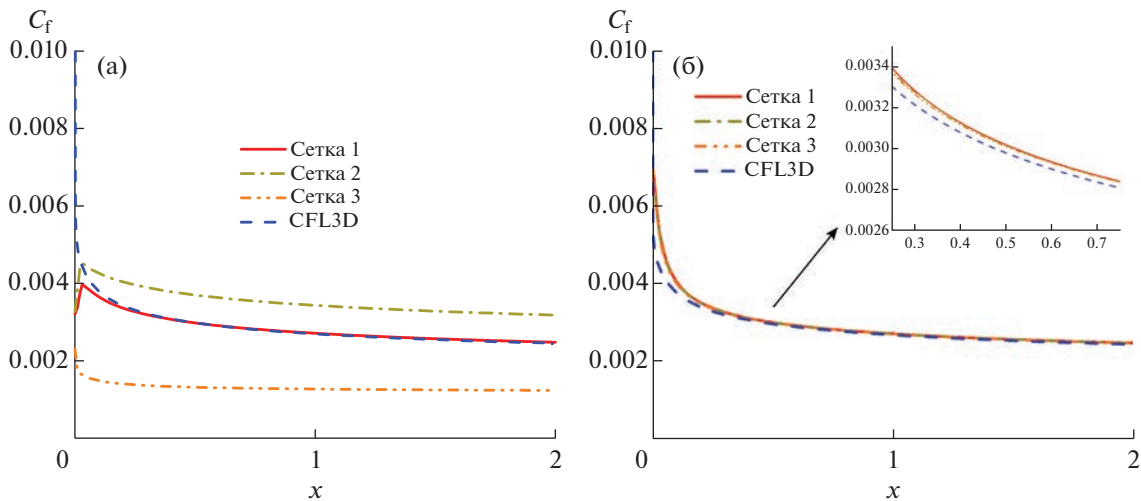
$$\tilde{f}_{Rei}(y^+) = (1 - \phi)f_{Rei}(y^+) + \phi f_{log}(y^+), \quad \phi = \tanh\left(\left(y^+/27\right)^4\right), \quad f_{log}(y^+) = \ln(y^+)/k + 21. \quad (28)$$

Параметры математической модели, определяющие течение, – число Рейнольдса и число Маха – равны $Re = 5 \times 10^6$ и $M = 0.2$. Задача решается в прямоугольной расчетной области в плоскости $XOY - [-0.3, 2] \times [0, 0.5]$ (фиг. 4). Поток движется слева направо, параллельно оси OX и набегает на бесконечно тонкую пластину с началом в точке $(0,0)$. При этом на левой границе расчетной области скорость потока постоянна и в безразмерном виде равна $(u)_{x=-0.3} = U_\infty = 1$. Здесь так же, как и для предыдущей задачи, на твердой стенке задаются нулевой поток температуры и граничное условие прилипания (модель Спаларта–Аллмараса) или условие (18) для поперечной и условие (19) для продольной составляющих вектора скорости (метод пенализированных пристеночных функций). Значение u_τ в (19) определяется из численного решения уравнения (22).

Численный расчет задачи проведен на последовательности структурированных декартовых сеток с различным пристеночным разрешением (см. табл. 8, где Δy_w^+ – нормированный поперечный размер пристеночной ячейки, остальные обозначения соответствуют использованным ранее при описании предыдущей задачи). Значение Δy_w^+ в таблице рассчитано по значению



Фиг. 4. Постановка задачи об обтекании пластины.



Фиг. 5. Коэффициент поверхностного трения на плоской пластине: (а) – модель Спаларта–Аллмараса, (б) – метод пенализированных пристеночных функций.

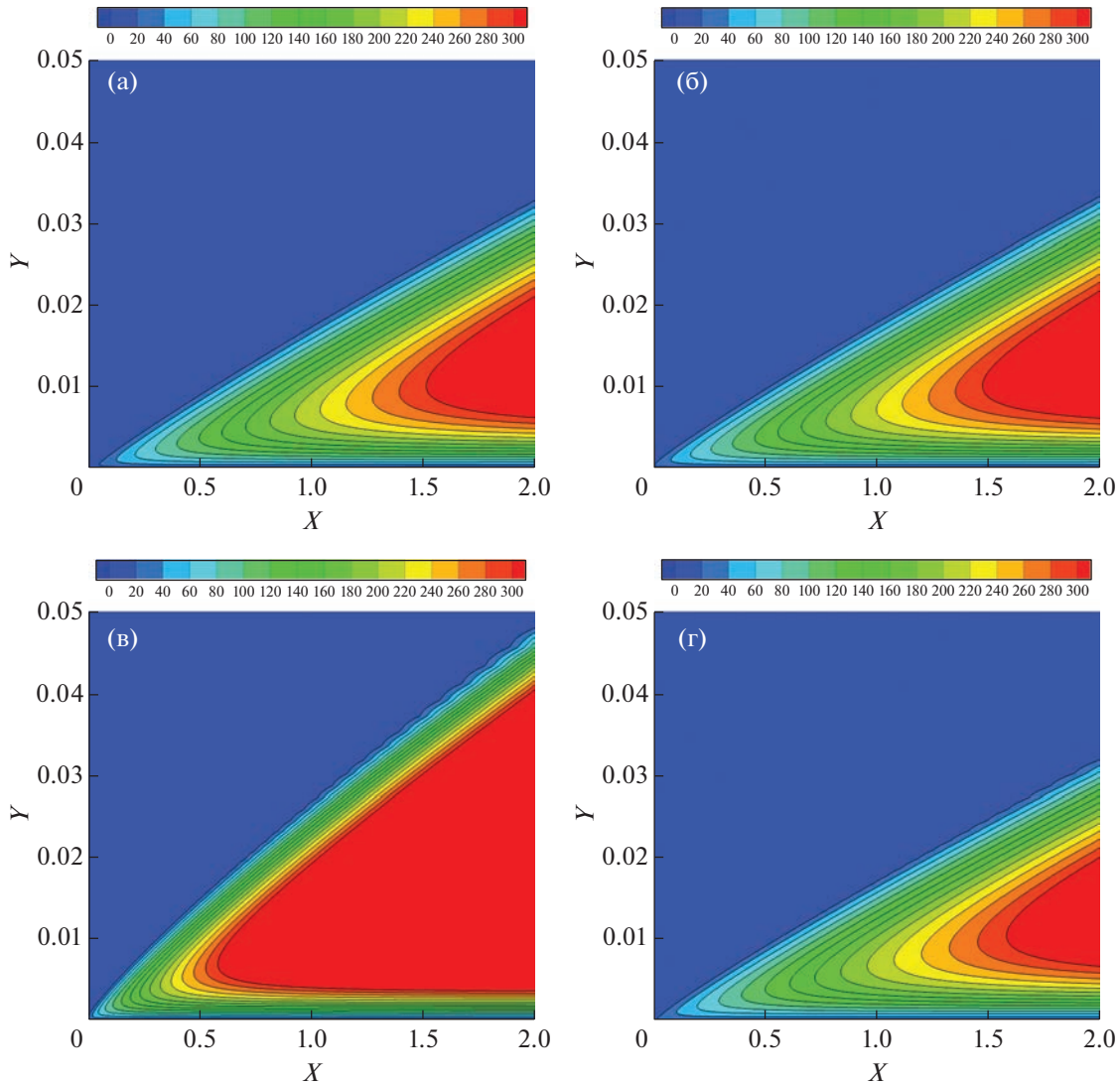
$u_\tau = 0.03688$, полученному из референсного решения (CFL3D) при $x = 0.97$. Размер ячеек в продольном направлении фиксирован – $\Delta x = 10^{-2}$, коэффициент разбега в поперечном направлении – $q = 1.2$, при этом размер ячеек в этом направлении ограничен величиной $\Delta y_{\max} = 10^{-2}$.

Для всех сеток задача решена в двух постановках: первая – пограничный слой на поверхности пластины моделируется уравнением Спаларта–Аллмараса, вторая – применяется метод пенализированных пристеночных функций. Во втором случае параметры уравнения для вычисления скорости трения на пластине для всех сеток одинаковы: коэффициенты пенализации $\eta_f = \eta_s = 10^{-2}$, толщина шивки $\sigma^+ = 58.2$, начальное приближение $u_\tau|_{l=0} = 4 \times 10^{-2}$.

На фиг. 5а представлены распределения коэффициента трения $C_f = \frac{\tau_w}{1/2\rho U_\infty^2}$, полученные при численном решении с применением модели Спаларта–Аллмараса. Видно, что результаты для сетки 1 хорошо согласуются с эталонным решением CFL3D, за исключением небольшой обла-

Таблица 8. Параметры сеток для расчетов задачи об обтекании пластины

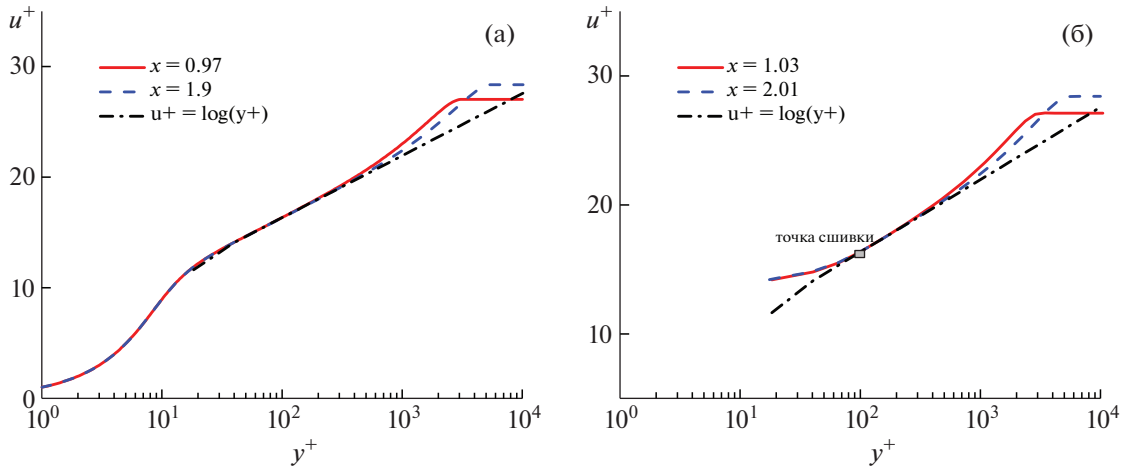
Сетка	N	Δy_w	$\Delta y_w^+(x = 0.97)$
1	19173	1.0×10^{-5}	1.84
2	17094	5.0×10^{-5}	9.22
3	16401	1.0×10^{-4}	18.44



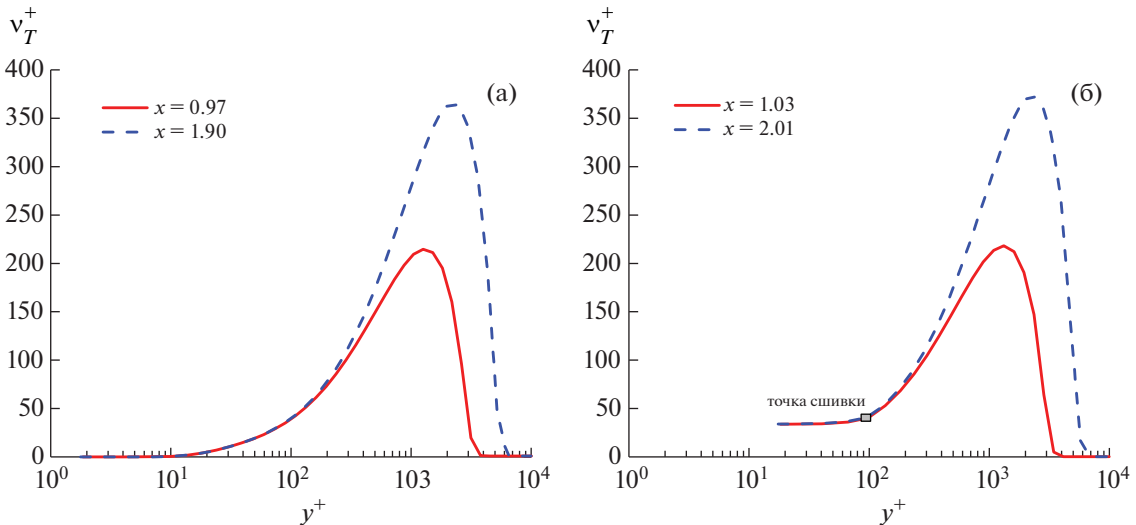
Фиг. 6. Линии уровня относительной турбулентной вязкости (v_T/v): (а) – код CFL3D, (б) – модель Спаларта–Алмараса (сетка 1), (в) – модель Спаларта–Алмараса (сетка 3), (г) – метод пенализированных пристеночных функций (сетка 3).

сти в начале пластины, вызванной сильным влиянием геометрической сингулярности на поведение регуляризованного решения в области передней кромки для разных граничных условий (условие переноса касательного напряжения (19) вместо условия прилипания). Расчеты с применением сеток более грубого пристеночного разрешения (сетки 2 и 3) показали неудовлетворительный результат, коэффициент трения существенно отличается от эталонного решения. На фиг. 5б те же зависимости приведены для решений, полученных с применением метода пенализированных пристеночных функций. Здесь уменьшение пристеночного сеточного разрешения не оказало заметного влияния на результаты расчетов, для всех сеток они хорошо согласуются с результатами CFL3D (за исключением окрестности точки $x = 0$).

Те же закономерности наблюдаются на графиках линий уровня относительной турбулентной вязкости $v_T^+ = v_T/v$ в пристеночной области, приведенных на фиг. 6. Расчеты с применением модели Спаларта–Алмараса (фиг. 6б) на сетке с разрешенным пограничным слоем (сетка 7) позволили получить распределение, согласующееся с результатами CFL3D (фиг. 6а), в отличие от результатов для сетки 3 (фиг. 6в). В то же время применение метода пенализированных пристеночных функций для расчетов на сетке 3 обеспечило хорошую точность результатов (фиг. 6г). Отметим лишь небольшой сдвиг линий уровня вправо, относительно распределений (а) и (б),



Фиг. 7. Профиль продольной скорости: (а) – модель Спаларта–Аллмараса (сетка 1), (б) – метод пенализированных пристеночных функций (сетка 3).



Фиг. 8. Профиль относительной турбулентной вязкости v_T^+ : (а) – модель Спаларта–Аллмараса (сетка 1), (б) – метод пенализированных пристеночных функций (сетка 3).

который объясняется фазовым сдвигом в распределении коэффициента трения на поверхности пластины при расчетах двумя методами с разными механизмами регуляризации решения в окрестности передней кромки пластины (см. фиг. 5).

На фиг. 7а приведены профили нормированной продольной скорости u^+ в сечениях $x = 0.97$ и $x = 1.9$ для расчета с применением модели Спаларта–Аллмараса на сетке с разрешенным пограничным слоем (сетка 1), на фиг. 7б – те же закономерности, но полученные уже методом пенализированных пристеночных функций на “грубой” сетке 3 для сечений $x = 1.03$ и $x = 2.01$. Справа от точки сшивки ($y^+ > \delta_{EL}^+ = 100$) профили в сечениях $x = 0.97$ и $x = 1.9$ хорошо согласуются с профилями в сечениях $x = 1.03$ и $x = 2.01$ соответственно. Сравнение проводится в сечениях, соответствующих значениям C_f : 0.00271 и 0.00247. Отметим, что отличие пространственного положения сравниваемых сечений обусловлено наличием относительного сдвига решений, вызванного различием в механизме регуляризации решений в области передней кромки пластины.

На фиг. 8а и 8б в тех же сечениях расчетной области сопоставлены профили относительной турбулентной вязкости v_T^+ . Слева от точки сшивки ($y^+ < \delta_{EL}^+$) при использовании метода пенали-

зированных пристеночных функций значение турбулентной вязкости практически постоянно. Это связано с завышенными относительно решения с помощью модели Спаларта–Аллмараса значениями продольной скорости в этом случае. Отметим, что несмотря на то что решение в этой области “экстраполировано” из внешнего решения переносом касательных напряжений из области сшивки на поверхность пластины, профили скорости и вязкости в численном решении справа от точки сшивки хорошо согласуются с решением задачи на основе модели Спаларта–Аллмараса.

ЗАКЛЮЧЕНИЕ

В работе представлен новый метод моделирования турбулентного пограничного слоя, применяемый при численном решении задач аэродинамики. С одной стороны, в нем так же, как и в классических подходах, используется идея сшивки внешнего решения с аналитическими пристеночными функциями, с другой – предлагается дифференциальная формулировка условия сшивки, которая позволяет избежать локальной интерполяции решения в точку сшивки и необходимости точного определения положения этой точки. Для переноса значений скорости трения на стенку используется характеристический метод штрафных функций. Метод сформулирован применительно к математической модели на основе усредненных по Рейнольдсу уравнений Навье–Стокса с замыканием уравнением Спаларта–Аллмараса.

Численная реализация метода не предполагает существенного усложнения вычислительного алгоритма. Дополнительное дифференциальное уравнение для расчета скорости трения на поверхности тела аппроксимировалось методом направленных разностей первого порядка, при этом сходимость по времени численного решения достигалась за несколько итераций.

Работоспособность метода продемонстрирована на примере решения двух задач – турбулентного течения в канале и обтекания плоской пластины. Результаты расчетов показали, что применение нового метода обеспечивает достаточную точность численного решения на сетках, в которых размер пристеночного шага более чем в 10 раз превышает ограничения, установленные для расчетов с применением модели Спаларта–Аллмараса (в которых точность моделирования пограничного слоя обеспечивается сеточным разрешением).

Дальнейшее развитие работы будет связано с несколькими направлениями. Во-первых, слабая формулировка условия сшивки делает его привлекательным для решения задач обтекания твердых тел с криволинейными границами, в том числе, с применением погруженных граничных условий. Во-вторых, особенности формулировки метода допускают его обобщение для применения в задачах с сильными отрывами, при этом предполагается использовать дифференциальные стандартные и неравновесные пристеночные функции.

СПИСОК ЛИТЕРАТУРЫ

1. *Moin P., Mahesh K.* Direct numerical simulation: A tool in turbulence research // *Ann. Rev. Fluid Mech.* 1998. V. 30. P. 539–578.
2. *Spalart P.R., Allmaras S.R.* A one equation turbulence model for aerodynamic flows // *AIAA J.* 1992. V. 94.
3. *Gatski T.B., Hussaini M.Y., Lumley J.L.* Simulation and Modeling of Turbulent Flows. Oxford, 1996.
4. *Durbin P.A., Reif B., Pettersson A.* Statistical Theory and Modeling for Turbulent Flows. Wiley, 2001.
5. *Wilcox D.C.* Formulation of the $k - \omega$ Turbulence Model Revisited // *AIAA J.* 2008. V. 46. 11. P. 2823–2838.
6. *Froehlich J., von Terzi D.* Hybrid LES/RANS methods for the simulation of turbulent flows // *Progress in Aerospace Sci.* 2008. JUL. V. 44. 5. P. 349–377.
7. *Xiao H., Jenny P.* A consistent dual-mesh framework for hybrid LES/RANS modeling // *J. Comp. Phys.* 2012. FEB 20. V. 231. 4. P. 1848–1865.
8. *Spalart P.R., Jou W.H., Strelets M. et al.* Comments on the feasibility of LES for wings, and on a hybrid RANS/LES approach // *First AFOSR Inter. Conf. on DNS/LES*, Ruston, Louisiana. V. 1. Greyden Press, Columbus, OH, 1997. P. 4–8.
9. *Spalart P.R., Deck S., Shur M.L., et al.* A new version of detached-eddy simulation, resistant to ambiguous grid densities // *Theoretic. and Comput. Fluid. Dynamic.* 2006. V. 20. 3. P. 181–195.
10. *Shur M.L., Spalart P.R., Strelets M.K., Travin A.K.* A hybrid RANS-LES approach with delayed-DES and wall-modelled LES capabilities // *Inter. J. of Heat and Fluid Flow.* 2008. V. 29. 6. P. 1638–1649.
11. *Patankar S.V., Spalding D.B.* Heat and Mass Transfer in Boundary Layers. Morgan-Grampia, 1968.
12. *Зайчик Л.И.* Пристеночные функции для моделирования турбулентного течения и теплообмена // *Теплофизика высоких температур.* 1997. Т. 35. 3. С. 391–396.

13. *Craft T.J., Gant S.E., Gerasimov A.V., et al.* Development and application of wall-function treatments for turbulent forced and mixed convection flows // *Fluid Dynamic. Res.* 2006. V. 38. 2. P. 127–144. Seiken Symposium. URL: <https://www.sciencedirect.com/science/article/pii/S0169598305000778>.
14. *Beaugendre H., Morency F.* Penalization of the Spalart–Allmaras turbulence model without and with a wall function: Methodology for a vortex in cell scheme // *Comput. and Fluid.* 2018. Jul. V. 170. P. 313–323. URL: <https://hal.inria.fr/hal-01963687>.
15. *Дубень А.П., Абалакин И.В., Цветкова В.О.* О граничных условиях на твердых стенках в задачах вязкого обтекания // *Матем. моделирование.* 2021. V. 32. 1. P. 79–98.
16. *Nichols R.H., Nelson C.C.* Wall function boundary conditions including heat transfer and compressibility // *AIAA J.* 2004. V. 42. 6. P. 1107–1114.
17. *Bodart J., Larsson J.* Wall-modeled large eddy simulation in complex geometries with application to high-lift devices // *Annual Research Briefs, Center for Turbulence Research, Stanford University.* 2011. P. 37–48.
18. *Beaugendre H., Morency F.* Penalization of the Spalart–Allmaras turbulence model without and with a wall function: Methodology for a vortex in cell scheme // *Comput. and Fluid.* 2018. JUL 15. V. 170. P. 313–323.
19. *Cai S.-G., Degrygn J., Boussuge J.-F., Sagaut P.* Coupling of turbulence wall models and immersed boundaries on Cartesian grids // *J. Comput. Phys.* 2021. V. 429. P. 109995. URL: <https://www.sciencedirect.com/science/article/pii/S0021999120307695>.
20. *Dhamankar N., Blaisdell G., Lyrantzis A.* Implementation of a wall-modeled sharp immersed boundary method in a high-order large eddy simulation tool for jet aeroacoustics // *54th AIAA Aerospace Sci. Meet.* 2016. 01.
21. *Brown-Dymkoski E., Kasimov N., Vasilyev O.V.* A characteristic based volume penalization method for general evolution problems applied to compressible viscous flows // *J. Comput. Phys.* 2014. Vol. 262. P. 344–357.
22. *Kasimov N., Dymkoski E., De Stefano G., Vasilyev O.V.* Galilean-invariant characteristic-based volume penalization method for supersonic flows with moving boundaries // *Fluids.* 2021. V. 6. 8. URL: <https://www.mdpi.com/2311-5521/6/8/293>.
23. *Kawai S., Larsson J.* Wall-modeling in large eddy simulation: length scales, grid resolution, and accuracy // *Phys. Fluid.* 2012. V. 24. 1. P. 015105.
24. *Kawai S., Larsson J.* Dynamic non-equilibrium wall-modeling for large eddy simulation at high Reynolds numbers // *Phys. Fluid.* 2013. V. 25. 1. P. 015105.
25. *Park G.I., Moin P.* An improved dynamic non-equilibrium wall-model for large eddy simulation // *Phys. Fluid.* 2014. V. 26. 1. P. 37–48.
26. *Абалакин И.В., Васильев О.В., Жданова Н.С., Козубская Т.К.* Метод характеристических штрафных функций для численного моделирования сжимаемых течений на неструктурированных сетках // *Ж. вычисл. матем. и матем. физ.* 2021. V. 61. 8. P. 1336–1352.
27. *Жданова Н.С., Абалакин И.В., Васильев О.В.* Расширение метода штрафных функций Бринкмана для сжимаемых течений вокруг подвижных твердых тел // *Матем. моделирование.* 2022. V. 34. 2. P. 41–57.
28. *Bardina J., Huang P., Coakley T., et al.* Turbulence modeling validation // *28th Fluid Dynamic. Conf.* 1997. P. 2121.
29. *Brown-Dymkoski E., Kasimov N., Vasilyev O.V.* Characteristic-based volume penalization method for arbitrary mach flows around solid obstacles // *Direct and Large-Eddy Simulation IX, Proceedings of the Ninth International ERCOFTAC Workshop on Direct and Large-Eddy Simulations.* Eds. J. Frohlich, H. Kuerten, B.J. Geurts, V. Armenio, Springer, 2015. P. 109–115.
30. *Gorobets A., Bakhvalov P.* Heterogeneous CPUx parallelization for high-accuracy scale-resolving simulations of compressible turbulent flows on hybrid supercomputers // *Comput. Phys. Communicat.* 2022. V. 271. P. 108231. URL: <https://www.sciencedirect.com/science/article/pii/S001046552100343X>.
31. *Gorobets A., Duben A.* Technology for supercomputer simulation of turbulent flows in the good new days of exascale computing // *Supercomput. Frontiers and Innovat.* 2021. Feb. V. 8, 4. P. 4–10. URL: <https://superfri.su-su.ru/index.php/superfri/article/view/400>.
32. *Bakhvalov P., Abalakin I., Kozubskaya T.* Edge-based reconstruction schemes for unstructured tetrahedral meshes // *Inter. J. Numer. Meth. Fluid.* 2016. V. 81. 6. P. 331–356.
33. *van der Vorst H.A.* BI-CGSTAB: A fast and smoothly converging variant of BI-CG for the solution of nonsymmetric linear systems // *SIAM J. Sci. Stat. Comput.* 1992. V. 13. 2. P. 631–644.
34. *Reichardt H.* Vollständige Darstellung der turbulenten Geschwindigkeitsverteilung in glatten Leitungend // *Zeitschrift fñr Angewandte Math. und Mech.* 1951. V. 31. 7. P. 208–219.
35. *Zanoun E.S.M.* Answers to Some Open Questions in Wall Bounded Laminar and Turbulent Shear Flows. Technische Fakultät der Universität Erlangen-Nürnberg., 2003. URL: <https://books.google.ru/books?id=IPVcxwEACAAJ>.
36. *Moser R.D., Kim J., Mansour N.N.* Direct numerical simulation of turbulent channel flow up to $Re_\tau Y0$ // *Phys. Fluids A.* 1999. V. 11. P. 943–945.
37. NASA Langley Research Center Turbulence Modeling Resource. URL: <https://turbmodels.larc.nasa.gov>.
38. *Knopp T.* On grid-independence of RANS predictions for aerodynamic flows using model-consistent universal wall-functions // *Proceed. of the European Conf. on Comput. Fluid Dynamic.* 2006.

МАТЕМАТИЧЕСКАЯ
ФИЗИКА

УДК 517.956

ПРИБЛИЖЕННОЕ РЕШЕНИЕ ОБРАТНОЙ ЗАДАЧИ
ДЛЯ ИНТЕГРОДИФФЕРЕНЦИАЛЬНОГО УРАВНЕНИЯ
ТЕПЛОПРОВОДНОСТИ С СИНГУЛЯРНЫМ ВОЗМУЩЕНИЕМ¹⁾

© 2023 г. А. М. Денисов^{1,*}

¹ 119999 Москва, Ленинские горы, МГУ, Россия

*e-mail: den@cs.msu.ru

Поступила в редакцию 10.09.2022 г.
Переработанный вариант 10.09.2022 г.
Принята к публикации 02.02.2023 г.

Для интегродифференциального уравнения теплопроводности с сингулярным возмущением рассматривается обратная задача, состоящая в определении граничного условия по дополнительной информации о решении начально-краевой задачи. Доказано, что приближенное решение обратной задачи может быть получено на основе использования конечного числа членов разложения решения начально-краевой задачи по малому параметру. Библ. 11.

Ключевые слова: интегродифференциальное уравнение теплопроводности, сингулярное возмущение, обратная задача, приближенное решение.

DOI: 10.31857/S0044466923050095, EDN: PKNZWI

В многообразии обратных задач для уравнений в частных производных можно выделить класс, в котором обратные задачи рассматриваются для сингулярно возмущенных уравнений. Методы приближенного решения обратных задач, основанные на замене исходного дифференциального уравнения сингулярно возмущенным, были предложены в [1] и получили затем развитие в [2–6] и ряде других работ. Обратные задачи для сингулярно возмущенных дифференциальных уравнений исследовались в [7–11].

В работе [11] для уравнения теплопроводности с сингулярным возмущением были предложены методы приближенного решения обратных задач, основанные на использовании разложения решения начально-краевой задачи по малому параметру. Эта статья посвящена применению этого подхода для приближенного решения обратной задачи для интегродифференциального уравнения теплопроводности с сингулярным возмущением.

Рассмотрим начально-краевую задачу для интегродифференциального уравнения теплопроводности с сингулярным возмущением

$$\varepsilon^2 u_t(x, t) = u_{xx}(x, t) + \varepsilon^2 \int_0^t K(t, \tau) u(x, \tau) d\tau, \quad (x, t) \in Q_{t_0}, \quad (1)$$

$$u(0, t) = \mu(t), \quad 0 \leq t \leq t_0, \quad (2)$$

$$u_x(\pi, t) = 0, \quad 0 \leq t \leq t_0, \quad (3)$$

$$u(x, 0) = 0, \quad 0 \leq x \leq \pi, \quad (4)$$

где $0 < \varepsilon \leq \varepsilon_0 < 1$, $Q_{t_0} = \{(x, t) : 0 < x < \pi, 0 < t \leq t_0\}$.

Далее, чтобы подчеркнуть зависимость решения задачи (1)–(4) от параметра ε , будем обозначать его через $u(x, t; \varepsilon)$.

¹⁾Работа выполнена при финансовой поддержке Минобрнауки РФ в рамках реализации программы Московского центра фундаментальной и прикладной математики (проект 075-15-2022-284).

Предположим, что функция $\mu \in C^2[0, T]$, $\mu(0) = 0$, а $K(t, \tau) \in C(\Delta_{t_0})$, где $\Delta_{t_0} = \{(t, \tau), 0 \leq \tau \leq t \leq t_0\}$. Из метода разделения переменных следует формула для решения задачи (1)–(4)

$$u(x, t; \varepsilon) = \mu(t) + \sum_{n=0}^{\infty} \frac{4}{\pi(2n+1)} T_n(t; \varepsilon) \sin\left(\frac{2n+1}{2}x\right), \quad (5)$$

где функции $T_n(t; \varepsilon)$ являются решениями задачи Коши для интегродифференциального уравнения

$$\varepsilon^2 T_n'(t; \varepsilon) = -\left(\frac{2n+1}{2}\right)^2 T_n(t; \varepsilon) + \varepsilon^2 \int_0^t K(t, \tau) T_n(\tau; \varepsilon) d\tau - \varepsilon^2 \left[\mu'(t) - \int_0^t K(t, \tau) \mu(\tau) d\tau \right], \quad 0 \leq t \leq t_0, \quad (6)$$

$$T_n(0, \varepsilon) = 0. \quad (7)$$

Интегрируя уравнение (6) с начальным условием (7), получим, что $T_n(t; \varepsilon)$ удовлетворяют интегральному уравнению

$$T_n(t; \varepsilon) = \int_0^t B(t, \tau; \varepsilon, n) T_n(\tau; \varepsilon) d\tau + \psi(t; \varepsilon, n), \quad 0 \leq t \leq t_0, \quad (8)$$

где

$$B(t, \tau; \varepsilon, n) = \int_{\tau}^t \exp\left(-\frac{(2n+1)^2}{4\varepsilon^2}(t-\theta)\right) K(\theta, \tau) d\theta, \quad (9)$$

$$\psi(t; \varepsilon, n) = -\int_0^t \exp\left(-\frac{(2n+1)^2}{4\varepsilon^2}(t-\tau)\right) \left[\mu'(\tau) - \int_0^{\tau} K(\tau, \theta) \mu(\theta) d\theta \right] d\tau. \quad (10)$$

Сформулируем обратную задачу. Пусть функция $K(t, \tau)$ и число ε заданы, а функция $\mu(t)$ неизвестна. Требуется определить $\mu(t)$, если задана дополнительная информация о решении задачи (1)–(4)

$$u(x_0, t; \varepsilon) = g(t; \varepsilon), \quad 0 \leq t \leq t_0, \quad (11)$$

где x_0 – заданное число, $x_0 \in (0, \pi]$, а $g(t; \varepsilon)$ – заданная функция.

Из формулы (5) и уравнения (8) следует, что можно получить разложение $u(x, t; \varepsilon)$ по малому параметру ε .

Рассмотрим вопрос о возможности построения приближенного решения обратной задачи на основе использования этого разложения.

Так как функция $T_n(t; \varepsilon)$ является решением уравнения (8), то для всех $n = 0, 1, 2, \dots$ справедлива оценка

$$\max_{[0, t_0]} |T_n(t; \varepsilon)| \leq c_1 \frac{\varepsilon^2}{(2n+1)^2}.$$

Здесь и далее через c обозначаются постоянные, не зависящие от ε и n .

Из предыдущего неравенства и формулы (5) следует, что для $u(x, t; \varepsilon)$ справедливо представление

$$u(x, t; \varepsilon) = \mu(t) + \varepsilon^2 w_0(x, t; \varepsilon), \quad (12)$$

где

$$\max_{Q_{t_0}} |w_0(x, t; \varepsilon)| \leq c_2.$$

Определив приближенное решение обратной задачи $\tilde{\mu}_0(t; \varepsilon)$ следующим образом $\tilde{\mu}_0(t; \varepsilon) = g(t; \varepsilon)$ и используя представление (12), имеем

$$\max_{[0, t_0]} |\mu(t) - \tilde{\mu}_0(t; \varepsilon)| = \varepsilon^2 \max_{[0, t_0]} |w_0(x_0, t; \varepsilon)| \leq c_2 \varepsilon^2.$$

Таким образом, при малых значениях ε функцию $\tilde{\mu}_0(t; \varepsilon)$ можно считать приближенным решением обратной задачи.

Рассмотрим вариант построения приближенного решения обратной задачи, использующий более высокие члены разложения $u(x, t; \varepsilon)$ по малому параметру.

Введем функцию $f_1(x) = x^2/2 - \pi x$.

Теорема 1. Пусть функции $K(t, \tau)$, $\mu(t)$ удовлетворяют следующим условиям:

$$K(t, \tau), K_t(t, \tau) \in C(\Delta_0), \quad \|K\|_{C(\Delta_0)} t_0 |f_1(x_0)| \varepsilon_0^2 < 1, \tag{13}$$

$\mu \in C^2[0, t_0]$, $\mu(0) = \mu'(0) = 0$, $\mu(t_0) = m_0$, где m_0 – заданное число. Тогда, если функция $\tilde{\mu}_1(t; \varepsilon)$ является решением задачи

$$\varepsilon^2 f_1(x_0) \tilde{\mu}'_1(t; \varepsilon) + \tilde{\mu}_1(t; \varepsilon) - \varepsilon^2 f_1(x_0) \int_0^t K(t, \tau) \tilde{\mu}_1(\tau; \varepsilon) d\tau = g(t; \varepsilon), \quad 0 \leq t \leq t_0, \tag{14}$$

$$\tilde{\mu}_1(t_0; \varepsilon) = m_0, \tag{15}$$

то

$$\max_{[0, t_0]} |\mu(t) - \tilde{\mu}_1(t; \varepsilon)| \leq c_3 \varepsilon^4. \tag{16}$$

Доказательство. Рассмотрим функции $T_n(t; \varepsilon)$, являющиеся решением интегрального уравнения (8). Используя резольвенту $R(t, \tau; \varepsilon, n)$ ядра $B(t, \tau; \varepsilon, n)$, имеем

$$T_n(t; \varepsilon) = \psi(t; \varepsilon, n) + \int_0^t R(t, \tau; \varepsilon, n) \psi(\tau; \varepsilon, n) d\tau, \quad 0 \leq t \leq t_0. \tag{17}$$

Интегрируя по частям интеграл в формуле (9), получим следующую оценку:

$$\max_{\Delta_0} |B(t, \tau; \varepsilon, n)| \leq c_4 \frac{\varepsilon^2}{(2n+1)^2},$$

из которой следует, что

$$\max_{\Delta_0} |R(t, \tau; \varepsilon, n)| \leq c_5 \frac{\varepsilon^2}{(2n+1)^2}. \tag{18}$$

Преобразуя формулу (10), имеем

$$\psi(t; \varepsilon, n) = -\frac{4\varepsilon^2}{(2n+1)^2} \left[\mu'(t) - \int_0^t K(t, \tau) \mu(\tau) d\tau \right] + w_1(t; \varepsilon, n), \tag{19}$$

где

$$w_1(t; \varepsilon, n) = \frac{4\varepsilon^2}{(2n+1)^2} \int_0^t \exp\left(-\frac{(2n+1)^2}{4\varepsilon^2}(t-\tau)\right) \left[\mu''(\tau) - K(\tau, \tau) \mu(\tau) - \int_0^\tau K_\tau(\tau, \theta) \mu(\theta) d\theta \right] d\tau, \tag{20}$$

а

$$\max_{[0, t_0]} |w_1(t; \varepsilon, n)| \leq c_6 \frac{\varepsilon^4}{(2n+1)^4}. \tag{21}$$

Из формул (17), (19), (20) и оценок (18), (21) следует, что

$$T_n(t; \varepsilon) = -\frac{4\varepsilon^2}{(2n+1)^2} \left[\mu'(t) - \int_0^t K(t, \tau) \mu(\tau) d\tau \right] + w_2(t; \varepsilon, n), \quad (22)$$

где

$$\max_{[0, t_0]} |w_2(t; \varepsilon, n)| \leq c_7 \frac{\varepsilon^4}{(2n+1)^4}. \quad (23)$$

Используя формулы (5), (22), оценку (23) и равенство

$$f_1(x) = -\sum_{n=0}^{\infty} \frac{16}{\pi(2n+1)^3} \sin\left(\frac{2n+1}{2}x\right),$$

получим, что для $u(x, t; \varepsilon)$ справедливо представление

$$u(x, t; \varepsilon) = \mu(t) + \varepsilon^2 f_1(x) \left[\mu'(t) - \int_0^t K(t, \tau) \mu(\tau) d\tau \right] + \varepsilon^4 w_3(x, t; \varepsilon), \quad (24)$$

где $\max_{\bar{Q}_0} |w_3(x, t; \varepsilon)| \leq c_8$.

Положив в равенстве (24) $x = x_0$ и используя условие (11), имеем

$$\mu(t) + \varepsilon^2 f_1(x_0) \left[\mu'(t) - \int_0^t K(t, \tau) \mu(\tau) d\tau \right] = g(t; \varepsilon) - \varepsilon^4 w_3(x_0, t; \varepsilon), \quad 0 \leq t \leq t_0. \quad (25)$$

Пусть $\check{\mu}_1(t; \varepsilon)$ – решение задачи (14), (15). Рассмотрим функцию $z(t; \varepsilon) = \check{\mu}_1(t; \varepsilon) - \mu(t)$. Из уравнения (14), равенства (25) и условия (15) получаем, что $z(t; \varepsilon)$ является решением задачи

$$z'(t; \varepsilon) + (\varepsilon^2 f_1(x_0))^{-1} z(t; \varepsilon) - \int_0^t K(t, \tau) z(\tau; \varepsilon) d\tau = \varepsilon^2 (f_1(x_0))^{-1} w_3(x_0, t; \varepsilon), \quad 0 \leq t \leq t_0, \quad z(t_0, \varepsilon) = 0.$$

Следовательно, $z(t; \varepsilon)$ является решением интегрального уравнения

$$z(t; \varepsilon) = -\int_t^{t_0} \exp\left(-\frac{t-\tau}{\varepsilon^2 f_1(x_0)}\right) \int_0^\tau K(\tau, \theta) z(\theta; \varepsilon) d\theta d\tau - \varepsilon^2 (f_1(x_0))^{-1} \int_t^{t_0} \exp\left(-\frac{t-\tau}{\varepsilon^2 f_1(x_0)}\right) w_3(x_0, \tau; \varepsilon) d\tau.$$

Используя это уравнение, условие (13) и оценку для функции $w_3(x_0, t; \varepsilon)$, имеем

$$\max_{[0, t_0]} |z(t; \varepsilon)| \leq c_3 \varepsilon^4.$$

Следовательно, справедлива оценка (16) и теорема 1 доказана.

Замечание 1. Очевидно, что при замене начального условия (15) на $\check{\mu}_1(0; \varepsilon) = 0$ решение уравнения (14) с этим условием не будет приближенным решением обратной задачи.

Покажем, что при дополнительных предположениях можно получить приближенное решение обратной задачи с оценкой погрешности порядка $O(\varepsilon^6)$.

Введем функцию

$$f_2(x) = \sum_{n=0}^{\infty} \frac{64}{\pi(2n+1)^5} \sin\left(\frac{2n+1}{2}x\right) = \frac{x^4}{24} - \frac{\pi x^3}{6} + \frac{\pi^3 x}{3}.$$

Рассмотрим следующую задачу для функции $y(t; \varepsilon)$:

$$\begin{aligned} &\varepsilon^4 f_2(x_0) y''(t; \varepsilon) + \varepsilon^2 f_1(x_0) y'(t; \varepsilon) + y(t; \varepsilon) = H(t; \varepsilon) + \\ &+ \varepsilon^4 f_2(x_0) \left[2K(t, t) y(t; \varepsilon) + \int_0^t \left(K_t(t, \tau) - K_\tau(t, \tau) - \int_\tau^t K(t, \theta) K(\theta, \tau) d\theta \right) y(\tau; \varepsilon) d\tau \right] + \\ &+ \varepsilon^2 f_1(x_0) \int_0^t K(t, \tau) y(\tau; \varepsilon) d\tau, \quad 0 \leq t \leq t_0, \end{aligned} \tag{26}$$

$$y(t_0; \varepsilon) = y_0, \quad y'(t_0; \varepsilon) = y_1, \tag{27}$$

где $H(t; \varepsilon)$ – заданная функция, а y_0 и y_1 – заданные числа.

Введем числа

$$a = -\frac{f_1(x_0)}{2f_2(x_0)} > 0, \quad b = \sqrt{\frac{4f_2(x_0) - (f_1(x_0))^2}{4(f_2(x_0))^2}}.$$

Лемма. Если выполнены следующие условия: $K \in C^1(\Delta_{t_0})$, $H \in C[0, t_0]$,

$$\frac{\varepsilon_0^2}{abf_2(x_0)} \left[\|K\|_{C(\Delta_{t_0})} (|f_1(x_0)|t_0 + 2\varepsilon^2 f_2(x_0)) \right] + \frac{\varepsilon_0^4}{ab} \left[(\|K_t\|_{C(\Delta_{t_0})} + \|K_\tau\|_{C(\Delta_{t_0})}) t_0 + \|K\|_{C(\Delta_{t_0})}^2 t_0^2 / 2 \right] < 1, \tag{28}$$

то существует единственная функция $y \in C^2[0, t_0]$, являющаяся решением задачи (26), (27).

Доказательство. Рассмотрим задачу Коши для функции $v(t; \varepsilon)$

$$\varepsilon^4 f_2(x_0) v''(t; \varepsilon) + \varepsilon^2 f_1(x_0) v'(t; \varepsilon) + v(t; \varepsilon) = Q(t), \quad 0 \leq t \leq t_0, \tag{29}$$

$$v(t_0, \varepsilon) = y_0, \quad v'(t_0, \varepsilon) = y_1. \tag{30}$$

Так как $f_1(x_0) < 0$, $f_2(x_0) > 0$ и $4f_2(x_0) - (f_1(x_0))^2 > 0$, то корни характеристического уравнения

$$\lambda^2 + \frac{f_1(x_0)}{\varepsilon^2 f_2(x_0)} \lambda + \frac{1}{\varepsilon^4 f_2(x_0)} = 0$$

равны $\lambda_{1,2} = a/(\varepsilon^2) \pm ib/(\varepsilon^2)$.

Решение задачи (29), (30) определяется формулой

$$\begin{aligned} v(t; \varepsilon) = &y_0 \exp\left(\frac{a}{\varepsilon^2}(t - t_0)\right) \cos\left(\frac{b}{\varepsilon^2}(t - t_0)\right) + \frac{y_1 \varepsilon^2 - a y_0}{b} E(t - t_0; \varepsilon) - \\ &- \frac{1}{\varepsilon^2 f_2(x_0) b} \int_t^{t_0} E(t - \tau; \varepsilon) Q(\tau) d\tau, \quad 0 \leq t \leq t_0, \end{aligned} \tag{31}$$

где

$$E(t - \tau; \varepsilon) = \exp\left(\frac{a}{\varepsilon^2}(t - \tau)\right) \sin\left(\frac{b}{\varepsilon^2}(t - \tau)\right).$$

Пусть функция $y(t; \varepsilon)$ является решением задачи (26), (27). Используя формулу (31), получим, что $y(t; \varepsilon)$ удовлетворяет интегральному уравнению

$$\begin{aligned} y(t; \varepsilon) = &y_0 \exp\left(\frac{a}{\varepsilon^2}(t - t_0)\right) \cos\left(\frac{b}{\varepsilon^2}(t - t_0)\right) + \frac{y_1 \varepsilon^2 - a y_0}{b} E(t - t_0; \varepsilon) - \\ &- \frac{1}{\varepsilon^2 f_2(x_0) b} \int_t^{t_0} E(t - \tau; \varepsilon) H(\tau; \varepsilon) d\tau - \\ &- \frac{1}{f_2(x_0) b} \int_t^{t_0} E(t - \tau; \varepsilon) \left[f_1(x_0) \int_0^\tau K(\tau, \theta) y(\theta; \varepsilon) d\theta + 2\varepsilon^2 f_2(x_0) K(\tau, \tau) y(\tau; \varepsilon) \right] d\tau - \end{aligned} \tag{32}$$

$$-\frac{\varepsilon^2}{b} \int_t^{t_0} E(t-\tau; \varepsilon) \int_0^\tau \left[K_\tau(\tau, \theta) - K_\theta(\tau, \theta) - \int_\theta^\tau K(\tau, \xi) K(\xi, \theta) d\xi \right] y(\theta; \xi) d\theta d\tau.$$

Уравнение (32) является уравнением Фредгольма II рода для функции $y(t; \varepsilon)$. Из условий леммы следует, что оно имеет единственное решение $y \in C^2[0, t_0]$, которое является решением задачи (26), (27). Лемма доказана.

Предположим, что известны значения функции $\mu(t)$ и ее производной при $t = t_0$: $\mu(t_0) = m_0$, $\mu'(t_0) = m_1$. Пусть $K \in C^1(\Delta_{t_0})$ и выполнено условие (28). Тогда из леммы следует, что существует единственная функция $\tilde{\mu}_2(t; \varepsilon) \in C^2[0, t_0]$, являющаяся решением задачи

$$\begin{aligned} \varepsilon^4 f_2(x_0) \tilde{\mu}_2''(t; \varepsilon) + \varepsilon^2 f_1(x_0) \tilde{\mu}_2'(t; \varepsilon) + \tilde{\mu}_2(t; \varepsilon) = g(t; \varepsilon) + \\ + \varepsilon^4 f_2(x_0) \left[2K(t, t) \tilde{\mu}_2(t; \varepsilon) + \int_0^t \left(K_t(t, \tau) - K_\tau(t, \tau) - \int_\tau^t K(t, \theta) K(\theta, \tau) d\theta \right) \tilde{\mu}_2(\tau; \varepsilon) d\tau \right] + \\ + \varepsilon^2 f_1(x_0) \int_0^t K(t, \tau) \tilde{\mu}_2(\tau; \varepsilon) d\tau, \quad 0 \leq t \leq t_0, \end{aligned} \quad (33)$$

$$\tilde{\mu}_2(t_0; \varepsilon) = m_0, \quad \tilde{\mu}_2'(t_0; \varepsilon) = m_1. \quad (34)$$

Теорема 2. *Предположим, что $\mu \in C^3[0, t_0]$, $\mu(0) = \mu'(0) = \mu''(0) = 0$, $\mu(t_0) = m_0$, $\mu'(t_0) = m_1$, $K \in C^2(\Delta_{t_0})$ и выполнено условие (28). Тогда*

$$\max_{[0, t_0]} |\tilde{\mu}_2(t; \varepsilon) - \mu(t)| \leq c_9 \varepsilon^6. \quad (35)$$

Доказательство. Преобразуем формулу (19) для свободного члена $\psi(t; \varepsilon, n)$ уравнения (8)

$$\begin{aligned} \psi(t; \varepsilon, n) = -\frac{4\varepsilon^2}{(2n+1)^2} \left[\mu'(t) - \int_0^t K(t, \tau) \mu(\tau) d\tau \right] + w_1(t; \varepsilon, n) = -\frac{4\varepsilon^2}{(2n+1)^2} \times \\ \times \left[\mu'(t) - \int_0^t K(t, \tau) \mu(\tau) d\tau \right] + \frac{16\varepsilon^4}{(2n+1)^4} \left[\mu''(t) - K(t, t) \mu(t) - \int_0^t K_t(t, \tau) \mu(\tau) d\tau \right] + w_4(t; \varepsilon, n), \end{aligned} \quad (36)$$

где

$$\max_{[0, t_0]} |w_4(t; \varepsilon, n)| \leq c_{10} \frac{\varepsilon^6}{(2n+1)^6}.$$

Для резольвенты $R(t, \tau; \varepsilon, n)$ ядра $B(t, \tau; \varepsilon, n)$ справедливо представление

$$R(t, \tau; \varepsilon, n) = B(t, \tau; \varepsilon, n) + R_1(t, \tau; \varepsilon, n), \quad \max_{\bar{Q}_{t_0}} |R_1(t, \tau; \varepsilon, n)| \leq c_{11} \frac{\varepsilon^4}{(2n+1)^4}.$$

Следовательно,

$$\begin{aligned} \int_0^t R(t, \tau; \varepsilon, n) \psi(\tau; \varepsilon, n) d\tau = -\frac{16\varepsilon^4}{(2n+1)^4} \int_0^t K(t, \tau) \left[\mu'(\tau) - \int_0^\tau K(\tau, \xi) \mu(\xi) d\xi \right] d\tau + w_5(t; \varepsilon, n) = \\ = -\frac{16\varepsilon^4}{(2n+1)^4} K(t, t) \mu(t) + \frac{16\varepsilon^4}{(2n+1)^4} \int_0^t \left[K_\tau(t, \tau) + \int_\tau^t K(t, \xi) K(\xi, \tau) d\xi \right] \mu(\tau) d\tau + w_5(t; \varepsilon, n), \end{aligned} \quad (37)$$

где

$$\max_{[0, t_0]} |w_5(t; \varepsilon, n)| \leq c_{12} \frac{\varepsilon^6}{(2n+1)^6}.$$

Учитывая формулы (36), (37), получаем следующее представление для функции $T_n(t; \varepsilon)$, являющейся решением интегрального уравнения (8)

$$T_n(t; \varepsilon) = -\frac{4\varepsilon^2}{(2n+1)^2} \left[\mu'(t) - \int_0^t K(t, \tau) \mu(\tau) d\tau \right] + \frac{16\varepsilon^4}{(2n+1)^4} \left[\mu''(t) - K(t, t) \mu(t) - \int_0^t K_t(t, \tau) \mu(\tau) d\tau \right] - \frac{16\varepsilon^4}{(2n+1)^4} \left[K(t, t) \mu(t) - \int_0^t \left[K_\tau(t, \tau) + \int_\tau^t K(t, \xi) K(\xi, \tau) d\xi \right] \mu(\tau) d\tau \right] + w_6(t; \varepsilon, n), \tag{38}$$

где

$$\max_{[0, t_0]} |w_6(t; \varepsilon, n)| \leq c_{13} \frac{\varepsilon^6}{(2n+1)^6}.$$

Из формул (5), (38) и определения функций $f_1(x)$ и $f_2(x)$ следует асимптотическое представление решения задачи (1)–(4)

$$u(x, t; \varepsilon) = \mu(t) + \varepsilon^2 f_1(x) \left[\mu'(t) - \int_0^t K(t, \tau) \mu(\tau) d\tau \right] + \varepsilon^4 f_2(x) \mu''(t) - \varepsilon^4 f_2(x) \left[2K(t, t) \mu(t) + \int_0^t \left[K_t(t, \tau) - K_\tau(t, \tau) - \int_\tau^t K(t, \xi) K(\xi, \tau) d\xi \right] \mu(\tau) d\tau \right] + w_7(x, t; \varepsilon),$$

где

$$\max_{Q_{t_0}} |w_7(x, t; \varepsilon)| \leq c_{14} \varepsilon^6.$$

Положив $x = x_0$ и используя условие (11), имеем

$$\varepsilon^4 f_2(x_0) \mu''(t) + \varepsilon^2 f_1(x_0) \mu'(t) + \mu(t) = g(t; \varepsilon) + \varepsilon^2 f_1(x_0) \int_0^t K(t, \tau) \mu(\tau) d\tau + 2\varepsilon^4 f_2(x_0) K(t, t) \mu(t) + \varepsilon^4 f_2(x_0) \int_0^t \left[K_t(t, \tau) - K_\tau(t, \tau) - \int_\tau^t K(t, \xi) K(\xi, \tau) d\xi \right] \mu(\tau) d\tau + w_7(x_0, t; \varepsilon), \quad 0 \leq t \leq t_0. \tag{39}$$

Введем функцию $z(t; \varepsilon) = \mu(t) - \mu_2(t; \varepsilon)$. Из уравнений (39), (33) и условий (34) следует, что функция $z(t; \varepsilon)$ является решением задачи

$$\begin{aligned} \varepsilon^4 f_2(x_0) z''(t; \varepsilon) + \varepsilon^2 f_1(x_0) z'(t; \varepsilon) + z(t; \varepsilon) &= w_7(x_0, t; \varepsilon) + \\ + \varepsilon^4 f_2(x_0) \left[2K(t, t) z(t; \varepsilon) + \int_0^t \left(K_t(t, \tau) - K_\tau(t, \tau) - \int_\tau^t K(t, \theta) K(\theta, \tau) d\theta \right) z(\tau; \varepsilon) d\tau \right] &+ \\ + \varepsilon^2 f_1(x_0) \int_0^t K(t, \tau) z(\tau; \varepsilon) d\tau, \quad 0 \leq t \leq t_0, & \\ z(t_0; \varepsilon) = 0, \quad z'(t_0; \varepsilon) = 0. & \end{aligned}$$

Аналогично доказательству леммы, используя формулу (31), получим, что $z(t; \varepsilon)$ удовлетворяет интегральному уравнению

$$\begin{aligned}
z(t; \varepsilon) = & -\frac{1}{\varepsilon^2 f_2(x_0) b} \int_t^{t_0} E(t - \tau; \varepsilon) w_7(x_0, \tau; \varepsilon) d\tau - \\
& -\frac{1}{f_2(x_0) b} \int_t^{t_0} E(t - \tau; \varepsilon) \left[f_1(x_0) \int_0^\tau K(\tau, \theta) z(\theta; \varepsilon) d\theta + 2\varepsilon^2 f_2(x_0) K(\tau, \tau) z(\tau; \varepsilon) \right] d\tau - \\
& -\frac{\varepsilon^2}{b} \int_t^{t_0} E(t - \tau; \varepsilon) \int_0^\tau \left[K_\tau(\tau, \theta) - K_\theta(\tau, \theta) - \int_\theta^\tau K(\tau, \xi) K(\xi, \theta) d\xi \right] z(\theta; \xi) d\theta d\tau.
\end{aligned} \tag{40}$$

Уравнение (40) является уравнением Фредгольма II рода для функции $z(t; \varepsilon)$. Для свободного члена этого уравнения справедлива оценка

$$\max_{|0, t_0|} \left| \frac{1}{\varepsilon^2 f_2(x_0) b} \int_t^{t_0} E(t - \tau; \varepsilon) w_7(x_0, \tau; \varepsilon) d\tau \right| \leq c_{15} \varepsilon^6.$$

Из этого неравенства, условия (28) и уравнения (40) следует оценка (35) и теорема 2 доказана.

Замечание 2. Очевидно, что при замене начальных условий (34) на

$$\tilde{\mu}_2(0; \varepsilon) = 0, \quad \tilde{\mu}'_2(0; \varepsilon) = 0,$$

решение уравнения (33) с этими условиями не будет приближенным решением обратной задачи.

Замечание 3. Используя более высокие члены разложения задачи (1)–(4) по малому параметру, можно получить приближенное решение с более высокой по порядку оценкой погрешности. Однако интегральное уравнение для приближенного решения будет очень громоздким.

Замечание 4. В случае, когда $K(t, \tau) = 0$, т.е. уравнение (1) представляет собой обычное уравнение теплопроводности, построенные приближенные решения определяются простыми явными формулами.

СПИСОК ЛИТЕРАТУРЫ

1. Латтес Р., Лионс Ж.-Л. Метод квазиобращения и его приложения. М.: Мир. 1970.
2. Иванов В.К. Задача квазиобращения для уравнения теплопроводности в равномерной метрике // Дифференц. ур-ния. 1972. Т. 8. № 4. С. 652–658.
3. Самарский А.А., Вабичевич П.Н. Численные методы решения обратных задач математической физики. Едиториал УРСС. Москва. 2004.
4. Короткий А.И., Цепелев И.А., Исмаил-заде А.Е. Численное моделирование обратных ретроспективных задач тепловой конвекции с приложениями к задачам геодинамики // Известия уральского университета. 2008. № 58. С. 78–87.
5. Табаринцева Е.В., Менихес Л.Д., Дрозин А.Д. О решении граничной обратной задачи методом квазиобращения // Вестник ЮУГУ, серия Математика. Механика, Физика. 2012. Вып. 6. С. 8–13.
6. Денисов А.М., Соловьева С.И. Численное решение обратных задач для гиперболического уравнения с малым параметром при старшей производной // Дифференц. ур-ния. 2018. Т. 54. № 7. С. 919–928.
7. Денисов А.М. Асимптотика решений обратных задач для гиперболических уравнений с малым параметром при старшей производной // Ж. вычисл. матем. и матем. физ. 2013. Т. 53. № 5. С. 744–752.
8. Belov Yury Ya., Kopylova Vera G. Determination of source function in composite type system of equations // Журн. СФУ. Сер. Матем. и физ. 2014. Т. 7. Вып. 3. С. 275–288.
9. Lukyanenko D.V., Shishlenin M.A., Volkov V.T. Asymptotic analysis of solving an inverse boundary value problem for a nonlinear singularly perturbed time-periodic reaction-diffusion-advection equation // J. Inverse and Ill posed Problems. 2019. V. 27. № 5. P. 745–758.
10. Lukyanenko D.V., Borzunov A.A., Shishlenin M.A. Solving coefficient inverse problems for a nonlinear singularly perturbed equations of the reaction-diffusion-advection type with data on the position of reaction front // Communication in Nonlinear Science Numerical Simulation. 2021. V. 99. 105824.
11. Денисов А.М. Приближенное решение обратных задач для уравнения теплопроводности с сингулярным возмущением // Ж. вычисл. матем. и матем. физ. 2021. Т. 61. № 12. С. 2040–2049.

УСВОЕНИЕ ДАННЫХ ДЛЯ ДВУМЕРНОГО УРАВНЕНИЯ АМБИПОЛЯРНОЙ ДИФФУЗИИ В МОДЕЛИ ИОНОСФЕРЫ ЗЕМЛИ¹⁾

© 2023 г. В. П. Дымников^{1,2}, Д. В. Кулямин^{1,2}, П. А. Останин^{1,3,*}, В. П. Шутяев^{1,3}

¹ 119333 Москва, ул. Губкина, 8, Институт вычислительной математики им. Г.И. Марчука РАН, Россия

² 129128 Москва, ул. Ростокинская, 9, Институт прикладной геофизики им. Е.К. Федорова, Россия

³ 141701 Долгопрудный, М.о., Институтский пер., 9, Московский физико-технический институт
(национальный исследовательский университет), Россия

*e-mail: ostanin.pavel@phystech.edu

Поступила в редакцию 11.08.2022 г.

Переработанный вариант 05.10.2022 г.

Принята к публикации 02.02.2023 г.

В работе рассматривается задача вариационного усвоения данных наблюдений для двумерной диффузионной модели F слоя ионосферы Земли ИВМ РАН. В качестве наблюдений рассматривались данные о полном интегральном электронном содержании вдоль заданных траекторий. Сформулирована общая постановка задачи в дифференциальной форме, исследована ее разрешимость. На основе регуляризованной постановки построен итерационный алгоритм решения задачи ассимиляции, показана его сходимость. Построена конечномерная аппроксимация и реализовано численное решение задачи, доказаны устойчивость и сходимость разностной схемы. На основе контрольных численных экспериментов исследовано качество восстановления полей распределения электронной концентрации, показано приемлемое восстановление слабозмущенного решения как для стационарной, так и для эволюционной постановок при вертикальных и наклонных траекториях интегрирования. Библ. 38. Фиг. 8.

Ключевые слова: ионосфера, амбиполярная диффузия, обратные задачи, вариационное усвоение данных, численное моделирование.

DOI: 10.31857/S0044466923050101, EDN: GFURSG

1. ВВЕДЕНИЕ

Известно, что знание основных характеристик земной ионосферы имеет существенное практическое значение, поскольку они определяют условия для распространения всех типов радиосигналов в околоземной среде, обеспечивающих бесперебойную работу систем дальней глобальной радиосвязи, спутниковых систем связи, радиолокации, а также навигационных систем глобального спутникового позиционирования [1]–[3]. Таким образом, по мере значительного роста использования систем навигации и спутниковой радиосвязи в социальной, экономической и оборонной сферах в последние десятилетия все острее признается потребность в надежной и точной системе анализа и прогнозирования состояния ионосферы и его изменчивости. Недостаточный на сегодняшний день уровень точности описания физических процессов и высокой изменчивости в верхней атмосфере и ионосфере Земли представляет собой фундаментальную сложность решения данной задачи. Результаты последних исследований показывают, что многие механизмы формирования изменчивости остаются совсем не изученными [4], [5], а для их детального описания, по-видимому, требуется комплексный подход, учитывающий внешние воздействия космической природы (солнечная активность и др.), электромагнитные и динамические процессы во всей атмосфере, а также нелинейное взаимодействие термосферы и ионосферы [6]–[8].

Поскольку в последние десятилетия произошло существенное увеличение объема и доступности глобальных данных наблюдений о параметрах ионосферы различного типа, это сильно стимулировало различные научные группы перейти к созданию систем усвоения данных для

¹⁾Работа выполнена при финансовой поддержке Минобрнауки РФ (соглашение № 075-15-2022-286), а также РФФИ (проект 20-11-20057, исследования в разд. 3).

ионосферы, а также к новому направлению по включению верхних слоев атмосферы и ионосферы в разрабатываемые модели Земной системы. Для решения таких задач в Институте вычислительной математики им. Г.И. Марчука (ИВМ РАН) [9] в течение последних лет уже разработаны новые версии глобальной модели общей циркуляции атмосферы (0–130 км) с включением нижних слоев термосферы и ионосферы [10–12], а также новые модели термосферы (90–500 км) [13] и F слоя ионосферы [14]. На их основе создана новая глобальная динамическая совместная модель термосферы и ионосферы [15].

Предлагаемая работа является первой из серии работ, посвященных обозначенной в начале задачи разработки перспективной прогностической системы четырехмерного вариационного усвоения данных для модели ионосферы Земли ИВМ РАН для решения проблемы прогноза глобального состояния ионосферы. Как известно, системы усвоения данных позволяют подготовить согласованные трехмерные распределения характеристик среды и начальное состояние, необходимое при решении задачи краткосрочного прогноза ее состояния. В силу небольшого времени предсказуемости первого рода для ионосферной динамики в F слое времена краткосрочного прогноза лежат в рамках суточного цикла (практическую значимость имеют времена от 15 мин до 4–6 ч). Успешное решение задачи ассимиляции данных для описания глобального состояния ионосферы стало возможным благодаря появлению и расширению ряда систем непрерывного мониторинга ее характеристик, отвечающих требованиям достаточной точности и глобальности покрытия, оперативности и доступности. В целом можно выделить два основных класса данных наблюдений ионосферы, отвечающих приведенным требованиям:

- Данные о полном содержании электронов (ПЭС/ТЕС) вдоль наклонных трасс связи наземных станций и спутников в фиксированные моменты времени (интеграл электронной плотности вдоль трассы радиосигнала станция–спутник), полученные с помощью обработки радиосигналов глобальных навигационных спутниковых систем (ГНСС/GNSS) или других спутниковых систем [16] (в первую очередь это данные систем GPS [16] и ГЛОНАСС [18]). Обе системы представляют собой большую группу космических аппаратов, обращающихся на среднем расстоянии от поверхности Земли 20000 км, наземный сегмент содержит широкую сеть стационарных приемников, расположенных в основном на суше. Станции принимают спутниковый сигнал и публикуют протоколы обмена данными (с частотой обновления 30 с), по которым можно определить ПЭС вдоль траектории сигнала с точностью до некоторой неопределенной аддитивной постоянной. Для исключения этой постоянной используются отдельные методы [19]–[21], которые в данной работе обсуждаться не будут. Таким образом, в общем доступе имеется оперативная непрерывная информация о наклонных значениях ПЭС с рядом ограничений, однако с широким покрытием и высокой частотой получения, которая с использованием дополнительных методов обработки и интерполяции может служить источником информации о глобальном состоянии ионосферы.

- Данные о профилях электронной плотности, главным образом полученные ионозондами, работающими в реальном времени и количество которых в мире непрерывно увеличивается [22], [23]. С помощью ионозондов на основе отражения радиосигнала удается в реальном времени (с частотой порядка 15 минут) восстанавливать вертикальный профиль электронной плотности вплоть до высоты ее максимума в месте расположения прибора, однако их возможности ограничены низкой автоматизацией обработки данных, неточностью при низких значениях концентрации в нижних слоях ионосферы.

В разрабатываемой в ИВМ РАН системе усвоения данных предполагается использование двух описанных видов данных непрерывного мониторинга ионосферы.

Разработка систем ассимиляции данных для ионосферы была начата за рубежом в конце 1990-х годов в США в рамках программы MURI с разработки двух разных ассимиляционных численных моделей ионосферы – Global Assimilation of Ionospheric Measurements с участием групп из Университета штата Юта (USU GAIM) [24] и Университета Южной Калифорнии (USC GAIM) [25], использующих методы фильтров Калмана для усвоения данных в физические модели ионосферы. В дальнейшем стали использовать ансамблевый подход и более развитые методы. В настоящее время версия модели USU GAIM на основе ансамблевого фильтра Гаусса-Маркова используется метеорологическим агентством ВВС США (AFWA). Несколько оперативных центров космической погоды внедрили методы усвоения данных для спецификации и прогноза характеристик ионосферы на основе ассимиляции информации в эмпирические модели: ряд центров использует сервис построения оперативных карт ТЕС Австралийской службы про-

гнозов ионосферы (IPS)², другими аналогичными примерами являются системы MIDAS-GPS [26] и EDAM-GPS [27]. Существует также ряд региональных систем усвоения данных, в основном созданных для объективного анализа и мониторинга ПЭС (к примеру US-TEC и др. [28]). В России попытка разработки ансамблевой системы усвоения данных для ионосферы на основе фильтра Калмана и американской модели STIP была предпринята в Центральной Аэрологической Обсерватории (ЦАО) Росгидромета [29].

Особенностью подхода к созданию системы усвоения данных для модели ионосферы, излагаемой в данной работе, являются использование четырехмерной вариационной ассимиляции данных наблюдений на основе оригинальной глобальной динамической модели F слоя ионосферы [14], [30] и ее детальное математическое исследование. В качестве первой версии системы усвоения рассматривается задача ассимиляции данных на основе системы уравнений амбиполярной диффузии в квазидвумерной постановке [14]. Подобная приближенная модель достаточно хорошо описывает состояние ионосферы средних и низких широт в спокойных условиях [3], [13], [14]. Данная постановка задачи усвоения имеет смысл, поскольку в модели F слоя ионосферы ИВМ РАН для интегрирования по времени используется метод расщепления по физическим процессам [15], позволяющий включить усвоение на одном из этапов расщепления в рамках реализации всей модели. Построение системы усвоения данных на основе такого подхода для полной трехмерной модели ионосферы является предметом следующих работ. Отметим, что в простейшей версии трехмерной модели, не включающей полного описания трехмерного переноса, можно использовать предложенный в данной статье подход практически без изменений, поскольку рассматриваемые уравнения записываются одинаково на всех долготах, а зависимость от долготы и времени входит в коэффициенты уравнений, их правые части, а также в данные наблюдений. В перспективе подобный подход позволит построить системы усвоения данных для совместных моделей термосферы – ионосферы.

Основная цель данной работы – разработка и реализация первой версии системы вариационного усвоения данных для диффузионной модели ионосферы [14]. Для получающейся при этом системы уравнений требуется построить ее аппроксимацию, исследовать разрешимость и сходимость, предложить алгоритм реализации, а также провести контрольные численные эксперименты для исследования точности предложенного алгоритма. В качестве данных наблюдений в работе рассматриваются данные о ПЭС.

Работа состоит из трех частей. Во втором разделе рассмотрена постановка задачи усвоения для простейшей двумерной модели ионосферы, в которой в качестве основного уравнения выбрано двумерное уравнение амбиполярной диффузии. Для этой модели сформулирована обобщенная постановка для задачи ассимиляции данных наблюдений в виде интегралов электронной концентрации по ряду независимых прямолинейных траекторий (ПЭС). Рассмотрены вопросы корректности и разрешимости данной задачи, сформулирован алгоритм ее дискретизации. В разд. 4 рассмотрены особенности численной реализации алгоритма для решения последовательности систем прямых и сопряженных уравнений (разностные схемы для диффузионного оператора прямой и сопряженной задач, сохраняющие направление магнитных силовых линий, способ аппроксимации оператора наблюдения), а также исследована сходимость разностных постановок. В разд. 5 приведены результаты контрольных численных экспериментов, оценивающих точность и эффективность построенного алгоритма и его реализации. Наконец, в заключении сформулированы основные результаты работы.

2. ПОСТАНОВКА ЗАДАЧИ

В качестве основной постановки модели ионосферы в данной работе рассмотрим двумерную модель динамики F слоя, учитывающую процессы ионизации, плазмохимии, а также амбиполярной диффузии и гравитационного оседания [14], [30].

Пусть z – высота, φ – широта, t – время, $\Omega = \{(z, \varphi) | z_b \leq z \leq z_t; -\pi/2 \leq \varphi \leq \pi/2\} \subset \mathbb{R}^2$ – сферический слой, $Q_T = (0, T) \times \Omega$. Рассмотрим уравнение амбиполярной диффузии в приближении тонкого сферического слоя:

² <http://www.ips.gov.au/Satellite/2/1>

$$\frac{\partial n}{\partial t} = \frac{1}{a^2 \cos \varphi} \frac{\partial}{\partial \varphi} \left[D \cos^2 I \frac{\partial n}{\partial \varphi} \cos \varphi \right] + \frac{\partial}{\partial z} \left[D \sin^2 I \frac{\partial n}{\partial z} \right] - \frac{1}{a \cos \varphi} \frac{\partial}{\partial \varphi} \left[D \sin I \cos I \frac{\partial n}{\partial z} \cos \varphi \right] -$$

$$- \frac{1}{a} \frac{\partial}{\partial z} \left[D \sin I \cos I \frac{\partial n}{\partial \varphi} \right] - \frac{1}{a \cos \varphi} \frac{\partial}{\partial \varphi} \left[\frac{D}{H} \sin I \cos I n_i \cos \varphi \right] + \frac{\partial}{\partial z} \left[\frac{D}{H} \sin^2 I n_i \right] + P - kn,$$

$$n|_{t=0} = n_0,$$

$$\left(D \sin^2 I \frac{\partial n}{\partial z} - \frac{1}{a} D \sin I \cos I \frac{\partial n}{\partial \varphi} + \frac{D}{H} \sin^2 I n \right) \Big|_{z=z_b} = F_{lb},$$

$$\left(D \sin^2 I \frac{\partial n}{\partial z} - \frac{1}{a} D \sin I \cos I \frac{\partial n}{\partial \varphi} + \frac{D}{H} \sin^2 I n \right) \Big|_{z=z_t} = F_{ub}.$$

Здесь $n(z, \varphi, t)$ – электронная концентрация, $I = (2\varphi)$ – угол магнитного склонения, $P(z, \varphi, t)$ – функция фотоионизации, $k(z, \varphi, t)$ – функция рекомбинации, $F_{ub}(\varphi, t)$ и $F_{lb}(\varphi, t)$ – потоки на верхней и нижней границах расчетной области, которые в данной работе мы будем полагать равными нулю. На полюсах (при $\varphi = \pm \frac{\pi}{2}$) эффективные коэффициенты диффузии равны нулю, и в задаче

ставится условие ограниченности производных $\frac{\partial n}{\partial \varphi}$ по широте.

Будем считать правую часть P равной $P_0 + U$, где P_0 – известная функция фотоионизации, а U – добавочное слагаемое, рассматриваемое в данной постановке задачи в качестве управления. Считаем также, что заданы данные наблюдений – конечный набор интегралов от решения вдоль заданных прямолинейных траекторий ℓ_k , соединяющих спутник с точкой на поверхности Земли (ТЕС – total electron content или ПЭС – полное электронное содержание). Для простоты будем вести интегрирование по тонким прямоугольникам Ω_k (аппроксимирующим прямые ℓ_k , вдоль которых рассчитывается сигнал для вычисления ТЕС-ов):

$$\int_{\Omega_k} n(z, \varphi, t) a \cos \varphi d\varphi dz = Tec_k(t) \quad \forall t \in (0, T), \quad k = 1, \dots, N. \quad (1)$$

Значения интегралов с течением времени меняются, поскольку функция $n(z, \varphi, t)$ зависит от времени, поэтому $Tec_k(t)$ также зависят от времени. Эти зависимости мы полагаем известными.

Рассмотрим задачу восстановления профилей по наблюдаемым ПЭС следующего вида: найти функцию $n(z, \varphi, t)$ и управление $U(z, \varphi, t)$, определенные в Q_T , такие, что n является решением уравнения амбиполярной диффузии с известным начальным условием и удовлетворяет данным наблюдений (1): известны интегралы по подобластям Ω_k и их зависимости от времени (ТЕС-и) (этих подобластей всего N).

Введем обозначения $K_1^2 = D \sin^2 I$, $K_2^2 = D \cos^2 I \cos^2 \varphi$, $K = \begin{pmatrix} K_1^2 & -K_1 K_2 \\ -K_1 K_2 & K_2^2 \end{pmatrix}$ – симметричная матрица коэффициентов диффузии, а также оператор $\frac{\partial}{\partial y} = \frac{1}{a \cos \varphi} \frac{\partial}{\partial \varphi}$ и градиент $\nabla = \begin{pmatrix} \partial/\partial z \\ \partial/\partial y \end{pmatrix}$.

Обозначим также вертикальные и горизонтальные скорости $u = \frac{D}{H} \sin^2 I$, $v = \frac{D}{H} \sin I \cos I$. Введем также для краткости обозначение $d\Omega = a \cos \varphi d\varphi dz$ – элемент площади. Тогда выписанную систему можно записать более кратко:

$$\left(\frac{\partial}{\partial t} - \nabla(K\nabla(\cdot)) - \frac{\partial}{\partial z}(u \cdot (\cdot)) - \frac{\partial}{\partial y}(v \cos \varphi \cdot (\cdot)) + k \right) n = P_0 + U,$$

$$n|_{t=0} = n_0, \quad \left(K_1^2 \frac{\partial n}{\partial z} - K_1 K_2 \frac{\partial n}{\partial y} - un \right) \Big|_{z=z_b} = 0, \quad \left(K_1^2 \frac{\partial n}{\partial z} - K_1 K_2 \frac{\partial n}{\partial y} - un \right) \Big|_{z=z_t} = 0,$$

$$\int_{\Omega_k} n(z, \varphi, t) d\Omega = Tec_k(t) \quad \forall t \in (0, T), \quad k = 1, \dots, N.$$

Уравнение амбиполярной диффузии-плазмохимии для модели ионосферы подробно рассмотрено в работе [14]. Отметим ряд его ключевых особенностей.

1. Величина коэффициента диффузии D меняется на 6–7 порядков при изменении высоты z от 100 до 500 км.

2. Уравнение модели фактически отражает баланс массы для заряженных частиц, и поэтому для решения этого уравнения необходимо использование консервативных схем.

3. Диффузия, описываемая этим уравнением, идет вдоль магнитных силовых линий. Это утверждение эквивалентно вырожденности в каждой точке тензора “коэффициентов диффузии”

$$\begin{pmatrix} K_1^2 & -K_1K_2 \\ -K_1K_2 & K_2^2 \end{pmatrix}.$$

Умножив уравнение скалярно на n и проинтегрировав по всей расчетной области, можно получить (при однородных краевых условиях и нулевых функциях P и k) интегральное соотношение, отражающее геометрические характеристики диффузионных процессов (под знаком двойного интеграла в правой части стоит квадрат производной электронной концентрации по направлению магнитной силовой линии):

$$\frac{1}{2} \frac{\partial}{\partial t} \iint_{\varphi, z} n^2 \cos \varphi d\varphi dz = - \iint_{\varphi, z} \left(K_1 \frac{\partial n}{\partial z} - \frac{K_2}{a \cos \varphi} \frac{\partial n}{\partial \varphi} \right)^2 \cos \varphi d\varphi dz \leq 0. \quad (2)$$

Корректная аппроксимация диффузионного оператора, обладающая разностным аналогом этого соотношения (и потому абсолютно устойчивая при подходящей аппроксимации по времени), уже изучена в работе [31]. Эта аппроксимация в данной работе будет использована при численном решении системы прямых и сопряженных уравнений.

3. ОБОБЩЕННАЯ ПОСТАНОВКА

Будем рассматривать функции $P_0, U \in \mathbb{L}_2(Q_T)$, $Tec_k(t) \in \mathbb{L}_2(0, T)$, а решения n из подпространства W в пространстве $\mathbb{L}_2(0, T; \mathbb{W}_2^1(\Omega))$, где W состоит из таких функций, что $\frac{\partial}{\partial t} n \in \mathbb{L}_2(0, T; (\mathbb{W}_2^1(\Omega))^*)$. Производная $\frac{\partial}{\partial t} n$ понимается в обобщенном смысле: для всякой $n \in \mathbb{L}_2(0, T; \mathbb{W}_2^1(\Omega))$ ее производная по времени есть элемент $D'((0, T); \mathbb{W}_2^1(\Omega))$, т.е. непрерывный оператор из пространства пробных (бесконечно гладких финитных) $\varphi \in D(0, T)$ в пространство $\mathbb{W}_2^1(\Omega)$. Его действие задается формулой $\left(\frac{\partial n}{\partial t}\right)(\varphi) = - \left\langle n, \left(\frac{\partial \varphi}{\partial t}\right) \right\rangle_{\mathbb{L}_2(0, T)}$. Ограничение $\frac{\partial}{\partial t} n \in \mathbb{L}_2(0, T; (\mathbb{W}_2^1(\Omega))^*)$ означает, что требуется, чтобы производная $\frac{\partial}{\partial t} n$ являлась функционалом из $(\mathbb{W}_2^1(\Omega))^*$ при почти всех t , и нормы этих функционалов (зависящие от параметра t) были интегрируемы с квадратом на $[0, T]$. Известно (см. [32], гл. 3, теорема 1.1), что все функции из W являются непрерывными как отображения $[0, T] \rightarrow \mathbb{L}_2(\Omega)$ при должном изменении на множестве меры нуль в отрезке $[0, T]$ (гарантируется, что такое изменение возможно).

Умножим скалярно первое уравнение на пробную функцию ψ из класса $\mathbb{W}_2^1(\Omega)$ и проинтегрируем по Ω . С помощью интегрирования по частям получим:

$$\begin{aligned} & \int_{\Omega} \frac{\partial n}{\partial t} \psi d\Omega + \int_{\Omega} \left(K_1^2 \frac{\partial n}{\partial z} \frac{\partial \psi}{\partial z} + K_2^2 \frac{1}{a \cos \varphi} \frac{\partial n}{\partial \varphi} \frac{1}{a \cos \varphi} \frac{\partial \psi}{\partial \varphi} - K_1 K_2 \frac{1}{a \cos \varphi} \frac{\partial n}{\partial \varphi} \frac{\partial \psi}{\partial z} - \right. \\ & \left. - K_1 K_2 \frac{1}{a \cos \varphi} \frac{\partial n}{\partial z} \frac{\partial \psi}{\partial \varphi} \right) d\Omega + \int_{\Omega} u n \frac{\partial \psi}{\partial z} d\Omega + \int_{\Omega} v n \frac{1}{a} \frac{\partial \psi}{\partial \varphi} d\Omega - \int_{\Omega} (P - kn) \psi d\Omega - \end{aligned}$$

$$\begin{aligned}
 & - \int_{-\pi/2}^{+\pi/2} \left(K_1^2 \frac{\partial n}{\partial z} - K_1 K_2 \frac{\partial n}{\partial y} - un \right) \Big|_{z=z_t} \psi(z_t, \varphi, t) a \cos \varphi d\varphi + \\
 & + \int_{-\pi/2}^{+\pi/2} \left(K_1^2 \frac{\partial n}{\partial z} - K_1 K_2 \frac{\partial n}{\partial y} - un \right) \Big|_{z=z_b} \psi(z_b, \varphi, t) a \cos \varphi d\varphi = 0.
 \end{aligned}$$

Последнее равенство, записанное для произвольной функции $\psi \in \mathbb{W}_2^1(\Omega)$, будет обобщенной постановкой исходной задачи, если подставить 0 вместо потоков

$$\left(K_1^2 \frac{\partial n}{\partial z} - K_1 K_2 \frac{\partial n}{\partial y} - un \right)$$

при $z = z_t$ и $z = z_b$: это исключит слагаемые с интегралами по широте на нижней и верхней границах.

Обозначим

$$\begin{aligned}
 \text{Diff}(n, \psi) = & \int_{\Omega} \left(K_1^2 \frac{\partial n}{\partial z} \frac{\partial \psi}{\partial z} + K_2^2 \frac{1}{a \cos \varphi} \frac{\partial n}{\partial \varphi} \frac{1}{a \cos \varphi} \frac{\partial \psi}{\partial \varphi} - \right. \\
 & \left. - K_1 K_2 \frac{1}{a \cos \varphi} \frac{\partial n}{\partial \varphi} \frac{\partial \psi}{\partial z} - K_1 K_2 \frac{1}{a \cos \varphi} \frac{\partial n}{\partial z} \frac{\partial \psi}{\partial \varphi} \right) d\Omega.
 \end{aligned}$$

Введем следующие билинейные формы и операторы:

$$\bullet \mathcal{B}(n, \psi) = \text{Diff}(n, \psi) + \left\langle un, \frac{\partial \psi}{\partial z} \right\rangle_{\mathbb{L}_2(\Omega)} + \left\langle \frac{1}{a} vn, \frac{\partial \psi}{\partial \varphi} \right\rangle_{\mathbb{L}_2(\Omega)} + \langle kn, \psi \rangle_{\mathbb{L}_2(\Omega)}.$$

Определим оператор B равенством $\mathcal{B}(n, \psi) = \langle Bn, \psi \rangle$, т.е. Bn – функционал, действующий на пробную функцию ψ . Оператор B действует из пространства $\mathbb{L}_2(0, T; (\mathbb{W}_2^1(\Omega)))$ в $\mathbb{L}_2(0, T; (\mathbb{W}_2^1(\Omega))^*)$, его область определения есть $D(B) = W$;

Наконец, определим оператор $L = \frac{\partial}{\partial t} + B : \mathbb{L}_2(0, T; (\mathbb{W}_2^1(\Omega))) \rightarrow \mathbb{L}_2(0, T; (\mathbb{W}_2^1(\Omega))^*)$, $D(L) = W$.

Отметим, что повторное интегрирование по частям в обобщенной постановке приводит к соотношениям, которые впоследствии будут использованы в качестве краевых условий сопряженной задачи (они имеют такой же вид, как краевые условия для n в прямой задаче, но без слагаемого $(-un)$):

$$K_1^2 \frac{\partial \psi}{\partial z} - K_1 K_2 \frac{\partial \psi}{\partial y} = K_1^2 \frac{\partial \psi}{\partial z} - \frac{K_1 K_2}{a \cos \varphi} \frac{\partial \psi}{\partial \varphi} = 0.$$

Физический смысл этого краевого условия тот же: поток вдоль магнитной силовой линии равен нулю.

Сопряженный оператор

$$L^* = -\frac{\partial}{\partial t} - \nabla(K\nabla(\cdot)) + u \frac{\partial}{\partial z} + \frac{1}{a} v \frac{\partial}{\partial \varphi} + k$$

определен на подпространстве

$$D(L^*) \subset \left\{ q \mid q|_{t=T} = 0, K_1^2 \frac{\partial \psi}{\partial z} - K_1 K_2 \frac{\partial \psi}{\partial y} \Big|_{z=z_b, z=z_t} = 0 \right\}.$$

• Оператор $C_{\text{tec}} : \mathbb{L}_2(Q_T) \rightarrow (\mathbb{L}_2(0, T))^N$, действие которого определяется формулой $C_{\text{tec}}n = \mathbf{Tec}(t)$, где $\mathbf{Tec}(t)$ – вектор-функция, компоненты которой есть $Tec_k(t)$, $k = 1, N$. Из билинейной формы

$$\langle C_{\text{tec}}n, \mathbf{f} \rangle_{(\mathbb{L}_2(0, T))^N} = \int_0^T \begin{pmatrix} \int_{\Omega_1} nd\Omega \\ \dots \\ \int_{\Omega_N} nd\Omega \end{pmatrix} \begin{pmatrix} f_1 \\ \dots \\ f_N \end{pmatrix} dt = \sum_{k=1}^N \int_0^T \int_{\Omega_k} nf_k d\Omega dt = \sum_{k=1}^N \int_{Q_T} n(z, \varphi, t) f_k(t) \chi_{\Omega_k}(z, \varphi) d\Omega_T$$

легко видеть, что действие оператора $C_{\text{tec}}^* : (\mathbb{L}_2(0, T))^N \rightarrow \mathbb{L}_2(Q_T)$ определяется формулой:

$$C_{\text{tec}}^* \mathbf{f} = \sum_{k=1}^N f_k(t) \chi_{\Omega_k}(z, \varphi),$$

где $\chi_{\Omega_k}(z, \varphi)$ есть характеристическая функция подобласти Ω_k .

Мы будем рассматривать действие C_{tec} лишь на функции из пространства W , а образы элементов $(\mathbb{L}_2(0, T))^N$ под действием C_{tec}^* рассматривать в том числе как элементы W^* .

Исходную задачу теперь можно сформулировать в виде

$$\langle Ln, \Psi \rangle = \langle P_0 + U, \Psi \rangle_{\mathbb{L}_2(\Omega)} \quad \forall \Psi \in W_2^1(\Omega) \quad \text{для п.в. } t \in [0, T]$$

$$C_{\text{tec}}n = \mathbf{Tec}(t).$$

Отметим, что функция потока F_{ub} на верхней границе расчетной области по существу неизвестна и была положена равной нулю. Одним из путей модификации предложенной модели может быть рассмотрение вектор-функции управления $\mathbf{V} = \begin{pmatrix} U \\ F_{ub} \end{pmatrix} \in \mathbb{L}_2(Q_T) \times \mathbb{L}_2([0, T] \times \Gamma)$ с тем, чтобы использовать данные наблюдений для восстановления не только функции U , но и функции F_{ub} .

3.1. Плотная разрешимость задачи, регуляризация и итерационный процесс

Рассмотрим задачу в операторной форме:

$$Ln = P_0 + U; \quad C_{\text{tec}}n = \mathbf{Tec}(t). \tag{3}$$

Если исключить из системы (3) переменную n , оставив только управление U , то останется задача вида $\mathcal{A}U = f$, где $\mathcal{A} = CL^{-1}$, $f = \mathbf{Tec} - CL^{-1}P_0$ (здесь и далее $C = C_{\text{tec}}$ – оператор наблюдений).

Утверждение 1. *Имеет место плотная разрешимость задачи $\mathcal{A}U = f$.*

Доказательство. Плотная разрешимость (плотность $\text{Im } \mathcal{A}$ в пространстве, куда действует \mathcal{A}) эквивалентна в силу теоремы Фредгольма $\text{Ker } \mathcal{A}^* = 0$.

Покажем, что уравнение $\mathcal{A}^*w = 0$ имеет только тривиальное решение. Будем при этом считать, что функции скоростей u и v , входящие в оператор L , таковы, что L имеет ограниченный обратный L^{-1} (при $u = v = 0$ это очевидно, поскольку функция рекомбинации в правой части прямого уравнения $k \geq 0$, и поэтому оператор L ограничен снизу). Тогда $\mathcal{A}^* = (L^{-1})^*C^*$, и уравнение $\mathcal{A}^*w = (L^{-1})^*C^*w = 0$ после обозначения $q = (L^{-1})^*C^*w$ принимает форму (краевые условия не указаны)

$$L^*q = C^*w, \quad q = 0,$$

или (в развернутой форме)

$$\left(-\frac{\partial}{\partial t} - \nabla(K\nabla(\cdot)) + u\frac{\partial}{\partial z} + \frac{1}{a}v\frac{\partial}{\partial \varphi} + k\right)q = \sum_{k=1}^N \chi_{\Omega_k}(z, \varphi)w_k(t),$$

$$q|_{\Omega} = 0.$$

Из $q|_{\Omega} = 0$ сразу следует, что $\sum_{k=1}^N \chi_{\Omega_k}(z, \varphi)w_k(t) = 0$. Будем считать, что Ω_k попарно различны. Тогда все функции $w_k(t)$ тождественно равны нулю, что и требовалось.

Замечание 1. Задача $\mathcal{A}U = f$ не является однозначно разрешимой.

В самом деле, однозначная разрешимость задачи эквивалентна $\text{Ker } \mathcal{A} = 0$. Несложно заметить, что пространство $\text{Ker } \mathcal{A}$ отлично от нулевого, поэтому решение задачи не является единственным.

Включим рассматриваемую задачу (3) в семейство задач оптимального управления вида

$$Ln_{\alpha} = P_{\alpha}; \quad J_{\alpha}(P_{\alpha}, n(P_{\alpha})) = \inf J_{\alpha}(U, n(U)), \quad (4)$$

где $P_{\alpha} = P_0 + U$, U – управление, а

$$J_{\alpha}(U, n(U)) = \alpha \|U\|_{\mathbb{L}_2(Q_T)}^2 + \|Cn(U) - \mathbf{Tec}\|_{\mathbb{L}_2(0,T)}^2 = \alpha \int_0^T \int_Q U^2 d\Omega dt + \int_0^T \sum_{k=1}^N \left(\int_{\Omega_k} n d\Omega - \mathbf{Tec}_k(t) \right)^2 dt.$$

Тем самым рассматривается регуляризация по Тихонову $\alpha U + \mathcal{A}^* \mathcal{A}U = \mathcal{A}^* f$ для исходной задачи $\mathcal{A}U = f$. При $\alpha > 0$ регуляризованная задача корректно разрешима, поскольку оператор $\alpha I + \mathcal{A}^* \mathcal{A}$ самосопряженный, а его спектр отделен от нуля.

Если же $\alpha = 0$, то такая задача оптимального управления эквивалентна задаче $\mathcal{A}^* \mathcal{A}U = \mathcal{A}^* f$ (симметризации Гаусса задачи $\mathcal{A}U = f$). Известно, что симметризация Гаусса для уравнения $\mathcal{A}U = f$ приводит к уравнению, равносильному уравнению $\mathcal{A}U = Vf$, где V – ортопроектор на $\overline{\text{Im } \mathcal{A}}$. Таким образом, переход от уравнения $\mathcal{A}U = f$ и последующее рассмотрение задач (4) связаны с последовательностью корректно разрешимых (при $\alpha > 0$) задач, не эквивалентных, вообще говоря, исходной задаче (3). При этом в предположении о компактности оператора \mathcal{A} известно [34], что при $\alpha \rightarrow +0$ решение регуляризованной системы U сходится к решению задачи при $\alpha = 0$, имеющему минимальную норму, т.е. к псевдорешению симметризованного уравнения $\mathcal{A}^* \mathcal{A}U = \mathcal{A}^* f$ (напомним, что $\text{Ker } \mathcal{A}$ нетривиально). Это значит, что при достаточно малых α решение регуляризованной задачи оптимального управления может рассматриваться как приближенное решение симметризованного уравнения. Запишем систему вариационных уравнений, выражающую условия оптимальности функционала J_{α} :

$$Ln_{\alpha} = P_{\alpha}; \quad L^*q_{\alpha} = C^*(Cn_{\alpha} - \mathbf{Tec}); \quad \alpha P_{\alpha} + q_{\alpha} = 0.$$

В “классической форме записи” эти уравнения имеют следующий вид:

$$\left(\frac{\partial}{\partial t} - \nabla(K\nabla(\cdot)) - \frac{\partial}{\partial z}(u \cdot (\cdot)) - \frac{\partial}{\partial y}(v \cos \varphi \cdot (\cdot)) + k\right)n_{\alpha} = P_{\alpha},$$

$$n_{\alpha}|_{t=0} = n_0,$$

$$\left(K_1^2 \frac{\partial n_{\alpha}}{\partial z} - K_1 K_2 \frac{\partial n_{\alpha}}{\partial y} - un_{\alpha}\right)\Big|_{z=z_b} = 0, \quad \left(K_1^2 \frac{\partial n_{\alpha}}{\partial z} - K_1 K_2 \frac{\partial n_{\alpha}}{\partial y} - un_{\alpha}\right)\Big|_{z=z_t} = 0;$$

$$\left(-\frac{\partial}{\partial t} - \nabla(K\nabla(\cdot)) + u\frac{\partial}{\partial z} + \frac{1}{a}v\frac{\partial}{\partial \varphi} + k\right)q_{\alpha} = \sum_{k=1}^N \chi_{\Omega_k}(z, \varphi) \left(\int_{\Omega_k} n_{\alpha} d\Omega - \mathbf{Tec}_k(t) \right),$$

$$q_{\alpha}|_{t=T} = 0,$$

$$\left(K_1^2 \frac{\partial q_{\alpha}}{\partial z} - K_1 K_2 \frac{\partial q_{\alpha}}{\partial y}\right)\Big|_{z=z_b} = 0, \quad \left(K_1^2 \frac{\partial q_{\alpha}}{\partial z} - K_1 K_2 \frac{\partial q_{\alpha}}{\partial y}\right)\Big|_{z=z_t} = 0;$$

$$\alpha P_{\alpha} + q_{\alpha} = 0.$$

Система прямых и сопряженных уравнений решается итерационно: рассматривается последовательность n_α^m, q_α^m и P_α^m , удовлетворяющая системе

$$\begin{aligned} & \left(\frac{\partial}{\partial t} - \nabla(K\nabla(\cdot)) - \frac{\partial}{\partial z}(u \cdot (\cdot)) - \frac{\partial}{\partial y}(v \cos \varphi \cdot (\cdot)) + k \right) n_\alpha^m = P_\alpha^m, \\ & n_\alpha^m \Big|_{t=0} = n_0, \\ & \left(K_1^2 \frac{\partial n_\alpha^m}{\partial z} - K_1 K_2 \frac{\partial n_\alpha^m}{\partial y} - u n_\alpha^m \right) \Big|_{z=z_b} = 0, \quad \left(K_1^2 \frac{\partial n_\alpha^m}{\partial z} - K_1 K_2 \frac{\partial n_\alpha^m}{\partial y} - u n_\alpha^m \right) \Big|_{z=z_t} = 0; \\ & \left(-\frac{\partial}{\partial t} - \nabla(K\nabla(\cdot)) + u \frac{\partial}{\partial z} + \frac{1}{a} v \frac{\partial}{\partial \varphi} + k \right) q_\alpha^m = \sum_{k=1}^N \chi_{\Omega_k}(z, \varphi) \left(\int_{\Omega_k} n_\alpha^m d\Omega - Tец_k(t) \right), \\ & q_\alpha^m \Big|_{t=T} = 0, \\ & \left(K_1^2 \frac{\partial q_\alpha^m}{\partial z} - K_1 K_2 \frac{\partial q_\alpha^m}{\partial y} \right) \Big|_{z=z_b} = 0, \quad \left(K_1^2 \frac{\partial q_\alpha^m}{\partial z} - K_1 K_2 \frac{\partial q_\alpha^m}{\partial y} \right) \Big|_{z=z_t} = 0; \\ & P_\alpha^{m+1} = P_\alpha^m - \tau_m (\alpha P_\alpha^m + q_\alpha^m). \end{aligned}$$

Отметим, что итерационные процессы для решения систем оптимальности, связанных с задачей ассимиляции данных наблюдений, подробно рассматриваются в [34].

Параметр τ_m в текущей постановке выбран постоянным. Относительно сходимости итерационного процесса известно следующее [33]:

Утверждение 2. При положительном параметре регуляризации $\alpha > 0$ итерационный процесс сходится при $\tau_m \leq \tau_{opt} = \frac{2}{2\alpha + C_0 + C_1}$, где отрезок $[C_0, C_1]$ содержит спектр оператора $\mathcal{A}^* \mathcal{A}$.

В частности, метод сходится при $\tau = \frac{2}{2\alpha + \|\mathcal{A}\|^2}$.

4. РАЗНОСТНЫЕ АППРОКСИМАЦИИ ОПЕРАТОРОВ ЗАДАЧИ

Приведем описание используемой конечно-разностной аппроксимации системы вариационных уравнений, выражающих условие оптимальности сеточного функционала J_α^h :

$$L^h n_\alpha^h = P_\alpha^h, \quad (L^h)^* q_\alpha^h = (C^h)^* (C^h n_\alpha^h - Tец^h); \quad \alpha P_\alpha^h + q_\alpha^h = 0. \tag{5}$$

Для аппроксимации диффузионного оператора ранее была разработана абсолютно устойчивая схема первого порядка точности, обладающая разностным аналогом соотношения

$$\frac{1}{2} \frac{\partial}{\partial t} \iint_{\varphi, z} n^2 \cos \varphi d\varphi dz = - \iint_{\varphi, z} \left(K_1 \frac{\partial n}{\partial z} - K_2 \frac{\partial n}{\partial \varphi} \right)^2 \cos \varphi d\varphi dz \leq 0,$$

имеющего место для уравнения амбиполярной диффузии при нулевых P, k и скорости u , а также нулевых краевых условиях. Подробное описание этой схемы приведено в работе [31]. Приведем эту схему для прямой задачи (для сопряженной задачи аппроксимация диффузионного оператора проводится аналогичным образом):

$$\begin{aligned} & \frac{\partial n_{ij}}{\partial t} = \left[(K_1^2)_{i+1/2, j} \frac{n_{i+1, j} - n_{ij}}{\Delta z^2} - (K_1^2)_{i-1/2, j} \frac{n_{ij} - n_{i-1, j}}{\Delta z^2} \right] + \\ & + \left[\frac{(K_2^2)_{i, j+1/2}}{a^2 \cos \varphi_{j+1/2}} \frac{n_{ij+1} - n_{ij}}{\Delta \varphi^2} - \frac{(K_2^2)_{i, j-1/2}}{a^2 \cos \varphi_{j-1/2}} \frac{n_{ij} - n_{ij-1}}{\Delta \varphi^2} \frac{\cos \varphi_{j-1}}{\cos \varphi_j} \right] + \end{aligned}$$

$$+ \left[-\frac{(K_1)_{i+1/2,j}(K_2)_{i,j+1/2} n_{i+1,j} - n_{ij}}{a \cos \varphi_{j+1/2} \Delta \varphi \Delta z} + \frac{(K_1)_{i+1/2,j-1}(K_2)_{i,j-1/2} n_{i+1,j-1} - n_{ij-1} \cos \varphi_{j-1}}{a \cos \varphi_{j-1/2} \Delta \varphi \Delta z} \cos \varphi_j \right] +$$

$$+ \left[-\frac{(K_1)_{i+1/2,j}(K_2)_{i,j+1/2} n_{ij+1} - n_{ij}}{a \cos \varphi_{j+1/2} \Delta \varphi \Delta z} + \frac{(K_1)_{i-1/2,j}(K_2)_{i-1,j+1/2} n_{i-1,j+1} - n_{i-1j}}{a \cos \varphi_{j+1/2} \Delta \varphi \Delta z} \right].$$

При использовании неявной схемы аппроксимации по времени (или схемы Кранка-Николсон) полученная разностная схема для решения двумерного уравнения является абсолютно устойчивой. На разностном уровне такая схема обладает аналогом приведенного интегрального соотношения:

$$\sum_{i=1}^{N_z} \sum_{j=1}^{N_\varphi} \left[\frac{\partial n_{ij}}{\partial t} \right] n_{ij} a \cos \varphi_j \Delta \varphi \Delta z =$$

$$= - \sum_{i=1}^{N_z} \sum_{j=1}^{N_\varphi} \left[(K_1)_{i+1/2,j} \frac{n_{i+1,j} - n_{i,j}}{\Delta z} - \frac{(K_2)_{i,j+1/2} n_{i,j+1} - n_{i,j}}{a \cos \varphi_{j+1/2} \Delta \varphi} \right]^2 a \cos \varphi_j \Delta \varphi \Delta z.$$

Наличие такого соотношения, с одной стороны, корректно отражает геометрию диффузионных процессов (которые идут вдоль магнитных силовых линий), а с другой – отражает диссипативность рассматриваемой задачи.

Для гиперболических слагаемых, отвечающих переносу, в прямой задаче используется стандартная аппроксимация центральными разностями.

Такая же схема применяется и для оператора диффузии в сопряженном уравнении (интегрирование по времени ведется “в обратном направлении”). Отметим, что аппроксимации всех операторов в сопряженной задаче выбираются таким образом, чтобы на разностном уровне соблюдалось тождество Лагранжа (в сферической метрике).

Аппроксимация интегралов вдоль прямых (интегралов по “тонким прямоугольникам”), фигурирующих в качестве условий наблюдения (наблюдаются ПЭС на отдельных траекториях) осуществляется следующим образом. Фиксируются точки (z_t, φ_t) и (z_b, φ_b) на верхней и нижней границах расчетной области – концы отрезка интегрирования. Этот отрезок разбивается на N_{int} частей, после чего интегрирование ведется методом прямоугольников. Узлы для вычисления интегральной суммы, вообще говоря, не попадают в узлы сетки, поэтому используется стандартная четырехточечная интерполяция (вычисляется ячейка сетки, в которую попадает соответствующий узел).

В результате для каждой фиксированной траектории, вдоль которой вычисляется ПЭС, формируется матрица коэффициентов c_{ij} , с которыми следует вычислять линейную форму $\sum_{i=1}^{N_z} \sum_{j=1}^{N_\varphi} c_{ij} n_{ij}$, аппроксимирующую интеграл $\int_{\Omega_k} n d\Omega$. При этом та же сеточная функция c_{ij} определяет и правую часть сопряженного уравнения: от фиксированной k -й траектории в правую часть сопряженного уравнения пойдет функция c_{ij} , умноженная на разность приближенно вычисленного ПЭС ($\sum_{i=1}^{N_z} \sum_{j=1}^{N_\varphi} c_{ij} n_{ij}$) и наблюдаемого значения ПЭС.

Приведем теперь ряд утверждений об аппроксимационных свойствах регуляризованной (т.е. при $\alpha > 0$) конечно-разностной задачи (5). Прежде всего докажем, что решение конечномерной системы (5) сходится к решению дифференциальной системы оптимальности (4). Поскольку построенные разностные схемы аппроксимируют операторы дифференциальной задачи, для этого остается доказать их устойчивость. Покажем, что она имеет место при условии $\frac{T^2}{\alpha} \|C^h\|^2 < 1$, где T – длина интервала времени, на котором решается задача усвоения. В случае, если константа $\frac{T^2}{\alpha} \|C^h\|^2$ окажется больше единицы, достаточно разбить отрезок времени $[0, T]$ на несколько подотрезков, на каждом из которых эта константа меньше единицы, и потому на каждом подотрезке рассмотренная схема устойчива.

Отметим, что исследования устойчивости разностных схем для систем прямых и сопряженных уравнений в различных постановках приведены в работе [35].

Утверждение 3. Разностная схема для системы оптимальности (5) устойчива при условии $\frac{T^2}{\alpha} \|C^h\|^2 < 1$.

Доказательство приведем для случая нулевой скорости u (в этом случае оператор L включает только производную по времени и диффузионный оператор B). Обозначим через B^h конечно-разностную аппроксимацию диффузионного оператора, соответствующую рассматриваемой схеме. Будем рассматривать сферическую сеточную норму (ℓ_2 -норму с весом $a \cos \varphi_j$): в этой метрике B^h симметричен и неотрицательно определен (см. [31]).

При нулевых начальных и краевых условиях система конечно-разностных прямых и сопряженных уравнений имеет вид:

$$\begin{aligned} n^{j+1} - n^j &= -\tau B^h n^{j+1} + \tau U^j, & n^0 &= n_0; \\ -(q^{j+1} - q^j) &= -\tau B^h q^j + \tau (C^h)^*(C^h n^j), & q^{N_t} &= 0; \\ \alpha U^j &= -q^j, \end{aligned}$$

(аппроксимация по времени – неявная). Здесь n , U и q – векторы (их размерность равна числу узлов пространственной сетки), а верхний индекс отвечает моменту времени. Отметим, что сопряженное уравнение (относительно переменной q) решается в обратном направлении по времени.

После исключения U с помощью последнего уравнения получаем:

$$\begin{aligned} (I + \tau B^h) n^{j+1} &= n^j - \frac{\tau}{\alpha} q^j; \\ (I + \tau B^h) q^j &= q^{j+1} + \tau (C^h)^*(C^h n^j). \end{aligned}$$

Введем в рассмотрение также оператор $R = (I + \tau B^h)^{-1}$. Поскольку в сферической метрике B^h обладает свойствами симметричности и неотрицательной определенности, $\|R\| \leq 1$. Систему можно переписать в виде

$$\begin{aligned} n^{j+1} &= R n^j - \frac{\tau}{\alpha} R q^j; \\ q^j &= R q^{j+1} + \tau R (C^h)^*(C^h n^j). \end{aligned}$$

Применяя рекурсивно первое уравнение системы, получаем $n^{j+1} = R^{j+1} n^0 - \frac{\tau}{\alpha} \sum_{k=0}^j R^{k+1} q^{j-k}$, откуда

$$\|n^{j+1}\| \leq \|n^0\| + \frac{\tau}{\alpha} \sum_{k=0}^j \|R\|^{k+1} \|q^{j-k}\| \leq \|n^0\| + \frac{(j+1)\tau}{\alpha} \max_{k=0, N_t} \|q^k\| \leq \|n^0\| + \frac{N_t \tau}{\alpha} \max_{k=0, N_t} \|q^k\|,$$

а тогда, поскольку правая часть не зависит от j , справедлива и оценка

$$\max_{j=0, N_t} \|n^j\| \leq \|n^0\| + \frac{N_t \tau}{\alpha} \max_{k=0, N_t} \|q^k\|.$$

Отметим, что $N_t \cdot \tau = T$ – длительность интервала времени, на котором решается система (5). Аналогичное рекурсивное применение второго уравнения самому к себе и последующая оценка норм приводят к похожему соотношению:

$$\|q^j\| \leq \|q^{N_t}\| + \tau (N_t - j) \|(C^h)^*(C^h)\| \max_{k=0, N_t} \|n^k\| \leq \|q^{N_t}\| + \tau N_t \|C^h\|^2 \max_{k=0, N_t} \|n^k\|,$$

откуда, поскольку правая часть не зависит от j ,

$$\max_{j=0, N_t} \|q^j\| \leq \|q^{N_t}\| + \tau N_t \|C^h\|^2 \max_{k=0, N_t} \|n^k\|.$$

Подставим оценку для $\max_{j=0, N_i} \|q^j\|$ в правую часть оценки для $\max_{j=0, N_i} \|n^j\|$ и получим

$$\begin{aligned} \max_{j=0, N_i} \|n^j\| &\leq \|n^0\| + \frac{T}{\alpha} \max_{k=0, N_i} \|q^k\| \leq \|n^0\| + \frac{T}{\alpha} \|q^{N_i}\| + \frac{T^2}{\alpha} \|C^h\|^2 \max_{j=0, N_i} \|n^j\| \Rightarrow \\ &\Rightarrow \left(1 - \frac{T^2}{\alpha} \|C^h\|^2\right) \max_{j=0, N_i} \|n^j\| \leq \|n^0\| + \frac{T}{\alpha} \|q^{N_i}\|. \end{aligned}$$

Замечание 2. Приведенное доказательство устойчивости остается справедливым и в случае, когда скорость u отлична от нулевой, но бездивергентна. Однако численные эксперименты показывают, что схема остается устойчивой и при ненулевой скорости u , не являющейся бездивергентной.

Относительно сходимости сеточных решений в численных экспериментах при устремлении параметра регуляризации α к нулю справедливо следующее:

Утверждение 4. Сеточные решения (n_α^h, P_α^h) при $\alpha \rightarrow +0$ сходятся, причем P_α^h сходится к псевдорешению $P^{h,+}$ конечно-разностной задачи $\mathcal{A}^h P^h = f^h$.

Доказательство. Пусть $\{v_1, \dots, v_n\}$ и $\{w_1, \dots, w_n\}$ – сингулярные ортонормированные базисы оператора \mathcal{A}^h (ортонормированные базисы из собственных векторов операторов $\mathcal{A}^h(\mathcal{A}^h)^*$ и $(\mathcal{A}^h)^*\mathcal{A}^h$), σ_k^2 – квадраты его сингулярных чисел, $\sigma_0 = 0$ – нулевое сингулярное число (оно присутствует, поскольку \mathcal{A}^h имеет ненулевое ядро). Решение регуляризованного уравнения имеет вид $P_\alpha^h = (\alpha I + (\mathcal{A}^h)^*\mathcal{A}^h)^{-1}(\mathcal{A}^h)^* f^h = \sum_{k=1}^m \frac{\sigma_k \langle f^h, w_k \rangle}{\alpha + \sigma_k^2} v_k$ (заметим, что оно не содержит компонент по векторам v_k , отвечающим $\sigma_0 = 0$). При $\alpha \rightarrow +0$ получаем в пределе сеточную функцию $\sum_{k=1}^m \frac{1}{\sigma_k} \langle f^h, w_k \rangle v_k$, совпадающую с псевдорешением $P^{h,+}$ задачи $\mathcal{A}^h P^h = f^h$, т.е. с решением задачи $(\mathcal{A}^h)^*\mathcal{A}^h P^h = (\mathcal{A}^h)^* f^h$ наименьшей нормы.

На практике сходимость итерационного процесса для конечно-разностной задачи наблюдается и при $\alpha = 0$, но параметр τ_m приходится выбирать на порядок меньше, чем в регуляризованном случае, даже при достаточно малых α .

Замечание 3. Псевдорешение (сеточная функция) $P^{h,+}$ аппроксимирует некоторое решение дифференциальной задачи $\mathcal{A}P = f$, при котором значение функционала $J_{\alpha=0}(P)$ близко к минимальному. При этом конечно-разностная задача $\mathcal{A}^h P^h = f^h$ плотно разрешимой не является (ядра \mathcal{A}^h и $(\mathcal{A}^h)^*$ ненулевые одновременно), поэтому минимальное значение сеточного функционала $J_{\alpha=0}^h(P^h)$ отлично от нуля и равно квадрату расстояния от сеточной функции f^h до образа \mathcal{A} . Заметим, однако, что при измельчении шага сетки нижняя грань множества значений J_α^h стремится к нулю, поскольку значение J_α^h при соответствующем выборе сеточных норм аппроксимирует значение J_α .

5. ЧИСЛЕННЫЕ ЭКСПЕРИМЕНТЫ

Для исследования качества восстановления распределения электронной концентрации проведем серию контрольных численных экспериментов, использующих в качестве данных наблюдений тестовые значения ПЭС с заданным относительным отклонением от решения прямой задачи. Проводится модельный эксперимент по усвоению данных наблюдений следующего вида: ищется численное решение невозмущенной прямой задачи на всем отрезке времени, затем оно изменяется заданным относительно малым возмущением, после чего вычисляются интегралы (ПЭС) по выбранному траекториям. В качестве прямой модели в данной работе используем уже рассмотренную в предыдущих работах [14], [30] диффузионную модель F слоя ионосферы, все параметры которой соответствуют стандартному состоянию невозмущенной ионосферы. Тем самым мы получим тестовые данные наблюдений (на всем отрезке времени) по возмущенному распределению электронной концентрации, вообще говоря, не являющемуся решением уравне-

ния амбиполярной диффузии. В ходе тестовых численных экспериментов будем восстанавливать с помощью системы ассимиляции решение по построенным возмущенным данным наблюдений. При этом в итерационном процессе для решения системы (5) на одной итерации решается прямое уравнение (“вперед по времени”), а затем сопряженное уравнение (“назад по времени”) на всем временном отрезке, после чего правая часть прямой задачи корректируется для перехода к следующей итерации.

Для всех численных экспериментов в качестве тестовых данных наблюдений будем использовать решения прямой задачи, возмущенные умножением на функцию $(1 + 0.1 \sin(\pi(z - 100)/400))$ (эта функция на верхней и нижней границах обращается в 1, а внутри расчетной области меняет исходное решение не более, чем на 10%). В качестве траекторий интегрирования рассмотрим по отдельности случаи отдельных вертикальных ПЭС и случаи непрерывного заполнения части расчетной области наклонными ПЭС.

Отметим, что во всех проведенных экспериментах 10 итераций достаточно для установления рассчитываемых профилей: на практике дальнейшие итерации не приводили к значимому изменению решения. При этом сходимость наблюдалась и при нулевом параметре регуляризации, но взятие достаточно малого положительного значения α позволяет на порядок увеличить параметр τ итерационного процесса без заметного изменения восстановленного численного решения.

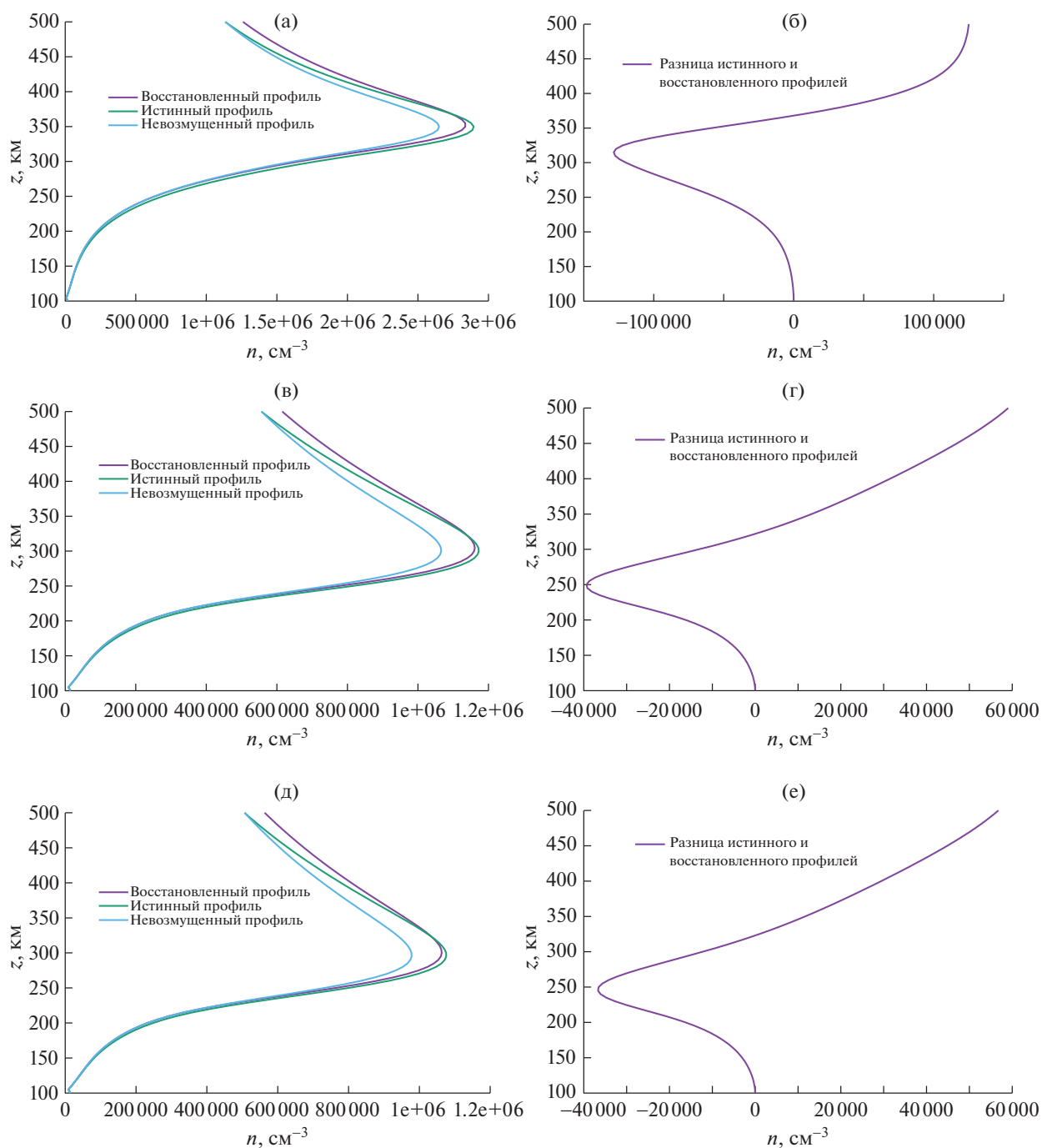
5.1. Восстановление профилей в дневном стационарном распределении

Для исследования качества восстановления стандартного стационарного пространственного распределения электронной концентрации рассмотрим постановку прямой задачи при фиксированной правой части, для которой имеет место сходимость к стационарному решению. В этой постановке на отрезке времени 24 ч рассматривается сходимость к дневному стационарному распределению электронной концентрации с нулевых начальных условий при постоянной функции фотоионизации P_0 . При такой постановке все параметры модели фиксированы во времени и соответствуют стандартному полуденному невозмущенному состоянию ионосферы в условиях весеннего равноденствия [14], [30]. Для прямой задачи при фиксированной правой части, отвечающей дневному невозмущенному состоянию системы, решение сходится к стационарному (характерное время сходимости на всей расчетной области с нулевых начальных условий – порядка 6 ч). Применим построенный алгоритм восстановления решения по вычисленным данным наблюдения и получим решение конечно-разностной задачи, дающее приближенное решение системы оптимальности (5). Цель такого контрольного эксперимента – проверка эффективной работы алгоритма и качества восстановления пространственного распределения электронной концентрации.

Для первого эксперимента для вычислений ПЭС используем строго вертикальные траектории. Будем в качестве данных наблюдений рассматривать интегралы возмущенного решения по вертикальным траекториям на ограниченном количестве точек (на фиксированных широтах: -71° , -53° , -35° , -17° , 0° , 17° , 35° , 53° и 71°).

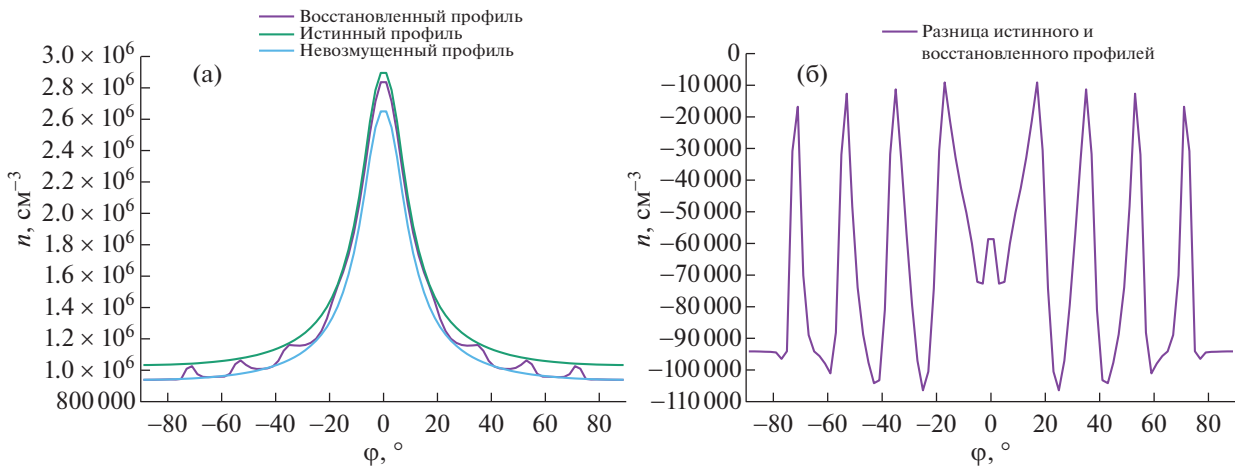
На фиг. 1 представлены результаты восстановления тестового вертикального профиля для области экватора, низких и средних широт (отметим, что при постановке предполагалось существование наблюдений на этих широтах). Эксперименты показали аналогичные результаты и для симметричных относительно экватора широт. Для всех широт интегральная ошибка много меньше (на 3 порядка), чем интеграл от восстановленного (или от “истинного”) профиля. Отметим, что ассимиляция смещает положение максимума электронной концентрации вверх, а отклонение восстановленного профиля от “истинного” по высоте составляет 4% вблизи максимума (такая ошибка по высоте расположена приблизительно на 40 км ниже максимума) и максимально вблизи верхней границы. При этом форма профиля в целом немного изменяется, что не противоречит постановке, поскольку ядро оператора \mathcal{A} ненулевое. Наибольшая ошибка восстановления вблизи верхней границы составляет около 10% и связана с выбором вида функции возмущения, на которую было умножено решение при подсчете модельных данных наблюдений: после умножения изменилась структура профилей, и возмущенное распределение электронной концентрации уже не является решением прямой задачи, а наклон профиля вблизи верхней границы не согласован с краевыми условиями для уравнения амбиполярной диффузии.

На фиг. 2 приведено распределение по широтам одного из ключевых практически значимых параметров ионосферы – величины максимума (по высоте z) электронной концентрации и ошибки восстановления. Видно, что с точки зрения максимальной по высоте электронной кон-



Фиг. 1. Вертикальные распределения электронной концентрации для широт 0° (а), 35° (в) и 53° (д), полученные решением только прямой задачи для модели ионосферы (голубая линия), решением задачи восстановления (фиолетовая линия) и ожидаемое по восстановлению тестовое “точное” решение (зеленая), а также профили абсолютной разницы восстановленного и точного решения для широт 0° (б), 35° (г) и 53° (е). В качестве тестовых данных наблюдений использованы данные о вертикальных ПЭС. Концентрации электронов приведены в см^{-3} .

центрации предложенный алгоритм с высокой точностью (ошибка в средних широтах, на которых были выбраны данные наблюдений, не превосходит 1% после ассимиляции по сравнению с исходным возмущением порядка 10%, а наименьшая точность $\sim 2\%$ получена на экваторе) восстанавливает искомое значение, причем усвоение данных по вертикали влияет на близлежащие



Фиг. 2. Широтные распределения максимума электронной концентрации по высоте (а), полученные решением только прямой задачи для модели ионосферы (голубая линия), решением задачи восстановления (фиолетовая линия) и ожидаемое по восстановлению тестовое “точное” решение (зеленая), а также разница восстановленного и точного широтных распределений максимума (б). Используемые данные наблюдений — вертикальные ПЭС. Концентрации электронов приведены в см^{-3} .

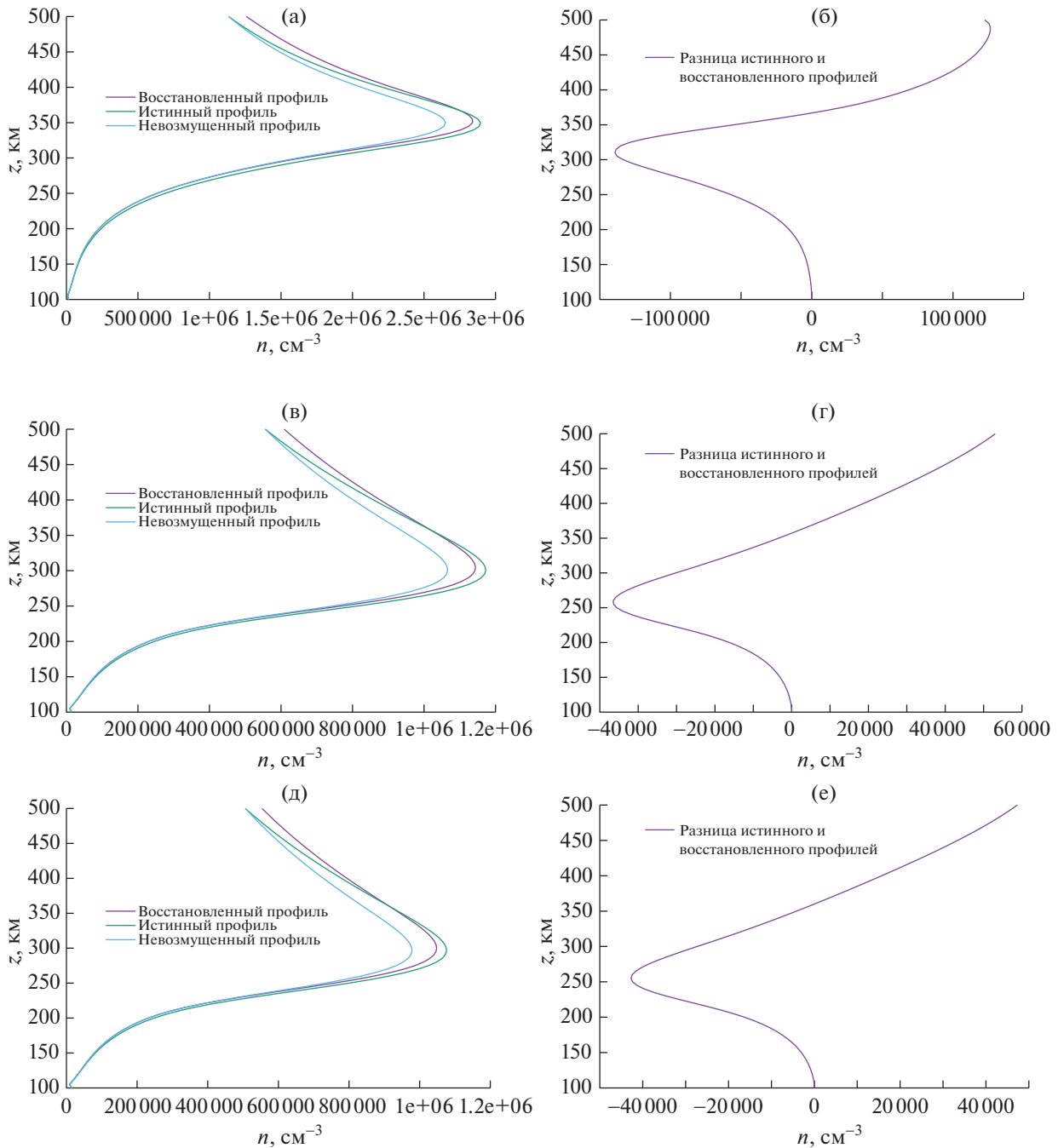
широты приблизительно на 2° по обе стороны от широты, где расположена траектория интегрирования.

Теперь рассмотрим эксперимент с наблюдаемыми данными — интегралами вдоль заданных наклонных прямолинейных траекторий, оставив прежними все остальные параметры модели. Заддим фиксированный угол наклона прямолинейных траекторий следующим образом: точку на высоте $z_a = 500$ км и широте φ соединим с точкой на высоте $z_b = 100$ км и широте $(\varphi - 35^\circ)$, а затем покроем максимально возможную часть области такими траекториями. При таком подходе вертикальные профили на широтах от -55° до 55° оказываются полностью покрыты траекториями интегрирования, поэтому вновь можно провести сравнение восстановленных и возмущенных профилей на этих широтах. На фиг. 3 приведены результаты восстановления профилей для тех же широт, что и в случае вертикальных профилей наблюдения.

Результаты восстановления имеют схожую структуру и величины ошибок со случаем вертикальных профилей. При этом ошибка восстановления профиля несколько уменьшилась вблизи верхней границы и увеличилась в районе максимума для средних широт. Это связано с тем, что наклонные траектории сильнее сконцентрированы вблизи верхней и нижней границ расчетной области по сравнению с вертикальными. Ниже на фиг. 4 приведены распределения по широтам для величины максимума (по высоте z) электронной концентрации, а также ошибка восстановления этой величины.

На широтах $[-90^\circ, -65^\circ]$ ассимиляция практически не повлияла на величину максимума, поскольку траектории интегрирования не затрагивали часть вертикального профиля, содержащую максимум. То же относится и к области в окрестности северного полюса. В остальных широтах восстановление максимума происходит приблизительно с одинаковой точностью (не более 3% по сравнению с исходным возмущением в 10%). Отметим, что абсолютная ошибка восстановления максимума на экваторе вдвое больше, чем в средних широтах, но и значение электронной концентрации в экваториальной зоне также приблизительно вдвое больше.

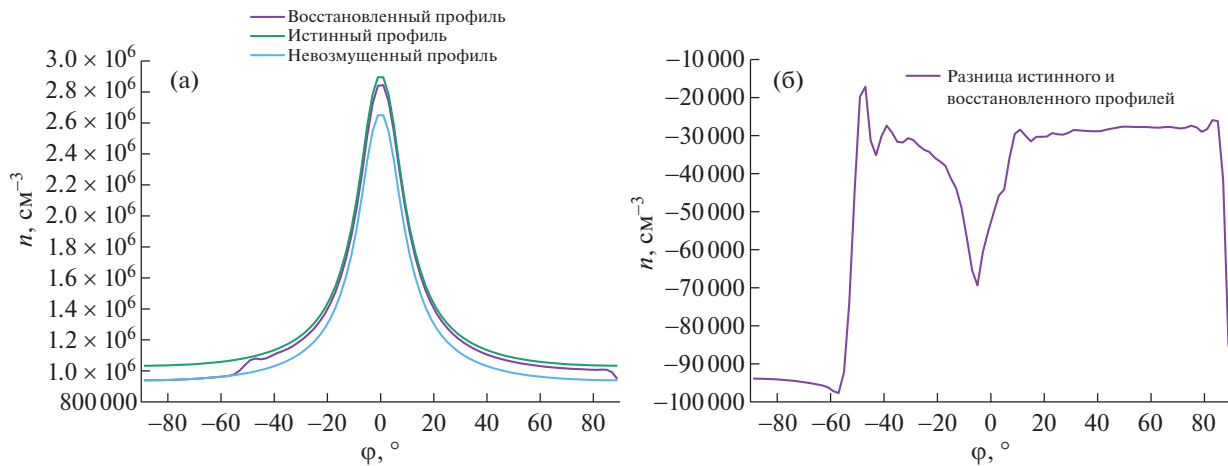
Поскольку в качестве тестовых данных наблюдений фигурируют лишь интегралы от возмущенного решения прямой задачи, но не само возмущенное решение, при модификации возмущенного профиля дополнительно без изменения интеграла данные никак не преобразуются, и поэтому восстановление приведет к тем же самым результатам. Высокое качество восстановления в обоих приведенных экспериментах по восстановлению стационарного решения обусловлено хорошей корреляцией возмущенного профиля и стационарного невозмущенного решения (положение максимума, промежутки убывания и возрастания скоррелированы).



Фиг. 3. Вертикальные распределения электронной концентрации для широт 0° (а), 35° (в) и 53° (д), полученные решением только прямой задачи для модели ионосферы (голубая линия), решением задачи восстановления (фиолетовая линия) и ожидаемое по восстановлению тестовое “точное” решение (зеленая), а также профили абсолютной разницы восстановленного и точного решения для широт 0° (б), 35° (г) и 53° (е). В качестве тестовых данных наблюдений использованы данные о наклонных ПЭС. Концентрации электронов приведены в см^{-3} .

5.2. Восстановление профилей в суточном ходе

Для более реалистичного эксперимента рассмотрим суточный ход на отрезке времени 48 ч, в котором начальные условия выбираются в соответствии с установившимся суточным циклом, фотоионизация P_0 меняется в зависимости от времени в соответствии с изменением зенитного угла Солнца [36], а остальные внешние параметры модели остаются фиксированными. Такая по-



Фиг. 4. Широтные распределения максимума электронной концентрации по высоте (а), полученные решением только прямой задачи для модели ионосферы (голубая линия), решением задачи восстановления (фиолетовая линия) и ожидаемое по восстановлению тестовое “точное” решение (зеленая), а также разница восстановленного и точного широтных распределений максимума (б). Используемые данные наблюдений — наклонные ПЭС. Концентрации электронов приведены в см^{-3} .

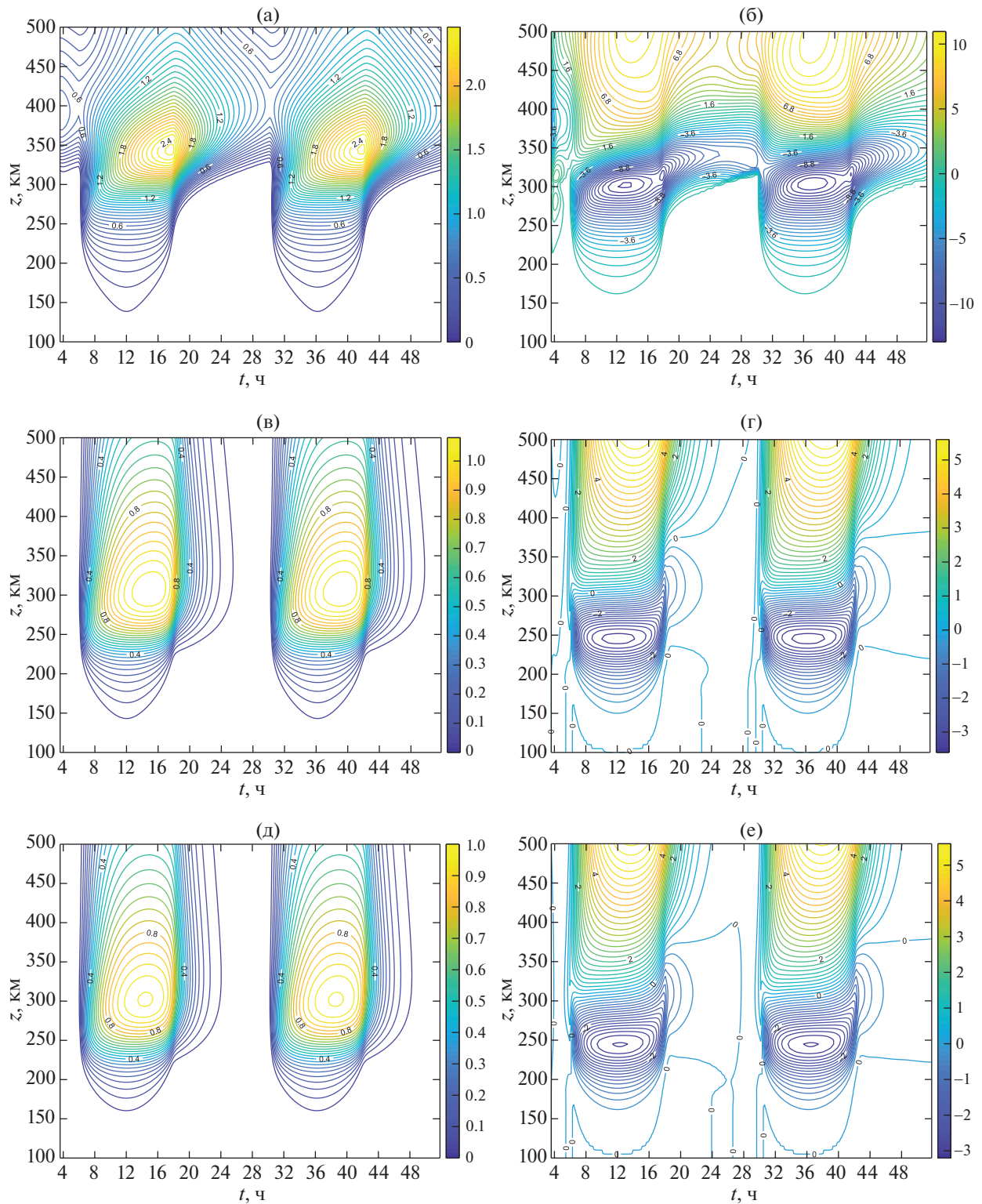
становка позволяет изучить характеристики восстановления в динамике. Как и в предыдущем эксперименте, данные считаются известными на каждом шаге по времени, а одной итерации процесса решения системы прямых и сопряженных уравнений соответствует решение прямого уравнения на всем временном интервале, а затем сопряженного уравнения на том же интервале времени в обратном направлении, после чего обновляется правая часть.

Исследуем качество ассимиляции данных наблюдений при моделировании суточного хода за время T , равное 2 сут. Цель эксперимента — проверка качества восстановления решения не только по пространству, но и во времени. Приведем результаты восстановления решения при задании вертикальных ПЭС на ограниченном количестве широт, аналогично предыдущим (см. п. 5.1) экспериментам (-71° , -53° , -35° , -17° , 0° , 17° , 35° , 53° и 71°). Восстановленное решение на широтах 0° , 35° и 53° , а также разница между возмущенным и восстановленным решениями на этих широтах представлены на графиках.

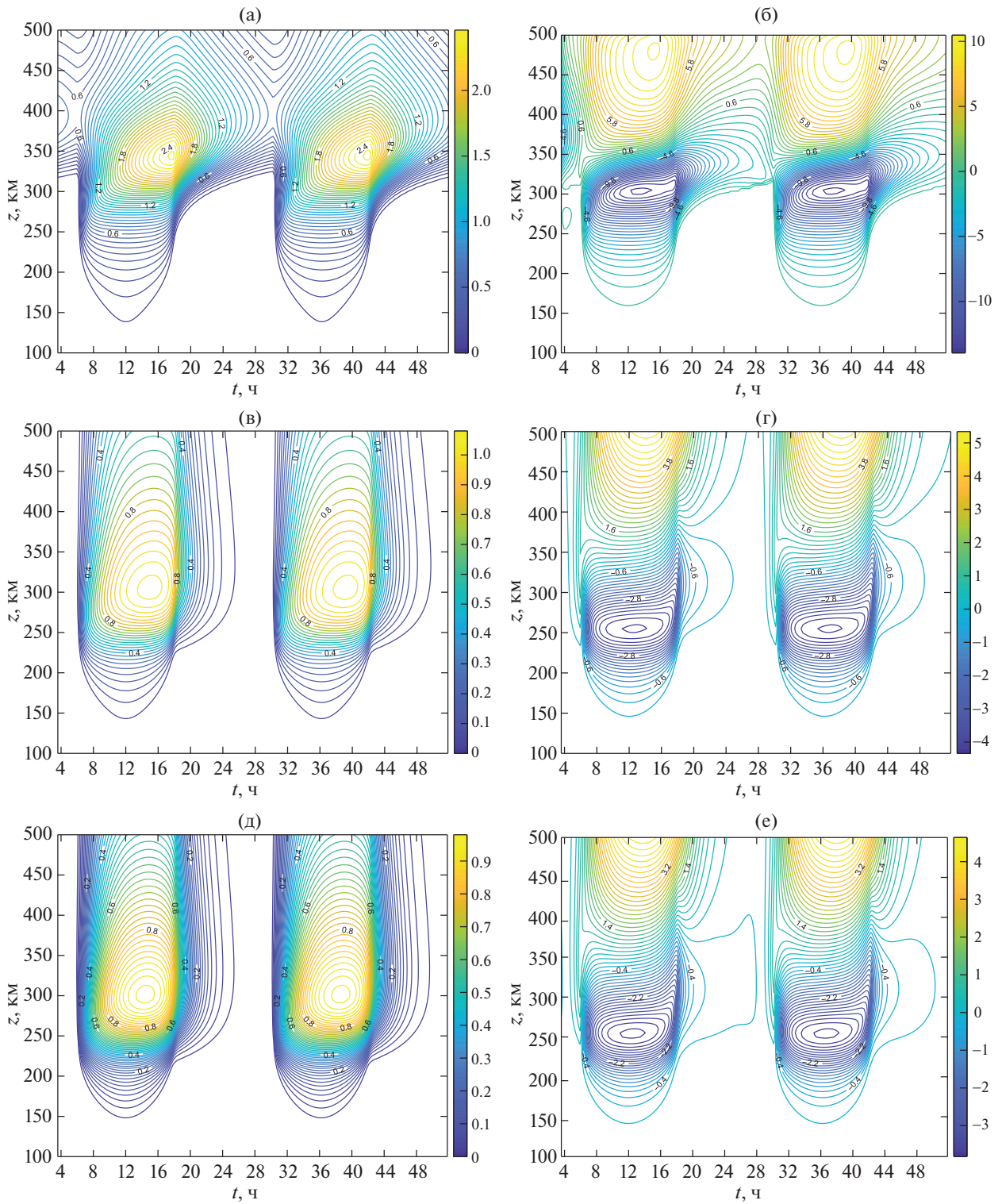
На фиг. 5 и фиг. 6 представлен временной ход вертикального распределения электронной концентрации, восстановленной в результате ассимиляции по данным о вертикальных (фиг. 5) и наклонных (фиг. 6) ПЭС на исследуемых широтах. Как отмечено выше, формирование тестовых данных наблюдений осуществлялось полностью аналогично описанному в п. 5.1 способу.

Представленные результаты показывают, что при ассимиляции данных для обоих типов ПЭС (вертикальных и наклонных) решение приближается к истинному возмущенному. Интегралы от полученного решения совпадают с выбранными данными наблюдений вплоть до третьего порядка. При этом максимум электронной концентрации немного сдвигается вверх по высоте, а наибольшая ошибка восстановления в окрестности максимума расположена приблизительно на 40 км ниже него и не превосходит 2% (на экваторе эта ошибка максимальна) по сравнению с исходной ошибкой 10% для невозмущенного решения. Во всей расчетной области ошибка восстановления максимальна вблизи верхней границы, как и в стационарном случае, в силу несогласованности возмущения с верхним краевым условием прямой задачи (восстановленное решение обязано ему удовлетворять, а возмущенное — заведомо не удовлетворяет). При этом вблизи верхней границы восстановленное решение оказывается несколько завышенным. Тем самым в динамике, при переменной по времени правой части, с точки зрения характеристик восстановленных профилей наблюдаются те же эффекты, что и для восстановленного дневного стационарного решения в п. 5.1. В суточном цикле наибольшие относительные ошибки восстановления ожидаемо соответствуют дневным условиям, ошибки в ночное время не превышают 1% для всех экспериментов.

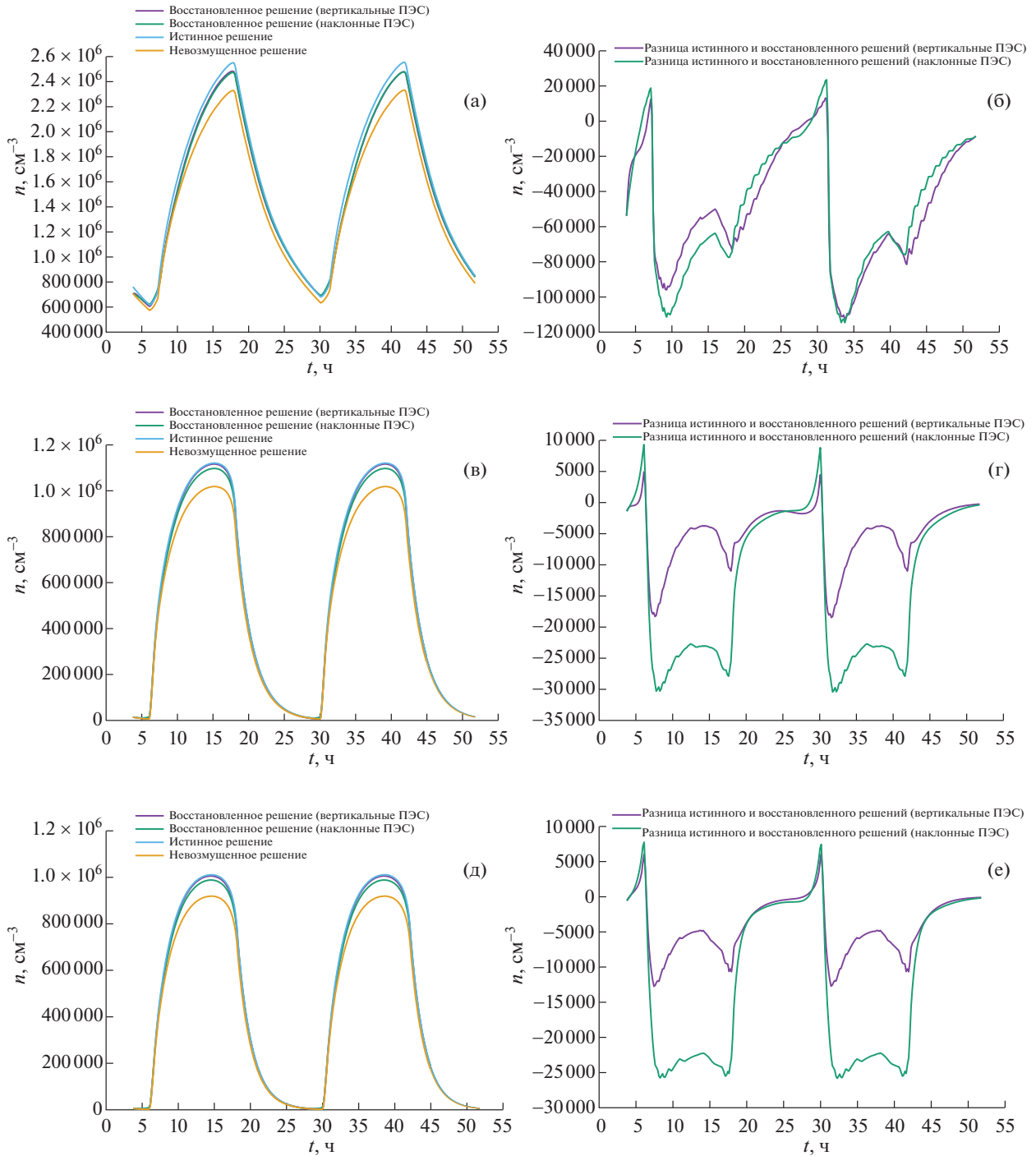
Рассмотрим отдельно восстановление величины максимальной (по высоте) электронной концентрации. На фиг. 7 приведены результаты восстановления этого максимума как для случая вертикальных траекторий интегрирования в данных наблюдений, так и для случая непрерывно-



Фиг. 5. Временной ход вертикальных распределений электронной концентрации для широт 0° (а), 35° (в) и 53° (д), полученный решением задачи восстановления, а также профиль абсолютной разницы восстановленного и точного решения для широт 0° (б), 35° (г) и 53° (е). В качестве тестовых данных наблюдений использованы данные о вертикальных ПЭС. Полудню соответствуют отметки 12 и 36 на оси времени. Изолинии проведены в 10^6 и 10^4 см^{-3} соответственно.

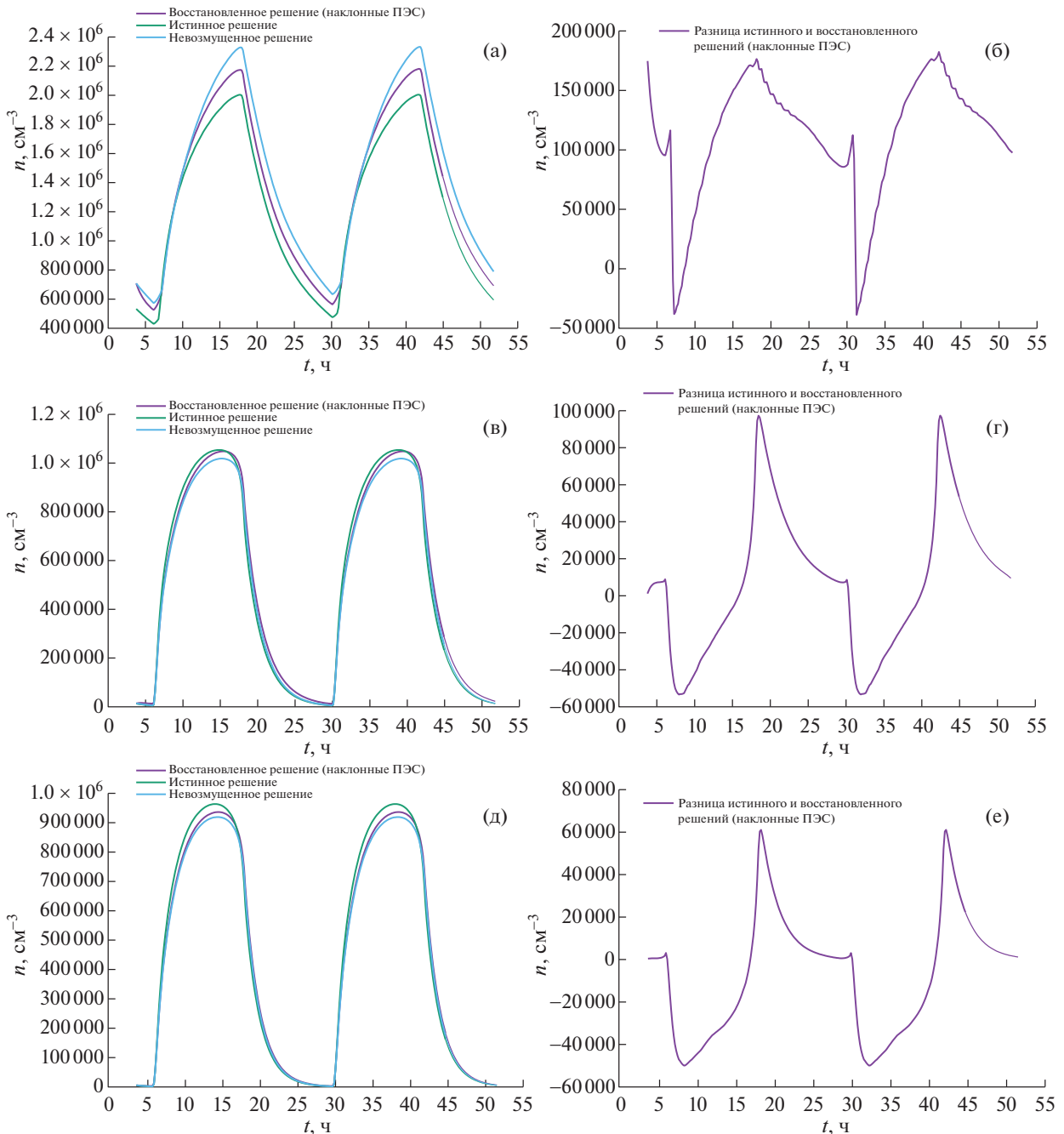


Фиг. 6. Временной ход вертикальных распределений электронной концентрации для широт 0° (а), 35° (в) и 53° (д), полученный решением задачи восстановления, а также профиль абсолютной разницы восстановленного и точного решения для широт 0° (б), 35° (г) и 53° (е). В качестве тестовых данных наблюдений использованы данные о наклонных ПЭС. Полудню соответствуют отметки 12 и 36 на оси времени. Изолинии проведены в 10^6 и 10^4 см^{-3} соответственно.



Фиг. 7. Временной ход максимума распределения электронной концентрации для широт 0° (а), 35° (в) и 53° (д), полученный решением задачи восстановления для случаев вертикальных траекторий (фиолетовая линия), наклонных траекторий (зеленая линия), ожидаемая “точная” зависимость максимума от времени (голубая линия) и зависимость максимума для невозмущенного решения прямой задачи (желтая линия), а также профиль абсолютной разницы восстановленного и точного значения максимума для вертикальных траекторий (фиолетовая линия) и наклонных траекторий (зеленая линия). Представлены широты 0° (б), 35° (г) и 53° (е).

го заполнения области наклонными траекториями: слева приведены результаты расчета максимума по высоте для обоих случаев, а справа приведены разности максимумов восстановленных решений и “истинного” (возмущенного) распределения для обоих случаев траекторий (вертикальных и наклонных).



Фиг. 8. Временной ход максимума распределения электронной концентрации для широт 0° (а), 35° (в) и 53° (д), полученный решением задачи восстановления для случая наклонных траекторий (фиолетовая линия), ожидаемая “точная” зависимость максимума от времени (зеленая линия) и зависимость максимума для невозмущенного решения прямой задачи (голубая линия), а также профиль абсолютной разницы восстановленного и точного значения максимума для наклонных траекторий (фиолетовая линия). Случай равенства нулю интеграла от возмущения решения по вертикали. Представлены широты 0° (б), 35° (г) и 53° (е).

Из представленных графиков видно, что ассимиляция существенно приближает максимальное значение электронной концентрации к “истинному” (ошибка составляет не более 2% для вертикальных ПЭС и 3% для наклонных ПЭС соответственно по сравнению с ошибкой в 10% для невозмущенного решения). При этом восстановление по наклонным траекториям оказывается несколько хуже, поскольку большая часть такой траектории расположена в окрестностях верхней и нижней границ расчетной области, в то время как максимум электронной концентрации расположен ближе к ее середине по оси z (на высоте порядка 300 км).

Отметим, что более качественное восстановление решения при вертикальных ПЭС было связано в том числе с конкретным видом возмущения. В самом деле, предположим, что решение было возмущено таким образом, чтобы интеграл по любой вертикальной траектории от данного возмущения равнялся нулю. Тогда в процессе усвоения по вертикальным ПЭС в правой части сопряженного уравнения стоит тождественный ноль, и начальное приближение в итерационном процессе для решения системы оптимальности (4) изменяться не будет. Тем не менее при выборе наклонных ПЭС в том же итерационном процессе решение уже будет меняться. Приведем результат восстановления максимума электронной концентрации для эксперимента, в котором данные наблюдений вычислены по решению прямой задачи $n(z, \varphi, t)$, возмущенному по правилу $n_{\text{new}}(z, \varphi, t) = n(z, \varphi, t) + 0.2 \sin(2\pi(z - 100)/400) \max_{z \in [100, 500]} n(z, \varphi, t)$. Нетрудно заметить, что интеграл по z на отрезке от 100 км до 500 км от добавочного слагаемого в правой части этой формулы равен нулю. Такое возмущение изменяет форму профиля распределения электронной концентрации по высоте, причем знаки возмущения различны в областях выше и ниже высоты $z = 300$ км. Ассимиляция данных приближала решение к возмущенному в части области выше максимума электронной концентрации (ошибка с 30% снижалась в этой подобласти до 15% в приэкваториальной зоне, с 15 до 3% вблизи широты 35° и с 15 до 7% в средних широтах) и мало изменяла профили ниже $z = 300$ км, поскольку коэффициент диффузии D принимает выше $z = 300$ км значения, на несколько порядков превосходящие соответствующие значения в нижней части расчетной области. Рассмотрим восстановление максимальных по z значений решения на различных широтах.

На фиг. 8 зависимость максимума электронной концентрации после усвоения при выборе вертикальных траекторий интегрирования не представлена, поскольку она совпадает с невозмущенным временным ходом этой величины. При этом использование данных наблюдений на наклонных траекториях позволяет улучшить восстановление неизвестного возмущенного распределения электронной концентрации. Отметим, что выбранное возмущение уменьшает значения функции $n(z, \varphi, t)$ в верхней части расчетной области и увеличивает в нижней, поэтому в “точном” возмущенном решении положение максимума электронной концентрации оказывается смещено по высоте. Это приводит к тому, что, в частности, максимальное значение на экваторе в возмущенном решении уменьшается, а в средних широтах увеличивается, и ошибки восстановления этой величины имеют характер, отличный от приведенного на фиг. 7 для эксперимента с ненулевым интегралом по вертикали от возмущения.

6. ЗАКЛЮЧЕНИЕ

В заключение кратко перечислим основные результаты данной работы.

1. Сформулирована постановка задачи вариационной ассимиляции данных наблюдений для разрабатываемой в ИВМ РАН динамической модели F слоя ионосферы Земли в двумерной диффузионной постановке [14], [30]. В качестве основного типа данных в данной работе рассматриваются интегралы электронной плотности (ПЭС) вдоль прямолинейных траекторий, соединяющих спутник с точкой на земной поверхности. В качестве управления в задаче оптимальности рассматривается неопределенная добавка к правой части (функции ионизации). Приведена полная постановка задачи, показана ее плотная разрешимость и рассмотрена ее регуляризация, на основе которой построен алгоритм решения с помощью итерационного метода, доказана сходимость сеточных решений к псевдорешению конечно-разностной задачи при стремлении параметра регуляризации к нулю.

2. Проведена дискретизация операторов задачи ассимиляции на основе методов, разработанных для решения прямой задачи [14], [30], показаны устойчивость и сходимость разработанного численного метода решения задачи оптимальности к решению дифференциальной задачи.

3. На основе контрольных численных экспериментов по восстановлению профилей по известным ПЭС для возмущений, максимум которых скоррелирован с максимумом невозмущенного распределения электронной плотности, показано высокое качество восстановления распределения электронной концентрации в областях с существованием данных наблюдений. Рассмотрены восстановление профилей в задаче о сходимости к дневному стационарному распределению электронной концентрации с нулевых начальных условий, а также восстановление профилей в суточном ходе. В обоих случаях алгоритм позволяет достаточно точно восстановить возмущенное тестовое решение (с отклонениями до 10%) с наибольшими ошибками 2–4%

(в дневных условиях, главным образом в утренние часы) в области максимума и верхней границы (что связано с особенностями функции возмущения).

Отметим, что в данной работе рассмотрена первая версия системы усвоения данных для модели ионосферы ИВМ РАН в двумерной диффузионной постановке [14]. В качестве промежуточного варианта трехмерной модели, не использующей трехмерного переноса, может быть рассмотрен предложенный в данной работе алгоритм, в котором в качестве данных наблюдений используются интегралы по трехмерным траекториям, не локализованным на заданной долготе. На следующих этапах работы на основе разработанных алгоритмов будет реализована система усвоения для трехмерной модели ионосферы, учитывающей трехмерный перенос. Один из вариантов обобщения предложенной в статье двумерной системы усвоения на трехмерный вариант заключается в использовании разработанного алгоритма на шаге расщепления трехмерной модели. В рамках трехмерной версии системы усвоения будет отдельно рассмотрена проблема ее эффективности и возможности использования в оперативном режиме для прогноза состояния ионосферы. Отдельно в данной модели будет реализовано усвоение другого типа данных – полных или частичных вертикальных профилей (получаемых с помощью ионозондов), что потребует изменение оператора наблюдений в постановке задачи и отдельного исследования точности и эффективности такой системы при различных конфигурациях данных наблюдений.

СПИСОК ЛИТЕРАТУРЫ

1. Дэвис К. Радиоволны в ионосфере. М.: Мир, 1973. 502 с.
2. Иванов-Холодный Г.С., Никольский Г.М. Солнце и ионосфера: коротковолновое излучение Солнца и его воздействие на ионосферу. М.: Наука, 1969. 455 с.
3. Kelley M.C. The Earth's Ionosphere. Academic Press, 1989. 480 p.
4. Deminov M.G., Deminova G.F., Zhrebtsov G.A., Polekh N.M. Statistical properties of variability of the quiet ionosphere F2-layer maximum parameters over Irkutsk under low solar activity // Advances in Space Research. 2011. V. 51. № 5. P. 702–711.
5. Ratovsky K.G., Medvedev A.V., Tolstikov M.V. Diurnal, seasonal and solar activity pattern of ionospheric variability from Irkutsk Digisonde data // Advances in Space Research. 2015. V. 55. № 8. P. 2041–2047.
6. Miller A., Schmidt H., Bunzel F. Vertical coupling of the middle atmosphere during sudden stratospheric warming events // J. Atmos. Sol.-Terr. Phys. 2013. V. 97. P. 11–21.
7. Yigit E., Koucki Knizova P., Georgieva K., Ward W. A review of vertical coupling in the Atmosphere–Ionosphere system: Effects of waves, sudden stratospheric warmings, space weather, and of solar activity // J. Atmos. Sol.-Terr. Phys. 2016. V. 141. P. 1–12.
8. Pedatella N.M. et al. Multimodel comparison of the ionosphere variability during the 2009 sudden stratosphere warming // J. of Geophysical Research: Space Phys. 2016. V. 121. № 7. P. 7204–7225.
9. Дымников В.П., Лыкозов В.Н., Володин Е. М. Mathematical simulation of earth system dynamics // Izvestiya. Atmospheric and Oceanic Phys. 2015. V. 51. № 3. P. 227–240.
10. Kulyamin D.V., Volodin E.M. INM RAS coupled atmosphere–ionosphere general circulation model INMAIM (0–130 km) // Rus. J. of Num. Analys. and Math. Model. 2018. V. 33. № 6. P. 351–357.
11. Кулямин Д.В., Дымников В.П. Моделирование климата нижней ионосферы // Известия Российской академии наук. Физика атмосферы и океана. 2015. Т. 51. № 3. С. 317–337.
12. Kulyamin D.V., Dymnikov V.P. Numerical modelling of coupled neutral atmospheric general circulation and ionosphere D region // Rus. J. of Num. Analys. and Math. Model. 2016. V. 31. № 3. P. 159–171.
13. Кулямин Д.В., Галин В.Я., Погорельцев А.И. Моделирование общей циркуляции термосферы с включением параметризации радиационных процессов // Метеорология и гидрология. 2015. Т. 40. № 6. С. 48–57.
14. Кулямин Д.В., Останин П.А., Дымников В.П. Моделирование F слоя земной ионосферы. Решение уравнений амбиполярной диффузии // Матем. моделирование. 2019. Т. 31. № 4. С. 57–74.
15. Дымников В.П., Кулямин Д.В., Останин П.А. Совместная модель глобальной динамики термосферы и ионосферы Земли // Известия Российской академии наук. Физика атмосферы и океана. 2020. Т. 56. № 3. С. 280–292.
16. Hofmann-Wellenhof B., Lichtenegger H., Wasle E. Global Navigation Satellite Systems: GPS, GLONASS, Galileo and more. Wien, New York: Springer, 2008. 545 p.
17. Parkinson B.W., Spilker J.J. Global Positioning System: Theory and Applications // Ciencia militar y naval UR. American Institute of Aeronautics & Astronautics 1996.
18. Глобальная навигационная спутниковая система ГЛОНАСС. Интерфейсный контрольный документ. М.: 1995.

19. *Afraimovich E.L., Astafyeva E.I., Demyanov V.V., Edemskiy I.K., Gavrilyuk N.S., Ishin A.B., et al.* A review of GPS/GLONASS studies of the ionospheric response to natural and anthropogenic processes and phenomena // *J Space Weather Space Clim.* 2013. V. 3. № A27.
20. *Rideout W., Coster A.* Automated GPS processing for global total electron content data // *GPS Solutions.* 2006. V. 10 (3) P. 219–228.
<https://doi.org/10.1007/s10291-006-0029-5>
21. *Brunini C., Meza A., Azpilicueta F., Zele M.A.V., Gende M., Daz A.* A New Ionosphere Monitoring Technology Based on GPS // *Astrophys. Space Sci.* 2004. V. 290. P. 415–429.
22. *Bibl K.* Evolution of the ionosonde // *Annali di Geophysica.* 1998. V. 41. № 5-6. P. 667–680.
23. *Reinisch B.W., Galkin I.A.* Global Ionospheric Radio Observatory (GIRO) // *Earth Planet Sp.* 2011. V. 63. P. 377–381.
<https://doi.org/10.5047/eps.2011.03.001>
24. *Schunk R.W., Scherliess L., Sojka J.J., Thompson D.C., Anderson D.N., Codrescu M., Minter C., Fuller-Rowell T.J., Heelis R.A., Hairston M., Howe B.* Global assimilation of ionospheric measurements (GAIM) // *Radio Sci.* 2004. V. 39. № RS1S02.
25. *Wang C., Hajj G., Pi X., Rosen I.G., Wilson B.* Development of the Global Assimilative Ionospheric Model // *Radio Sci.* 2004. V. 39. № RS1S06.
<https://doi.org/10.1029/2002RS00854>
26. *Dear R., Mitchell C.* GPS interfrequency biases and total electron content errors in ionospheric imaging over Europe // *Radio Sci.* 2006. V. 41. № RS6007.
<https://doi.org/10.1029/2005RS003269>
27. *Angling M., Shaw J., Shukla A., Cannon P.* Development of an HF selection tool based on the Electron Density Assimilative Model near-real-time ionosphere // *Radio Sci.* 2009. V. 44. № RS0A13.
<https://doi.org/10.1029/2008RS004022>
28. *Fuller-Rowell T. et al.* US-TEC: a new data assimilation product from the space environment center characterizing the ionospheric total electron content using real-time GPS data // *Radio Sci.* 2006. V. 41.
<https://doi.org/10.10292005RS003393>
29. *Khattatov B. et al.* Ionospheric nowcasting via assimilation of GPS measurements of ionospheric electron content in a global physics-based time-dependent model // *Q. J. R. Meteorol. Soc.* 2005. V. 131. P. 3543–3559.
30. *Ostanin P.A., Kulyamin D.V., Dymnikov V.P.* Numerical modelling of the Earth's ionosphere F region // *IOP Conf. Series: Earth and Environmental Science.* 2017. V. 96. № 1. P. 012011(1)–012011(11).
31. *Ostanin P.A.* On the approximation of the diffusion operator in the ionosphere model with conserving the direction of geomagnetic field // *Rus. J. of Num. Analys. and Math. Model.* 2022. V. 37. № 1. P. 25–39.
32. *Лионс Ж.-Л.* Оптимальное управление системами, описываемыми уравнениями с частными производными. М.: МИР, 1972. 416 с.
33. *Шутяев В.П.* Операторы управления и итерационные алгоритмы в задачах вариационного усвоения данных. М.: Наука, 2001. 239 с.
34. *Агошков В.И.* Методы оптимального управления и сопряженных уравнений в задачах математической физики. М.: ИВМ РАН, 2016. 244 с.
35. *Пармузин Е.И., Шутяев В.П.* О численных алгоритмах решения одной задачи об усвоении данных // *Ж. вычисл. матем. и матем. физ.* 1997. V. 37. № 7. С. 816–827.
36. *Schunk R.W., Nagy A.F.* IONOSPHERES Physics, Plasma Physics, and Chemistry. New York, United States: Cambridge University Press, 2009. 628 p.

ВОЛНОВОДНАЯ МОДЕЛЬ РАЗВИТОГО ТУРБУЛЕНТНОГО ПОГРАНИЧНОГО СЛОЯ

© 2023 г. В. А. Жаров^{1,*}, И. И. Липатов¹, Р. С. Селим¹

¹ 141701 Долгопрудный, М.о., Институтский пер., 9, МФТИ, Россия

*e-mail: v_zharov@mail.ru

Поступила в редакцию 08.11.2022 г.
Переработанный вариант 08.11.2022 г.
Принята к публикации 02.02.2023 г.

В работе представлено исследование развитого турбулентного пограничного слоя, возникающего при обтекании вязкой несжимаемой жидкостью пластины под нулевым углом атаки и с нулевым продольным градиентом давления. Для описания турбулентного пограничного слоя использован волноводный подход, в котором турбулентные пульсации связаны с волнами Толлмина–Шлихтинга, находящимися в состоянии трехволнового резонанса. Для исследования исходной нелинейной системы уравнений предложена оценка гидродинамических величин, которая не нарушает общепринятый подход в пограничном слое, но приводит к появлению нового малого параметра: отношению толщины потери импульса пограничного слоя к длине затухания наименее затухающей моды волн Толлмина–Шлихтинга. На основе метода многих масштабов получены уравнения для когерентной и стохастической частей пульсаций. Определены дисперсионные характеристики волн наименее затухающей моды на профиле средней продольной скорости развитого турбулентного пограничного слоя; проанализированы условия множественного трехволнового резонанса этой моды волн Толлмина–Шлихтинга. Для когерентной части пульсаций проведено сравнение пульсационных характеристик с известными численными результатами. Библ. 31. Фиг. 7. Табл. 1.

Ключевые слова: несжимаемая вязкая жидкость, турбулентный пограничный слой, волны Толлмина–Шлихтинга, когерентные структуры.

DOI: 10.31857/S0044466923050241, **EDN:** GGJGQN

ВВЕДЕНИЕ

На основе большого объема экспериментальных данных в работах [1–5] были проанализированы физические процессы в ламинарных и турбулентных пограничных слоях (ТПС). В них показано, что развитый турбулентный пограничный слой содержит организованные вихревые структуры, определяющие многие физические свойства этих течений. Большой фактический материал, полученный в последние годы экспериментально о структуре турбулентного пограничного слоя, содержится в работах [6–9]. Краткое изложение этих работ содержится в лекции № 7 Ю.С. Качанова [10]. Математическая интерпретация особенностей этих течений предложена в работах [11–13].

В случае ТПС в данной работе рассматриваются модели, основанные на волноводном представлении динамики пульсаций. Полученная таким образом модель очень хорошо описывает свойства развитого турбулентного пограничного слоя (ТПС) и поэтому заслуживает серьезного обоснования, которое может быть достигнуто путем привлечения теории динамических систем [14]. Для этого удобно воспользоваться формулировкой известной теоремы Биркгофа–Хинчина, которая была дана в работе [15]. Она дает условия существования средних величин по времени для динамических систем определенного вида. Далее приведена теоретическая постановка задачи описания когерентной структуры [16] в развитом ТПС и показано, что с точки зрения теоремы Биркгофа–Хинчина результаты, полученные в [16], вполне обоснованы.

Для определения волноводных свойств ТПС в данной работе используется спектр собственных значений и собственных функций уравнения Орра–Зоммерфельда на профиле развитого ТПС [17].

Теоретическая часть задачи связана с решением уравнения для когерентной структуры, некоторым приближением которого является совокупность гармоник в состоянии множественного трехволнового резонанса. Ее динамика описывается автономной системой обыкновенных дифференциальных уравнений. Динамический инвариант этой системы уравнений можно положить равным 1, что позволяет применить теорему Биркгофа–Хинчина. В результате можно утверждать, что определенные ранее в [16] средние по времени величины квадратов модулей амплитуд, описывающих когерентную структуру, конечны и допускают корректную численную оценку. При этом когерентная структура, по своим свойствам воспроизводимости, вполне соответствует обнаруженной в эксперименте [10] детерминированной турбулентности (DeTu).

1. В работе [16] по аналогии с волноводной моделью [18], [19] из уравнений Навье–Стокса получено нелинейное уравнение (1) для фурье-составляющей вертикальной скорости волн Толлмина–Шлихтинга (Т–Ш), описывающей пульсации в пограничном слое на пластине под нулевым углом атаки при отсутствии продольного градиента давления в одномодовом приближении. Для этого поле скорости разлагалось на две составляющие: среднюю по времени и флуктуирующую части. Уравнения для когерентной составляющей отделяются от уравнений для случайных пульсаций в явной форме за счет наличия трехволнового резонанса волн Толлмина–Шлихтинга дискретного спектра, т.е. имеем в итоге тройное разложение [20] поля скоростей турбулентного течения. При этом для некогерентной части получается, в общем случае, система “кинетических” уравнений, которая при определенных предположениях может быть сведена к одному.

Пусть d – характерный поперечный масштаб течения, L – продольный. Поскольку при оценке величин в исходных уравнениях сравниваются силовые характеристики, то $d \sim \delta^{**}$, δ^{**} – толщина потери импульса. Обозначим через U и V средние продольную и поперечную компоненты скорости, через u, v, w, p – пульсационные компоненты скорости и давление, U_∞ – скорость набегающего потока, ν – коэффициент кинематической вязкости. Введем безразмерные величины: $\bar{U} = U/U_\infty$, $\varepsilon^2 \bar{V} = V/U_\infty$, $\varepsilon \bar{u}_i = u_i/U_\infty$, $\varepsilon \bar{p} = p/\rho U_\infty^2$, $\bar{x}_i = x_i/d$, $\bar{t} = tU_\infty/d$, $X = x/L$, $T = tU_\infty/L$, $\{u_i\} = (u, v, w)$, $\{x_i\} = (x, y, z)$, $i = 1, 2, 3$, $R = U_\infty d/\nu$, $\varepsilon^2 = d/L \ll 1$. Тогда уравнения для \bar{U} и \bar{V} в приближении пограничного слоя запишутся в виде

$$\frac{\partial \bar{U}}{\partial T} + \bar{U} \frac{\partial \bar{U}}{\partial X} + \bar{V} \frac{\partial \bar{U}}{\partial \bar{y}} = \frac{\partial}{\partial \bar{y}} \left(-\overline{uv} + \frac{1}{\varepsilon^2 R} \frac{\partial \bar{U}}{\partial \bar{y}} \right), \quad \frac{\partial \bar{U}}{\partial X} + \frac{\partial \bar{V}}{\partial \bar{y}} = 0,$$

$$\overline{u\bar{v}} = \lim_{t_1 \rightarrow \infty} \frac{1}{t_1} \int_0^{t_1} \left(\lim_{t_0 \rightarrow \infty} \frac{1}{t_0} \int_0^{t_0} \langle \bar{u}\bar{v} \rangle dt_0 \right) dt_1, \quad t_0 = t/(d/U_\infty), \quad t_1 = \varepsilon t_0$$

(здесь $\langle \dots \rangle$ – среднее по ансамблю), а для пульсаций \bar{u}_i, \bar{p} , $i = 1, 2, 3$,

$$\frac{\partial \bar{u}_i}{\partial \bar{t}} + \bar{U} \frac{\partial \bar{u}_i}{\partial \bar{x}_1} + f_i = -\frac{\partial \bar{p}}{\partial \bar{x}_i} + \frac{1}{R} \nabla^2 \bar{u}_i + \varepsilon \zeta_i - \varepsilon^2 Q_i + o(\max[\varepsilon^2, 1/R]),$$

$$\frac{\partial \bar{u}_i}{\partial \bar{x}_i} = 0, \quad \mathbf{f} = \left(u_2 \frac{\partial U}{\partial y}, 0, 0 \right),$$

$$Q_1 = \bar{u} \frac{\partial \bar{U}}{\partial X} + \bar{V} \frac{\partial \bar{u}}{\partial \bar{y}}, \quad Q_2 = \bar{v} \frac{\partial \bar{V}}{\partial \bar{y}} + \bar{V} \frac{\partial \bar{v}}{\partial \bar{y}} = \frac{\partial \bar{v}\bar{V}}{\partial \bar{y}}, \quad Q_3 = \bar{V} \frac{\partial \bar{w}}{\partial \bar{y}}.$$

В уравнениях пульсаций использованы следующие обозначения: $\zeta_i = \partial(\overline{u_i u_j} - \bar{u}_i \bar{u}_j)/\partial \bar{x}_j$, $i, j = 1, 2, 3$ (по одинаковым индексам предполагается суммирование). В качестве граничных условий для сформулированной системы уравнений примем: $\bar{u}_i = 0$, $i = 1, 2, 3$, при $\bar{y} = 0, \infty$, $\bar{p} = 0$ при $\bar{y} = \infty$. Относительно числа Рейнольдса R предполагается, что $R \gg 1$, $\varepsilon^2 R \gg 1$.

Уравнения для пульсаций приводятся к системе из двух уравнений: уравнению Орра–Зоммерфельда для вертикальной компоненты скорости и уравнению Сквайра для вертикальной компоненты завихренности. Благодаря тому, что разложение по собственным функциям линейных частей указанных уравнений обладает полнотой, можно решение задачи представить в виде рядов по собственным функциям.

Исключая давление из уравнений для пульсаций, получим для вертикальных компонентов пульсационной скорости \bar{v} и завихренности $\bar{\eta}$ соответственно

$$\begin{aligned} \left(\frac{\partial}{\partial \bar{t}} + \bar{U} \frac{\partial}{\partial \bar{x}}\right) \Delta \bar{v} - \frac{\partial^2 \bar{U}}{\partial \bar{y}^2} \frac{\partial \bar{v}}{\partial \bar{x}} - \frac{1}{R} \Delta^2 \bar{v} &= \varepsilon D_{Osj} \zeta_j - \varepsilon^2 D_{Osj} Q_j, \\ \left(\frac{\partial}{\partial \bar{t}} + \bar{U} \frac{\partial}{\partial \bar{x}}\right) \bar{\eta} + \frac{d\bar{U}}{d\bar{y}} \frac{\partial \bar{v}}{\partial \bar{z}} - \frac{1}{R} \Delta \bar{\eta} &= \varepsilon D_{Sj} \zeta_j - \varepsilon^2 D_{Sj} Q_j, \\ (D_{Osj}) &= \left(-\frac{\partial^2}{\partial \bar{x} \partial \bar{y}}, \frac{\partial^2}{\partial \bar{x}^2} + \frac{\partial^2}{\partial \bar{z}^2}, -\frac{\partial^2}{\partial \bar{z} \partial \bar{y}}\right), \quad (D_{Sj}) = \left(\frac{\partial}{\partial \bar{z}}, 0, -\frac{\partial}{\partial \bar{x}}\right). \end{aligned}$$

При этом компоненты \bar{u} , \bar{w} , которые остались неизвестными, определяются из соотношений

$$\frac{\partial \bar{u}}{\partial \bar{x}} + \frac{\partial \bar{w}}{\partial \bar{z}} = -\frac{\partial \bar{v}}{\partial \bar{y}}, \quad \frac{\partial \bar{u}}{\partial \bar{z}} - \frac{\partial \bar{w}}{\partial \bar{x}} = \bar{\eta}.$$

Систему уравнений для вертикальных компонентов скорости и завихренности можно упростить, выразив продольную и поперечную компоненты через вертикальную скорость. Приближенно вклад вертикальной завихренности можно учесть в виде индуцированной части решения уравнения Сквайра для вертикальной завихренности в предположении отсутствия резонансов между модами Т–Ш и модами уравнения Сквайра. Ищем решение в виде фурье-представления для вертикальной скорости пульсаций

$$v_{\omega k}(\mathbf{r}, t) = A_k(t_1, t_2) \varphi_{\omega k}(y) e^{-i\omega(\mathbf{k})t_0 + i\mathbf{k} \times \mathbf{r}_0}, \quad \mathbf{k} = (\alpha, \beta), \quad \mathbf{r} = (x, z).$$

$$\frac{\partial A_k}{\partial t} = -\varepsilon \int \bar{H}_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2} A_{\mathbf{k}_1} A_{\mathbf{k}-\mathbf{k}_1} d\mathbf{k}_1 + \varepsilon^2 \left(\Omega_k A_k - \int \bar{F}_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3} A_{\mathbf{k}_1} A_{\mathbf{k}_2} A_{\mathbf{k}-\mathbf{k}_1-\mathbf{k}_2} d\mathbf{k}_1 d\mathbf{k}_2 \right) + o(\varepsilon^2), \quad (1)$$

$$\Omega_k = \text{Im}(\bar{\omega}_l(\mathbf{k})) + h_k, \quad \mathbf{k}_3 = \mathbf{k} - \mathbf{k}_1 - \mathbf{k}_2.$$

Здесь $v_k(\mathbf{r}, t)$ – фурье-компонент вертикальной скорости пульсаций, A_k – фурье-компонент амплитуды вертикальной скорости пульсаций, $\varphi(y)$ – нормированная по норме L_2 наименее затухающая собственная функция дискретного спектра уравнения Орра–Зоммерфельда с профилем развитого турбулентного пограничного слоя, $\bar{\omega}_l(\mathbf{k}) = \omega_l(\mathbf{k}) / \max_{\mathbf{k}} |\omega_l(\mathbf{k})|$, $\omega(\mathbf{k})$ – собственная частота, $\omega(\mathbf{k}) = i\omega_l(\mathbf{k}) + \omega_R(\mathbf{k})$, $i^2 = -1$, $\bar{H}_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2}$, $\bar{F}_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3}$ – матричные элементы, определяемые через интегралы от собственных функций [16], h_k – члены, обусловленные малыми продольными градиентами компонентов средней скорости ТПС.

Уравнение (1) содержит малый параметр $\varepsilon^2 = \frac{\delta^{**}}{L}$, где δ^{**} – толщина потери импульса, L – характерный продольный масштаб длины. Для асимптотического согласования членов уравнения вводится соотношение:

$$\frac{\delta^{**}}{L} \sim \frac{\delta^{**} |\text{Im}(\omega^{**}(\mathbf{k}, \delta^{**}))|}{U_\infty}. \quad (2)$$

Условие (2), с физической точки зрения, показывает, что учитывается затухание волн Т–Ш на масштабе L .

В связи с наличием параметра ε появляется иерархия масштабов времени:

$$\tau_0 = \frac{\delta^{**}}{U_\infty}, \quad \tau_1 = \frac{\tau_0}{\varepsilon}, \quad \tau_2 = \frac{\tau_0}{\varepsilon^2}, \quad \dots,$$

которые определяют безразмерные переменные $t_0 = t/\tau_0$, $t_1 = t/\tau_1$, $t_2 = t/\tau_2, \dots$

Когерентная часть пульсаций определяется с помощью соотношения

$$A_k = A_k^c + A_k', \quad A_k^c = \langle A_k \rangle. \quad (3)$$

Далее представим скорость в пограничном слое в виде тройной суммы [20]:

$$\mathbf{u} = \mathbf{U} + \mathbf{u}^c + \mathbf{u}', \quad \mathbf{U} = \frac{1}{t_1} \int_0^{t_1} \left(\frac{1}{t_0} \int_0^{t_0} \langle \mathbf{u} \rangle dt_0 \right) dt_1, \tag{4}$$

$$\langle \mathbf{u}' \rangle = 0, \quad \frac{1}{t_1} \int_0^{t_1} \left(\frac{1}{t_0} \int_0^{t_0} \langle \mathbf{u}^c \rangle dt_0 \right) dt_1 = 0, \quad \mathbf{u} = (u, v, w).$$

Здесь \mathbf{U} – средняя скорость, \mathbf{u}^c – когерентная составляющая, \mathbf{u}' – стохастическая составляющая. Амплитуды волн, в соответствии с принципом тройной декомпозиции, также представляются в виде суммы когерентной и некогерентной частей, для которых получается система уравнений, содержащих малый параметр ϵ . Его наличие позволяет искать решение уравнений (1) для амплитуд в виде разложения по этому параметру.

Для решения уравнения (1) воспользуемся методом многих масштабов [21]. Представим производную по времени амплитуды $A_{\mathbf{k}}^c$, в виде разложения по параметру ϵ :

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial t_0} + \epsilon \frac{\partial}{\partial t_1} + \epsilon^2 \frac{\partial}{\partial t_2} + \dots, \quad A_{\mathbf{k}}^c = A_{\mathbf{k}}^{c(0)} + \epsilon A_{\mathbf{k}}^{c(1)} + \epsilon^2 A_{\mathbf{k}}^{c(2)} + \dots$$

Собирая члены при одинаковых степенях ϵ , в итоге получаем последовательность из трех уравнений. Разрешая эти уравнения последовательно в масштабах t_0, t_1 и исключая секулярные члены, получим из первых двух:

$$A_{\mathbf{k}}^{c(0)} = A_{\mathbf{k}}^{c(0)}(t_1, t_2), \quad \frac{\partial A_{\mathbf{k}}^{c(0)}}{\partial t_1} = - \int H_{\mathbf{k}\mathbf{k}_1\mathbf{k}_2} \Xi(\Omega_3) A_{\mathbf{k}_1}^{c(0)} A_{\mathbf{k}_2}^{c(0)} d\mathbf{k}_1, \tag{5}$$

$$\mathbf{k}_2 = \mathbf{k} - \mathbf{k}_1, \quad \Omega_3 = \omega(\mathbf{k}) - \omega(\mathbf{k}_1) - \omega(\mathbf{k}_2), \quad \Xi(x) = 1, \quad x = 0, \quad \Xi(x) = 0, \quad x \neq 0.$$

Отметим свойство уравнения (5) для $A_{\mathbf{k}}^{c(0)}$ в масштабе t_1 . Оно инвариантно относительно замены переменных

$$A_{\mathbf{k}}^{c(0)}(t_1, t_2) = \phi(t_2) \tilde{A}_{\mathbf{k}}^{c(0)}(t'_1), \quad t'_1 = \phi(t_2) t_1. \tag{6}$$

В итоге решение для когерентной части представляется в виде (множественный трехволновой резонанс [19]):

$$A_{\mathbf{k}}^{c(0)} = a_1 \delta(\mathbf{k} - \mathbf{k}_1^0) + \sum_{l=1}^m \left\{ a_2^l \delta(\mathbf{k} - \mathbf{k}_2^{0l}) + a_3^l \delta(\mathbf{k} - \mathbf{k}_3^{0l}) \right\} +$$

$$+ a_1^* \delta(\mathbf{k} + \mathbf{k}_1^0) + \sum_{l=1}^m \left\{ a_2^{*l} \delta(\mathbf{k} + \mathbf{k}_2^{0l}) + a_3^{*l} \delta(\mathbf{k} + \mathbf{k}_3^{0l}) \right\}, \tag{7}$$

$$\omega_{\mathbf{k}_1^0}^{(0)} = \omega_{\mathbf{k}_2^{0l}}^{(0l)} + \omega_{\mathbf{k}_3^{0l}}^{(0l)}, \quad \mathbf{k}_1^0 = \mathbf{k}_2^{0l} + \mathbf{k}_3^{0l}, \quad \omega_{-\mathbf{k}}^{(0)} = -\omega_{\mathbf{k}}^{*(0)}, \quad \varphi_{-\mathbf{k}}^{(0)} = \varphi_{\mathbf{k}}^{*(0)}, \quad l = 1, \dots, m,$$

где $\mathbf{k}_i^{(0l)}$, $i = 1, 2, 3, l = 1, \dots, m$, – множество векторов, удовлетворяющих уравнениям трехволнового резонанса, l – номера субгармоник, m – число субгармоник, $\delta(x)$ – дельта-функция Дирака. В итоге, в случае множественного трехволнового резонанса получаем автономную систему обыкновенных дифференциальных уравнений [12]:

$$\frac{da_1}{dt_1} = \sum_{l=1}^m \Lambda_1^l a_2^l a_3^l, \quad \frac{da_2^l}{dt_1} = \Lambda_2^l a_1 a_3^{l*}, \quad \frac{da_3^l}{dt_1} = \Lambda_3^l a_1 a_2^{l*}, \quad l = 1, 2, \dots, m, \tag{8}$$

где $\Lambda_j^l, j = 1, 2, 3, l = 1, 2, \dots, m$ – коэффициенты, определяемые интегралами от произведений собственных функций уравнения Орра–Зоммерфельда, принадлежащих одной моде.

Умножая эти уравнения на соответствующие комплексно сопряженные амплитуды вместе с некоторыми произвольными весовыми множителями $q_s^l, l = 1, 2, \dots, m, s = 1, 2, 3$, и полагая

$$\left(q_1 \Lambda_1 + q_2^l \Lambda_2^{l*} + q_3^l \Lambda_3^{l*} \right) = \left(q_1 \Lambda_1^* + q_2^l \Lambda_2^l + q_3^l \Lambda_3^l \right) = 0, \quad l = 1, \dots, m,$$

получим систему алгебраических уравнений для пар величин

$$q'_{12} = \frac{q'_2}{q_1}, \quad q'_{13} = \frac{q'_3}{q_1}, \quad l = 1, \dots, m. \tag{9}$$

Поведение решения системы (6) дифференциальных уравнений находится в зависимости от значений величин q'_{12} и q'_{13} , которые определяются по коэффициентам Λ_i , $i = 1, 2, 3$. В случае положительности q'_{12} и q'_{13} динамика системы финитна и удовлетворяет динамическому инварианту:

$$I^{(m)}(t_1) = |a_1|^2 + \sum_l \left(q_{12}^{(l)} |a_2^{(l)}|^2 + q_{13}^{(l)} |a_3^{(l)}|^2 \right) = \text{const}, \tag{10}$$

здесь сумма берется по всем субгармоникам в состоянии трехволнового резонанса с основной гармоникой. В случае $q_{12}^{(j)} > 0$, $q_{13}^{(j)} > 0$, инвариант $I^{(m)}(t_1)$ можно преобразовать к уравнению n -мерной ($n = 2 + 8m$) сферы.

Воспользуемся теперь преобразованием (6): $a_i = \phi(t_2) \tilde{a}_i(\phi(t_2)t_1, t_2)$, ϕ – действительная функция от t_2 , которое оставляет неизменным уравнение для амплитуд a_i в масштабе t_1 . В результате применения этого преобразования константу в выражении (10) можно положить равной 1. В этом случае относительно системы уравнений (8) выполняются все условия теоремы Биркгофа–Хинчина в формулировке [15].

2. Приведем здесь теорему Биркгофа–Хинчина в формулировке [15], так как в этой формулировке непосредственно видно соответствие с поставленной задачей.

Пусть состояние рассматриваемой системы представлено точкой P на замкнутом n -мерном многообразии, а координаты $a_1, a_2, a_3, \dots, a_n$ точки P удовлетворяют следующей автономной системе дифференциальных уравнений:

$$\frac{da_i}{dt} = A_i(a_1, a_2, \dots, a_n), \quad i = 1, 2, \dots, n. \tag{11}$$

Уравнения (10) в предположении однозначности их решения определяют для каждого t преобразование T_t любого подмножества E многообразия $M^n : T_t E = E'$, где E' – множество всех точек P' , в которые передвинутся точки P множества E за промежуток времени t . Интегральным инвариантом называется такая функция множества $I(E)$, для которой всегда $I(T_t E) = I(E)$, где $I(M^n) = 1$. Формулировка теоремы имеет вид [15].

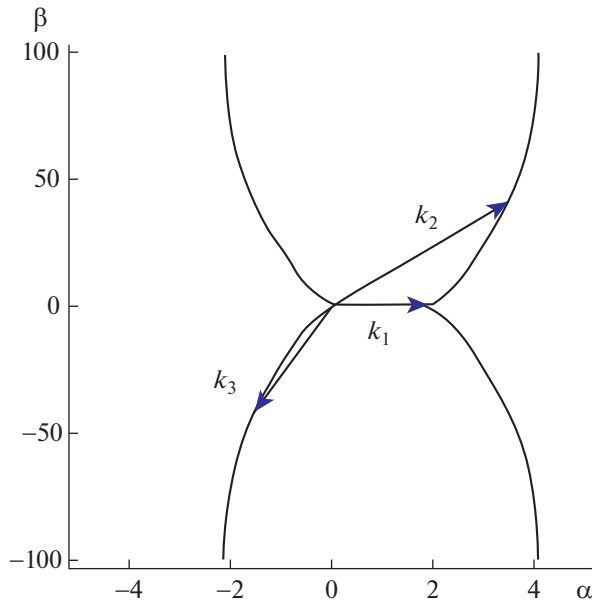
Так что, для любой вещественнозначной функции $f(P)$, определенной на M^n и суммируемой относительно $I(E)$, существует и конечен для всех P , за исключением, возможно, не более чем множества U , для которого $I(U) = 0$ предел:

$$\lim_{c \rightarrow \infty} \frac{1}{c} \int_0^c f(T_t P) dt = \psi(P). \tag{12}$$

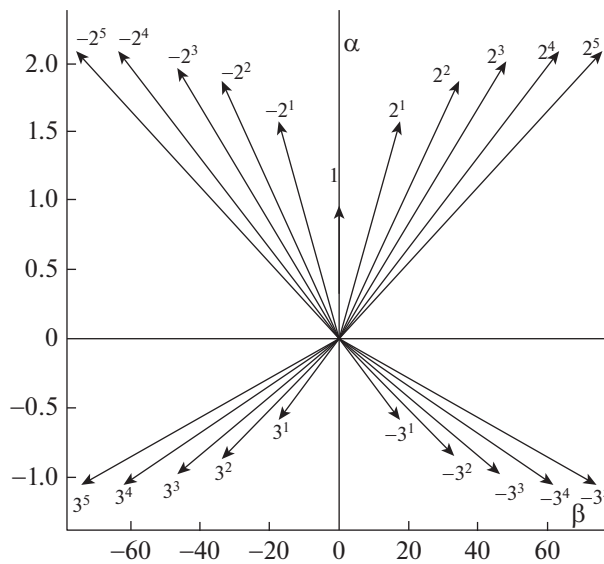
В предыдущем разделе получена автономная система дифференциальных уравнений (8), решение которой удовлетворяет инварианту (10), нормированному на 1. Это и определяет применимость теоремы Биркгофа–Хинчина [10] к данному случаю. Таким образом, динамика автономной системы осуществляется на поверхности n -мерной (далее $n = 42$) единичной сферы.

3. Теорема Биркгофа–Хинчина гарантирует существование средних по времени значений квадратов амплитуд гармоники и конечного множества субгармоник. В данной статье представлены результаты численного определения этих величин с помощью спектральных характеристик волн Толлмина–Шлихтинга (Т–Ш) на профиле средней продольной скорости развитого ТПС [17].

На фиг. 1 представлена кривая 3 волнового резонанса для $\mathbf{k}_1 = (2, 0)$, $\mathbf{k}_2 = (\alpha_2(\beta_2), \beta_2)$, $\mathbf{k}_3 = \mathbf{k}_1 - \mathbf{k}_2$, $\text{Re}_\delta = 10^4$. Кривая представляет решение системы уравнений $\omega(\mathbf{k}_1) = \omega(\mathbf{k}_2) + \omega(\mathbf{k}_3)$, $\mathbf{k}_1 = \mathbf{k}_2 + \mathbf{k}_3$, для наименее затухающей моды волн Т–Ш.

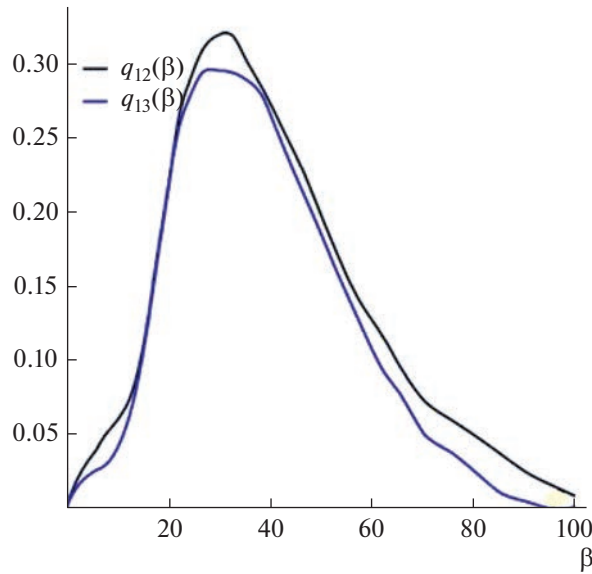


Фиг. 1. Кривая трехволнового резонанса для наименее затухающей моды волн Г–Ш, $Re\delta = 10^4$.



Фиг. 2. Совокупность волновых векторов множественного трехволнового резонанса, $Re\delta = 10^4$.

Фигура 2 показывает, что с продольной волной Г–Ш с волновым вектором \mathbf{k}_0 вступают в резонанс субгармоники с \mathbf{k}_1 и \mathbf{k}_2 , которые имеют волновые вектора, расположенные практически по нормали к плоскости симметрии среднего течения, т.е. волны с продольной завихренностью. Следует обратить внимание на то, что на фиг. 2 по горизонтали отложены единицы, а по вертикали – сотни. Фигура 2 дает геометрическое представление о гармонике и субгармониках в пространстве волновых чисел. Совокупность векторов субгармоник берется по всем резонансным тройкам и симметричным к ним относительно горизонтальной оси. Выбор набора субгармоник проводится из области пространства волновых векторов, в которой множители $q_{12}^{(j)}$, $q_{13}^{(j)}$ положительны (фиг. 3).



Фиг. 3. Зависимость значений множителей $q_{12}(\beta)$ и $q_{13}(\beta)$ на кривой трехволнового резонанса от волнового числа.

На фиг. 3 представлено поведение $q_{12}(\beta)$, $q_{13}(\beta)$ на кривой трехволнового резонанса от поперечного волнового числа β . Графическое совпадение кривых в области $11 < \beta < 21$ не дает представления о небольшом, но вполне устойчивом различии этих величин.

Система уравнений (8) для множественного трехволнового резонанса решалась численно с помощью пакета прикладных программ. Начальные условия для системы (8) задавались точкой, взятой из множества точек, статистически равномерно распределенных на n -мерной ($n = 42$) единичной сфере. В реальности эти граничные условия формируются в нелинейной области перехода. Однако окончательные представления об этом процессе являются еще предметом исследований [22], [23].

Численные результаты показывают, что за очень большой промежуток времени ($0 \leq t \leq 1000$) инвариант меняется в шестом знаке. В результате можно найти средние значения квадратов амплитуд гармоники и субгармоник за достаточно длительную реализацию во времени численно.

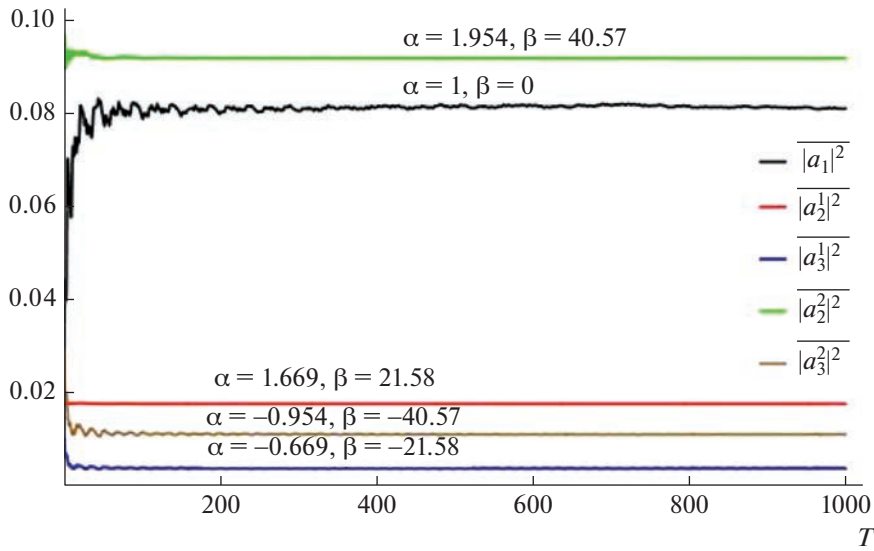
На фиг. 4 представлена зависимость средних квадратов модулей основной гармоники и двух амплитуд субгармоник множественного трехволнового резонанса, описываемого системой уравнений (8), вычисленных за конечный интервал времени T . Средние величины определены как функции интервала осреднения:

$$\overline{|a_l^s|^2} = \left(|a_l^s|^2 \right)_{t_1} = \frac{1}{T} \int_0^T |a_l^s(t_1)|^2 dt_1, \quad l = 1, 2, 3, \quad s = 1, 2, \dots, m.$$

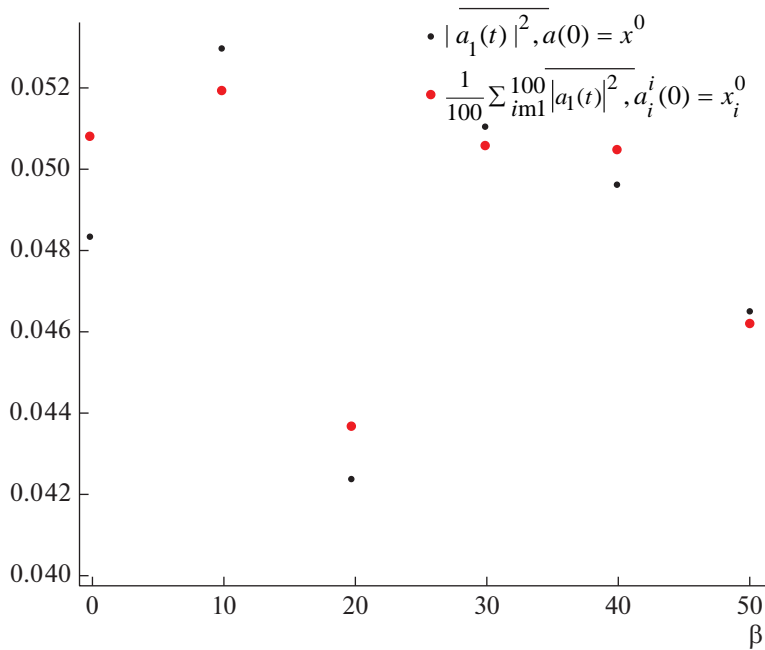
Здесь представлены типичные зависимости от времени амплитуды $\left(|a_1(t_1)|^2 \right)_{t_1}$ основной гармоники при $\beta = 0$ и двух амплитуд субгармоник $\left(|a_1^1(t_1)|^2 \right)$, $\left(|a_2^1(t_1)|^2 \right)$ при $\beta = (50.12, 75.8)$ соответственно, из набора множественного трехволнового резонанса первой моды.

На фиг. 5 представлено сравнение среднего квадрата модулей амплитуд гармоники и субгармоник по времени со средними по времени, усредненными по случайному набору начальных данных (100 вариантов).

В табл. 1 приведено сравнение между средним значением по времени и усредненным по пространству при пяти гармониках.



Фиг. 4. Временные зависимости средних значений квадрата амплитуд основной гармоники $a_1(t)$, $\beta = 0$ и $a_1^1(t)$, $a_2^1(t)$, и двух субгармоник $a_{12}(t)$, $a_{22}(t)$, при $\beta = (-50.12, -75.8)$, где $\left(|a_i^s|^2 \right)_{t_1}$ – средние по времени в масштабе τ_1 значения квадрата модуля амплитуд гармоники и субгармоник.



Фиг. 5. Сравнение средних квадратов модулей амплитуд по времени (интервал осреднения $T = 1000$) и по случайному набору начальных данных (100 вариантов) – определяются синими и желтыми точками соответственно.

Сравнение выше полученного результата с усреднением по пространству $\left(|a_i|^2 \right)_S = \frac{1}{\sigma(S)} \int_S |a_i|^2 ds \cong 0.047619$, S есть n -мерная единичная сфера ($n = 42$), $\sigma(S)$ – площадь сферы S , ds – элемент площади), показывает, что эти средние близки друг к другу, поэтому можно предположить, что выполняется условие эргодичности. Однако проверка этого предположения тре-

Таблица 1. Сравнение между средним значением по времени и усредненным по пространству при пяти гармониках

l	$\overline{ a_l ^2}$, $a(t=0) = x^0$ – Случайная точка на единичной сфере	$\frac{1}{100} \sum_{i=1}^{100} \overline{ a_i^{(l)} ^2}$, $a_i^{(l)}(t=0) = x_i^0$ $l = 1, \dots, 5$	Δ
$l = 1$	0.0529904	0.0519488	9%
$l = 2$	0.0423804	0.0436673	8%
$l = 3$	0.0510327	0.0505745	6.2%
$l = 4$	0.049629	0.0505255	5.1%
$l = 5$	0.046456	0.0461685	3%

будет дальнейших более точных вычислений при получении решения системы (8) и численного определения средних по времени. Возможно также, что необходимо рассматривать большее число субгармоник.

4. Рассмотренная модель ТПС является моделью с самосогласованным полем, так как за основу берется известное распределение средней продольной скорости. Поэтому первая проверка модели должна отвечать на вопрос, получаем ли мы исходное распределение средней продольной скорости. Такая проверка была проведена в работе [26], в которой показано, что экспериментально определенное касательное напряжение вместе с уравнениями Рейнольдса и реологическими соотношениями модели дают правильное распределение продольной скорости, совпадающее с экспериментальным [27].

Дальнейшая проверка связана с определением средних компонент тензора напряжений. Решение задачи представимо в виде $\mathbf{u} = \bar{\mathbf{U}} + \mathbf{u}_c + \mathbf{u}'$ ($\mathbf{u} = (u, v, w)$), $\bar{\mathbf{U}} = (\bar{U}, \bar{V}, 0)$, $\mathbf{u}' = (u', v', w')$ тройной декомпозиция [20]), где $\bar{\mathbf{U}}$ – среднее значение по времени и ансамблю и соответствующую флуктуирующую часть $\mathbf{u}_c + \mathbf{u}'$ где \mathbf{u}_c – когерентная и \mathbf{u}' – стохастическая составляющие турбулентных пульсаций. Для упрощения записи далее индекс u когерентной составляющей опущен. Стохастические величины полагаются равными нулю. Ниже используется определение (4) средних [11], с помощью которого получают формулы для средних пульсаций продольной скорости и компонент тензора напряжений.

Используя уравнение непрерывности и приближенное выражение для индуцированного решения уравнения Сквайра нормальной составляющей завихренности $\eta_{\mathbf{k}\omega}$, можно получить компоненты Фурье продольной скорости и завихренности.

$$u_{\omega\mathbf{k}} = \frac{i}{\mathbf{k}^2} \left(\alpha \frac{dv_{\omega\mathbf{k}}}{dy} - \beta \eta_{\omega\mathbf{k}} \right), \quad \eta_{\omega\mathbf{k}} = \frac{i}{\mathbf{k}^2} \left(\frac{\beta v_{\omega\mathbf{k}}}{-\omega(\mathbf{k}) + \alpha \bar{U}(y)} \frac{d\bar{U}}{dy} \right), \tag{13}$$

$$v_{\omega\mathbf{k}} = A_{\mathbf{k}}^{c(0)}(t_1) \delta(\omega - \omega_{T-S}(\mathbf{k})) e^{i\omega(\mathbf{k})t_0 - i\mathbf{k} \cdot \mathbf{r}_0} \varphi_0(\mathbf{k}, y). \tag{14}$$

Здесь $A_{\mathbf{k}}^{c(0)}$ определяется формулой (7), ω – собственное значение спектральной задачи для уравнения Орра–Зоммерфельда, v, η – вертикальные компоненты пульсаций скорости и завихренности, $\mathbf{k} = (\alpha, \beta)$, α, β – соответственно продольное и поперечное волновые числа, $A_{\mathbf{k}}^{c(0)}$ – фурье-компонент амплитуды вертикальной составляющей скорости в развитом турбулентном пограничном слое, \mathbf{k}_1 – волновое число основной гармоники, $\mathbf{k}_j^s, j = 2, 3; s = 1, 2, \dots, m$, – волновые числа субгармоник, $\delta(x)$ – дельта-функция Дирака, $\varphi_0(\mathbf{k}, y)$ – собственная функция наименее затухающей моды уравнения Орра–Зоммерфельда.

Суммирование осуществляется по всем субгармоникам в состоянии множественного трехволнового резонанса, где $a_i^s, i = 1, 2, 3, s = 1, 2, \dots, m$, – амплитуды гармоники и субгармоник в состоянии трехволнового резонанса. Амплитуды этих волн удовлетворяют системе дифференциальных уравнений (8). Система амплитуд $a_i^s, i = 1, 2, 3, s = 1, 2, \dots, m$ удовлетворяет инварианту,

представляющему сумму квадратов модулей амплитуд этих волн, умноженных на весовые множители. С помощью преобразования (6)

$$a_i(\tilde{t}_1, t_2) = \phi(t_2) \tilde{a}_i(\phi(t_2)t_1, t_2), \quad i = 1, 2, 3, \quad \phi(t_2) = \sqrt{\tilde{\lambda}(t_2)/\tilde{\Phi}(t_2)},$$

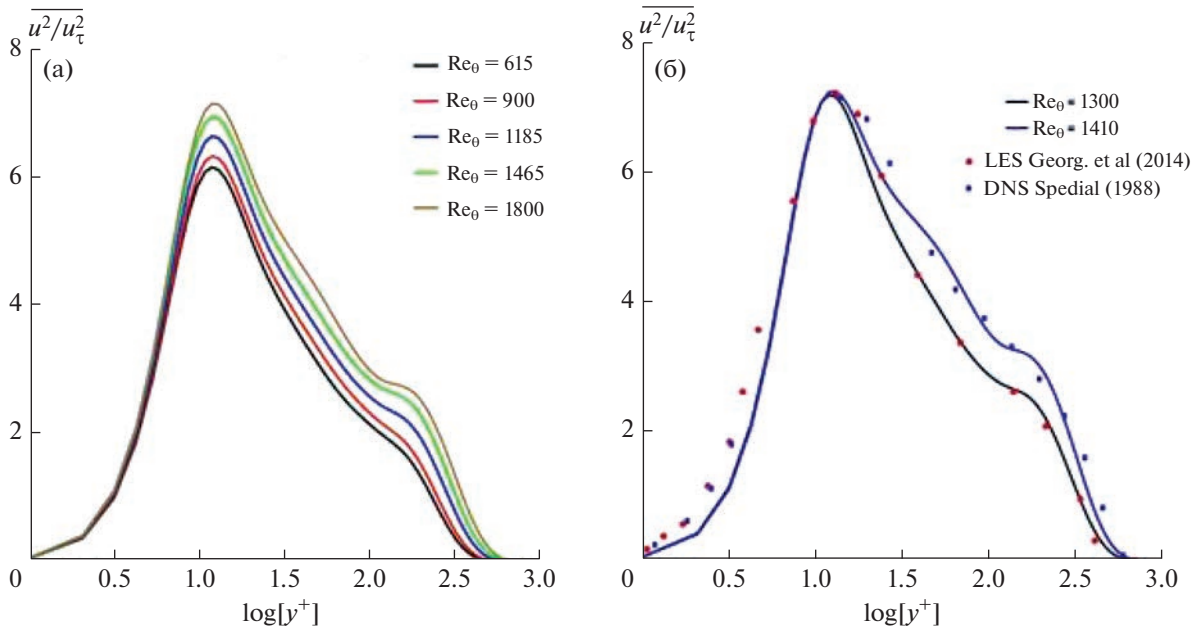
и линейной замены переменных инвариант приводится к единице, здесь $\tilde{\lambda}$ и $\tilde{\Phi}$ определены в работе [12] и вычисляются с помощью статистики 4-го порядка в статистически стационарном случае.

Из уравнений (13), (14) вычисляются квадраты модулей амплитуд волн, и усреднение с помощью (13) по (t_0, t_1) дает компоненты пульсационных скоростей:

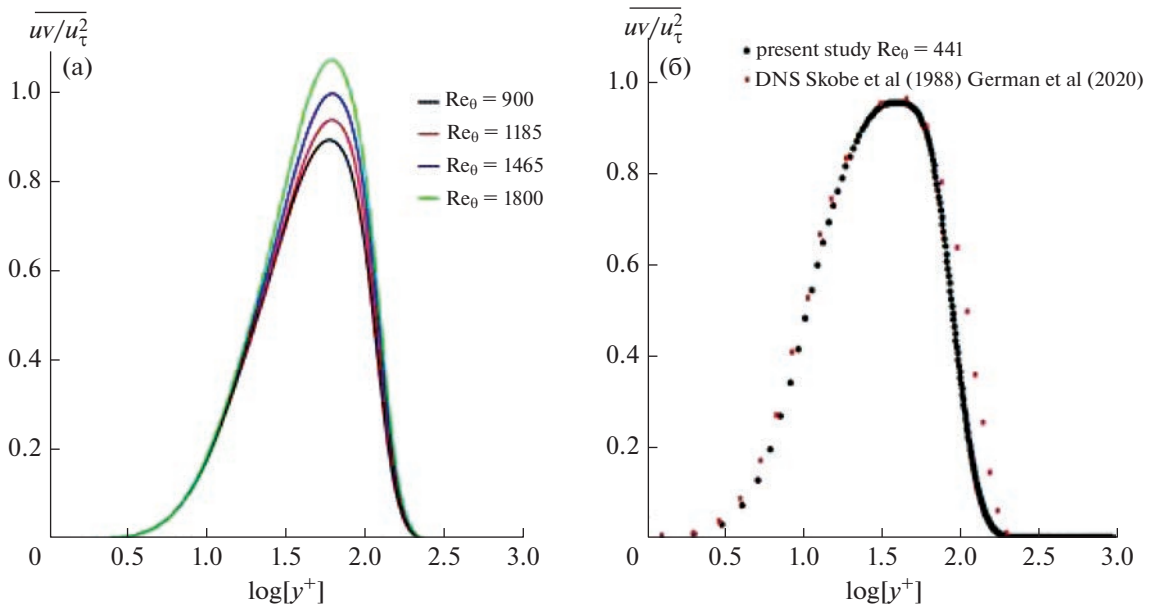
$$\begin{aligned} \overline{u^2} &= \frac{\phi^2}{(2\pi)^3} \left(f_1(\mathbf{k}_1, y) + 2f_2(\mathbf{k}_l^s, y) \frac{dU}{dy} + 2f_3(\mathbf{k}_l^s, y) \left(\frac{dU}{dy} \right)^2 \right), \quad (15) \\ f_1(\mathbf{k}_1, y) &= -\frac{(|a_1|^2)_{t_1} \alpha_1^2}{|\mathbf{k}_1|^4} |\varphi_1'(\mathbf{k}_1, y)|^2 - 2 \sum_{s=1}^m \sum_{l=2}^3 \frac{(|a_l^s|^2)_{t_1}}{q_{l,l}^s} (\alpha_l^s)^2 \frac{|\varphi_l'(\mathbf{k}_l^s, y)|^2}{|\mathbf{k}_l^s|^4}, \\ f_2(\mathbf{k}_l^s, y) &= -\frac{(|a_1|^2)_{t_1} \alpha_1 \beta_1^2}{|\mathbf{k}_1|^4} \left(\frac{\varphi_1^*(\mathbf{k}_1, y) \varphi_1(\mathbf{k}_1, y)}{\omega_r(\mathbf{k}_1) - \alpha_1 U(y) + i\omega_i(\mathbf{k}_1)} + \frac{\varphi_1'(\mathbf{k}_1, y) \varphi_1'^*(\mathbf{k}_1, y)}{\omega_r(\mathbf{k}_1) - \alpha_1 U(y) - i\omega_i(\mathbf{k}_1)} \right) - \\ &\quad - \sum_{s=1}^m \sum_{l=2}^3 \frac{(|a_l^s|^2)_{t_1} \alpha_l^s (\beta_l^s)^2}{q_{l,l}^s |\mathbf{k}_l^s|^4} \left(\frac{\varphi_l'^*(\mathbf{k}_l^s, y) \varphi_l'(\mathbf{k}_l^s, y)}{\omega_r(\mathbf{k}_l^s) - \alpha_l^s U(y) + i\omega_i(\mathbf{k}_l^s)} + \frac{\varphi_l'(\mathbf{k}_l^s, y) \varphi_l'^*(\mathbf{k}_l^s, y)}{\omega_r(\mathbf{k}_l^s) - \alpha_l^s U(y) - i\omega_i(\mathbf{k}_l^s)} \right), \\ f_3(\mathbf{k}_l^s, y) &= -\frac{(|a_1|^2)_{t_1} \beta_1^4}{|\mathbf{k}_1^s|^4 \left((\omega_r(\mathbf{k}_1) - \alpha_1 U(y))^2 + \omega_i^2(\mathbf{k}_1) \right)} |\varphi_1(\mathbf{k}_1, y)|^2 - \\ &\quad - \sum_{s=1}^m \sum_{l=2}^3 \frac{(|a_l^s|^2)_{t_1} (\beta_l^s)^4}{q_{l,l}^s |\mathbf{k}_l^s|^4 \left((\omega_r(\mathbf{k}_l^s) - \alpha_l^s U(y))^2 + \omega_i^2(\mathbf{k}_l^s) \right)} |\varphi_l(\mathbf{k}_l^s, y)|^2, \\ \tau = -\overline{uv} &= \frac{\phi^2}{(2\pi)^3} \left(g_1(\mathbf{k}_1, y) + 2 \sum_{s=1}^m \sum_{l=2}^3 \frac{(|a_l^s|^2)_{t_1}}{|\mathbf{k}_l^s|^2 q_{l,l}^s} \left(g_2(\mathbf{k}_l^s, y) + g_3(\mathbf{k}_l^s, y) \frac{dU}{dy} \right) \right), \quad (16) \\ g_1(\mathbf{k}_1, y) &= \frac{(|a_1|)_{t_1} \alpha_1}{|\mathbf{k}_1|} \operatorname{Im} \left(\varphi_1^*(\mathbf{k}_1, y) \varphi_1'(\mathbf{k}_1, y) \right), \quad g_2(\mathbf{k}_l^s, y) = \alpha_l^s \operatorname{Im} \left(\varphi_l'^*(\mathbf{k}_l^s, y) \varphi_l'(\mathbf{k}_l^s, y) \right), \\ g_3(\mathbf{k}_l^s, y) &= \frac{(\beta_l^s)^2 |\varphi_l(\mathbf{k}_l^s, y)|^2 \omega_i(\mathbf{k}_l^s)}{\left((\omega_r(\mathbf{k}_l^s) - \alpha_l^s U(y))^2 + \omega_i^2(\mathbf{k}_l^s) \right)}, \end{aligned}$$

где $\varphi'(\mathbf{k}_l^s, y) = \frac{d\varphi(\mathbf{k}_l^s, y)}{dy}$, $q_{l,l}^s$, $l = 2, 3, \dots$ – весовые множители (8), суммирование по s распространяется на все субгармоники, которые лежат на кривой трехволнового резонанса.

Здесь $(|a_l^s|^2)_{t_1}$ – средние по времени в масштабе τ_1 значения квадратов модулей амплитуд гармоники и субгармоник. (Подчеркнем, что время t_0 входит в определение вертикальной скорости, и, соответственно, амплитуд a_l^s , только через множитель $e^{i\omega t_0 - i\mathbf{k} \cdot \mathbf{r}_0}$, поэтому усреднение по t_0 при-



Фиг. 6. (а) – профили величины $\overline{u^2}/u_\tau^2$ при $\mathbf{k}_1 = (0.5, 0)$, $Re_0 = 615 - 1800$. (б) – сравнение $\overline{u^2}/u_\tau^2$: черная, синяя линия – настоящее исследование при $\mathbf{k}_1 = (0.5, 0)$, $Re_0 = 1300, 1410$; (“●”) – LES [28], при $Re_0 = 1300$; (“■”) – DNS [29], $Re_0 = 1410$.



Фиг. 7. (а) – Напряжение Рейнольдса при $\mathbf{k}_1 = (0.5, 0)$, отнесенное к квадрату динамической скорости, при различных значениях числа Рейнольдса: (—) $Re_0 = 900$; (—) $Re_0 = 1185$; (—) $Re_0 = 1465$. (—) $Re_0 = 1800$. (б) – Сравнение напряжения при $\mathbf{k}_1 = (0.5, 0)$, с результатами DNS: (“●”) – настоящее исследование; (“■”) [30, 31] при $Re_0 = 441$.

водит к тому, что в окончательных формулах остаются только квадраты модулей указанных выше амплитуд.) Усреднение по времени в масштабе τ_1 производится по формуле (13) с $T = 1000$.

Члены в правой части формул (15), (16) зависят от значений $q_{l\nu}^s$, средних значений амплитуд гармоник и субгармоник по времени, определенных выше, профиля средней продольной ско-

рости, его производной по y , собственных значений и собственных функций. Действительная и мнимая части собственных значений наименее затухающей моды обозначаются соответственно $\text{Re}[\omega] = \omega_r(k)$, $\text{Im}[\omega] = \omega_i(k)$. Кроме того, в формулах (6)–(8) учтена симметрия по β .

На фиг. 6а, б приведено поведение квадрата модуля продольной пульсации (6) в ТПС при пяти различных числах Рейнольдса от $\text{Re}_0 = 615$ до $\text{Re}_0 = 1800$. Полученные результаты относительно $\overline{u^2}/u_r^2$ хорошо согласуются как с моделированием методом крупных вихрей [28], так и с данными прямого численного моделирования [29] по всей толщине пограничного слоя.

Напряжения сдвига Рейнольдса показаны на фиг. 7а при различных значениях числа Рейнольдса. Кроме того, видно хорошее согласие с данными прямого численного моделирования [30] при $\text{Re}_0 = 441$ (7б).

ВЫВОДЫ

В работе представлена волноводная модель развитого турбулентного пограничного слоя, которая получена с некоторыми физическими допущениями из уравнений Навье–Стокса. К параметрам, определяющим пограничный слой, добавлена длина затухания наименее затухающей моды дискретного спектра уравнения Орра–Зоммерфельда. В итоге в системе определяющих масштабов возникает малый параметр, и к исходной системе уравнений Навье–Стокса оказалось возможным применение метода многих масштабов. Асимптотический метод позволяет приближенно решить полученные уравнения (т.е. корректно замкнуть стохастическую задачу) и получить решения исходной задачи в виде величин, удовлетворяющих системе дифференциальных уравнений для организованной части и интегро-дифференциального уравнения для стохастической части пульсаций. Этот результат качественно согласуется с имеющимися в настоящее время экспериментальными наблюдениями по DeTu. Кроме того, полученные уравнения позволяют вычислить средний тензор пульсационных напряжений и, при решении осредненных уравнений гидродинамики (уравнений Рейнольдса) с полученными реологическими связями, восстановить профиль средней продольной скорости в пограничном слое. Таким образом, волноводная модель, дополненная решением полученной системы уравнений для стохастического компонента пульсаций, вполне может воспроизводить среднюю динамику газа в турбулентном пограничном слое.

СПИСОК ЛИТЕРАТУРЫ

1. *Kachanov Y.S.* Physical mechanisms of laminar boundary layer transition // *Annu. Rev. Fluid Mech.* 1994. V. 26. P. 411–482.
2. *Реник Е.У., Соседко Ю.П.* Исследование прерывистой структуры течения в пристенной области турбулентного пограничного слоя. Турбулентные течения. – М.: Наука, 1974, 226 с.
3. *Robinson S.K.* Coherent motions in the turbulent boundary layer // *Annu. Rev. Fluid Mech* 1991. V. 23. P. 601–639.
4. *Бойко А.В., Грек Г.Р., Довгаль А.В., Козлов В.В.* Возникновение турбулентности в пристенных течениях. Новосибирск: Наука, 1999, 328 с.
5. *Белоцерковский О.М., Хлопков Ю.И., Жаров В.А., Горелов С.Л., Хлопков А.Ю.* Организованные структуры в турбулентных течениях. Анализ экспериментальных работ по турбулентному пограничному слою. М.: МФТИ, 2009, 302 с.
6. *Borodulin V.I., Kachanov Y.S., Roschekhtayev A.P.* The deterministic wall turbulence is possible // *Advances in Turbulence XI. Proceedings of 11th EUROMECH European Turbulence Conference, June 25–28, 2007, Porto, Portugal* // J.M.L.M. Palma and A. Silva Lopes, eds. – Heidelberg: Springer, 2007. P. 176–178.
7. *Borodulin V.I., Kachanov Y.S., Roschekhtayev A.P.* Experimental detection of deterministic turbulence // *Journal of Turbulence*. 2011. V. 12. № 23. P. 1–34.
8. *Borodulin V.I., Kachanov Y.S.* On the reproducibility of instantaneous and statistical characteristics of the deterministic turbulence // *Theoretical and Applied Mechanics Letters* // 2014. V. 4, 062004.
9. *Borodulin V.I., Kachanov Y.S.* Experimental Study of Reproducibility of Instantaneous Structure of the Deterministic Wall Turbulence // *Proceedings of 8th Intl. Symposium on Turbulence and Shear Flow Phenomena-TS-FP-8. E.N.S.M.A. Poitiers, France. 2013. P. 1–6.*
10. Онлайн-курс лекций д.ф.-м.н., профессора Юрия Семеновича Качанова “ИСТОКИ ТУРБУЛЕНТНОСТИ. ФИЗИЧЕСКАЯ ПРИРОДА”. Лекция 7. 11.03.2022: “Детерминированная турбулентность – новый подход к исследованию турбулентности”. <https://www.youtube.com/channel/UCyT-kQLOsbkb5tecI-STYAbQ>

11. *Жаров В.А.* О волновой теории развитого турбулентного пограничного слоя // Ученые записки ЦАГИ. 1986. Т. XVII, № 5. С. 28–38.
12. *Жаров В.А.* Волноводная модель когерентной и стохастической составляющих развитого турбулентного пограничного слоя. // Труды ЦАГИ. 2014, вып. 2731. С. 3–50.
13. *Жаров В.А., Липатов И.И., Селим Р.С.* Волноводная модель организованных структур в турбулентном пограничном слое на пластине с нулевым продольным градиентом давления. // Ученые записки ЦАГИ. 2020. Т. VI. № 6. С. 51–59.
14. *Каток А.Б., Хасселблат Б.* Введение в современную теорию динамических систем. М.: “ФАКТОРИАЛ”, 1999. С. 768.
15. *Колмогоров А.Н.* Избранные труды Математика и Механика. М: Наука, 1985, 569 с.
16. *Жаров В.А.* Модельное представление когерентной структуры в развитом турбулентном пограничном слое // Ученые записки ЦАГИ. 2014. Т. XLV. № 5. С. 33–46.
17. *Musker A.J.* Explicit expression for the smooth wall velocity distribution in turbulent boundary layer // AIAA Journal. 1979. V. 17 (6). P. 655–657.
18. *Landahl M.T.* A wave-guide model for turbulent shear flow // J. Fluid Mech. 1967. V. 29. Pt. 3. P. 441–459.
19. *Zelman M.B.* Tollmien–Schlichting–wave resonant mechanism for subharmonic–type transition // J. Fluid Mech. 1993. V. 252. P. 449–478.
20. *Hussain A.K.M.F., Reynolds W.C.* The mechanics of an organized wave in turbulent shear flow. Part 2. Experimental results // J. Fluid Mech. 1972. V. 54. P. 241–261.
21. *Nayfeh A.H.* Perturbation Methods. // Wiley. 1973. 425 p. (Найфэ А. Методы возмущений. М.: Наука. 1986. 454 с.)
22. *Zhang W., Liu P., Guo H.* Conditional Sampling and Wavelet Analysis in Early Stage of Step-Generated Transition // AIAA J. 2018. V. 56. P. 2471–2477.
23. *Pushpender K.S., Tapan K.S.* Effect of frequency and wave number on the three-dimensional routes of transition by wall excitation // Phys. Fluids. 2019. V. 31. P. 64–107.
24. *Mathematica 5.0 User’s Guide.* Wolfram Research, 2003. P. 1301.
25. *Селим Р.С.* Собственные моды уравнения Орра-Зоммерфельда в развитом турбулентном пограничном слое // Труды МАИ. 2019. Вып № 109, URL: <http://trudymai.ru/published.php?ID=111352>.
26. *Жаров В.А., Ле Ван Ха.* Построение профиля продольной скорости по заданному касательному напряжению в развитом турбулентном пограничном слое на пластине. Ученые записки ЦАГИ, 2016. № 8. Т. XLVII. С. 50–60.
27. *Klebanoff P.S.* Characteristics of turbulence in a boundary layer with zero pressure gradient. NACA Rep. 1247 (1955) (Шлихтинг Г. Теория пограничного слоя. М.: Физматлит. 1974. 712 с.)
28. *Spalart P.R.* Direct simulation of a turbulent boundary layer up to $Re_\theta = 14000$ // J. Fluid Mech. 1988. V. 137. P. 61–98.
29. *Jimenez J., Hoyas S., Simens M., Mizuno Y.* Turbulent boundary layer and channels at moderate Reynold number // J. Fluid Mech. 2010. V. 667. P. 335–360.
30. *Skote M., Henningson Dan S., and Henkes R.A.W.M.* Direct numerical simulation of self-similar turbulent boundary layers in adverse pressure gradients // Flow Turbulence and Combustion. 1998. V.60. № 1. P. 47–85.
31. *German S., Guillermo A.* Reynolds shear stress modeling in turbulent boundary layers’ subject to very strong Favorable Pressure Gradient // Computers and Fluids. 2020. <https://doi.org/10.1016/j.compfluid.2020.104494>

МАТЕМАТИЧЕСКАЯ
ФИЗИКА

УДК 517.988

МЕТОД КВАЗИРЕШЕНИЙ И ПРОБЛЕМА ГЛОБАЛЬНОЙ
МИНИМИЗАЦИИ ФУНКЦИОНАЛА НЕВЯЗКИ
УСЛОВНО КОРРЕКТНЫХ ОБРАТНЫХ ЗАДАЧ

© 2023 г. М. Ю. Кокурин^{1,*}

¹ 424001 Йошкар-Ола, пл. Ленина, 1, Марийский государственный университет, Россия

*e-mail: kokurinm@yandex.ru

Поступила в редакцию 20.11.2021 г.

Переработанный вариант 12.11.2022 г.

Принята к публикации 02.02.2023 г.

Рассматривается класс условно корректных задач, характеризуемый гёльдеровой оценкой условной устойчивости на выпуклом компакте в гильбертовом пространстве. Оператор прямой задачи и правая часть уравнения заданы с погрешностями, близость производных точного и возмущенного операторов не предполагается. Исследуются свойства выпуклости и одноэкстремальности функционала невязки метода квазирешений. Для этого функционала устанавливается, что каждая его стационарная точка на множестве условной корректности, не слишком далекая от искомого решения исходной обратной задачи, лежит в малой окрестности решения. Даны оценки диаметра указанной окрестности в терминах погрешностей входных данных. Показано, что эта окрестность является аттрактором итераций метода проекции градиента, и получены оценки скорости сходимости итераций к аттрактору. Устанавливается необходимость используемой оценки условной устойчивости для существования итерационных процессов с указанными свойствами. Библ. 16.

Ключевые слова: обратная задача, условно корректная задача, метод квазирешений, глобальная оптимизация, оценка точности, эффект кластеризации, метод проекции градиента, аттрактор, скорость сходимости.

DOI: 10.31857/S0044466923050137, **EDN:** PKJKKS

1. ПОСТАНОВКА ЗАДАЧИ

В работе рассматриваются нелинейные обратные задачи, описываемые операторными уравнениями

$$F(x) = f, \quad x \in D, \quad (1.1)$$

где $F : H_1 \rightarrow H_2$ – оператор прямой задачи, предполагаемый дифференцируемым по Фреше на выпуклом компактном множестве $D \subset H_1$; H_1, H_2 – вещественные гильбертовы пространства. В приложениях множество D определяет априорные ограничения на искомый элемент $x \in H_1$. Этот элемент описывает подлежащий определению набор параметров исследуемой модели. Элемент $f \in H_2$ описывает результаты наблюдения и выполняет роль входных данных для обратной задачи реконструкции параметров модели. Всюду далее считаем, что оператор F инъективен на множестве D и $f \in F(D)$. Обозначим через $x^* = F^{-1}(f)$ искомое решение задачи (1.1). Ниже будем предполагать, что производная F' удовлетворяет условию Липшица

$$\|F'(x) - F'(y)\|_{L(H_1, H_2)} \leq L \|x - y\|_{H_1} \quad \forall x, y \in D. \quad (1.2)$$

Из (1.2) следует, что

$$\|F(x) - F(y)\|_{H_2} \leq M \|x - y\|_{H_1}, \quad \|F'(x)\|_{L(H_1, H_2)} \leq M \quad \forall x, y \in D. \quad (1.3)$$

В (1.3) можно положить $M = \|F'(\bar{x})\|_{L(H_1, H_2)} + Ld_0$, $\bar{x} \in D$. Здесь и далее $d_0 = \text{diam}_{H_1}(D)$,

$$\text{diam}_X(\Omega) \triangleq \sup\{\|x - y\|_X : x, y \in \Omega\}, \quad \Omega \subset X.$$

Через $\|\cdot\|_X$ обозначаем норму в гильбертовом или банаховом пространстве X , $(\cdot, \cdot)_X$ есть скалярное произведение в гильбертовом пространстве X . Следствием (1.2) и интегральной теоремы о конечных приращениях (см. [1, § 32]) является представление

$$\begin{aligned} F'(x)(x - y) &= F(x) - F(y) + G(x, y), \\ \|G(x, y)\|_{H_2} &\leq \frac{1}{2}L\|x - y\|_{H_1}^2, \quad x, y \in D. \end{aligned} \quad (1.4)$$

Ввиду неизбежных погрешностей измерения, элемент f и оператор F , как правило, бывают заданы приближенно, так что вместо них доступны аппроксимации $\tilde{f} \in H_2$ и $\tilde{F} : H_1 \rightarrow H_2$. Конечномерные аппроксимации F , используемые при построении численно реализуемых процедур решения (1.1), в ряде случаев также удобно рассматривать в качестве приближенных операторов \tilde{F} . Типичные требования к качеству аппроксимации сводятся к тому, что (см. [2, гл. 2, § 4])

$$\|\tilde{f} - f\|_{H_2} \leq \delta, \quad (1.5)$$

оператор \tilde{F} дифференцируем по Фреше и, кроме того,

$$\|\tilde{F}(x) - F(x)\|_{H_2} \leq h, \quad (1.6)$$

$$\|\tilde{F}'(x) - F'(x)\|_{L(H_1, H_2)} \leq h \quad \forall x \in D. \quad (1.7)$$

Величины δ, h в (1.5)–(1.7) характеризуют уровень погрешности в задании элемента f и оператора F соответственно.

На практике условие (1.7), предполагающее равномерную близость производных исходного и возмущенного операторов, часто оказывается излишне ограничительным. Например, при исследовании проекционных схем аппроксимации уравнения (1.1) вводятся операторы $\tilde{F}(x) = F(\mathcal{P}_N x)$, где $\{\mathcal{P}_N\}$ есть семейство ортопроекторов из H_1 на конечномерные подпространства \mathcal{H}_N :

$$\mathcal{H}_N \subset \mathcal{H}_{N+1}, \quad N = 1, 2, \dots; \quad \overline{\bigcup_{N=1}^{\infty} \mathcal{H}_N} = \mathcal{H}_1.$$

В этом случае условие (1.6) имеет место с $h = M\gamma_N(D)$, где

$$\gamma_N(D) = \max\{\|(E - \mathcal{P}_N)x\|_{H_1} : x \in D\}.$$

При этом $\lim_{N \rightarrow \infty} \gamma_N(D) = 0$ в силу компактности множества D . Однако, если оператор F не является вполне непрерывным, то условие (1.7), вообще говоря, выполняется лишь при $h \geq h_1 > 0$ с фиксированной постоянной h_1 , в чем убеждает пример $F(x) = x$. Эти соображения мотивируют введение следующего ослабленного по сравнению с (1.7) требования:

$$\|\tilde{F}'(x)\|_{L(H_1, H_2)} \leq M_0, \quad \|(\tilde{F}'(x) - F'(x))(x - x^*)\|_{H_2} \leq \mu \quad \forall x \in D. \quad (1.8)$$

Здесь μ – мера возмущения производной оператора F , предполагаемая достаточно малой. Нетрудно видеть, что условие (1.8) является следствием (1.7), при этом можно принять $\mu = d_0 h$. Пусть теперь $\tilde{F}(x) = F(\mathcal{P}_N x)$, и неравенства (1.2), (1.3) имеют место для всех $x, y \in O_{R_0}(0) \supset D$:

$$O_R(x) \triangleq \{y \in H_1 : \|x - y\|_{H_1} \leq R\}.$$

Тогда (1.8) выполнено с $M_0 = M$ и $\mu = \mu_N = (Ld_0 + 2M)\gamma_N(D)$, при этом $\lim_{N \rightarrow \infty} \mu_N = 0$. В дальнейшем будем считать, что $h + \mu + \delta \leq \Delta_0$ для некоторого фиксированного Δ_0 .

Характерное свойство многих прикладных обратных задач, в частности, обратных задач математической физики (см. [3–6]) состоит в том, что обратный оператор F^{-1} как отображение из H_2 в H_1 не является непрерывным в точках из $F(D)$ в том смысле, что из сходимости

$$\lim_{n \rightarrow \infty} \|f_n - f\|_{H_2} = 0, \quad f_n \in H_2, \quad f \in F(D),$$

не следует сходимость элементов $F^{-1}(f_n)$, если они определены, к $F^{-1}(f)$ в метрике H_1 . В этом случае уравнение (1.1) является некорректной по Адамару задачей. Широко распространенным подклассом некорректных обратных задач (1.1) является класс условно корректных (корректных по А.Н. Тихонову) на D задач (см. [3], [4], [7], [8]). Условная корректность задачи (1.1) на множестве D по определению означает, что оператор F^{-1} относительно непрерывен на $F(D)$, т.е. равенство

$$\lim_{n \rightarrow \infty} \|f_n - f\|_{H_2} = 0, \quad f_n, f \in F(D),$$

влечет $\lim_{n \rightarrow \infty} \|F^{-1}(f_n) - F^{-1}(f)\|_{H_1} = 0$. Здесь требуется, чтобы аппроксимирующие $f \in F(D)$ элементы f_n также лежали в $F(D)$. В наших предположениях (1.1) условно корректна ввиду компактности D .

Классический подход к построению устойчивых процедур аппроксимации решений некорректных задач связан с понятием регуляризирующего оператора (алгоритма). Ниже обратимся к методу квазирешений В.К. Иванова (см. [3], [9]). Регуляризирующий оператор, отвечающий методу квазирешений, имеет вид

$$R(\tilde{F}, \tilde{f}) = \arg \min_{x \in D} \tilde{J}(x), \quad \tilde{J}(x) = \frac{1}{2} \|\tilde{F}(x) - \tilde{f}\|_{H_2}^2. \quad (1.9)$$

Таким образом, в качестве приближения к искомому решению x^* выбирается элемент из D , доставляющий глобальный минимум функционалу \tilde{J} на множестве D . Ввиду непрерывности оператора \tilde{F} и компактности D , множество \tilde{X}^* решений задачи (1.9) не пусто. При этом каждый элемент $x = \tilde{x}^* \in \tilde{X}^*$ удовлетворяет необходимому условию минимума (см. [10, гл. 1, § 2])

$$(\tilde{J}'(\tilde{x}), \tilde{x} - x)_{H_1} \leq 0 \quad \forall x \in D. \quad (1.10)$$

Градиент функционала \tilde{J} имеет вид

$$\tilde{J}'(x) = \tilde{F}'^*(x)(\tilde{F}(x) - \tilde{f}). \quad (1.11)$$

Точки $\tilde{x} \in D$, удовлетворяющие условию (1.10), будем называть стационарными точками функционала \tilde{J} на D .

Схема (1.9), как и близкая к ней схема А.Н. Тихонова (см. [9]), труднореализуема численно в случае произвольного гладкого оператора \tilde{F} . Причина состоит в том, что задача (1.9), вообще говоря, многоэкстремальная и, следовательно, трудноразрешима стандартными итерационными методами. Это же относится и к конечномерным аппроксимациям (1.9), используемым при построении численно реализуемых методов решения (1.1). Непосредственное применение к указанным задачам классических итерационных методов минимизации в лучшем случае приводит лишь к некоторой стационарной точке, близость которой к глобальному минимуму в (1.9) априори не гарантируется. Отсутствие свойств выпуклости минимизируемого функционала в (1.9) создает дополнительные трудности при исследовании поведения стандартных процедур минимизации (см., например, [10, гл. 1, § 4]). При сделанных выше предположениях большинство таких процедур гарантируют лишь сходимость подходящих подпоследовательностей к соответствующим стационарным точкам. Тем не менее ниже будет показано, что для широкого класса задач за счет обозримых дополнительных условий на их характеристики может быть обеспечена определенная регулярность итераций метода проекции градиента. Нам понадобится требование равномерной условной корректности задачи (1.1).

Условие 1. Имеет место степенная оценка условной устойчивости задачи (1.1) с показателем $p \in [1, 2]$:

$$\|F(x) - F(y)\|_{H_2} \geq m \|x - y\|_{H_1}^p \quad \forall x, y \in D; \quad m > 0. \quad (1.12)$$

Условие 1 выполняется для многих обратных задач математической физики (см., например, [5], [6]).

Далее будет установлено, что любая не слишком удаленная от x^* стационарная точка задачи (1.9) на самом деле лежит в малой окрестности искомого решения x^* . Тем самым, малая окрестность x^* содержит кластер стационарных точек (1.9), удаленный от других стационарных точек

этой задачи. Будет дана оценка диаметра этой окрестности в терминах уровней погрешности h , μ , δ . Кроме того покажем, что указанная окрестность является аттрактором итераций метода проекции градиента и получим оценки скорости сходимости итерационной последовательности к этому аттрактору в зависимости от значений p .

План работы следующий. В разд. 2 в предположении $h = 0$ исследуются свойства локальной сильной выпуклости и одноэкстремальности функционала невязки (1.9) в зависимости от значений показателя p из условия 1. В разд. 3 приводятся примеры образования кластеров стационарных точек \tilde{J} на D в окрестности решения при возмущениях оператора, удовлетворяющих условиям (1.7) и (1.8). Раздел 4 посвящен доказательству основных результатов о локализации кластеров стационарных точек и сходимости к ним итераций метода проекции градиента. В разд. 5 устанавливается необходимость условия 1 для существования итерационных процессов с указанными свойствами.

2. СЛУЧАЙ ТОЧНО ЗАДАННОГО ОПЕРАТОРА

Изучим свойства локальной выпуклости и одноэкстремальности функционала задачи (1.9). Содержательные результаты здесь удается получить лишь в случае точно заданного оператора, т.е. при $\tilde{F} = F$. В большинстве случаев приходится также предполагать, что $\tilde{f} = f$. Рассмотрим отдельно два возможных случая, отличающихся значением показателя p в условии (1.12).

1) Пусть $p = 1$. Убедимся, что при $h = 0$ и достаточно малом δ функционал \tilde{J} является сильно выпуклым в окрестности точного решения x^* уравнения (1.1). Воспользуемся известным критерием сильной выпуклости (см. [10, гл. 1, § 2]). Для произвольных $x, y \in D$ с учетом (1.4) имеем

$$\begin{aligned} (\tilde{J}'(x) - \tilde{J}'(y), x - y)_{H_1} &= (F'^*(x)(F(x) - \tilde{f}) - F'^*(y)(F(y) - \tilde{f}), x - y)_{H_1} = \\ &= (F(x) - F(y), F'(x)(x - y))_{H_2} + (F(y) - \tilde{f}, (F'(x) - F'(y))(x - y))_{H_2} = \\ &= (F(x) - F(y), F(x) - F(y) + G(x, y))_{H_2} + (F(y) - \tilde{f}, (F'(x) - F'(y))(x - y))_{H_2} \geq \\ &\geq \|F(x) - F(y)\|_{H_2}^2 - \frac{1}{2}L\|F(x) - F(y)\|_{H_2}\|x - y\|_{H_1}^2 - L\|F(y) - \tilde{f}\|_{H_2}\|x - y\|_{H_1}^2. \end{aligned}$$

Отсюда с использованием (1.12) при $p = 1$ и (1.3), (1.5) получаем

$$\begin{aligned} (\tilde{J}'(x) - \tilde{J}'(y), x - y)_{H_1} &\geq \|F(x) - F(y)\|_{H_2}^2 - \frac{L}{2m^2}\|F(x) - F(y)\|_{H_2}^3 - L(\|F(y) - f\|_{H_2} + \delta) \times \\ &\times \|x - y\|_{H_1}^2 \geq \|F(x) - F(y)\|_{H_2}^2 \left(1 - \frac{LM}{2m^2}\|x - y\|_{H_1}\right) - L(M\|y - x^*\|_{H_1} + \delta)\|x - y\|_{H_1}^2. \end{aligned} \tag{2.1}$$

Пусть теперь $x, y \in D \cap O_R(x^*)$. Тогда из (2.1) следует

$$(\tilde{J}'(x) - \tilde{J}'(y), x - y)_{H_1} \geq \|F(x) - F(y)\|_{H_2}^2 \left(1 - \frac{LMR}{m^2}\right) - L(MR + \delta)\|x - y\|_{H_1}^2. \tag{2.2}$$

Предположим, что

$$2LMR + L\delta \leq m^2 - \nu, \quad \nu > 0. \tag{2.3}$$

Используя (1.12), из (2.2) и (2.3) получаем

$$(\tilde{J}'(x) - \tilde{J}'(y), x - y)_{H_1} \geq (m^2 - 2LMR - L\delta)\|x - y\|_{H_1}^2 \geq \nu\|x - y\|_{H_1}^2.$$

Тем самым доказана следующая теорема, устанавливающая локальную сильную выпуклость \tilde{J} в окрестности решения x^* .

Теорема 1. Пусть $h = 0$, выполняются условие 1 при $p = 1$ и условия (1.5), (2.3). Тогда функционал \tilde{J} является сильно выпуклым на $D \cap O_R(x^*)$.

Следствие 1. В условиях теоремы 1 функционал \tilde{J} имеет не более одной стационарной точки в $D \cap O_R(x^*)$.

Условие (2.3) выполняется с $R = (4LM)^{-1}(m^2 - \nu)$, если $\delta \leq (2L)^{-1}(m^2 - \nu)$, $0 < \nu < m^2$.

Пусть \tilde{x}^* есть произвольная точка глобального минимума \tilde{J} на D . Тогда $\|F(\tilde{x}^*) - \tilde{f}\|_{H_2} \leq \|F(x) - \tilde{f}\|_{H_2}$ для всех $x \in D$. Полагая здесь $x = x^*$ и используя (1.12), получаем

$$\|\tilde{x}^* - x^*\|_{H_1} \leq \frac{2\delta}{m}. \quad (2.4)$$

Следствие 2. Пусть выполнены условия теоремы 1 и условие

$$\delta \leq \frac{1}{2} Rm.$$

Тогда множество $\tilde{X}^* = \{\tilde{x}^*\}$ решений задачи (1.9) одноточечное. В области сильной выпуклости $D \cap O_R(x^*)$ функционала \tilde{J} лежит единственная его стационарная точка \tilde{x}^* и имеет место оценка (2.4).

2) Пусть $p \in (1, 2]$. В этом случае локальная сильная выпуклость функционала невязки уже не может быть гарантирована даже при $\tilde{f} = f$. Тем не менее в случае точных данных для функционала

$$J(x) = \frac{1}{2} \|F(x) - f\|_{H_2}^2 \quad (2.5)$$

имеет место

Теорема 2. Пусть $h = \delta = 0$ и выполняется условие 1.

1. Предположим, что $p \in (1, 2)$. Тогда для всех $x \in D \cap O_R(x^*)$ при

$$R = \frac{1}{M} \left(\frac{2(1-\varepsilon)m^{2/p}}{L} \right)^{p/(2-p)}, \quad \varepsilon \in (0, 1), \quad (2.6)$$

справедливо

$$(J'(x), x - x^*)_{H_1} \geq \varepsilon m^2 \|x - x^*\|_{H_1}^{2p}. \quad (2.7)$$

2. Пусть $p = 2$ и выполняется условие

$$L \leq 2(1-\varepsilon)m, \quad \varepsilon \in (0, 1). \quad (2.8)$$

Тогда для всех $x \in D$ с указанным значением p имеет место оценка (2.7).

Доказательство. Для произвольных $x \in D$, $p \in (1, 2]$ с использованием (1.4) получаем

$$(J'(x), x - x^*)_{H_1} = (F(x) - f, F'(x)(x - x^*))_{H_2} \geq \|F(x) - f\|_{H_2}^2 - \frac{1}{2} L \|F(x) - f\|_{H_2} \|x - x^*\|_{H_1}^2.$$

Отсюда с учетом (1.12) следует

$$(J'(x), x - x^*)_{H_1} \geq \|F(x) - f\|_{H_2}^2 \left(1 - \frac{L}{2m^{2/p}} \|F(x) - f\|_{H_2}^{(2-p)/p} \right). \quad (2.9)$$

Используя (1.3), (2.6), (2.8), на основании (2.9) заключаем, что при выполнении условий теоремы справедливо

$$(J'(x), x - x^*)_{H_1} \geq \varepsilon \|F(x) - f\|_{H_2}^2. \quad (2.10)$$

Требуемая оценка непосредственно следует из (1.12), (2.10). Теорема доказана.

Точка $x^* = F^{-1}(f)$, доставляющая глобальный минимум функционалу J на D , является стационарной для J . Следующее утверждение устанавливает отсутствие других стационарных точек J в окрестности x^* .

Следствие 3. В условиях п. 1 теоремы 2 функционал J не имеет в $D \cap O_R(x^*)$ других стационарных точек кроме x^* . В условиях п. 2 теоремы 2 x^* является единственной стационарной точкой J на всем множестве D .

Для доказательства заметим, что если \hat{x} — стационарная точка функционала (2.5), то согласно (1.10) $(J'(\hat{x}), \hat{x} - x^*)_{H_1} \leq 0$. Требуемое утверждение непосредственно следует теперь из (2.7).

В следующем примере покажем, что результат следствия 3 не допускает непосредственного обобщения на случай $p > 2$ и, кроме того, условие (2.8) не может быть ослаблено.

Пример 1. Рассмотрим инъективное отображение $F : D \rightarrow \mathbb{R}$, $D = [-a, b]$, $a, b > 0$, $F(x) = b^p - (b - x)^p$, $p \geq 2$. Непосредственно устанавливается оценка

$$|F(x) - F(y)| \geq |x - y|^p, \quad x, y \in D, \tag{2.11}$$

неулучшаемая по порядку $|x - y|$. Кроме того,

$$|F'(x) - F'(y)| \leq p(p - 1)(a + b)^{p-2} |x - y|, \quad x, y \in D. \tag{2.12}$$

Пусть $f = 0$, тогда $x^* = 0$. Функция $J(x) = 0.5F^2(x)$ имеет на D стационарные точки $\hat{x}^{(1)} = 0$, $\hat{x}^{(2)} = b$. Как видно из (2.11), (2.12), в этом примере оценки (1.2), (1.3) и (1.12) выполняются с $L = p(p - 1)(a + b)^{p-2}$, $M = p(a + b)^{p-1}$ и $m = 1$ соответственно. Пусть вначале $p > 2$. Выбирая $R = b$ произвольно малым и варьируя $a > 0$, видим, что окрестность $O_R(x^*) \cap D$ содержит две стационарные точки $\hat{x}^{(1)} = 0$, $\hat{x}^{(2)} = b$ функции J на D . Таким образом, следствия 1–3 непосредственно не обобщаются на случай, когда условие 1 выполнено с показателем $p > 2$. Полагая $p = 2$, имеем $L = 2$, так что условие (2.8) не выполнено ни при каком $\varepsilon > 0$. Вновь для любого $R = b > 0$ имеем две стационарные точки в $O_R(x^*) \cap D$. Следовательно, условие (2.8) в п. 2 следствия 3 не допускает ослабления.

3. МЕТОД ПРОЕКЦИИ ГРАДИЕНТА И КЛАСТЕРИЗАЦИЯ СТАЦИОНАРНЫХ ТОЧЕК ФУНКЦИОНАЛА НЕВЯЗКИ

Далее нас будет интересовать поведение метода проекции градиента

$$x_{n+1} = P_D(x_n - \rho \tilde{J}'(x_n)), \quad x_0 \in D, \quad \rho > 0, \tag{3.1}$$

в применении к задаче (1.9) в предположении, что $h + \delta > 0$. Здесь P_D – оператор метрического проектирования из H_1 на D . Анализ итераций (3.1) предположим несколько примеров, иллюстрирующих вид \tilde{J} и множества стационарных точек \tilde{J} на D в случае приближенного оператора. Следующий пример показывает, что если $p = 1$, то при $h > 0$ в отличие от случая $h = 0$ (теорема 1) сильная выпуклость \tilde{J} на D в наших предположениях может отсутствовать.

Пример 2. Положим $F(x) = x$, $\tilde{F}(x) = x + hn^{-1} \sin(nx)$, $\tilde{f} = f = 0$, $D = [0, 1]$. Условия (1.2), (1.12) очевидно выполнены с $L = 0$, $m = 1$, $p = 1$. Условия (1.6), (1.7) также выполнены. Имеем

$$\tilde{J}''(x) = (1 + h \cos(nx))^2 - hn(x + hn^{-1} \sin(nx)) \sin(nx).$$

Отсюда следует, что при $x = x_m = (2n)^{-1}\pi + 2\pi n^{-1}m$, $m \in \mathbb{Z}$, будет $\tilde{J}''(x_m) = 1 - h(\pi/2 + h) - 2\pi mh$. Выбирая $h = n^{-1/2}$, $m = [n^{3/4}]$, $n \in \mathbb{N}$, получаем $0 < x_m \leq (2n)^{-1}\pi + 2\pi n^{-1/4}$. Поэтому точки x_m лежат в D и сколь угодно близки к $x^* = 0$ при достаточно больших n . Кроме того, для указанных n с подходящей константой $\kappa_0 > 0$ имеем $\tilde{J}''(x_m) \leq -\kappa_0 n^{1/4} < 0$. Таким образом, сильная выпуклость \tilde{J} нарушается в сколь угодно малой окрестности решения x^* . Пример 2 закончен.

Эффект разрушения сильной выпуклости функционала невязки при $p = 1$ проиллюстрируем также ниже в примере 4.

Если $p \in (1, 2]$, то при возмущениях вида (1.6), (1.7) множество стационарных точек \tilde{J} на D может иметь достаточно сложный вид.

Пример 3. В координатном гильбертовом пространстве

$$H_1 = \left\{ \sum_{i=1}^{\infty} x_i e_i : \sum_{i=1}^{\infty} x_i^2 < \infty \right\},$$

определенном ортонормированном базисом $\{e_i\}_{i=1}^\infty$, введем оператор $F : H_1 \rightarrow H_2 = H_1 \times \mathbb{R}^N$:

$$F(x) = (Ax; G(x)), \quad G(x) = (x_1^2, \dots, x_N^2),$$

где $A^* = A \geq O$ – самосопряженный вполне непрерывный оператор с нулевым ядром и полным набором собственных векторов $\{e_i\}_{i=1}^\infty$. Обозначим через $\{\lambda_i\}_{i=1}^\infty$ соответствующие положительные собственные значения, $\lim_{i \rightarrow \infty} \lambda_i = 0$. Норма в H_2 определяется равенством $\|(x; y)\|_{H_2}^2 = \|x\|_{H_1}^2 + \|y\|_{\mathbb{R}^N}^2$, $(x; y) \in H_2$. Положим также $f = (0; 0)$,

$$D = \{x = A^q v, \|v\|_{H_1} \leq d\}, \quad q \geq 1, \quad d > 0.$$

Ясно, что D компактно, оператор F инъективен на D и решением задачи (1.1) является элемент $x^* = 0$. В координатном представлении оператор A имеет вид

$$Ax = \sum_{i=1}^\infty \lambda_i x_i e_i, \quad x = \sum_{i=1}^\infty x_i e_i.$$

Будем считать, что λ_i малы при $1 \leq i \leq N, i \geq M + 1$, где $M > N$, и

$$\max_{1 \leq i \leq N} \lambda_i \leq \max_{i \geq M+1} \lambda_i.$$

Определим приближенный оператор

$$\tilde{F}(x) = (\tilde{A}x; G(x)), \quad \tilde{A}x = \sum_{i=N+1}^M \lambda_i x_i e_i.$$

В данном случае

$$\begin{aligned} \max_{x \in D} \|\tilde{F}(x) - F(x)\|_{H_2} &\leq \|\tilde{A} - A\|_{L(H_1, H_1)} d (\max_{i \geq 1} \lambda_i)^q, \\ \|\tilde{F}'(x) - F'(x)\|_{L(H_1, H_2)} &= \|\tilde{A} - A\|_{L(H_1, H_1)}. \end{aligned}$$

Кроме того, $\|\tilde{A} - A\|_{L(H_1, H_1)} = \max_{i \geq M+1} \lambda_i$. Таким образом, условия (1.6), (1.7) выполняются с $h = \sigma \max_{i \geq M+1} \lambda_i$, $\sigma = \max\{1, d(\max_{i \geq 1} \lambda_i)^q\}$. Положим $\tilde{f} = (0; \delta/\sqrt{N}, \dots, \delta/\sqrt{N})$, так что $\|\tilde{f} - f\|_{H_2} = \delta$. Непосредственно проверяется, что условие Липшица (1.2) выполнено с $L = 2$. Для проверки условия 1 запишем

$$\|F(x) - F(y)\|_{H_2} = (\|A(x - y)\|_{H_1}^2 + \|G(x) - G(y)\|_{\mathbb{R}^N}^2)^{1/2} \geq \|A(x - y)\|_{H_1}. \tag{3.2}$$

Поскольку $x, y \in D$, имеем $x = A^q u, y = A^q v$, где $\|u\|_{H_1}, \|v\|_{H_1} \leq d$. Таким образом,

$$x - y = A^q w, \quad \|w\|_{H_1} = \|u - v\|_{H_1} \leq 2d.$$

Используя неравенство моментов (см. [11, гл. 2, § 8]), далее получаем

$$\|x - y\|_{H_1} = \|A^q w\|_{H_1} \leq \|A^{q+1} w\|_{H_1}^{q/(q+1)} \|w\|_{H_1}^{1/(q+1)} \leq \|A(x - y)\|_{H_1}^{q/(q+1)} (2d)^{1/(q+1)}.$$

Отсюда с учетом (3.2) следует

$$\|F(x) - F(y)\|_{H_2} \geq (2d)^{-1/q} \|x - y\|_{H_1}^{(q+1)/q}, \quad x, y \in D.$$

Видим, что условие 1 выполняется с $m = (2d)^{-1/q}$ и показателем $p = (q + 1)/q$. При соответствующем выборе $q \geq 1$ величина p принимает любые значения в $(1, 2]$. Функционал \tilde{J} в рассматриваемом случае имеет вид

$$\tilde{J}(x) = \frac{1}{2} \sum_{i=N+1}^M \lambda_i^2 x_i^2 + \frac{1}{2} \sum_{i=1}^N (x_i^2 - \delta/\sqrt{N})^2.$$

Поскольку

$$(\tilde{J}'(x), \eta)_{H_1} = \sum_{i=N+1}^M \lambda_i^2 x_i \eta_i + 2 \sum_{i=1}^N (x_i^2 - \delta/\sqrt{N}) x_i \eta_i, \quad \eta \in H_1,$$

имеем следующий набор множеств в H_1 , на которых $\tilde{J}'(x) = 0$:

$$\mathcal{M}(\bar{x}_1, \dots, \bar{x}_N) = \left\{ (\bar{x}_1, \dots, \bar{x}_N, 0, \dots, 0, x_{M+1}, x_{M+2}, \dots) : \sum_{i=M+1}^{\infty} x_i^2 < \infty \right\}.$$

Здесь $\bar{x}_i \in \{0, \pm\sqrt{\delta/\sqrt{N}}\}$, $1 \leq i \leq N$, выбираются произвольно. Таким образом, градиент \tilde{J}' обращается в нуль на 3^N непересекающихся множествах $\mathcal{M}(\bar{x}_1, \dots, \bar{x}_N)$. Нетрудно указать условия, при которых все эти множества имеют общие точки с D . Достаточное условие для $\mathcal{M}(\bar{x}_1, \dots, \bar{x}_N) \cap D \neq \emptyset$ для всех 3^N наборов $\{\bar{x}_1, \dots, \bar{x}_N\}$ имеет вид

$$\frac{\delta}{\sqrt{N}} \sum_{i=1}^N \frac{1}{\lambda_i^{2q}} \leq \Delta^2 < d^2. \tag{3.3}$$

Здесь множество стационарных точек \tilde{J} на D содержит семейство \hat{X}^* из 3^N дизъюнктивных компактов коразмерности M , полученных параллельным переносом компакта

$$\mathcal{H} = \left\{ x \in H_1 : x_i = 0, 1 \leq i \leq M, \sum_{i=M+1}^{\infty} \frac{x_i^2}{\lambda_i^{2q}} \leq d^2 - \Delta^2 \right\} \tag{3.4}$$

в точки $(\bar{x}_1, \dots, \bar{x}_N, 0, 0, \dots)$. Оценим расстояние найденных фрагментов множества стационарных точек \tilde{J} на D до решения x^* . Используя (3.3), (3.4), получаем, что для любой точки $x \in \hat{X}^*$ выполняется

$$\|x\|_{H_1}^2 = \sum_{i=1}^N x_i^2 + \sum_{i=M+1}^{\infty} x_i^2 \leq \sqrt{N}\delta + \sum_{i=M+1}^{\infty} x_i^2 \leq \sigma^{-2q} \Delta^2 h^{2q} + \sigma^{-2q} (d^2 - \Delta^2) h^{2q} = \sigma^{-2q} d^2 h^{2q}.$$

Таким образом, для всех $h \in (0, h_0]$ с фиксированным h_0 имеет место оценка

$$\text{dist}_{H_1}(x^*, \hat{X}^*) \leq dh^q \leq C_1 h^{1/p}. \tag{3.5}$$

Пример 3 закончен.

Через C_1, C_2, \dots обозначаем фиксированные постоянные, не зависящие от h, μ, δ . В (3.5) и далее используется обозначение $\text{dist}_{H_1}(x, \Omega) = \inf \{\|x - y\|_{H_1} : y \in \Omega\}$, $\Omega \subset H_1$.

Следующий пример иллюстрирует появление обширного кластера стационарных точек в окрестности решения в случае $p = 1$ в условиях (1.6), (1.8).

Пример 4. В пространстве H_1 из примера 3 рассмотрим компакт

$$D = \left\{ x = \sum_{i=1}^{\infty} x_i e_i : |x_i| \leq \frac{1}{2^i}, i = 1, 2, \dots \right\}$$

и тождественное отображение $F(x) = x$, $x \in H_1 = H_2$. В данном случае (1.2) и условие 1 выполнены очевидным образом, при этом $m = 1, p = 1, L = 0$. Положим $f = 0$, тогда $x^* = 0$. В качестве приближенных данных выберем

$$\tilde{F}(x) = \sum_{i=1}^N x_i e_i + \sum_{i=N+1}^{\infty} x_i^2 e_i, \quad \tilde{f} = \sum_{i=N+1}^{\infty} \delta_i e_i, \quad 0 < \delta_i \leq \frac{1}{4^i}, \quad i \geq N + 1.$$

Тогда

$$\|\tilde{f} - f\|_{H_1} \leq \delta = \frac{1}{4^N \sqrt{15}}.$$

Кроме того, при $x \in D$ справедливо

$$\|\tilde{F}(x) - F(x)\|_{H_1}^2 = \sum_{i=N+1}^{\infty} (x_i^2 - x_i)^2 \leq \sum_{i=N+1}^{\infty} \frac{1}{2^{2(i-1)}} = \frac{1}{3 \times 4^{N-1}}.$$

Таким образом, условие (1.6) выполняется при $h = 1/(2^{N-1}\sqrt{3})$. Оценим величину μ в условии (1.8). Имеем

$$(\tilde{F}'(x) - F'(x))(x - x^*) = \sum_{i=N+1}^{\infty} x_i(2x_i - 1)e_i,$$

$$\|(\tilde{F}'(x) - F'(x))(x - x^*)\|_{H_1}^2 = \sum_{i=N+1}^{\infty} x_i^2(2x_i - 1)^2 \leq 4 \sum_{i=N+1}^{\infty} x_i^2 \leq \frac{1}{3 \times 4^{N-1}}.$$

Поэтому можно принять $\mu = h$. В рассматриваемом случае

$$\tilde{J}(x) = \frac{1}{2} \sum_{i=1}^N x_i^2 + \frac{1}{2} \sum_{i=N+1}^{\infty} (x_i^2 - \delta_i)^2, \quad (\tilde{J}'(x), \eta)_{H_1} = \sum_{i=1}^N x_i \eta_i + 2 \sum_{i=N+1}^{\infty} (x_i^2 - \delta_i) x_i \eta_i, \quad \eta \in H_1. \quad (3.6)$$

Из (3.6) следует, что $\tilde{J}'(x) = 0$ на множестве точек

$$\hat{X}^* = \left\{ x = \sum_{i=N+1}^{\infty} \bar{x}_i e_i : \bar{x}_i \in \{0, \pm\sqrt{\delta_i}\}, i \geq N+1 \right\},$$

имеющем мощность континуума. Указанные точки являются стационарными для \tilde{J} на D . Непосредственно проверяется, что

$$\text{dist}_{H_1}(x^*, \hat{X}^*) \leq \frac{1}{2^N \sqrt{3}} = \frac{1}{2} h. \quad (3.7)$$

Пример 4 закончен.

Отсутствие свойства (сильной) выпуклости минимизируемого функционала значительно затрудняет анализ итераций (3.1) и других итерационных методов решения задачи (1.9) (см. [10]). Известные утверждения о скорости сходимости этих методов по норме получены в предположении о сильной выпуклости целевого функционала. Теоремы 1, 2 утверждают единственность стационарной точки функционала (1.9) в окрестности решения x^* при $h = 0$ или $h = \delta = 0$ в шаре, диаметр которого определяется параметрами исходной задачи (1.1). Примеры 3, 4 показывают, что при замене точных данных (F, f) их приближенными аналогами (\tilde{F}, \tilde{f}) указанное свойство локальной одноэкстремальности в общем случае не имеет места. В целом распределение стационарных точек функционала невязки в случае приближенных данных приобретает следующий вид. Вместо единственной точки глобального минимума функционала J будем иметь кластер стационарных точек \tilde{J} в окрестности x^* . Существенно, что все точки кластера находятся в шаре $O_{r(h,\mu,\delta)}(x^*)$ радиуса

$$r(h, \mu, \delta) = O((h + \mu + \delta)^{1/p}). \quad (3.8)$$

Отмеченное свойство кластеризации в случае $h = \mu = 0$ было ранее обосновано в [12], [13]. Для доказательства утверждения в общем виде ниже покажем, что множество $B_{r(h,\mu,\delta)} = O_{r(h,\mu,\delta)}(x^*) \cap D$ является аттрактором итераций (3.1) в том смысле, что

$$\lim_{n \rightarrow \infty} \text{dist}_{H_1}(x_n, B_{r(h,\mu,\delta)}) = 0. \quad (3.9)$$

Одновременно будет получена оценка скорости сходимости в (3.9), заменяющая классические оценки скорости сильной сходимости градиентных итераций в сильно выпуклых задачах. Основное условие выполнения равенства (3.9) имеет вид $x_0 \in O_{r_0}(x^*) \cap D$, где величина $r_0 = r_0(L, M_0, m)$ определяется параметрами задачи (1.1). Поскольку последовательность (3.1) стационарна по n при старте из любой стационарной точки функционала \tilde{J} , из (3.9) следует, что все стационарные точки задачи (1.9), принадлежащие $O_{r_0}(x^*) \cap D$, лежат исключительно в аттракторе $B_{r(h,\mu,\delta)}$ диаметра $O((h + \mu + \delta)^{1/p})$. Совокупность описанных результатов определяет принцип

кластеризации стационарных точек задачи (1.9) в случае возмущений общего вида. Заметим, что оценки (3.5), (3.7) из примеров 3, 4 согласуются по порядку с размером кластера, указанным в (3.8).

Перейдем к обоснованию анонсированных выше утверждений об асимптотических свойствах итераций (3.1).

4. ОЦЕНКА АТТРАКТОРА ИТЕРАЦИЙ МЕТОДА ПРОЕКЦИИ ГРАДИЕНТА

В этом разделе будем считать выполненными условие 1 и соотношения (1.5), (1.6), (1.8). Исследуем сходимость итераций (3.1). Используя свойство

$$\|P_D(x) - P_D(y)\|_{H_1} \leq \|x - y\|_{H_1}, \quad \forall x, y \in H_1,$$

из (3.1) получаем

$$\|x_{n+1} - x^*\|_{H_1}^2 \leq \|x_n - x^*\|_{H_1}^2 - 2\rho(\tilde{J}'(x_n), x_n - x^*)_{H_1} + \rho^2 \|\tilde{J}'(x_n)\|_{H_1}^2. \quad (4.1)$$

С учетом (1.11) запишем

$$\begin{aligned} (\tilde{J}'(x_n), x_n - x^*)_{H_1} &= (F(x_n) - f, F'(x_n)(x_n - x^*))_{H_2} + (\tilde{F}(x_n) - F(x_n) + f - \tilde{f}, F'(x_n)(x_n - x^*))_{H_2} + \\ &+ (\tilde{F}(x_n) - F(x_n) + f - \tilde{f}, (\tilde{F}'(x_n) - F'(x_n))(x_n - x^*))_{H_2} + (F(x_n) - f, (\tilde{F}'(x_n) - F'(x_n))(x_n - x^*))_{H_2}. \end{aligned} \quad (4.2)$$

Для первого скалярного произведения в правой части (4.2) в силу (1.4) имеем оценку

$$\begin{aligned} (F(x_n) - f, F'(x_n)(x_n - x^*))_{H_2} &= (F(x_n) - f, F(x_n) - f + G(x_n, x^*))_{H_2} \geq \\ &\geq \|F(x_n) - f\|_{H_2}^2 - \frac{1}{2}L\|F(x_n) - f\|_{H_2} \|x_n - x^*\|_{H_1}^2. \end{aligned} \quad (4.3)$$

Сумма остальных скалярных произведений с учетом (1.4), (1.5), (1.8) оценивается сверху величиной

$$(h + \mu + \delta)\|F(x_n) - f\|_{H_2} + \frac{1}{2}L(h + \delta)\|x_n - x^*\|_{H_1}^2 + \mu(h + \delta). \quad (4.4)$$

Из условия 1 следует, что

$$\|x_n - x^*\|_{H_1}^2 \leq C_2\|F(x_n) - f\|_{H_2}, \quad C_2 = (d_0^{p-2}m)^{-1}.$$

Поэтому сумма первого и второго слагаемых в (4.4) мажорируется величиной

$$C_3(h + \mu + \delta)\|F(x_n) - f\|_{H_2} \leq \frac{1}{2}\|F(x_n) - f\|_{H_2}^2 + C_4(h + \mu + \delta)^2. \quad (4.5)$$

Здесь $C_3 = 1 + C_2L/2$, $C_4 = C_3^2/2$. Из (4.2)–(4.5) следует

$$(\tilde{J}'(x_n), x_n - x^*)_{H_1} \geq \frac{1}{2}\|F(x_n) - f\|_{H_2}^2 - \frac{1}{2}L\|F(x_n) - f\|_{H_2} \|x_n - x^*\|_{H_1}^2 - (C_4 + 1)(h + \mu + \delta)^2. \quad (4.6)$$

С использованием (1.5), (1.6), (1.8), (1.11) находим

$$\|\tilde{J}'(x_n)\|_{H_1} \leq \|\tilde{F}'(x_n)\|_{L(H_1, H_2)} \|\tilde{F}(x_n) - \tilde{f}\|_{H_2} \leq M_0(\|F(x_n) - f\|_{H_2} + h + \delta),$$

поэтому

$$\|\tilde{J}'(x_n)\|_{H_1}^2 \leq 2M_0^2\|F(x_n) - f\|_{H_2}^2 + 2M_0^2(h + \delta)^2. \quad (4.7)$$

Объединяя неравенства (4.1), (4.6), (4.7), получаем оценку

$$\begin{aligned} \|x_{n+1} - x^*\|_{H_1}^2 &\leq \|x_n - x^*\|_{H_1}^2 - \rho\|F(x_n) - f\|_{H_2}^2 + L\rho\|F(x_n) - f\|_{H_2} \|x_n - x^*\|_{H_1}^2 + \\ &+ 2M_0^2\rho^2\|F(x_n) - f\|_{H_2}^2 + 2(C_4 + 1)\rho(h + \mu + \delta)^2 + 2M_0^2\rho^2(h + \delta)^2 \leq \\ &\leq \|x_n - x^*\|_{H_1}^2 - \rho\|F(x_n) - f\|_{H_2}^2 + L\rho\|F(x_n) - f\|_{H_2} \|x_n - x^*\|_{H_1}^2 + \\ &+ 2M_0^2\rho^2\|F(x_n) - f\|_{H_2}^2 + K(\rho)\rho(h + \mu + \delta)^2, \quad K(\rho) = 2(C_4 + 1) + 2M_0^2\rho. \end{aligned}$$

Отсюда с учетом (1.12) получаем

$$\begin{aligned} \|x_{n+1} - x^*\|_{H_1}^2 &\leq \|x_n - x^*\|_{H_1}^2 - \rho \|F(x_n) - f\|_{H_2}^2 + \frac{L\rho}{m^{2/p}} \|F(x_n) - f\|_{H_2}^{(2+p)/p} + \\ &+ 2M_0^2\rho^2 \|F(x_n) - f\|_{H_2}^2 + K(\rho)\rho(h + \mu + \delta)^2. \end{aligned} \quad (4.8)$$

Перепишем (4.8) в виде

$$\begin{aligned} \|x_{n+1} - x^*\|_{H_1}^2 &\leq \|x_n - x^*\|_{H_1}^2 - \\ &- \rho \|F(x_n) - f\|_{H_2}^2 \left(1 - \frac{L}{m^{2/p}} \|F(x_n) - f\|_{H_2}^{(2-p)/p} - 2M_0^2\rho \right) + K(\rho)\rho(h + \mu + \delta)^2. \end{aligned} \quad (4.9)$$

Сделаем индуктивное предположение о том, что

$$\frac{L}{m^{2/p}} \|F(x_n) - f\|_{H_2}^{(2-p)/p} + 2M_0^2\rho \leq \frac{1}{2}. \quad (4.10)$$

Тогда (4.9) вместе с (1.12) влечет оценку

$$\|x_{n+1} - x^*\|_{H_1}^2 \leq \|x_n - x^*\|_{H_1}^2 - \frac{1}{2}\rho m^2 \|x_n - x^*\|_{H_1}^{2p} + C_5\rho(h + \mu + \delta)^2, \quad C_5 = K(1/(4M_0^2)). \quad (4.11)$$

Для удобства далее будем считать выполненными следующие условия, достаточные для выполнения (4.10):

$$\|x_n - x^*\|_{H_1} \leq \omega \triangleq \frac{1}{M_0} \left(\frac{m^{2/p}}{4L} \right)^{p/(2-p)}, \quad p \in [1, 2); \quad \frac{L}{m} \leq \frac{1}{4}, \quad p = 2; \quad (4.12)$$

$$0 < \rho \leq \frac{1}{8M_0^2}. \quad (4.13)$$

Покажем теперь, что если $p \in [1, 2)$, то неравенство $\|x_n - x^*\|_{H_1} \leq \omega$ из (4.12) при подходящих дополнительных условиях выполняется также для $x = x_{n+1}$. Рассмотрим два возможных случая.

1) Пусть $p \in (1, 2)$. Нетрудно убедиться, что функция

$$\varphi(\lambda) = \lambda - \frac{1}{2}\rho m^2 \lambda^p$$

монотонно возрастает при $\lambda \in [0, \tilde{\lambda}]$, где

$$\tilde{\lambda} = \left(\frac{2}{\rho m^2} \right)^{1/(p-1)}.$$

Предположим, что

$$\rho < \frac{2}{\rho m^2 \omega^{2p-2}}. \quad (4.14)$$

Тогда $\|x_n - x^*\|_{H_1}^2 \leq \omega^2 \leq \tilde{\lambda}$, поэтому

$$\varphi(\|x_n - x^*\|_{H_1}^2) \leq \varphi(\omega^2). \quad (4.15)$$

2) Пусть $p = 1$. Тогда при выполнении (4.14) функция $\varphi(\lambda)$ монотонно возрастает при $\lambda \geq 0$. Следовательно, неравенство (4.15) также имеет место.

В дополнение предположим, что

$$h + \mu + \delta \leq \frac{m\omega^p}{\sqrt{2C_5}}. \quad (4.16)$$

При выполнении (4.14), (4.16) из (4.11) следует

$$\|x_{n+1} - x^*\|_{H_1}^2 \leq \varphi(\omega^2) + C_5\rho(h + \mu + \delta)^2 = \omega^2 - \frac{1}{2}\rho m^2 \omega^{2p} + C_5\rho(h + \mu + \delta)^2 \leq \omega^2.$$

Таким образом, при выполнении условия

$$\|x_0 - x^*\|_{H_1} \leq \omega \tag{4.17}$$

будем иметь $\|x_n - x^*\|_{H_1} \leq \omega$ для всех $n = 1, 2, \dots$. Поэтому оценка (4.11) справедлива для всех $n = 0, 1, \dots$. Следующая теорема подытоживает проведенные рассуждения.

Теорема 3. Пусть выполняются условие 1 и условия (1.5), (1.6), (1.8), (4.13). Предположим, что в случае $p = 2$ выполняется $L/m \leq 1/4$, а в случае $p \in [1, 2)$ имеют место (4.14), (4.16), (4.17). Тогда для $a_n = \|x_n - x^*\|_{H_1}^2$ выполняется оценка

$$a_{n+1} \leq a_n - \frac{1}{2} \rho m^2 a_n^p + C_5 \rho (h + \mu + \delta)^2, \quad n = 0, 1, \dots \tag{4.18}$$

Покажем теперь, что множество $B_{r(h, \mu, \delta)} = O_{r(h, \mu, \delta)}(x^*) \cap D$ с $r(h, \mu, \delta)$ вида (3.8) является аттрактором итераций (3.1) и дадим оценку скорости сходимости в соотношении (3.9). В случае $p = 1$ из (4.18) следует

$$a_{n+1} \leq \left(1 - \frac{1}{2} \rho m^2\right) a_n + C_5 \rho (h + \mu + \delta)^2. \tag{4.19}$$

Применяя к оценке (4.19) лемму 1 из [14, гл. 2, § 2], получаем следующее утверждение.

Теорема 4. Пусть $p = 1$, выполняются условие 1 и условия (1.5), (1.6), (1.8), (4.13), (4.14), (4.16), (4.17). Тогда

$$\|x_n - x^*\|_{H_1}^2 \leq \frac{2C_5}{m^2} (h + \mu + \delta)^2 + \|x_0 - x^*\|_{H_1}^2 \left(1 - \frac{1}{2} \rho m^2\right)^n, \quad n = 0, 1, \dots \tag{4.20}$$

Следствие 4. В условиях теоремы 4 справедливо

$$\text{dist}_{H_1}(x_n, B_{r(h, \mu, \delta)}) \leq \|x_0 - x^*\|_{H_1} \left(1 - \frac{1}{2} \rho m^2\right)^{n/2}, \quad r(h, \mu, \delta) = \frac{\sqrt{2C_5}}{m} (h + \mu + \delta).$$

Пусть теперь $p \in (1, 2]$. Могут представиться три случая.

1) Для всех $n = 0, 1, \dots$ справедливо

$$a_n \geq C_6 (h + \mu + \delta)^{2/p}, \quad C_6 = \left(\frac{2C_5}{m^2}\right)^{1/p}. \tag{4.21}$$

В этом случае $C_5 \rho (h + \mu + \delta)^2 \leq \rho m^2 a_n^p / 2$, поэтому (4.18) влечет $0 \leq a_{n+1} \leq a_n$, $n = 0, 1, \dots$. Следовательно, существует предел $\bar{a} = \lim_{n \rightarrow \infty} a_n$. Переходя в (4.18) к пределу, получаем $\bar{a} \leq C_6 (h + \mu + \delta)^{2/p}$. Поэтому из (4.21) следует

$$\lim_{n \rightarrow \infty} a_n = \bar{a} = C_6 (h + \mu + \delta)^{2/p}. \tag{4.22}$$

2) Существует такой номер n_0 , что

$$a_n \geq C_6 (h + \mu + \delta)^{2/p}, \quad n = 0, \dots, n_0; \quad a_n \leq C_6 (h + \mu + \delta)^{2/p}, \quad n \geq n_0 + 1.$$

Как и в случае 1), получаем $a_{n+1} \leq a_n$ для всех $n = 0, 1, \dots, n_0$. Вместо (4.22) здесь имеем $\limsup_{n \rightarrow \infty} a_n \leq C_6 (h + \mu + \delta)^{2/p}$.

3) Для некоторого номера n_1 выполняется

$$a_{n_1} \leq C_6 (h + \mu + \delta)^{2/p}, \quad a_{n_1+1} > C_6 (h + \mu + \delta)^{2/p}.$$

Тогда из (4.18) следует

$$a_{n_1+1} \leq a_{n_1} + C_5 \rho (h + \mu + \delta)^2 \leq C_7 (h + \mu + \delta)^{2/p}, \quad C_7 = C_6 + C_5 \rho \Delta_0^{(2p-2)/p}. \tag{4.23}$$

Как и в случае 1), $a_{n_1+2} \leq a_{n_1+1}$, поэтому $a_{n_1+2} \leq C_7 (h + \mu + \delta)^{2/p}$. Далее, если $a_{n_1+2} > C_6 (h + \mu + \delta)^{2/p}$, то, как и выше, $a_{n_1+3} \leq a_{n_1+2} \leq C_7 (h + \mu + \delta)^{2/p}$. Если же $a_{n_1+2} \leq C_6 (h + \mu + \delta)^{2/p}$, то, аналогично (4.23),

$a_{n+3} \leq C_7(h + \mu + \delta)^{2/p}$. Поэтому в случае 3) для всех номеров $n \geq n_1$ выполняется $a_n \leq C_7(h + \mu + \delta)^{2/p}$.

Проведенный анализ показывает, что во всех случаях 1)–3) справедливо

$$\limsup_{n \rightarrow \infty} a_n \leq C_7(h + \mu + \delta)^{2/p}.$$

Более того, последовательность a_n является монотонно убывающей до момента достижения ею величины $C_7(h + \mu + \delta)^{2/p}$, а после этого момента $a_n \leq C_7(h + \mu + \delta)^{2/p}$. Таким образом, при $r(h, \mu, \delta) = C_8(h + \mu + \delta)^{1/p}$, $C_8 = \sqrt{C_7}$ имеет место (3.9).

Для получения оценки скорости сходимости в (3.9) зафиксируем произвольно $\kappa \in (0, 1/2)$ и заметим, что условие

$$a_n \geq C_9(h + \mu + \delta)^{2/p}, \quad C_9 = \left(\frac{2C_5}{(1 - 2\kappa)m^2} \right)^{1/p} \quad (4.24)$$

влечет

$$-\frac{1}{2} \rho m^2 a_n^p + C_5 \rho (h + \mu + \delta)^2 \leq -\kappa \rho m^2 a_n^p.$$

Поэтому ввиду (4.18) имеем

$$a_{n+1} \leq a_n - \kappa \rho m^2 a_n^p. \quad (4.25)$$

Следовательно, если (4.24) выполняется для номеров $n = 0, 1, \dots, l$, то для этих же номеров n справедлива оценка (4.25). Используя [14, гл. 2, § 2, лемма 6], получаем

$$a_n \leq \frac{a_0}{(1 + (p-1)a_0^{p-1} \kappa \rho m^2 n)^{1/(p-1)}}, \quad n = 0, 1, \dots, l. \quad (4.26)$$

Применим оценку (4.26) к номерам n таким, что $a_n > C_{10}(h + \mu + \delta)^{2/p}$, $C_{10} = \max\{C_7, C_9\}$. Получаем следующую оценку скорости сходимости итераций (3.1) к аттрактору $B_{r(h, \mu, \delta)}$, $r(h, \mu, \delta) = C_{11}(h + \mu + \delta)^{1/p}$, $C_{11} = \sqrt{C_{10}}$.

Теорема 5. Пусть выполняются условие 1 и условия (1.5), (1.6), (1.8). Предположим, что в случае $p \in (1, 2]$ выполняются (4.13), (4.14), (4.16), (4.17), а при $p = 2$ имеет место $L/m \leq 1/4$. Тогда

$$\limsup_{n \rightarrow \infty} \|x_n - x^*\|_{H_1} \leq C_8(h + \mu + \delta)^{1/p}.$$

Кроме того, справедлива оценка

$$\|x_n - x^*\|_{H_1} \leq \max \left\{ \frac{\|x_0 - x^*\|_{H_1}}{\left((1 + (p-1)\|x_0 - x^*\|_{H_1}^{2p-2} \kappa \rho m^2 n)^{1/(2p-2)} \right)}, C_{11}(h + \mu + \delta)^{1/p} \right\}. \quad (4.27)$$

Следствие 5. В условиях теоремы 5 справедливо

$$\text{dist}_{H_1}(x_n, B_{r(h, \mu, \delta)}) \leq \frac{\|x_0 - x^*\|_{H_1}}{\left((1 + (p-1)\|x_0 - x^*\|_{H_1}^{2p-2} \kappa \rho m^2 n)^{1/(2p-2)} \right)},$$

$$r(h, \mu, \delta) = C_{11}(h + \mu + \delta)^{1/p}.$$

Следствия 4 и 5 приводят к следующему результату о локализации кластера стационарных точек функционала (1.9) на множестве D .

Теорема 6. Пусть выполняются условие 1 и условия (1.5), (1.6), (1.8), (4.16). Тогда все стационарные точки функционала \tilde{J} на D , принадлежащие $O_\omega(x^*) \cap D$, лежат в $B_{r(h, \mu, \delta)}$, $r(h, \mu, \delta) = O((h + \mu + \delta)^{1/p})$.

Теорема 6 уточняет более грубую оценку размера кластера стационарных точек в условиях (1.5)–(1.7) (см. [15]), согласно которой кластер лежит в шаре $O_{r(h,\delta)}(x^*)$ радиуса $r(h,\delta) = O(\sqrt{h} + \delta)^{1/p}$.

Теоремы 4, 5 не содержат информации о сходимости последовательности $\{x_n\}$ к какому-либо пределу. Положительные утверждения такого рода получаются при наложении на приближенный оператор \tilde{F} и шаговый множитель ρ подходящих дополнительных условий. Предположим, что производная \tilde{F}' , наряду с F' , удовлетворяет условию Липшица (1.2). В этом случае градиент \tilde{J}' также удовлетворяет условию Липшица:

$$\|\tilde{J}'(x) - \tilde{J}'(y)\|_{H_1} \leq \Lambda \|x - y\|_{H_1} \quad \forall x, y \in D.$$

Из теоремы 1.4.4 (см. [10]) следует, что при выполнении дополнительного условия

$$0 < \rho < \frac{2}{\Lambda} \tag{4.28}$$

для итераций (3.1) справедливо $\lim_{n \rightarrow \infty} \|x_{n+1} - x_n\|_{H_1} = 0$. С другой стороны, в силу компактности D любое бесконечное подмножество последовательности $\{x_n\}$ имеет сильно сходящуюся подпоследовательность $\{x_{n_k}\}$, $\lim_{k \rightarrow \infty} \|x_{n_k} - \bar{x}\|_{H_1} = 0$. Тогда и $\lim_{k \rightarrow \infty} \|x_{n_k+1} - \bar{x}\|_{H_1} = 0$. Полагая в (3.1) $n = n_k$ и переходя к пределу при $k \rightarrow \infty$, заключаем, что $\bar{x} = P_D(\bar{x} - \rho \tilde{J}'(\bar{x}))$, т.е. \bar{x} есть стационарная точка \tilde{J} на D . Тем самым получаем следующее уточнение теорем 4 и 5.

Теорема 7. Пусть выполняются условия теорем 4, 5, производная \tilde{F}' удовлетворяет условию Липшица (1.2) и выполнено (4.28). Тогда последовательность $\{x_n\}$ имеет предельные точки, и каждая такая точка \bar{x} является стационарной для \tilde{J} на D . При этом $\bar{x} \in B_{r(h,\mu,\delta)}$ с $r(h,\mu,\delta)$, указанными в следствиях 4 и 5.

5. О НЕОБХОДИМОСТИ УСЛОВИЯ 1

Покажем необходимость условия 1 для теорем 4, 5 и их следствий. Ради упрощения изложения примем $h = \mu = 0$. Согласно следствиям 4 и 5, при выполнении надлежащих условий на характеристики задачи (1.1) и параметры процесса (3.1) имеет место равенство (3.9), т.е.

$$\lim_{n \rightarrow \infty} \text{dist}_{H_1}(x_n, B_{r(\delta)}) = 0, \quad r(\delta) = C_{12} \delta^{1/p}. \tag{5.1}$$

Здесь $x_n = x_n(\tilde{f})$ – порождаемая (3.1) последовательность. Одним из условий для (5.1) является $\|x_0 - x^*\|_{H_1} \leq \omega$, величина ω определена в (4.12). Это условие выполнено для любого $x_0 \in D$, если $\text{diam}_{H_1}(D) \leq \omega$. Оценки скорости сходимости из теорем 4 и 5 позволяют указать момент останова итераций $n = n(\delta)$ так, что оператор $R_\delta(\tilde{f}) = x_{n(\delta)}(\tilde{f})$ определяет регуляризующий алгоритм для задачи (1.1). Это означает, что

$$\lim_{\delta \rightarrow 0} \Delta_D(\delta) = 0, \quad \Delta_D(\delta) \triangleq \sup_{x \in D, \|\tilde{f} - F(x)\|_{H_2} \leq \delta} \|R_\delta(\tilde{f}) - x\|_{H_1}. \tag{5.2}$$

Величина $\Delta_D(\delta)$ есть равномерная на D оценка точности алгоритма R_δ на входных данных, известных с погрешностью δ . Используя оценки скорости сходимости (4.20), (4.27), нетрудно выбрать момент останова $n = n(\delta)$ так, что в условиях теорем 4, 5 для погрешности $\Delta_D(\delta)$ справедлива оценка

$$\Delta_D(\delta) \leq C_{13} \delta^{1/p}. \tag{5.3}$$

Отвлекаясь от оценок скорости сходимости $\{x_n\}$ к аттрактору $B_{r(\delta)}$, на основании (5.1) констатируем, что оператор итерационного перехода

$$T(\tilde{f})(x) = P_D(x - \rho \tilde{J}'(x)), \quad \tilde{J}(x) = \frac{1}{2} \|F(x) - \tilde{f}\|_{H_2}^2,$$

обеспечивает асимптотическую регуляризацию задачи (1.1) в смысле предельного соотношения

$$\limsup_{n \rightarrow \infty} \|T^n(\tilde{f})(x_0) - x\|_{H_1} \leq r(\delta) \quad \forall x_0 \in D, \quad \forall x \in D, \quad \|\tilde{f} - F(x)\|_{H_2} \leq \delta. \quad (5.4)$$

Соотношение (5.4) формализует факт устойчивой сходимости итераций $x_{n+1} = T(\tilde{f})(x_n)$ к аттрактору $B_{r(\delta)}$. Однако, если $T(\tilde{f}) : D \rightarrow H_1$ — произвольное отображение со свойством (5.4), то при отсутствии равномерной по \tilde{f} оценки скорости сходимости в (5.4) отображение $\tilde{f} \rightarrow T^n(\tilde{f})$ ни с каким фиксированным правилом останова $n = n(\delta)$ в общем случае не порождает регуляризующий алгоритм для (1.1) в смысле (5.2). Тем не менее следующая теорема утверждает, что без привлечения условия 1 построение итерационных операторов $T(\tilde{f})$, обладающих достаточно слабым свойством асимптотической регуляризации (5.4), невозможно, каково бы ни было $p \geq 1$.

Теорема 8. Пусть D — связное компактное множество,

$$\delta_0 = \text{diam}_{H_2}(F(D)),$$

$F : D \rightarrow H_2$ — инъективный оператор. Пусть для некоторого оператора $T(\tilde{f}) : D \rightarrow H_1$ при всех $\delta \in (0, \delta_0]$ выполняется (5.4) с $r(\delta) = C\delta^{1/p}$, $p \geq 1$. Тогда

$$\|F(x) - F(y)\|_{H_2} \geq m_0 \|x - y\|_{H_1}^p \quad \forall x, y \in D; \quad m_0 = (2C)^{-p}. \quad (5.5)$$

Доказательство. Зафиксируем $x_1^*, x_2^* \in D$ так, что для $f_j = F(x_j^*)$, $j = 1, 2$, выполняется $\|f_1 - f_2\|_{H_2} = \delta \in (0, \delta_0]$. Поскольку $F(D)$ компактно и связно и $\delta \leq \text{diam}_{H_2}(F(D))$, элементы $f_j \in F(D)$, $j = 1, 2$, со свойством $\|f_1 - f_2\|_{H_2} = \delta$ существуют. Поэтому $x_j^* = F^{-1}(f_j)$ также определены для всех $\delta \in (0, \delta_0]$. Рассмотрим уравнение (1.1) с $f = f_1$, решением которого является $x^* = x_1^*$. В качестве приближенной правой части \tilde{f} выберем $\tilde{f} = f_1$. В данном случае $\|\tilde{f} - f\|_{H_2} = 0 \leq \delta$. Рассмотрим также задачу (1.1) с $f = f_2$ и приближенным элементом $\tilde{f} = f_1$. Ее точным решением является $x^* = x_2^*$. В этом случае $\|\tilde{f} - f\|_{H_2} = \delta \leq \delta$. Согласно (5.4), для произвольного $x_0 \in D$ и соответствующих положительных последовательностей $\{\alpha_n\}$, $\{\beta_n\}$ выполняется

$$\begin{aligned} \|T^n(f_1)(x_0) - x_1^*\|_{H_1} &\leq C\delta^{1/p} + \alpha_n, & \lim_{n \rightarrow \infty} \alpha_n &= 0, \\ \|T^n(f_1)(x_0) - x_2^*\|_{H_1} &\leq C\delta^{1/p} + \beta_n, & \lim_{n \rightarrow \infty} \beta_n &= 0. \end{aligned} \quad (5.6)$$

Из (5.6) следует

$$\|x_1^* - x_2^*\|_{H_1} \leq 2C\delta^{1/p} + \alpha_n + \beta_n, \quad n = 0, 1, \dots$$

Переходя здесь к пределу при $n \rightarrow \infty$, получаем

$$\|x_1^* - x_2^*\|_{H_1} \leq 2C\delta^{1/p} = 2C\|F(x_1^*) - F(x_2^*)\|_{H_2}^{1/p}. \quad (5.7)$$

Поскольку любые пары точек (x_1^*, x_2^*) соответствуют некоторым $f_j = F(x_j^*)$, $j = 1, 2$, со свойством $\|f_1 - f_2\|_{H_2} \leq \delta$, $\delta \in (0, \delta_0]$, неравенство (5.7) имеет место для любых $x_1^*, x_2^* \in D$. Соотношение (5.5) непосредственно следует из (5.7). Теорема доказана.

Замечание 1. Повторяя доказательство теоремы 8 с $\alpha_n = \beta_n = 0$ и заменой $T^n(f_1)(x_0)$ на $R_\delta(f_1)$, убеждаемся, что условие 1 является также необходимым для существования регуляризующего алгоритма с оценкой (5.3). В случае $p = 1$ этот факт был ранее отмечен в [16].

Полагая $p = 1$ и объединяя теоремы 1 и 8, с учетом замечания 1 получаем следующий вариант необходимого критерия регуляризуемости задачи (1.1) с оценкой точности $O(\delta)$.

Следствие 6. Пусть D – связное компактное множество, $F: D \rightarrow H_2$ – инъективный оператор, $\delta_0 = \text{diam}_{H_2}(F(D)) \leq ((2C)^{-2p} - \nu)/(2L)$, $\nu > 0$. Предположим, что для некоторого регуляризующего алгоритма $R_\delta(\tilde{f})$ выполняется оценка $\Delta_D(\delta) \leq C\delta$, $\delta \in (0, \delta_0]$. Тогда для любого $f \in F(D)$ функционал невязки \tilde{J} является сильно выпуклым в $D \cap O_R(x^*)$, $R = (4LM)^{-1}((2C)^{-2p} - \nu)$, $x^* = F^{-1}(f)$.

Теорема 8 показывает, что свойство асимптотической регуляризуемости задачи (1.1) в смысле (5.4) и формально более сильное свойство ее регуляризуемости в смысле классического определения (5.2) с оценкой (5.3) по существу эквивалентны при любом $p \geq 1$.

СПИСОК ЛИТЕРАТУРЫ

1. Треногин В.А. Функциональный анализ. М.: Наука, 1980.
2. Бакушинский А.Б., Кокурин М.Ю. Алгоритмический анализ нерегулярных операторных уравнений. М.: ЛЕНАНД, 2012.
3. Лаврентьев М.М., Романов В.Г., Шишатский С.П. Некорректные задачи математической физики и анализа. М.: Наука, 1980.
4. Кабанихин С.И. Обратные и некорректные задачи. Новосибирск: Сиб. науч. изд-во, 2008.
5. Романов В.Г. Устойчивость в обратных задачах. М.: Научный мир, 2004.
6. Isakov V. Inverse Problems for Partial Differential Equations. N.Y.: Springer, 2006.
7. Кокурин М.Ю. Об условно корректных и обобщенно корректных задачах // Ж. вычисл. матем. и матем. физ. 2013. Т. 53. № 6. С. 857–866.
8. Kokurin M.Yu. On a characteristic property of conditionally well-posed problems // J. Inv. Ill-Posed Probl. 2015. V. 23. № 3. P. 245–262.
9. Тихонов А.Н., Леонов А.С., Ягола А.Г. Нелинейные некорректные задачи. М.: Наука, Физматлит, 1995.
10. Васильев Ф.П. Методы решения экстремальных задач. М.: Наука, 1981.
11. Красносельский М.А., Вайникко Г.М., Забрейко П.П., Рунтцкий Я.Б., Стеценко В.Я. Приближенное решение операторных уравнений. М.: Наука, 1969.
12. Kokurin M.Yu. On stable finite dimensional approximation of conditionally well-posed inverse problems // Inv. Probl. 2016. V. 32. № 10. 105007.
13. Kokurin M.Yu. Stable gradient projection method for nonlinear conditionally well-posed inverse problems // J. Inv. Ill-posed Probl. 2016. V. 24. № 3. P. 323–332.
14. Поляк Б.Т. Введение в оптимизацию. М.: Наука, 1983.
15. Кокурин М.Ю. О кластеризации стационарных точек функционалов невязки условно-корректных обратных задач // Сиб. журн. вычисл. матем. 2018. Т. 21. № 4. С. 393–406.
16. Леонов А.С. О возможности получения линейных оценок точности приближенных решений обратных задач // Изв. вуз. Матем. 2016. № 10. С. 29–35.

УДК 517.95

ЗАДАЧА СЛОЖНОГО ТЕПЛООБМЕНА С УСЛОВИЯМИ
ТИПА КОШИ НА ЧАСТИ ГРАНИЦЫ¹⁾© 2023 г. П. Р. Месенев¹, А. Ю. Чеботарев^{2,3,*}¹ 690922 Владивосток, о. Русский, п. Аякс, 10, ДВФУ,
Дальневосточный центр математических исследований, Россия² 690041 Владивосток, ул. Радио, 7, ИПМ ДВО РАН, Россия³ 675027 Благовещенск, шоссе Игнатьевское, 21, АмГУ, Россия

*e-mail: cheb@iam.dvo.ru

Поступила в редакцию 09.08.2022 г.
Переработанный вариант 16.11.2022 г.
Принята к публикации 02.02.2023 г.

Рассмотрена краевая задача для стационарных уравнений сложного теплообмена с незадачным краевым условием для интенсивности излучения на части границы и условием переопределения на другой части границы. Предложен оптимизационный метод решения указанной задачи и представлен анализ соответствующей задачи граничного оптимального управления. Показано, что последовательность решений экстремальных задач сходится к решению задачи с условиями типа Коши. Эффективность алгоритма проиллюстрирована численными примерами. Библ. 27. Фиг. 2.

Ключевые слова: уравнения радиационно-кондуктивного теплообмена, диффузионное приближение, краевая задача с условиями типа Коши, задача оптимального управления.

DOI: 10.31857/S0044466923050162, EDN: GFVZBT

1. ПОСТАНОВКА ЗАДАЧИ

Рассмотрим следующую систему полулинейных эллиптических уравнений, которая моделирует радиационный и диффузионный (сложный) теплообмен в ограниченной липшицевой области $\Omega \subset \mathbb{R}^3$ с границей $\Gamma = \partial\Omega$:

$$-\alpha\Delta\theta + b\kappa_a(|\theta|^3 - \varphi) = 0, \quad -\alpha\Delta\varphi + \kappa_a(\varphi - |\theta|^3) = 0, \quad x \in \Omega. \quad (1)$$

Через θ и φ здесь обозначены температура и усредненная по всем направлениям интенсивность теплового излучения. Положительные параметры a , b , κ_a и α , описывающие свойства среды, являются заданными (см. [1]).

Пусть граница области состоит из двух участков, $\Gamma\partial\Omega := \bar{\Gamma}_1 \cup \bar{\Gamma}_2$, так что Γ_1, Γ_2 измеримы и имеют положительную меру, $\Gamma_1 \cap \Gamma_2 = \emptyset$. На всей границе Γ задается тепловой поток q_b :

$$a\partial_n\theta = q_b, \quad x \in \Gamma. \quad (2)$$

Для задания краевого условия для интенсивности излучения требуется знать функцию γ , описывающую отражающие свойства границы (см. [1]). В случае, если эта функция неизвестна на части границы Γ_2 , краевое условие для интенсивности излучения на Γ_2 не ставится, а в качестве условия переопределения на Γ_1 , в дополнение к условию на φ , задается температурное поле θ_b :

$$\alpha\partial_n\varphi + \gamma(\varphi - \theta_b^4) = 0, \quad \theta = \theta_b \quad x \in \Gamma_1. \quad (3)$$

Здесь через ∂_n обозначаем производную в направлении внешней нормали \mathbf{n} .

¹⁾Работа выполнена при финансовой поддержке РФФИ (код проекта 20-01-00113), ИПМ ДВО РАН (НИОКТР АААА-А20-120120390006-0) и Минобрнауки России (проект 122082400001-8).

Разрешимость задачи (1)–(3) в точной постановке является открытой проблемой. Поэтому в настоящей работе предлагается оптимизационный метод решения задачи, который заключается в рассмотрении граничного оптимального управления для эквивалентной системы эллиптических уравнений. Как будет показано ниже, задача оптимального управления является аппроксимацией задачи (1)–(3). Для постановки задачи управления введем новую неизвестную функцию $\psi = a\theta + \alpha b\varphi$. Складывая первое уравнение в (1) со вторым, умноженным на b , заключаем, что ψ – гармоническая функция. Исключая φ из первого уравнения в (1) и используя краевые условия (2), (3), получаем краевую задачу

$$-a\Delta\theta + g(\theta) = \frac{\kappa_a}{\alpha}\psi, \quad \Delta\psi = 0, \quad x \in \Omega, \tag{4}$$

$$a\partial_n\theta = q_b, \text{ на } \Gamma, \quad \alpha\partial_n\psi + \gamma\psi = r, \quad \theta = \theta_b \text{ на } \Gamma_1. \tag{5}$$

Здесь $g(\theta) = b\kappa_a|\theta|^3 + \frac{a\kappa_a}{\alpha}\theta$, $r = \alpha b\gamma\theta_b^4 + \alpha q_b + a\gamma\theta_b$.

Сформулируем задачу оптимального управления, которая аппроксимирует задачу (4), (5). Задача состоит в отыскании тройки $\{\theta_\lambda, \psi_\lambda, u_\lambda\}$ такой, что

$$J_\lambda(\theta, u) = \frac{1}{2} \int_{\Gamma_1} (\theta - \theta_b)^2 d\Gamma + \frac{\lambda}{2} \int_{\Gamma_2} u^2 d\Gamma \rightarrow \inf, \quad -a\Delta\theta + g(\theta) = \frac{\kappa}{\alpha}\psi, \quad \Delta\psi = 0, \quad x \in \Omega, \tag{6}$$

$$a\partial_n\theta + s\theta = q_b + s\theta_b, \quad \alpha\partial_n\psi + \gamma\psi = r \text{ на } \Gamma_1, \quad a\partial_n\theta = q_b, \quad \alpha\partial_n\psi = u \text{ на } \Gamma_2.$$

Здесь $\lambda, s > 0$ – регуляризирующие параметры. Отметим, что введение параметра $s > 0$ связано с численным решением задачи (6) и позволяет более эффективно находить решение нелинейной краевой задачи с заданным u . Представленные ниже теоретические результаты справедливы и в случае $s = 0$.

Интерес к граничным задачам сложного теплообмена связан с важными инженерными приложениями (см. [2]). Применение оптимизационного метода в обратных задачах для моделей теплопроводности с заданным переопределением на части границы рассмотрено, например, в [3], [4].

Нелинейные модели сложного теплообмена в рамках P_1 -приближения для уравнения переноса теплового излучения достаточно полно изучены. В работах [5–7] представлен анализ различных прямых краевых задач для системы (1). Задачи оптимального управления для уравнений сложного теплообмена изучены в [8–11]. Обратные задачи с конечномерным переопределением и задачи с данными типа Коши на границе для температуры рассмотрены в [12–17]. В [18–22] представлены интересные результаты анализа краевых задач сложного теплообмена без использования P_1 -приближения.

2. РАЗРЕШИМОСТЬ ЗАДАЧИ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

Рассмотрим пространства $H = L^2(\Omega)$, $V = H^1(\Omega) = W_2^1(\Omega)$, V' – пространство, сопряженное с V . Пространство H отождествляем с пространством H' так, что $V \subset H = H' \subset V'$. Через U обозначаем пространство управлений $L^2(\Gamma_2)$. Стандартную норму в H обозначаем $\|\cdot\|$, (f, v) – значение функционала $f \in V'$ на элементе $v \in V$, совпадающее со скалярным произведением в H , если $f \in H$.

Будем предполагать, что исходные данные удовлетворяют условиям

(i) $a, b, \alpha, \kappa_a, \lambda, s = \text{Const} > 0$;

(ii) $0 < \gamma_0 \leq \gamma \in L^\infty(\Gamma_1)$, $\theta_b, r \in L^2(\Gamma_1)$, $q_b \in L^2(\Gamma)$.

Определим операторы $A_{1,2} : V \rightarrow V'$, $B_1 : L^2(\Gamma_1) \rightarrow V'$, $B_2 : U \rightarrow V'$, используя равенства, справедливые для любых $y, z \in V$, $f, v \in L^2(\Gamma_1)$, $h, w \in U$:

$$(A_1 y, z) = a(\nabla y, \nabla z) + s \int_{\Gamma_1} yz d\Gamma, \quad (A_2 y, z) = \alpha(\nabla y, \nabla z) + \int_{\Gamma_1} \gamma yz d\Gamma,$$

$$(B_1 f, v) = \int_{\Gamma_1} f v d\Gamma, \quad (B_2 h, w) = \int_{\Gamma_2} h w d\Gamma.$$

Заметим, что билинейная форма $(A_1 y, z)$ определяет скалярное произведение в пространстве V и норма $\|z\|_V = \sqrt{(A_1 z, z)}$ эквивалентна стандартной норме V . Кроме того, определены непрерывные обратные операторы $A_{1,2}^{-1}: V' \mapsto V$. Для $y \in V$, $f \in L^2(\Gamma_1)$, $h \in V'$ справедливы неравенства

$$\|y\| \leq K_0 \|y\|_{V'}, \quad \|B_1 f\|_{V'} \leq K_1 \|f\|_{L^2(\Gamma_1)}, \quad \|B_2 h\|_{V'} \leq K_2 \|h\|_{U'}. \quad (7)$$

Здесь постоянные $K_j > 0$ зависят только от Ω , Γ_1 , Γ_2 и параметров a, s .

Используя введенные операторы, слабую формулировку краевой задачи, на решениях которой минимизируется функционал (6), нетрудно записать в виде

$$A_1 \theta + g(\theta) = \frac{\kappa_a}{\alpha} \psi + f_1, \quad A_2 \psi = f_2 + B_2 u, \quad (8)$$

где $f_1 = B_1(q_b + s\theta_b) + B_2 q_b$, $f_2 = B_1 r$.

Для формализации задачи оптимального управления определим оператор ограничений $F(\theta, \psi, u): V \times V \times U \rightarrow V' \times V'$:

$$F(\theta, \psi, u) = \left\{ A_1 \theta + g(\theta) - \frac{\kappa_a}{\alpha} \psi - f_1, A_2 \psi - f_2 - B_2 u \right\}.$$

Задача (P_λ) . Найти тройку $\{\theta_\lambda, \psi_\lambda, u_\lambda\} \in V \times V \times U$ такую, что

$$J_\lambda(\theta, u) = \frac{1}{2} \|\theta - \theta_b\|_{L^2(\Gamma_1)}^2 + \frac{\lambda}{2} \|u\|_{U'}^2 \rightarrow \inf, \quad F(\theta, \psi, u) = 0. \quad (9)$$

Лемма 1. Пусть выполняются условия (i), (ii), $u \in U$. Тогда существует единственное решение системы (8) и при этом

$$\|\theta\|_V \leq \frac{K_0^2 \kappa_a}{\alpha} \|\psi\|_{V'} + \|f_1\|_{V'}, \quad \|\psi\|_{V'} \leq \|A_2^{-1}\| (\|f_2\|_{V'} + K_2 \|u\|_{U'}). \quad (10)$$

Доказательство. Из второго уравнения в (8) следует, что $\psi = A_2^{-1}(f_2 + B_2 u)$ и поэтому

$$\|\psi\|_{V'} \leq \|A_2^{-1}\| (\|f_2\|_{V'} + K_2 \|u\|_{U'}).$$

Однозначная разрешимость первого уравнения (8) с монотонной нелинейностью хорошо известна (см., например, [23]). Умножим скалярно это уравнение на θ , отбросим неотрицательное слагаемое $(g(\theta), \theta)$ и оценим правую часть, используя неравенство Коши–Буняковского. Тогда, с учетом неравенств (7), получаем

$$\|\theta\|_V^2 \leq \frac{\kappa_a}{\alpha} \|\psi\| \|\theta\| + \|f_1\|_{V'} \|\theta\|_V \leq \frac{K_0^2 \kappa_a}{\alpha} \|\psi\|_{V'} \|\theta\|_V + \|f_1\|_{V'} \|\theta\|_V.$$

В результате получаем оценки (10).

Теорема 1. При выполнении условий (i), (ii) существует решение задачи (P_λ) .

Доказательство. Обозначим через j_λ точную нижнюю грань целевого функционала J_λ на множестве $u \in U$, $F(\theta, \psi, u) = 0$ и рассмотрим последовательности такие, что $u_m \in U$, $\theta_m \in V$, $\psi_m \in V$, $J_\lambda(\theta_m, u_m) \rightarrow j_\lambda$,

$$A_1 \theta_m + g(\theta_m) = \frac{\kappa_a}{\alpha} \psi_m + f_1, \quad A_2 \psi_m = f_2 + B_2 u_m. \quad (11)$$

Из ограниченности последовательности u_m в пространстве U следуют, на основании леммы 1, оценки

$$\|\theta_m\|_V \leq C, \quad \|\psi_m\|_{V'} \leq C, \quad \|\theta_m\|_{L^6(\Omega)} \leq C.$$

Здесь через $C > 0$ обозначена наибольшая из постоянных, ограничивающих соответствующие нормы и не зависящих от m . Переходя при необходимости к подпоследовательностям, заключаем, что существует тройка $\{\hat{u}, \hat{\theta}, \hat{\psi}\} \in U \times V \times VP$

$$u_m \rightarrow \hat{u} \text{ слабо в } U, \theta_m, \psi_m \rightarrow \hat{\theta}, \hat{\psi} \text{ слабо в } V, \text{ сильно в } L^4(\Omega). \tag{12}$$

Заметим также, что $\forall v \in V$ имеем

$$|([\theta_m]^4 - [\hat{\theta}]^4, v)| \leq 2 \|\theta_m - \hat{\theta}\|_{L^4(\Omega)} \|v\|_{L^4(\Omega)} \left(\|\theta_m\|_{L_6(\Omega)}^3 + \|\hat{\theta}\|_{L_6(\Omega)}^3 \right). \tag{13}$$

Результаты о сходимости (12) и неравенство (13) позволяют перейти к пределу в (11). В итоге получим

$$A_1 \hat{\theta} + g(\hat{\theta}) = \frac{\kappa_a}{\alpha} \hat{\psi} + f_1, \quad A_2 \hat{\psi} = f_2 + B_2 \hat{u}. \tag{14}$$

Поскольку целевой функционал слабо полунепрерывен снизу, то $j_\lambda \leq J_\lambda(\hat{\theta}, \hat{u}) \leq \liminf J_\lambda(\theta_m, u_m) = j_\lambda$, и поэтому тройка $\{\hat{\theta}, \hat{\psi}, \hat{u}\}$ есть решение задачи (P_λ) .

3. УСЛОВИЯ ОПТИМАЛЬНОСТИ ПЕРВОГО ПОРЯДКА

Воспользуемся принципом Лагранжа для гладко-выпуклых экстремальных задач (см. [24]). Невырожденность условий оптимальности гарантируется условием, что образ производной оператора ограничений $F(y, u)$, где $y = \{\theta, \psi\} \in V \times V$, совпадает с пространством $V' \times V'$. Последнее означает, что линейная система

$$A_1 \xi + g'(\theta) \xi - \frac{\kappa_a}{\alpha} \eta = q_1, \quad A_2 \eta = q_2$$

разрешима для всех $\theta \in V, q_1, q_2 \in V'$. Здесь $g'(\theta) = 4b\kappa_a|\theta|^3 + \kappa_a/\alpha$. Из второго уравнения получаем $\eta = A_2^{-1}q_2$. Разрешимость первого уравнения при известном $\eta \in V$ очевидным образом следует из леммы Лакса–Мильграма. Отметим, что справедливость остальных условий принципа Лагранжа также очевидна.

Функция Лагранжа задачи (P_λ) имеет вид

$$L(\theta, \psi, u, p_1, p_2) = J_\lambda(\theta, u) + \left(A_1 \theta + g(\theta) - \frac{\kappa_a}{\alpha} \psi - f_1, p_1 \right) + (A_2 \psi - f_2 - B_2 u, p_2),$$

где $p = \{p_1, p_2\} \in V \times V$ – сопряженное состояние.

Пусть $\{\hat{\theta}, \hat{\psi}, \hat{u}\}$ – решение задачи (P_λ) . Вычислив производные Гато функции Лагранжа по θ, ψ и u , получаем в силу принципа Лагранжа (см. [24], теорема 1.5) следующие равенства $\forall v \in V, w \in U$:

$$(B_1(\hat{\theta} - \theta_b), v) + (A_1 v + g'(\hat{\theta})v, p_1) = 0, \quad -\frac{\kappa_a}{\alpha}(v, p_1) + (A_2 v, p_2) = 0, \tag{15}$$

$$\lambda(B_2 \hat{u}, w) - (B_2 w, p_2) = 0. \tag{16}$$

Из условий (15), (16) вытекают уравнения для сопряженного состояния, которые вместе с уравнениями (14) для оптимальной тройки дают систему оптимальности задачи (P_λ) .

Теорема 2. Пусть выполняются условия (i), (ii). Если $\{\hat{\theta}, \hat{\psi}, \hat{u}\}$ – решение задачи (P_λ) , то существует единственная пара $\{p_1, p_2\} \in V \times V$ такая, что

$$A_1 p_1 + g'(\hat{\theta})p_1 = -B_1(\hat{\theta} - \theta_b), \quad A_2 p_2 = \frac{\kappa_a}{\alpha} p_1, \quad \lambda \hat{u} = p_2|_{\Gamma_2}. \tag{17}$$

Замечание 1. Если рассмотреть приведенный целевой функционал $\tilde{J}_\lambda(u) = J_\lambda(\theta(u), u)$, где $\theta(u)$ – компонента решения задачи (8), соответствующая управлению $u \in U$, то градиент функционала $\tilde{J}_\lambda(u)$ равен $\tilde{J}'_\lambda(u) = \lambda u - p_2$. Здесь p_2 – компонента решения сопряженной системы (17), где $\hat{\theta} := \theta(u)$.

4. АППРОКСИМАЦИЯ РЕШЕНИЯ ЗАДАЧИ (1)–(3)

Определим слабое решение задачи (1)–(3). Пару $\{\theta, \varphi\} \in V \times V$ будем называть слабым решением задачи (1)–(3), если найдется функция $q \in L^2(\Gamma_2)$ такая, что для всех $v \in V$ справедливы равенства

$$a(\nabla\theta, \nabla v) + b\kappa_a(|\theta|\theta^3 - \varphi, v) = \int_{\Gamma} q_b v d\Gamma, \quad (18)$$

$$\alpha(\nabla\varphi, \nabla v) + \int_{\Gamma_1} \gamma\varphi v d\Gamma + \kappa_a(\varphi - |\theta|\theta^3, v) = \int_{\Gamma_1} \gamma\theta_b^4 v d\Gamma + \int_{\Gamma_2} q v d\Gamma \quad (19)$$

и при этом $\theta|_{\Gamma_1} = \theta_b$.

Покажем, что если существует единственное слабое решение задачи (1)–(3), то решения задачи (P_λ) при $\lambda \rightarrow +0$ аппроксимируют решение задачи (1)–(3).

Теорема 3. Пусть выполняются условия (i), (ii) и существует единственное слабое решение $\{\theta, \varphi\}$ задачи (1)–(3). Если $\{\theta_\lambda, \psi_\lambda, u_\lambda\}$ – решение задачи (P_λ) для $\lambda > 0$, то при $\lambda \rightarrow +0$

$$\theta_\lambda \rightarrow \theta_*, \quad \frac{1}{\alpha b}(\psi_\lambda - a\theta_\lambda) \rightarrow \varphi_* \text{ слабо в } V, \text{ сильно в } H.$$

Доказательство. Умножим равенство (18) на α , (19) на αb и сложим равенства. Тогда, полагая $\psi = a\theta + \alpha b\varphi$, $u = \alpha b q + \alpha q_b|_{\Gamma_2}$, получаем

$$\alpha(\nabla\psi, \nabla v) + \int_{\Gamma_1} \gamma\psi v d\Gamma = \int_{\Gamma_1} r v d\Gamma + \int_{\Gamma_2} u v d\Gamma.$$

Здесь $r = \alpha b \gamma \theta_b^4 + \alpha q_b + a \gamma \theta_b$. Поэтому $A_2\psi = f_2 + B_2u$.

Из (18), с учетом условия $\theta|_{\Gamma_1} = \theta_b$, выводим равенство $A_1\theta + g(\theta) = \frac{\kappa_a}{\alpha}\psi + f_1$. Таким образом, тройка $\{\theta, \psi, u\} \in V \times V \times U$ является допустимой для задачи (P_λ) и, следовательно,

$$J_\lambda(\theta_\lambda, u_\lambda) = \frac{1}{2}\|\theta_\lambda - \theta_b\|_{L^2(\Gamma_1)}^2 + \frac{\lambda}{2}\|u_\lambda\|_U^2 \leq J_\lambda(\theta, u) = \frac{\lambda}{2}\|u\|_U^2.$$

Тогда

$$\|u_\lambda\|_U^2 \leq \|u\|_U^2, \quad \|\theta_\lambda - \theta_b\|_{L^2(\Gamma_1)}^2 \rightarrow 0, \quad \lambda \rightarrow +0.$$

Из ограниченности последовательности u_λ в пространстве U следуют, на основании леммы 1, оценки

$$\|\theta_\lambda\|_V \leq C, \quad \|\psi_\lambda\|_V \leq C,$$

где постоянная $C > 0$ не зависит от λ . Поэтому можно выбрать последовательность $\lambda \rightarrow +0$ такую, что

$$u_\lambda \rightarrow u_* \text{ слабо в } U, \quad \theta_\lambda, \psi_\lambda \rightarrow \theta_*, \psi_* \text{ слабо в } V, \text{ сильно в } H, L^4(\Omega), \quad (20)$$

$$\theta_\lambda|_{\Gamma_1} \rightarrow \theta_*|_{\Gamma_1} \text{ сильно в } L^2(\Gamma_1).$$

Результаты о сходимости (20) позволяют перейти к пределу при $\lambda \rightarrow +0$ в уравнениях для $\theta_\lambda, \psi_\lambda, u_\lambda$, и тогда

$$A_1\theta_* + g(\theta_*) = \frac{\kappa_a}{\alpha}\psi_* + f_1, \quad A_2\psi_* = f_2 + B_2u_*, \quad \theta_*|_{\Gamma_1} = \theta_b. \quad (21)$$

Полагая $\varphi_* = \frac{1}{\alpha b}(\psi_* - a\theta_*)$, из первого уравнения в (21), учитывая, что $\theta_*|_{\Gamma_1} = \theta_b$, выводим равенство

$$a(\nabla\theta_*, \nabla v) + b\kappa_a(|\theta_*|\theta_*^3 - \varphi_*, v) = \int_{\Gamma} q_b v d\Gamma \quad \forall v \in V. \quad (22)$$

Далее, из второго уравнения в (21), с учетом полученного равенства (22), следует

$$\alpha(\nabla\varphi_*, \nabla v) + \int_{\Gamma_1} \gamma\varphi_* v d\Gamma + \kappa_a(\varphi_* - |\theta_*|\theta_*^3, v) = \int_{\Gamma_1} \gamma\theta_b^4 v d\Gamma + \int_{\Gamma_2} q_* v d\Gamma, \quad (23)$$

где $q_* = \frac{1}{\alpha b} u_* - \frac{1}{b} q_b$. Поэтому пара $\{\theta_*, \varphi_*\}$ – слабое решение задачи (1)–(3), которое в силу единственности совпадает с $\{\theta, \varphi\}$. Кроме того, условие единственности решения гарантирует, что сходимости (20) имеют место для любой последовательности $\lambda \rightarrow +0$.

Замечание 2. В случае, если решение задачи (1)–(3) неединственно, то из доказательства теоремы 3 следует существование последовательности $\lambda \rightarrow +0$, для которой будет указанная сходимость к одному из решений.

5. ЧИСЛЕННЫЕ ПРИМЕРЫ

Представим итерационный алгоритм решения задачи оптимального управления. Пусть $\tilde{J}_\lambda(u) = J_\lambda(\theta(u), u)$, где $\theta(u)$ – компонента решения задачи (1)–(3), соответствующая управлению $u \in U$.

В соответствии с (17) градиент функционала $\tilde{J}_\lambda(u)$ равен

$$\tilde{J}'_\lambda(u) = \lambda u - p_2.$$

Здесь p_2 – соответствующая компонента сопряженного состояния из системы (17), где $\hat{\theta} := \theta(u)$.

Алгоритм градиентного спуска

1. Выбираем значение градиентного шага ε .
2. Выбираем количество итераций N .
3. Выбираем начальное приближение для управления $u_0 \in U$.
4. **for** $k \leftarrow 0, 1, \dots, N$ **do**
5. Для данного u_k рассчитываем состояние $y_k = \{\theta_k, \varphi_k\}$ – решение задачи (1), (2).
6. Рассчитываем значение целевого функционала $J_\lambda(\theta_k, u_k)$.
7. Рассчитываем сопряженное состояние $p_k = \{p_{1k}, p_{2k}\}$ из уравнений (15), где $\hat{\theta} := \theta_k, \hat{u} = u_k$.
8. Пересчитываем управление $u_{k+1} = u_k - \varepsilon(\lambda u_k - p_2)$

Значение параметра ε выбирается эмпирически таким образом, чтобы значение $\varepsilon(\lambda u_k - p_2)$ являлось существенной поправкой для u_{k+1} . Количество итераций N выбирается достаточным для выполнения условия $J_\lambda(\theta_k, u_k) - J_\lambda(\theta_{k+1}, u_{k+1}) < \delta$, где $\delta > 0$ определяет точность расчетов.

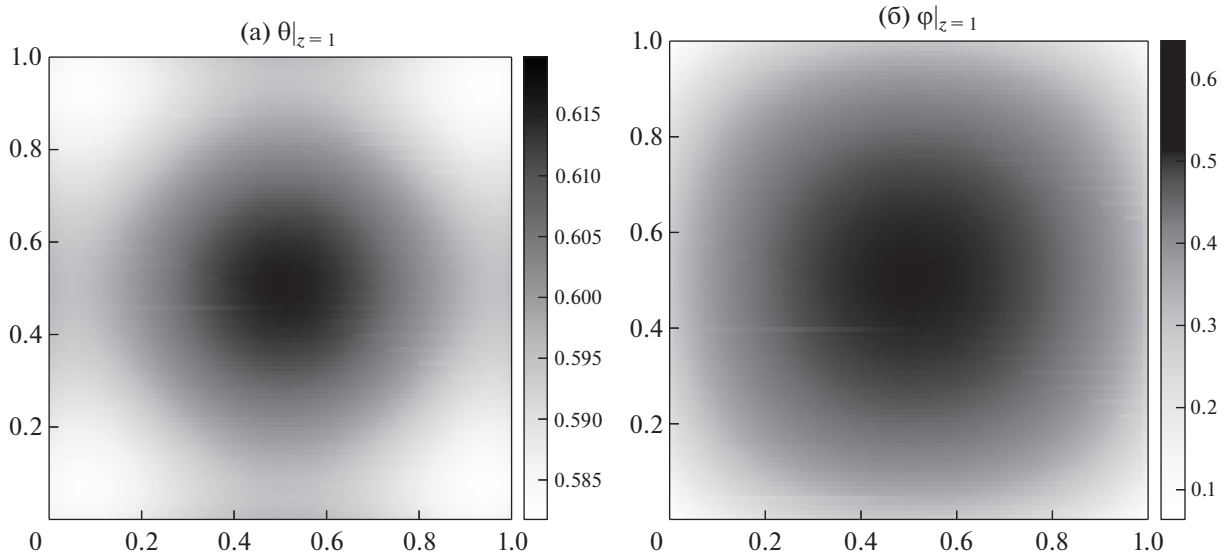
Для анализа работы алгоритма проводились серии вычислительных экспериментов при различных малых значениях регуляризирующего параметра λ , которые подтверждают эффективность предложенного оптимизационного алгоритма нахождения решения краевой задачи сложного теплообмена с данными Коши на части границы. Численное решение прямой задачи с заданным управлением проводилось методом Ньютона для линеаризации задачи и ее решения методом конечных элементов. Решение сопряженной системы, которая является линейной при заданной температуре, не вызывает трудностей. Для численного моделирования использовался солвер (см. FEniCS [25], [26]). Исходный код экспериментов можно найти по ссылке [27].

Примеры, рассмотренные ниже, иллюстрируют работоспособность предложенного алгоритма при малых, что важно, значениях параметра регуляризации $\lambda \leq 10^{-12}$.

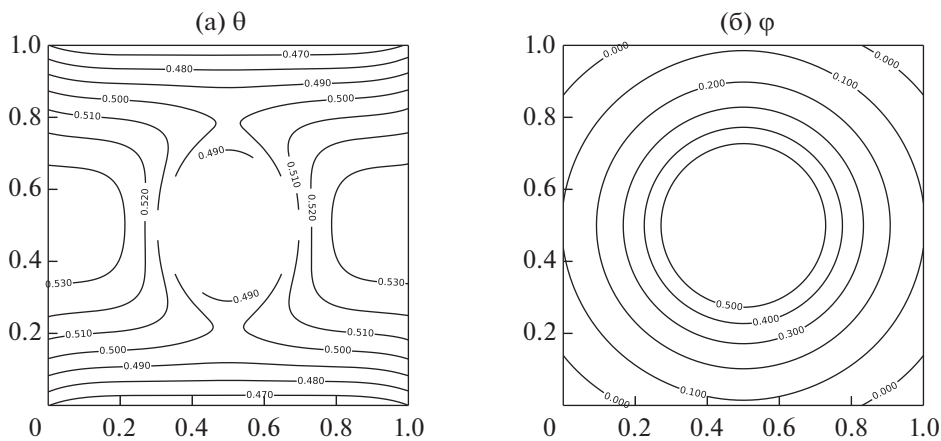
Пример 1. Рассмотрим куб $\Omega = (x, y, z) : 0 < x, y, z < l$ с границей $\Gamma = \Gamma_1 \cup \Gamma_2$, где

$$\Gamma_1 = \{(x, y, z), 0 \leq x, y \leq l, z = 0, l\}, \quad \Gamma_2 = \partial\Omega \setminus \Gamma_1.$$

Будем считать, что $l = 1$ см, $a = 0.6$ см²/с, $b = 0.025$ см/с, $\kappa_a = 1$ см⁻¹, $\alpha = 0.3$ см. Указанные параметры соответствуют стеклу (см. [9]). Параметр регуляризации $\lambda = 10^{-12}$.



Фиг. 1. Результаты первого эксперимента.



Фиг. 2. Результаты второго эксперимента.

Пусть граничные данные задачи (2), (3) имеют вид $q_b = 0.5$ и $\theta_b = 0.1 + z/2$. Задачу решаем, используя предложенный оптимизационный алгоритм с параметром регуляризации $\lambda = 10^{-12}$ и начальным приближением управления $u_0 = 0$.

На фиг. 1а и 1б представлены полученные поля θ и ϕ на грани куба $z = 1$. Начальное значение целевого функционала 0.025 и через сто итераций становится равным 5×10^{-5} .

Пример 2. Рассмотрим двумерный случай, где область Ω является квадратом с круговой полостью:

$$\Omega = \{(x, y) : 0 < x, y < l, (x - l/2)^2 + (y - l/2)^2 > \rho^2\}, \quad l = 1 \text{ см}, \quad \rho = 0.15 \text{ см}.$$

Условия для интенсивности излучения неизвестны на внутренней границе $\Gamma_2 = \{(x, y) : (x - l/2)^2 + (y - l/2)^2 = \rho^2\}$, а условие переопределения ставится на $\Gamma_1 = \Gamma \setminus \Gamma_2$. Параметры среды такие же, как в примере 1, а граничные данные q_b и θ_b положим равными

$$\theta_b = 0.5, \quad q_b = \begin{cases} 0.2, & \text{если } x \in \Gamma_1 \\ -0.2, & \text{если } x \in \Gamma_2. \end{cases}$$

На фиг. 2а и 2б представлены полученные поля θ и φ . Начальное значение целевого функционала 0.045 и через тридцать итераций становится равным 6.2×10^{-5} .

Представленные численные примеры демонстрируют успешную работу предложенного алгоритма численного решения граничной задачи (1)–(3).

СПИСОК ЛИТЕРАТУРЫ

1. *Kovtanyuk A.E., Chebotarev A.Yu., Botkin N.D., Hoffmann K.-H.* The unique solvability of a complex 3D heat transfer problem // *J. Math. Anal. Appl.* 2014. V. 409. № 2. P. 808–815.
2. *Алифанов О.М., Ненарокомов А.В., Артюхин Е.А.* Обратные задачи в исследовании сложного теплообмена. М., Янус-К. 2009. 297 с.
3. *Алифанов О.М.* Определение тепловых нагрузок из решения нелинейной обратной задач // *Теплофизика высоких температур.* 1977. Т. 15. Вып. 3. С. 598–605.
4. *Алифанов О.М., Ненарокомов А.В.* Трехмерная граничная обратная задача теплопроводности // *Теплофизика высоких температур.* 1999. Т. 37. № 2. С. 231–238.
5. *Ковтанюк А.Е., Чеботарев А.Ю.* Стационарная задача сложного теплообмена // *Ж. вычисл. матем. и матем. физ.* 2014. Т. 54. № 4. С. 711–719.
6. *Kovtanyuk A.E., Chebotarev A.Y., Botkin N.D., Hoffmann K.-H.* Solvability of P1 approximation of a conductive-radiative heat transfer problem // *Appl. Math. Comput.* 2014. V. 249. P. 247–252.
7. *Kovtanyuk A.E., Chebotarev A.Yu., Botkin N.D., Hoffmann K.-H.* Unique solvability of a steady-state complex heat transfer model // *Commun. Nonlinear Sci. Numer. Simul.* 2015. V. 20. № 3. P. 776–784.
8. *Pinnau R.* Analysis of optimal boundary control for radiative heat transfer modeled by SP_1 -system // *Commun. Math. Sci.* 2007. V. 5. № 4. P. 951–969.
9. *Grenkin G.V., Chebotarev A.Yu., Kovtanyuk A.E., Botkin N.D., Hoffmann K.-H.* Boundary optimal control problem of complex heat transfer model // *J. Math. Anal. Appl.* 2016. V. 433. № 2. P. 1243–1260.
10. *Kovtanyuk A.E., Chebotarev A.Yu., Botkin N.D., Hoffmann K.-H.* Optimal boundary control of a steady-state heat transfer model accounting for radiative effects // *J. Math. Anal. Appl.* 2016. V. 439. № 2. P. 678–689.
11. *Chebotarev A.Yu., Kovtanyuk A.E., Grenkin G.V., Botkin N.D., Hoffmann K.-H.* Nondegeneracy of optimality conditions in control problems for a radiative-conductive heat transfer model // *Appl. Math. Comput.* 2016. V. 289. P. 371–380.
12. *Chebotarev A.Yu., Grenkin G.V., Kovtanyuk A.E., Botkin N.D., Hoffmann K.-H.* Inverse problem with finite overdetermination for steady-state equations of radiative heat exchange // *J. Math. Anal. Appl.* 2018. V. 460. № 2. P. 737–744.
13. *Chebotarev A.Y., Kovtanyuk A.E., Botkin N.D.* Problem of radiation heat exchange with boundary conditions of the Cauchy type // *Commun. Nonlinear Sci. and Numeric. Simulat.* 2019. V. 75. P. 262–269.
14. *Chebotarev A.Yu., Pinnau R.* An inverse problem for a quasi-static approximate model of radiative heat transfer // *J. Math. Anal. Appl.* 2019. V. 472. № 1. P. 314–327.
15. *Гренкин Г.В., Чеботарев А.Ю.* Обратная задача для уравнений сложного теплообмена // *Ж. вычисл. матем. и матем. физ.* 2019. Т. 59. № 8. С. 1420–1430.
16. *Месенев П.Р., Чеботарев А.Ю.* Граничная обратная задача для уравнений сложного теплообмена // *Дальневост. матем. журн.* 2018. Т. 18. № 1. С. 75–84.
17. *Месенев П.Р., Чеботарев А.Ю.* Анализ оптимизационного метода решения задачи сложного теплообмена с граничными условиями типа Коши // *Ж. вычисл. матем. и матем. физ.* 2022. Т. 62. № 1. С. 36–44.
18. *Amosov A.A.* Unique solvability of stationary radiative-conductive heat transfer problem in a system of semitransparent bodies // *J. Math. Sci.* 2017. V. 224. № 5. P. 618–646.
19. *Amosov A.A., Крытов N.E.* On a nonstandard boundary value problem arising in homogenization of complex heat transfer problems // *J. Math. Sci.* 2020. V. 244. P. 357–377.
20. *Amosov A.* Unique solvability of a stationary radiative-conductive heat transfer problem in a system consisting of an absolutely black body and several semitransparent bodies // *Math. Meth. Appl. Sci.* 2021. V. 44. № 13. P. 10703–10733.
21. *Amosov A.A.* Unique solvability of the stationary complex heat transfer problem in a system of gray bodies with semitransparent inclusions // *J. Math. Sci. (United States).* 2021. V. 255. Iss. 4. P. 353–388.
22. *Amosov A.* Nonstationary radiative-conductive heat transfer problem in a semitransparent body with absolutely black inclusions // *Mathematics* 2021. V. 9. № 13. P. 1471.
23. *Fučik S., Kufner A.* Nonlinear differential equations. Amsterdam—Oxford—New York, Elsevier, 1980.
24. *Fursikov A.V.* Optimal control of distributed systems // *Theory and Appl., Am. Math. Soc.*, 2000.
25. *Alnaes M.S., Blechta J., Hake J., Johansson A., Kehlet B., Logg A., Richardson C., Ring J., Rognes M.E., Wells G.N.* The FEniCS Project Version 1.5 Archive of Numerical Software, vol. 3, 2015.
26. *Logg A., Wells G.N.* DOLFIN: Automated Finite Element Computing ACM Transactions on Mathematical Software, vol. 37, 2010.
27. https://github.com/mesenev/articles_src

О СХЕМЕ РУСАНОВА ТРЕТЬЕГО ПОРЯДКА ТОЧНОСТИ ДЛЯ МОДЕЛИРОВАНИЯ ПЛАЗМЕННЫХ КОЛЕБАНИЙ¹⁾

© 2023 г. Е. В. Чижонков^{1,*}

¹ 119899 Москва, Ленинские горы, МГУ им. М.В. Ломоносова, Россия

*e-mail: chizhonk@mech.math.msu.su

Поступила в редакцию 25.05.2022 г.

Переработанный вариант 02.12.2022 г.

Принята к публикации 02.02.2023 г.

Для моделирования нерелятивистских колебаний холодной плазмы предложена модификация известной схемы Русанова, имеющей третий порядок точности. Ранее для подобных расчетов в эйлеровых переменных были известны только схемы первого и второго порядков точности. Для тестовой задачи с гладким решением проведено исследование погрешностей построенной схемы, а также – сравнение с погрешностями схемы Мак-Кормака. Для задачи о свободных плазменных колебаниях, инициированных коротким мощным лазерным импульсом, приведены результаты численных экспериментов по сохранению энергии и дополнительной функции для обеих схем, а также – по точности электронной плотности в центре области. Сделан вывод о теоретическом превосходстве схемы Русанова, хотя для практических вычислений более приспособлена схема Мак-Кормака. В первую очередь это касается расчетов “долгоживущих” процессов и колебаний холодной плазмы, близких к реальным. Теоретическое исследование аппроксимации и устойчивости вместе с экспериментальным наблюдением за количественными характеристиками погрешности для наиболее чувствительных величин существенно повышает достоверность вычислений. Библ. 20. Фиг. 3. Табл. 4.

Ключевые слова: численное моделирование, плазменные колебания, схемы Русанова и Мак-Кормака, порядок точности разностной схемы, законы сохранения, гиперболические системы уравнений.

DOI: 10.31857/S0044466923050083, EDN: GFRVIX

ВВЕДЕНИЕ

При расчетах гиперболических законов сохранения порядок точности метода играет важную роль [1]. Сфокусируем внимание на схемах сквозного счета, которые допускают естественное обобщение на задачи пространственной размерности более единицы. Традиционными требованиями к таким схемам являются монотонность в окрестности разрыва решения и порядок аппроксимации выше первого на гладкой части решения. Это привело к тому, что точное или приближенное решение классической задачи Римана (задачи Коши с кусочно постоянными начальными данными) для дивергентной формы уравнений является основой большинства современных алгоритмов численного решения для соответствующих постановок задач [1]. Следует отметить, что в задачах, связанных с плазменными колебаниями, постановка классической задачи Римана не имеет никакого физического смысла, так как начальная разрывная функция электрического поля уже означает бесконечную концентрацию заряда в точках разрыва. Подобная ситуация обозначается термином “опрокидывание” (см. [2], стр. 12–15). Иными словами, схемы сквозного счета, предложенные изначально для гиперболических законов сохранения в дивергентной форме, требуют специальных модификаций при использовании их для расчетов плазменных колебаний, так как в этом случае часть уравнений имеет дивергентную форму, а часть – нет. Более того, само решение уравнений с частными производными, характеризующими, например, простейшую гидродинамическую модель холодной плазмы, имеет следующую специфику: стартуя с гладких начальных данных, оно либо остается глобально гладким во времени и пространстве, либо на конечном промежутке времени приходит к градиентной катастрофе, т.е.

¹⁾Работа выполнена при финансовой поддержке Минобрнауки РФ в рамках реализации программы Московского центра фундаментальной и прикладной математики по соглашению 075-15-2022-284.

когда само решение остается ограниченным, а его пространственные производные (градиенты) становятся сингулярными.

Обозначим актуальность развития схем сквозного счета для подобных постановок. Они напрямую связаны с приложением результатов, удостоенных Нобелевской премии по физике 2018 года. Приведем следующие примеры практически важных задач этой тематики: лазерное ускорение электронов и ионов, быстрое зажигание термоядерного синтеза, ядерные реакции в луче лазера, синхротронное и субмиллиметровое излучение и пр. Основным методом их исследования является численное моделирование; поэтому необходимым являются разработка и обоснование соответствующих алгоритмов для систем дифференциальных уравнений, составляющих рабочую математическую модель. Численному моделированию колебаний в холодной плазме, а также — кильватерных волн, возбуждаемых коротким мощным лазерным импульсом, посвящена монография [2].

В работе [3] впервые были предложены соответствующие модификации классических схем второго порядка точности: Мак-Кормака и Лакса—Вендрофа. При этом было показано, что первая из них имеет несомненные практические преимущества перед второй, хотя в теоретических аспектах они весьма схожи. В настоящей работе рассматривается модификация на случай расчета плазменных колебаний схемы В.В. Русанова третьего порядка точности [4], которая стала широко известной благодаря публикациям [5], [6]. Сама идея модификации проста: схема Русанова, в силу ориентации на расчет разрывных решений, излагается в терминах конечных разностей (а не аппроксимаций производных!); каждая разность представляет собой по определению интеграл либо может быть представлена в виде суммы интегралов от функций, входящих в уравнения не дивергентной формы; поэтому для замены “разности—интеграла” надо использовать квадратурную формулу с подходящей оценкой погрешности. Для схем второго порядка точности можно было ограничиться квадратурной формулой прямоугольников с узлом в центральной точке, для схемы Русанова этого недостаточно, нужна формула парабол (Симпсона) [7].

Статья имеет следующую структуру. В первом разделе приведена простейшая (плоская пространственно одномерная) система уравнений, описывающая колебания холодной плазмы в эйлеровых переменных. В рассматриваемом случае она имеет гиперболический тип, однако ее форма записи имеет как дивергентное уравнение, так и не дивергентное. Также в первом разделе указаны аналитические свойства решения дифференциальной задачи: теорема существования и единственности, законы сохранения, нечетность решения при нечетности начальных данных, формула электронной плотности на оси симметрии области. Во втором разделе приведена модификация классической схемы Русанова третьего порядка точности. В третьем разделе, самом объемном, приведены результаты расчетов. Сначала они иллюстрируют порядок точности схемы Русанова на гладком специально сконструированном решении неоднородной задачи; для сравнения приводятся аналогичные расчеты по схеме Мак-Кормака. Затем демонстрируется применение схемы Русанова на расчете плазменных колебаний, инициированных коротким мощным лазерным импульсом, что также сопровождается сравнением с результатами схемы Мак-Кормака. В заключении систематизированы результаты проведенных исследований.

1. ПОСТАНОВКА ЗАДАЧИ И ЕЕ СВОЙСТВА

1.1. Уравнения и начальные условия

Будем считать плазму нерелятивистской электронной жидкостью, пренебрегая рекомбинационными эффектами и движением ионов. Тогда в рамках одной из простейших моделей плазмы, которую часто называют уравнениями гидродинамики “холодной” плазмы (см., например, [8], [9]), ее плоские одномерные колебания можно описать безразмерной системой уравнений

$$\frac{\partial V}{\partial \theta} + V \frac{\partial V}{\partial \rho} = -E, \quad \frac{\partial E}{\partial \theta} + V \frac{\partial E}{\partial \rho} = V, \quad (1)$$

где V — скорость электронов, E — электрическое поле. К системе (1) обычно добавляют уравнение

$$N(\rho, \theta) = 1 - \frac{\partial E(\rho, \theta)}{\partial \rho}, \quad (2)$$

характеризующее безразмерную плотность электронов N . Формула (2) является частным случаем теоремы Гаусса [10], которая в дифференциальной размерной форме имеет вид

$\operatorname{div} \mathbf{E} = 4\pi e(n - n_0)$. Вывод уравнений (1), (2) можно найти в различных источниках (см., например, [2], [3], [16]).

Ниже мы будем изучать в полуплоскости $\{(\rho, \theta) : \rho \in \mathbb{R}, \theta > 0\}$ решение задачи Коши для (1) с начальными условиями

$$V(\rho, 0) = V_0(\rho), \quad E(\rho, 0) = E_0(\rho), \quad \rho \in \mathbb{R}. \quad (3)$$

Система (1) относится к гиперболическому типу. Хорошо известно, что для таких систем существует локально по времени единственное решение задачи Коши того же класса, что и начальные данные, в нашем случае это C^2 . Также известно, что для таких систем потеря решением гладкости происходит по одному из следующих сценариев: либо сами компоненты решения в течение конечного времени обращаются в бесконечность, либо они остаются ограниченными, но в бесконечность обращаются их производные [11]. Последняя возможность реализуется, например, для однородных законов сохранения, к которым относятся уравнения газовой динамики, где возникновение особенности соответствует образованию ударной волны.

Следует отметить, что моделирование нерелятивистских колебаний холодной плазмы можно проводить также на основе системы уравнений Эйлера–Пуассона, которая имеет смешанный тип: гиперболический и эллиптический одновременно [12]. Однако при таком подходе отсутствует расщепление уравнений, аналогичное (1), (2), что автоматически увеличивает примерно в полтора раза объем вычислительной работы при дальнейшем использовании любого численного алгоритма.

1.2. Свойства решения

Напомним основные свойства решения задачи (1), (3), впервые полученные в работе [13] (подробные доказательства приведены в [14]).

Теорема существования и единственности (Для рассматриваемой постановки принципиально важной является теорема существования глобального по времени решения [13], [14]). Пусть начальные данные (3) принадлежат классу $C^2(\mathbb{R})$. Для существования и единственности непрерывно дифференцируемого по обеим переменным 2π -периодического по времени при всех $\theta > 0$ решения $V(\theta, \rho)$, $E(\theta, \rho)$ задачи (1), (3) необходимо и достаточно, чтобы в каждой точке $\rho \in (-\infty, \infty)$ было выполнено неравенство

$$(V_0'(\rho))^2 + 2E_0'(\rho) - 1 < 0. \quad (4)$$

Если же существует хотя бы одна точка ρ , для которой выполняется неравенство, противоположное (4), то в течение конечного времени производные решения обращаются в бесконечность.

Следствие. Если необходимое и достаточное условие существования решения (4) выполнено в начальный момент времени, то оно сохраняется во времени и пространстве, т.е. для произвольной точки (ρ, θ) справедливо

$$\left(\frac{\partial V(\rho, \theta)}{\partial \rho}\right)^2 + 2\frac{\partial E(\rho, \theta)}{\partial \rho} - 1 < 0. \quad (5)$$

Даже если же условие (4) не выполнено, то сами функции $V(\rho, \theta)$ и $E(\rho, \theta)$ все равно остаются ограниченными, в силу равенства $V^2 + E^2 = V_0^2(\rho_0) + E_0^2(\rho_0) = \text{const}$, справедливого для произвольной характеристики системы (1), стартующей из точки ρ_0 . Отсюда следует, что разрушение решения может иметь только вид градиентной катастрофы. В частности, нарушение условия (4) приводит к обращению электронной плотности в бесконечность, т.е. к опрокидыванию колебаний.

Следует отметить, что, в силу соотношения (2), из (5) вытекают строгая положительность и отделенность от нуля функции электронной плотности:

$$N(\rho, \theta) > \frac{1}{2}.$$

Таким образом, из приведенной информации следует, что решение задачи (1) с начальными условиями (3) полностью характеризуется поведением функций $V(\rho, \theta)$, $E(\rho, \theta)$ и их пространственных производных (градиентов).

Законы сохранения. Дифференциальная форма записи закона сохранения полной энергии электрического поля и электронов плазмы в рассматриваемом случае имеет вид [3], [15]

$$\frac{\partial}{\partial \theta}(E^2 + NV^2) + \frac{\partial}{\partial \rho}(NV^3) = 0. \quad (6)$$

После интегрирования этого соотношения по области $\rho \in (-\infty, \infty)$ при использовании однородных условий на бесконечности получим сохранение во времени величины

$$Eh(\theta) = \int_{-\infty}^{\infty} [E^2(\rho, \theta) + N(\rho, \theta)V^2(\rho, \theta)] d\rho \equiv \text{const}. \quad (7)$$

В работе [3] также приведен закон сохранения некоторой вспомогательной функции. В дифференциальной форме он имеет вид

$$\frac{\partial}{\partial \theta} N(E^2 + V^2) + \frac{\partial}{\partial \rho} NV(E^2 + V^2) = 0. \quad (8)$$

Из уравнения (8) при использовании однородных условий на бесконечности следует сохранение во времени величины

$$Af(\theta) = \int_{-\infty}^{\infty} N(\rho, \theta) [E^2(\rho, \theta) + V^2(\rho, \theta)] d\rho. \quad (9)$$

Обратим внимание, что соотношение (8) не противоречит закону сохранения энергии (6): разница между соответствующими выражениями есть просто утверждение теоремы Гаусса (2), умноженное на подходящую комбинацию функций.

Нечетность решения. Предположим, что начальные условия (3) нечетны относительно оси симметрии области, т.е. прямой $\rho = 0$. Покажем, что в этом случае решение системы (1) будет оставаться нечетным в течение всего времени его существования.

Рассмотрим решение $V(\theta, \rho)$, $E(\theta, \rho)$ задачи (1) с нечетными начальными условиями (3). Одновременно с ним рассмотрим пару функций $\hat{V}(\theta, \rho) = -V(\theta, -\rho)$, $\hat{E}(\theta, \rho) = -E(\theta, -\rho)$. При $\theta = 0$, в силу нечетности (3), справедливы равенства

$$\hat{V}(0, \rho) = -V(0, -\rho) = V(0, \rho), \quad \hat{E}(0, \rho) = -E(0, -\rho) = E(0, \rho).$$

Кроме того, функции $\hat{V}(\theta, \rho)$, $\hat{E}(\theta, \rho)$ удовлетворяют уравнениям

$$\frac{\partial \hat{V}}{\partial \theta} + \hat{V} \frac{\partial \hat{V}}{\partial \rho} = -\hat{E}, \quad \frac{\partial \hat{E}}{\partial \theta} + \hat{V} \frac{\partial \hat{E}}{\partial \rho} = \hat{V}, \quad (10)$$

что несложно проверяется непосредственной подстановкой в (1).

В результате имеем, что пары функций $V(\theta, \rho)$, $E(\theta, \rho)$ и $\hat{V}(\theta, \rho)$, $\hat{E}(\theta, \rho)$ удовлетворяют одинаковым уравнениям и одинаковым начальным условиям. Отсюда, в силу Теоремы 1 о существовании и единственности решения системы (1) с начальными условиями (3), получаем совпадение функций в произвольный момент времени θ

$$\hat{V}(\theta, \rho) = V(\theta, \rho), \quad \hat{E}(\theta, \rho) = E(\theta, \rho).$$

Теперь, вспоминая определения функций \hat{V} , \hat{E} получаем нечетность решения V , E для любого θ , если начальные данные были нечетными.

Это свойство решения исходной задачи позволяет уменьшить область определения решения в два раза (было $\rho \in (-\infty, \infty)$, стало $\rho \in (0, \infty)$). При этом, вследствие свойства нечетности на оси $\rho = 0$ симметрии, формируются граничные условия:

$$V(\theta, 0) = E(\theta, 0) = 0 \quad \forall \theta.$$

Формула для плотности на оси симметрии. Продолжим изучение решения системы (1) с нечетными начальными данными (3). Учитывая его нечетность, выпишем соответствующее разложение в ряд Тейлора в окрестности прямой $\rho = 0$:

$$V(\theta, \rho) = \rho \left. \frac{\partial V(\theta, \rho)}{\partial \rho} \right|_{\rho=0} + O(\rho^3), \quad E(\theta, \rho) = \rho \left. \frac{\partial E(\theta, \rho)}{\partial \rho} \right|_{\rho=0} + O(\rho^3).$$

Нас интересует поведение производных на оси симметрии $\rho = 0$

$$W(\theta) = \left. \frac{\partial V(\theta, \rho)}{\partial \rho} \right|_{\rho=0}, \quad D(\theta) = \left. \frac{\partial E(\theta, \rho)}{\partial \rho} \right|_{\rho=0},$$

чтобы с помощью общей формулы (2) получить частную формулу

$$N(\rho = 0, \theta) = 1 - \left. \frac{\partial E(\theta, \rho)}{\partial \rho} \right|_{\rho=0} = 1 - D(\theta).$$

После подстановки разложений Тейлора в систему (1) и учета условия $\rho = 0$ приходим к вспомогательной задаче для функций $W(\theta)$ и $D(\theta)$ с вещественными начальными данными:

$$\begin{aligned} W' &= -D - W^2, & D' &= (1 - D)W, \\ W(0) &= \beta, & D(0) &= \alpha. \end{aligned} \quad (11)$$

Эти начальные данные есть производные по ρ начальных функций из (3), взятые при $\rho = 0$.

Справедливо следующее (см. [16])

Утверждение. Пусть величина $s = (1 - \alpha)/\sqrt{\alpha^2 + \beta^2}$ удовлетворяет условию $s > 1$, тогда решение (11) имеет вид

$$W(\theta) = \frac{\cos(\theta + \theta_2)}{s + \sin(\theta + \theta_2)}, \quad D(\theta) = \frac{\sin(\theta + \theta_2)}{s + \sin(\theta + \theta_2)}, \quad (12)$$

где

$$\cos \theta_2 = \frac{\beta}{\sqrt{\alpha^2 + \beta^2}}, \quad \sin \theta_2 = \frac{\alpha}{\sqrt{\alpha^2 + \beta^2}}. \quad (13)$$

В дальнейшем будет полезен частный случай этого решения, когда $W(0) = \beta = 0$. Это дает: $\theta_2 = \pi/2$, $s = (1 - \alpha)/\alpha$, следовательно,

$$N(\rho = 0, \theta) = \frac{1 - \alpha}{1 - \alpha(1 - \cos \theta)}, \quad \alpha = E'_0(\rho = 0). \quad (14)$$

Функция электронной плотности N традиционно является очень важной для описания процессов в холодной плазме, при этом – наиболее сильно изменяющейся в сравнении с V и E . Предполагается использовать формулу (14) для анализа точности численного метода решения задачи (1), (3).

2. СХЕМА РУСАНОВА ТРЕТЬЕГО ПОРЯДКА ТОЧНОСТИ

2.1. Формулы для дивергентной записи уравнения

Пусть исходное уравнение имеет вид

$$\frac{\partial \mathbf{U}}{\partial \theta} + \frac{\partial \mathbf{F}(\mathbf{U})}{\partial \rho} = \mathbf{S}(\mathbf{U}, \rho, \theta), \quad (15)$$

где \mathbf{U} , \mathbf{F} , \mathbf{S} – вектор-функции, рассматриваемые в полуплоскости $\{(\rho, \theta) : \theta \geq 0, \rho \in \mathbb{R}\}$, и в момент времени $\theta = 0$ заданы начальные условия

$$\mathbf{U}(\rho, \theta = 0) = \mathbf{U}^0(\rho), \quad \rho \in \mathbb{R}. \quad (16)$$

Будем считать, что нас интересует приближенное решение задачи Коши, определенной соотношениями (15), (16), про которое известно, что оно существует, единственно и обладает достаточной гладкостью.

Определим дискретизацию независимых переменных с помощью постоянных параметров τ и h так, что

$$\theta^n = n\tau, \quad n \geq 0, \quad \rho_i = ih, \quad i = 0, \pm 1, \pm 2, \dots$$

и будем обозначать зависимую переменную $U(\rho, \theta)$ в узле сетки (ρ_i, θ^n) как U_i^n . Стандартная схема Русанова третьего порядка точности, предложенная в [4], состоит из трех этапов вычислений и может быть представлена как

$$U_{i+1/2}^{(1)} = (U_{i+1}^n + U_i^n)/2 - \frac{\tau}{3h}(F_{i+1}^n - F_i^n) + \frac{\tau}{3}(S_{i+1}^n + S_i^n)/2, \quad (17)$$

$$U_i^{(2)} = U_i^n - \frac{2\tau}{3h}(F_{i+1/2}^{(1)} - F_{i-1/2}^{(1)}) + \frac{2\tau}{3}(S_{i+1/2}^{(1)} + S_{i-1/2}^{(1)})/2, \quad (18)$$

$$U_i^{n+1} = U_i^n - \frac{3\tau}{8h}(F_{i+1}^{(2)} - F_{i-1}^{(2)}) + \tau(S_i^n + 3S_i^{(2)})/4 - \frac{\tau}{24h}(-2F_{i+2}^n + 7F_{i+1}^n - 7F_{i-1}^n + 2F_{i-2}^n) - w_{32}(U_{i+2}^n - 4U_{i+1}^n + 6U_i^n - 4U_{i-1}^n + U_{i-2}^n), \quad (19)$$

где верхний индекс обозначает промежуточный шаг ((1) или (2)) или номер временного слоя (n или $n+1$). Обратим внимание, что значения функций, помеченные индексом (1), относятся к моменту времени $\theta^n + \tau/3$, поэтому величина $S_{i+1/2}^{(1)}$ вычисляется как $S(U_{i+1/2}^{(1)}, \rho_i + h/2, \theta^n + \tau/3)$.

Аналогично – индекс (2) относится к моменту времени $\theta^n + 2\tau/3$, поэтому, например, величина $S_i^{(2)}$ вычисляется как $S(U_i^{(2)}, \rho_i, \theta^n + 2\tau/3)$.

Отметим также, что последнее слагаемое в (19) не связано с точностью метода, так как пропорционально разностному оператору четвертого порядка. Зато оно влияет на устойчивость. В частности, в [4–6] для модельного уравнения $u_t + cu_x = 0$, $c = \text{const}$ показано, что устойчивость обеспечивается выполнением неравенств

$$|v| \leq 1, \quad 4v^2 - v^4 \leq 24w_{32} \leq 3, \quad \text{где } v = c\tau/h. \quad (20)$$

Приведенная выше схема хорошо известна и давно используется в вычислительной практике. Ее свойства представлены в популярных монографиях, посвященных численному анализу и математическому моделированию (см., например, [18], [19]). Напомним, что она на гладких решениях имеет порядок аппроксимации $O(\tau^3 + h^4/\tau)$ и условие устойчивости $\tau = O(h)$, поэтому ее традиционно называют “схемой третьего порядка точности”.

2.2. Модификация на основе формулы Симпсона

Как правило, разностные методы формулируются для дивергентной формы уравнений (15), поэтому определенный интерес представляет “обход” недивергентной записи уравнения, описывающего динамику электрического поля,

$$\frac{\partial E}{\partial \theta} + V \frac{\partial E}{\partial \rho} = V.$$

Определим с этой целью в произвольный требуемый момент времени θ , в данном случае играющий роль параметра, вспомогательную функцию $\varphi(\rho, \theta)$ соотношениями

$$\frac{d\varphi}{d\rho} = V \frac{\partial E}{\partial \rho}, \quad \varphi(a, \theta) = \varphi_a(\theta). \quad (21)$$

Заметим, что при достаточной гладкости определяющих правую часть функций $V(\rho, \theta)$ и $E(\rho, \theta)$, функция $\varphi(\rho, \theta)$ существует, единственна и обладает гладкостью не меньшей, чем $E(\rho, \theta)$.

Обратим также внимание, что в формулах (17)–(19) участвуют не просто значения φ в узлах сетки (либо с целыми, либо с дробными узлами), а разности этих значений, например, $(\varphi_{i+1} - \varphi_i)$, представляющие собой по определению интегралы вида

$$\varphi_{i+1} - \varphi_i \equiv \int_{\rho_i}^{\rho_{i+1}} \frac{d\varphi}{d\rho} d\rho = \int_{\rho_i}^{\rho_{i+1}} V \frac{\partial E}{\partial \rho} d\rho. \quad (22)$$

В целях аппроксимации интеграла в (22) воспользуемся квадратурной формулой Симпсона (парабол) [7], т.е.

$$\int_a^b f(x) dx = \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right] + O((b-a)^5).$$

Учитывая, что в (22) длина отрезка интегрирования равна h , имеем

$$\varphi_{i+1} - \varphi_i = \frac{h}{6} \left[\left(V \frac{\partial E}{\partial \rho} \right)_{\rho=\rho_i} + 4 \left(V \frac{\partial E}{\partial \rho} \right)_{\rho=\rho_{i+1/2}} + \left(V \frac{\partial E}{\partial \rho} \right)_{\rho=\rho_{i+1}} \right] + O(h^5).$$

Теперь, вспоминая, что при условии устойчивости вида $\tau = O(h)$ порядок точности метода Русанова на конечном отрезке интегрирования есть $O(\tau^3 + h^4/\tau)$ (см. [18], [19]), получаем, что локальная погрешность на одном шаге по времени, равная $O(\tau^4 + h^4)$, не ухудшится вследствие замены точного значения интеграла его аппроксимацией по формуле Симпсона, в силу погрешности последней порядка $O(h^5)$. Вышесказанное означает, что на гладких решениях разница в погрешности аппроксимации для дивергентной и недивергентной форм записи исходных уравнений не должна быть существенной.

Однако реализация такого подхода, впервые предложенного в работе [3], требует дополнительной вычислительной работы. В частности, для вычисления промежуточных величин

$$V(\rho_{i+1/2}, \theta), \quad \frac{\partial E}{\partial \rho}(\rho_{i+1/2}, \theta), \quad \frac{\partial E}{\partial \rho}(\rho_i, \theta)$$

необходимы формулы высокого порядка аппроксимации. Приведем их для полноты описания алгоритма:

$$y'(x) = \frac{-y(x+2h) + 8y(x+h) - 8y(x-h) + y(x-2h)}{12h} + O(h^4), \quad (23)$$

$$y'(x+h/2) = \frac{-y(x+2h) + 27y(x+h) - 27y(x) + y(x-h)}{24h} + O(h^4), \quad (24)$$

$$y(x+h/2) = \frac{-y(x+2h) + 9y(x+h) + 9y(x) - y(x-h)}{16} + O(h^4). \quad (25)$$

Конечно, от приближаемых величин здесь требуется повышенная гладкость: формулы (23) и (24) требуют $y(x) \in C^{(5)}$, а формула (25) — $y(x) \in C^{(4)}$. Заметим также, что вычисление электронной плотности по формуле (2) проводится с помощью соотношения (23).

В заключение раздела следует обратить внимание, что предложенная модификация известного метода никак не влияет на его условие устойчивости в классической версии — (20), которое было получено на основе как спектрального признака устойчивости и метода замороженных коэффициентов, так и метода дифференциального приближения (см. [18], [19]). Это связано с тем, что источникные члены в уравнениях плазменных колебаний изначально линейны, а линеаризация и дивергентной формы уравнений, и недивергентной, порождает одинаковые выражения при замораживании переменных коэффициентов.

Отметим, что обсуждаемая модификация схемы Русанова (17)–(19) на основе формулы Симпсона теоретически безупречна для расчета гладких решений, хотя вычислительные затраты при ее реализации представляются чрезмерными. Тем не менее упрощение модификации за счет использования более простой (и менее точной!) формулы прямоугольников, как это сделано в работе [3], не представляется возможным потому, что подобная замена понизит порядок точно-

сти по пространству до второго, точнее – до $O(h^3 / \tau)$, а это сразу приведет к потере смысла формул (17)–(19).

3. ЧИСЛЕННЫЕ ЭКСПЕРИМЕНТЫ

3.1. Тестовые расчеты с гладким решением

Рассмотрим задание решения в окрестности прямой $\rho = 0$ с помощью функции

$$F(x) = \alpha x \exp\left\{-2\frac{x^2}{\rho_*}\right\}, \quad \alpha = \left(\frac{a_*}{\rho_*}\right)^2, \quad (26)$$

где параметры ρ_* и a_* характеризуют масштаб области локализации и максимальную величину $F_{\max} = a_*^2 / (\rho_* 2\sqrt{e}) \approx 0.3a_*^2 / \rho_*$. Здесь $e = 2.71828\dots$ – основание натурального логарифма. Вид функции (26) выбран из соображений, что подобное возмущение электрического поля может возбуждаться в разреженной плазме ($\omega_l \gg \omega_p$) лазерным импульсом с частотой ω_l при его фокусировке в линию (этого можно добиться при использовании цилиндрической линзы, см. детали в [20]).

Построим с помощью (26) в области $[-d, d]$, $d = 5\rho_*$ тестовое гладкое решение системы уравнений

$$\frac{\partial V}{\partial \theta} + V \frac{\partial V}{\partial \rho} + E = G_V, \quad \frac{\partial E}{\partial \theta} + V \frac{\partial E}{\partial \rho} - V = G_E. \quad (27)$$

Правые части уравнений – функции $G_V(\rho, \theta)$ и $G_E(\rho, \theta)$ – определим с помощью точного решения

$$V(\rho, \theta) = F(\rho - \beta\rho_* \sin \theta) \sin \theta, \quad E(\rho, \theta) = F(\rho + \beta\rho_* \sin \theta) \cos \theta, \quad (28)$$

где для определенности зафиксированы параметры:

$$\beta = 0.5, \quad \rho_* = 0.6, \quad \alpha = 0.25.$$

Недостающие начальные и граничные условия берутся из формул (28), а контроль вычислений, кроме искомых функций $V(\rho, \theta)$ и $E(\rho, \theta)$, дополнительно осуществляется за электронной плотностью

$$N(\rho, \theta) = 1 - F_x(\rho + \beta\rho_* \sin \theta) \cos \theta, \quad (29)$$

так как она является наименее гладкой из всех интересующих физических величин.

Применим для численного решения системы (27) модификацию схемы Русанова на основе формулы Симпсона из п. 2.2 с целью экспериментального наблюдения за порядком точности. Результаты расчетов при $\tau = h$ ($|v| \leq 0.1$) и $w_{32} = 2 \times 10^{-3}$ приведены в табл. 1.

Колонка $\text{Err}(V)$ содержит величины погрешностей для скорости V , посчитанные по формуле $\max_{\rho, \theta} |V_i^n - V(\rho_i, \theta^n)|$ для области $\rho \in [-d, d]$, $\theta \in [0, 20.1]$, а колонка $\text{Rel}(V)$ – отношение погрешностей, взятых при $2h$ и h . Аналогичный смысл имеют колонки таблицы, характеризующие погрешности и их отношения для функций электрического поля E и плотности электронов N .

Таблица 1. Погрешности схемы Русанова для решения с $\alpha = 0.25$

$1/h$	$\text{Err}(V)$	$\text{Rel}(V)$	$\text{Err}(E)$	$\text{Rel}(E)$	$\text{Err}(N)$	$\text{Rel}(N)$
200	0.14907×10^{-6}	–	0.13879×10^{-6}	–	0.42601×10^{-5}	–
400	0.17971×10^{-7}	8.30	0.16715×10^{-7}	8.30	0.53217×10^{-6}	8.01
800	0.22035×10^{-8}	8.16	0.20492×10^{-8}	8.16	0.65903×10^{-7}	8.08
1600	0.27267×10^{-9}	8.08	0.25357×10^{-9}	8.08	0.81895×10^{-8}	8.05
3200	0.33927×10^{-10}	8.04	0.31571×10^{-10}	8.03	0.10267×10^{-8}	7.98

Таблица 2. Погрешности схемы Мак-Кормака для решения с $\alpha = 0.25$

$1/h$	$\text{Err}(V)$	$\text{Rel}(V)$	$\text{Err}(E)$	$\text{Rel}(E)$	$\text{Err}(N)$	$\text{Rel}(N)$
200	0.22015×10^{-4}	—	0.21057×10^{-4}	—	0.88364×10^{-3}	—
400	0.55348×10^{-5}	3.98	0.52656×10^{-5}	4.00	0.23465×10^{-3}	3.77
800	0.13860×10^{-5}	3.99	0.13161×10^{-5}	4.00	0.59433×10^{-4}	3.95
1600	0.34669×10^{-6}	4.00	0.32905×10^{-6}	4.00	0.14916×10^{-4}	3.98
3200	0.86687×10^{-7}	4.00	0.82267×10^{-7}	4.00	0.37327×10^{-5}	4.00

Смысл отношений весьма прост: как только погрешность решения начинает характеризоваться первым (главным) членом асимптотики, отношение должно быть похоже на 2^s , где s – порядок точности метода. Более точно, пусть для схемы Русанова, в силу порядка точности $s = 3$, справедливо

$$\text{Err}(V) = c_3 h^3 + c_4 h^4 + \dots,$$

где постоянные c_i , $i = 3, 4, \dots$, не зависят от параметров дискретизации, тогда если $\text{Rel}(V) = 8.16$, то шаг $h = 1/800$ еще слишком велик, чтобы правильно характеризовать главный член погрешности $c_3 h^3$. С другой стороны, шаг $h = 1/3200$ уже достаточно хорош, так как для него $\text{Rel}(V) = 8.04$, т.е. влияние младших членов погрешности на асимптотику не превышает 0.5%.

Главный вывод, который можно сделать на основе табл. 1: численные эксперименты наглядно демонстрируют теоретический (третий) порядок точности схемы Русанова, причем ее модификация на случай недивергентной формы уравнений этот порядок не ухудшает. Однако для преобладания в погрешности главного члена асимптотики требуется использование достаточно малых параметров дискретизации.

С целью сравнения со схемой Русанова приведем в табл. 2 результаты аналогичных расчетов для схемы Мак-Кормака, модификация которой для уравнений недивергентной формы опубликована в [3]. Все параметры задачи и смысл значений колонок полностью совпадают с аналогичными для схемы Русанова.

Схема Мак-Кормака имеет второй порядок точности ($s = 2$), т.е. погрешность для конкретной функции можно охарактеризовать формулой

$$\text{Err}(V) = \hat{c}_2 h^2 + \hat{c}_3 h^3 + \dots,$$

где постоянные \hat{c}_i , $i = 2, 3, \dots$, не зависят от параметров дискретизации.

Из данных табл. 2 следует, что влияние на главные члены погрешностей искомых функций $V(\rho, \theta)$ и $E(\rho, \theta)$, не превосходящее 0.5%, достигается уже при $h = 1/800$, т.е. эффект преобладания достигается существенно быстрее, чем для схемы Русанова. Это легко объяснить ростом производных решения, т.е. доминированием величин c_3 над \hat{c}_2 , c_4 над \hat{c}_3 и т.д.

В целом для задачи с гладким решением $V(\rho, \theta)$ и $E(\rho, \theta)$ поведение погрешностей для схем различных порядков точности в табл. 1 и 2 вполне естественно. Их абсолютные значения для более точной схемы на несколько порядков меньше аналогов для менее точной схемы, что напрямую связано с различными объемами вычислительной работы при реализации этих схем. Экспериментальные данные свидетельствуют, что при одинаковых сеточных параметрах расчет по схеме Русанова длится примерно в 3.95 дольше, чем по схеме Мак-Кормака. Правда, оптимизация именно на предмет сокращения объема вычислительной работы для каждой из схем целенаправленно не производилась, но трехкратное увеличение вычислений в схеме Русанова по сравнению со схемой Мак-Кормака можно считать правдоподобным.

Менее очевидной является динамика погрешности для электронной плотности N , так как она полностью характеризуется порядком точности схемы. Плотность представляет собой констатив за вычетом градиента электрического поля E , поэтому формальный порядок точности градиента функции должен быть на единицу меньше порядка точности самой функции, однако расчеты свидетельствуют об одинаковых порядках. Это можно объяснить в рамках гипотезы о главном члене погрешности, т.е. при приближенном вычислении гладкой функции главный член

погрешности также является гладкой функцией. Иными словами, получаемое сеточное решение E_i^n является следом на сетке (проекцией) суммы достаточно гладких функций

$$E_i^n = E(\rho_i, \theta^n) + h^3 P(\rho_i, \theta^n) + o(h^3),$$

где $E(\rho, \theta)$ — гладкое решение дифференциального уравнения, а $P(\rho, \theta)$ — гладкий множитель, характеризующий главный член погрешности. В этом случае использование формулы (23) с погрешностью $O(h^4)$ для приближенного вычисления электронной плотности N приведет к сохранению порядка точности схемы Русанова. Аналогичные рассуждения справедливы и для величин меньшего порядка, порождаемых схемой Мак-Кормака и приведенных в табл. 2. Краткий вывод, связанный с поведением погрешности электронной плотности N , таков: порядок точности в обеих схемах такой же, как для $V(\rho, \theta)$ и $E(\rho, \theta)$, хотя сами абсолютные значения погрешностей примерно на порядок больше.

3.2. Расчеты с менее гладким решением

Рассмотрим традиционную постановку для моделирования нерелятивистских колебаний в холодной плазме [2]. Положим, что скорость электронов в начальный момент времени ($\theta = 0$) равна нулю

$$V_0(\rho) \equiv V(\rho, \theta = 0) = 0, \quad (30)$$

а колебания возбуждаются в начальный момент времени электрическим полем следующего вида:

$$E_0(\rho) \equiv E(\rho, \theta = 0) = F(\rho). \quad (31)$$

Для однозначного задания начальных условий зафиксируем параметры: $\rho_* = 0.6$, $a_* = 0.414$. В результате получим $\alpha = 0.4761$. Обратим внимание, что условия (30), (31) удовлетворяют критерию (3), т.е. единственное 2π -периодическое по времени решение уравнений (1) глобально существует. Однако оно является менее гладким, чем рассмотренное в предыдущем разделе. Например, электронная плотность при $\alpha = 0.4761$ колеблется в диапазоне $[0.52, 10.96]$ (границы указаны с двумя верными знаками после десятичной точки), т.е. изменяется на периоде примерно в 20 раз. Тогда как при $\alpha = 0.25$ она изменяется за период всего в два раза.

Заметим, что на больших расстояниях от прямой $\rho = 0$, в силу начального условия (31), плазменные колебания не возбуждаются. Поэтому будем считать, что выполнены следующие условия:

$$V(\rho \rightarrow \pm\infty, \theta) = 0, \quad E(\rho \rightarrow \pm\infty, \theta) = 0. \quad (32)$$

Однако в целях численного моделирования плазменных колебаний расчетную область необходимо ограничить; определим ее по переменной ρ как отрезок $[-d, d]$, на концах которого следует задать искусственные граничные условия. Обсуждению их построения посвящен разд. 3.6 в [2], здесь же мы ограничимся “обрезанием” бесконечной области с помощью однородных граничных условий первого рода:

$$V(\pm d, \theta) = E(\pm d, \theta) = 0.$$

Конечно, параметр d следует выбирать достаточно большим. В силу экспоненциального затухания функции $E_0(\rho)$, достаточно положить $d = 4.5\rho_*$. В этом случае имеем $\exp\{-d^2/\rho_*^2\} \approx 2.5768 \times 10^{-18}$. Это означает, что при вычислениях с двойной точностью величина скачка начальной функции E_0 в точках $\rho = \pm d$ соизмерима с машинной точностью, т.е. с обычной погрешностью округления данных. Другими словами, при численном моделировании колебаний эффект обрезания начальных условий замечен совершенно не будет, что полностью соответствует понятию “искусственной границы”. Учитывая сказанное, параметр d , характеризующий искусственную границу, положим $d = 4.5\rho_*$. Ограничение расчетной области по переменной θ примем как отрезок $[0, 40\pi]$, считая, что 20 периодов достаточно для наблюдения за решением.

Вычисления проведем на последовательности вложенных сеток: $\tau = h \in \{1/1600, 1/3200, 1/6400, 1/12800\}$, удовлетворяющих условию устойчивости используемых численных алгоритмов. Основным интерес будут представлять эксперименты с расчетами по схеме Русанова третье-

Таблица 3. Погрешности схемы Русанова для решения с $\alpha = 0.4761$

$1/h$	$\text{Err}(Ax)$	$\text{Rel}(Ax)$	$\text{Err}(En)$	$\text{Rel}(En)$	$\text{Err}(Af)$	$\text{Rel}(Af)$
1600	0.17070×10^0	—	0.23101×10^{-4}	—	0.23101×10^{-4}	—
3200	0.74069×10^{-1}	2.30	0.37742×10^{-5}	6.12	0.37742×10^{-5}	6.12
6400	0.22495×10^{-1}	3.29	0.52576×10^{-6}	7.18	0.52576×10^{-6}	7.18
12800	0.45987×10^{-2}	4.89	0.67610×10^{-7}	7.78	0.67609×10^{-7}	7.78

го порядка точности, для сравнения будут использоваться вычисления схеме Мак-Кормака, имеющей второй порядок точности [3].

В данном случае аналитическое решение уравнений (1) в рассматриваемой области отсутствует, поэтому будем наблюдать за сохранением величин $En(\theta)$ (7) и $Af(\theta)$ (9), а также — за погрешностью формулы (14), характеризующей динамику электронной плотности на оси симметрии области.

Сделаем замечание о влиянии ограничения области по переменной ρ на сохранение величин $En(\theta)$ и $Af(\theta)$. В начальный момент времени $\theta = 0$ для начальных условий (30), (31) несложно вычислить точно

$$En(0) = Af(0) = \frac{a_*^4 \sqrt{\pi}}{\rho_* 16}.$$

Затем, используя составную формулу трапеций [7], вычислим приближенное значение тех же величин, но по ограниченной области $[-d, d]$:

$$En^d(0) = \int_{-d}^d E_0^2(\rho) d\rho, \quad Af^d(0) = \int_{-d}^d N_0(\rho) E_0^2(\rho) d\rho.$$

Относительная ошибка, возникающая вследствие замены ∞ на d с соответствующими знаками, для различных пространственных шагов h не превышала по порядку величины 10^{-14} , что дает основание пренебречь эффектом обрезания области.

Применим для численного решения системы (1) с начальными условиями (30), (31) модификацию схемы Русанова с целью экспериментального наблюдения за точностью решения. Результаты расчетов при $\tau = h$ ($|v| \leq 0.1$) и $w_{32} = 2 \times 10^{-3}$ приведены в табл. 3.

Колонка $\text{Err}(Ax)$ содержит величины относительных погрешностей для значений электронной плотности N на оси симметрии ($\rho = 0$), а колонка $\text{Rel}(Ax)$ — отношение этих погрешностей, взятых при $2h$ и h . В свою очередь, колонка $\text{Err}(En)$ содержит величины максимумов по времени для относительных погрешностей значений энергии En , т.е. результатов расчетов по формуле $\max_{\theta^n} |En(\theta^n) - En(0)| / En(0)$ для области $\rho \in [-d, d]$, $\theta \in [0, 40\pi]$. Для вычисления приближенных значений $En(\theta^n)$ использовалась составная формула трапеций [7]. Колонка $\text{Rel}(En)$ содержит отношения погрешностей, взятых при $2h$ и h . Аналогичный смысл имеют колонки таблицы, характеризующие относительные погрешности и их отношения для функции Af .

Главной особенностью результатов из табл. 3 является иллюстрация чрезвычайно медленного доминирования главного члена асимптотики для погрешностей рассматриваемых функций. Причиной этого может быть в разложении локальной погрешности

$$\text{Err}(Ax) = c_3 h^3 + c_4 h^4 + c_5 h^5 + \dots$$

сильный рост постоянных, т.е. $c_{k+1} \gg c_k$, $k \geq 3$, связанных с недостаточной гладкостью функции электронной плотности N в окрестности прямой $\rho = 0$. Ситуация с интегральными характеристиками En и Af заметно лучше, но отношения погрешностей, характеризующих порядок точности схемы Русанова, весьма далеки от их аналогов для гладкого решения (см. табл. 1).

С целью сравнения со схемой Русанова приведем в табл. 4 результаты аналогичных расчетов для схемы Мак-Кормака, где сохранены все параметры задачи и смысл значений колонок.

Таблица 4. Погрешности схемы Мак-Кормака для решения с $\alpha = 0.4761$

$1/h$	$\text{Err}(Ax)$	$\text{Rel}(Ax)$	$\text{Err}(En)$	$\text{Rel}(En)$	$\text{Err}(Af)$	$\text{Rel}(Af)$
1600	0.41861×10^{-1}	—	0.62113×10^{-5}	—	0.37771×10^{-4}	—
3200	0.12439×10^{-1}	3.36	0.15588×10^{-5}	3.98	0.96216×10^{-5}	3.93
6400	0.32984×10^{-2}	3.77	0.39010×10^{-6}	4.00	0.24177×10^{-5}	3.98
12800	0.83870×10^{-3}	3.93	0.97564×10^{-7}	4.00	0.60520×10^{-6}	3.99

Легко заметить, что асимптотический показатель скорости сходимости для схемы Мак-Кормака, равный двум, при уменьшении параметров дискретизации наблюдается гораздо быстрее, т.е. эффект насыщения погрешности достигается при меньших h . Соответственно, для разложения погрешности

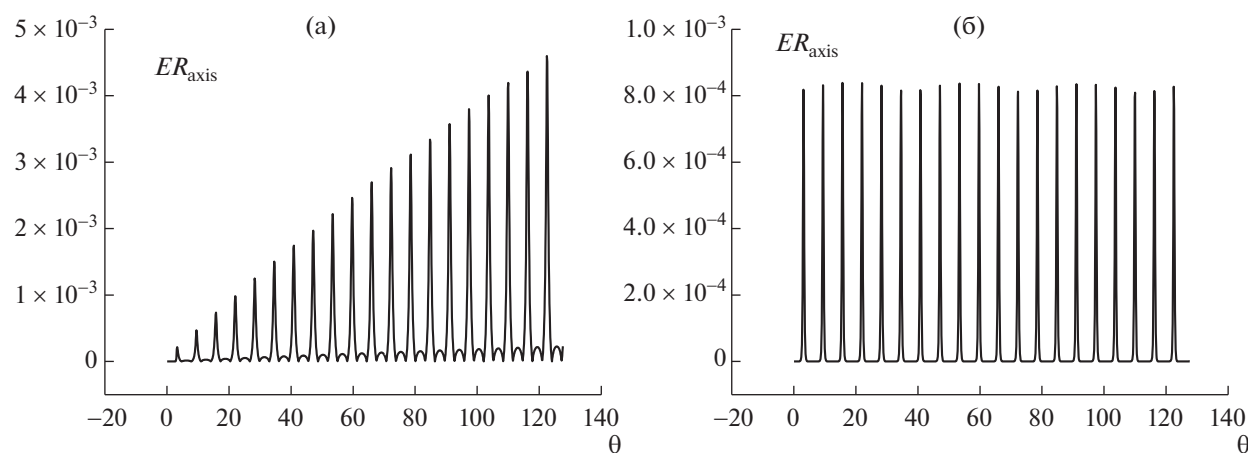
$$\text{Err}(Ax) = \hat{c}_2 h^2 + \hat{c}_3 h^3 + \hat{c}_4 h^4 + \dots;$$

здесь можно предположить, что рост постоянных \hat{c}_k также имеет место, но он не столь значителен, как для схемы Русанова. Весьма любопытным является то, что сами погрешности $\text{Err}(Ax)$ при всех использованных параметрах τ и h для схемы Мак-Кормака меньше, чем для схемы Русанова. Видимо, это объясняется большой разницей постоянных при главных членах погрешностей в схемах, т.е. $\hat{c}_2 \ll c_3$.

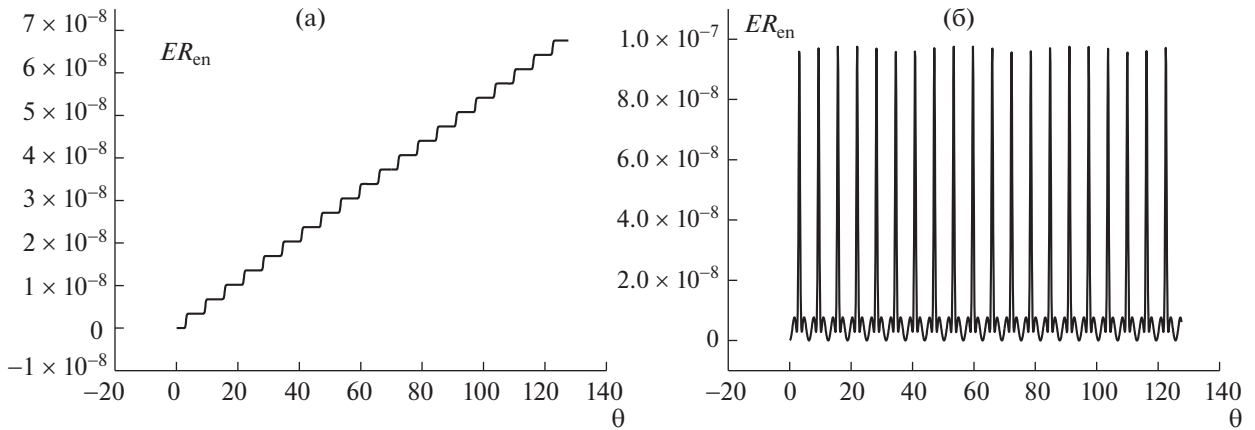
Таким образом, сравнение схем Мак-Кормака и Русанова на менее гладком решении (с $\alpha = 0.4761$) указывает, что схема Русанова точнее в теоретическом плане, а для практики предпочтительнее схема Мак-Кормака. Иными словами, преимущество схемы Русанова станет заметно только при достаточно малых параметрах дискретизации, возможно недостижимых для реальных расчетов колебаний холодной плазмы. Заметим, что параметры задачи ρ_* и a_* в обсуждаемых численных экспериментах не являются критическими, т.е. градиенты решения еще весьма далеки от сингулярных. Формально дифференциальная задача еще хороша, а посчитать ее по схеме Русанова уже трудно.

В завершение процедуры сравнения двух рассматриваемых схем рассмотрим качественную динамику их погрешностей на больших временах (порядка 20 периодов); все иллюстрации сделаны на основании расчетов при $\tau = h = 1/12800$.

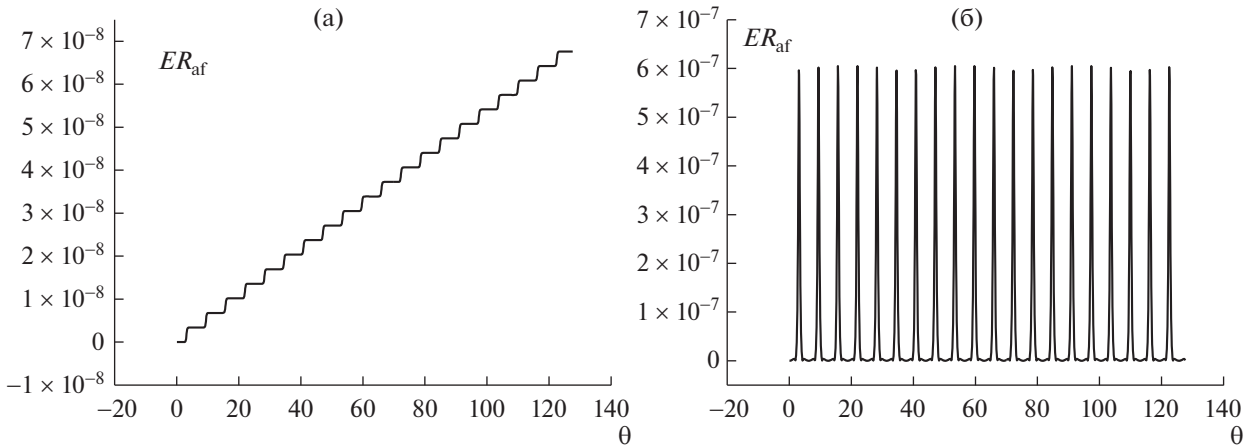
На фиг. 1 представлена посчитанная по обеим схемам динамика относительной погрешности плотности электронов в центре области. Легко заметить качественное отличие в поведении погрешности для представленных схем. Для схемы Мак-Кормака погрешность является равномерно ограниченной величиной, изменяющейся в течение каждого периода пропорционально функции электронной плотности. При этом само ограничивающее значение убывает строго в



Фиг. 1. Динамика относительной погрешности плотности электронов в центре области $N(\rho = 0, \theta)$; а – для схемы Русанова, б – для схемы Мак-Кормака.



Фиг. 2. Динамика относительной погрешности энергии $En(\theta)$; а – для схемы Русанова, б – для схемы Мак-Кормака.



Фиг. 3. Динамика относительной погрешности функции $Af(\theta)$; а – для схемы Русанова, б – для схемы Мак-Кормака.

соответствии с теорией как величина второго порядка малости, т.е. при уменьшении параметров τ и h в два раза оно убывает в четыре раза. В свою очередь, для схемы Русанова из графика наблюдаем, что верхняя огибающая колеблющейся погрешности представляет собой линейную по времени функцию. Учитывая, что свойства аппроксимации и устойчивости имеют асимптотический характер, т.е. справедливы при достаточно малых параметрах дискретизации, можно предположить, что линейный рост погрешности при больших значениях времени наблюдения за решением приведет к отсутствию близости точного и приближенного решений задачи. Напомним также, что сами ограничивающие погрешности величины различаются примерно в 5 раз в пользу схемы Мак-Кормака.

На фиг. 2 представлена посчитанная по обеим схемам динамика относительной погрешности энергии. Как и при анализе погрешности аксиального решения для плотности электронов, погрешность энергии, посчитанная по схеме Мак-Кормака, ведет себя в полном соответствии с теорией: при уменьшении параметров дискретизации она квадратично убывает и, кроме того, – равномерно ограничена по времени. Расчеты, проведенные по схеме Русанова, также близки к теоретическим, а сами значения погрешности меньше, чем для схемы Мак-Кормака. Однако здесь для схемы Русанова также наблюдается линейный рост погрешности в зависимости от длины интервала времени, что нельзя никак отнести к достоинству схемы.

На фиг. 3 представлена посчитанная по обеим схемам динамика относительной погрешности дополнительной функции $Af(\theta)$ (9). Как и при анализе погрешностей плотности и энергии, погрешность дополнительной функции, посчитанная по схеме МакКормака, ведет себя практиче-

ски идеально: при уменьшении параметров дискретизации она квадратично убывает и, кроме того, — равномерно ограничена по времени. Расчеты, проведенные по схеме Русанова, также согласуются с теорией, однако относительная погрешность, как и в предыдущих расчетах, линейно растет во времени. Можно также отметить, что погрешность дополнительной функции $Af(\theta)$ для схемы Русанова асимптотически стремится к погрешности энергии $En(\theta)$: на малых временах они немного отличаются, зато на больших практически совпадают.

На основании проведенных расчетов можно сделать следующие выводы:

1) при проведении многопериодного моделирования традиционных (не очень гладких!) плазменных колебаний схема Мак-Кормака обладает несомненными преимуществами по сравнению со схемой Русанова как качественного (равномерное ограничение всех видов погрешностей), так и количественного характера (при доступных параметрах дискретизации погрешность важнейшей физической характеристики — электронной плотности — меньше);

2) схемы более высокого порядка точности имеют преимущества на более гладких решениях;

3) для повышения достоверности вычислений, кроме стандартных теоретических исследований аппроксимации и устойчивости, необходимо использовать контроль как за частными аналитическими решениями, так и за сохранением энергии и других дополнительных функций.

ЗАКЛЮЧЕНИЕ

В настоящей работе для моделирования нерелятивистских колебаний холодной плазмы предложена модификация классической схемы Русанова, имеющей третий порядок точности. Отметим, что ранее для расчетов в эйлеровых переменных были известны только схемы первого [15] и второго [3] порядков точности. Для тестовой задачи с гладким решением проведено исследование погрешностей построенной схемы, а также — сравнение с погрешностями схемы Мак-Кормака [3]. Для задачи о свободных плазменных колебаниях, инициированных коротким мощным лазерным импульсом, приведены результаты численных экспериментов по сохранению энергии и дополнительной функции для обеих схем, а также по точности электронной плотности в центре области (при $\rho = 0$). На основании проведенных расчетов можно сделать вывод о теоретическом превосходстве схемы Русанова, хотя для практических вычислений следует рекомендовать схему Мак-Кормака. В первую очередь это касается расчетов “долгоживущих” процессов и колебаний холодной плазмы, близких к реальным. Кроме этого, подтвердим точку зрения, что теоретическое исследование аппроксимации и устойчивости вместе с экспериментальным наблюдением за количественными характеристиками погрешности для наиболее чувствительных величин существенно повышает достоверность вычислений.

СПИСОК ЛИТЕРАТУРЫ

1. Куликовский А.Г., Погорелов Н.В., Семенов А.Ю. Математические вопросы численного решения гиперболических систем уравнений. 2-е изд., испр. и доп. М.: Физматлит, 2012. С. 46–100.
2. Чижонков Е.В. Математические аспекты моделирования колебаний и кильватерных волн в плазме. М.: Физматлит, 2018. С. 12–240.
3. Чижонков Е.В. О схемах второго порядка точности для моделирования плазменных колебаний // Вычисл. методы и программирование. 2020. Т. 21. С. 115.
4. Русанов В.В. Разностные схемы третьего порядка точности для сквозного счета разрывных решений // Докл. АН СССР. 1968. Т. 180. № 6. С. 1303.
5. Rusanov V.V. On difference schemes of third order of accuracy for non-linear hyperbolic systems // J. Comput. Phys. 1970. V. 5. № 3. P. 507.
6. Burstein S.Z., Mirin A.A. Third order difference methods for hyperbolic equations // J. Comput. Phys. 1970. V. 5. № 3. P. 547.
7. Бахвалов Н.С., Корнев А.А., Чижонков Е.В. Численные методы. Решения задач и упражнения. Серия Классический университетский учебник. Учеб. пособие для вузов. 2-е изд., испр. и дополн. М.: Лаборатория знаний, 2016. С. 95–96.
8. Александров А.Ф., Богданкевич Л.С., Рухадзе А.А. Основы электродинамики плазмы. М.: Высш. школа, 1988. С. 102–113.
9. Гинзбург В.Л., Рухадзе А.А. Волны в магнитоактивной плазме. М.: Наука, 1975. С. 112–124.
10. Davidson R.C. Methods in nonlinear plasma theory. New York: Academic Press, 1972. P. 33–53.
11. Dafermos C.M. Hyperbolic Conservation Laws in Continuum Physics. The 4th Edition, Berlin-Heidelberg: Springer, 2016. P. 221–225.

12. *Engelberg S., Liu H., Tadmor E.* Critical Thresholds in Euler – Poisson Equations // *Indiana University Math. J.* 2001. V. 50. P. 109.
13. *Розанова О.С., Чижонков Е.В.* О существовании глобального решения одной гиперболической задачи // Докл. АН. Математика, информатика, процессы управления. 2020. Т. 492. № 1. С. 97.
14. *Rozanova O.S., Chizhonkov E.V.* On the conditions for the breaking of oscillations in a cold plasma // *Z. Angew. Math. Phys.* 2021. V. 72. № 13. P. 1.
15. *Фролов А.А., Чижонков Е.В.* О применении закона сохранения энергии в модели холодной плазмы // Ж. вычисл. матем. и матем. физ. 2020. Т. 60. № 3. С. 503.
16. *Розанова О.С., Чижонков Е.В.* Об аналитическом и численном решении одномерных уравнений холодной плазмы // Ж. вычисл. матем. и матем. физ. 2021. Т. 61. № 9. С. 1508.
17. *MacCormack R.W.* The effect of viscosity in hypervelocity impact cratering // *J. Spacecr. Rockets.* 2003. V. 40. № 5. P. 757.
18. *Шокин Ю.И., Яненко Н.Н.* Метод дифференциального приближения. Применение к газовой динамике. Новосибирск: Наука, 1985. С. 251–252.
19. *Андерсон Д., Таннехилл Дж., Плетчер Р.* Вычислительная гидромеханика и теплообмен. Т. 1. М.: Мир, 1990. С. 179–180.
20. *Sheppard C.J.R.* Cylindrical lenses – focusing and imaging: a review [Invited] // *Applied Optics.* 2013. V. 52. № 4. P. 538.