

# СОДЕРЖАНИЕ

---

---

Том 58, номер 7, 2022

---

---

## ОБЫКНОВЕННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

- Об одном способе математического моделирования электрических синапсов  
*С. Д. Глызин, А. Ю. Колесов* 867
- О кратности собственных значений дифференциального оператора четвёртого порядка на графе  
*Р. Ч. Кулаев, А. А. Уртаева* 882
- 

## УРАВНЕНИЯ С ЧАСТНЫМИ ПРОИЗВОДНЫМИ

- Обратная задача с переопределением третьего рода для абстрактного дифференциального уравнения второго порядка  
*И. В. Тихонов, М. Алмохамед* 890
- 

## ЧИСЛЕННЫЕ МЕТОДЫ

- Численное решение задачи Коши для интегро-дифференциального уравнения второго порядка  
*П. Н. Вабищевич* 912
- О решении уравнения смешанного типа в неограниченной области  
*М. П. Галанин, Д. Л. Сорокин, А. Р. Ухова* 921
- О свойствах одного разностного метода решения двухфазной задачи Стефана  
*А. О. Гусев, О. В. Щерица, О. С. Мажорова* 930
- Компактные и монотонные разностные схемы для обобщённого уравнения Фишера  
*П. П. Матус, Б. Д. Утебаев* 947
- Интервальный подход к решению задач параметрической идентификации динамических систем  
*А. Ю. Морозов, Д. Л. Ревизников* 962
- Валидация вычислительного алгоритма на основе разрывного метода Галёркина для релаксационной модели Баера–Нуницато  
*Р. Р. Полежаина, М. В. Алексеев, Е. Б. Савенков* 977
- Оценки точности проекционного метода со стабилизатором дробной гладкости в задаче восстановления волнового фронта по его наклонам  
*А. В. Разгулин, Н. Г. Ирошников, А. В. Ларичев, С. А. Турганбаев, Т. Е. Романенко* 995
- 
-



===== ОБЫКНОВЕННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ =====

УДК 517.929

## ОБ ОДНОМ СПОСОБЕ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ЭЛЕКТРИЧЕСКИХ СИНАПСОВ

© 2022 г. С. Д. Глызин, А. Ю. Колесов

Вводится в рассмотрение новая математическая модель нейронной сети с электрическими связями, представляющая собой сингулярно возмущённую систему дифференциально-разностных уравнений с запаздыванием. Излагаются некоторые методы исследования вопросов о существовании и устойчивости в этой системе релаксационных периодических движений. В случае диффузионных и симметричных полносвязных нейронных сетей устанавливается реализуемость в соответствующих математических моделях известного феномена буферности.

DOI: 10.31857/S0374064122070019, EDN: CDXLYG

**1. Постановка задачи.** Остановимся сначала на выборе базовой математической модели функционирования отдельного нейрона. Следуя работам [1, 2], в качестве такой возьмём скалярное нелинейное дифференциально-разностное уравнение вида

$$\dot{u} = \lambda f(u(t-1))u \quad (1.1)$$

для мембранного потенциала  $u = u(t) > 0$ . Здесь параметр  $\lambda > 0$ , характеризующий скорость протекания электрических процессов в системе, предполагается большим, точка означает дифференцирование по  $t$ , а функция  $f(u) \in C^2(\mathbb{R}_+)$ ,  $\mathbb{R}_+ = \{u \in \mathbb{R} : u \geq 0\}$ , обладает свойствами

$$f(0) = 1; \quad f(u) + a, \quad uf'(u), u^2 f''(u) = O(1/u) \quad \text{при } u \rightarrow +\infty, \quad (1.2)$$

где  $a = \text{const} > 0$ . Примером такой функции является

$$f(u) = (1-u)/(1+u/a). \quad (1.3)$$

Уравнение (1.1), представляющее собой некоторую модификацию известного уравнения Хатчинсона [3], было предложено и исследовано в статье [4]. В ней показано, что при всех  $\lambda \gg 1$  оно допускает экспоненциально орбитально устойчивый релаксационный цикл  $u_*(t, \lambda) > 0$ ,  $u_*(0, \lambda) \equiv 1$  периода  $T_*(\lambda)$ , удовлетворяющий предельным соотношениям

$$\lim_{\lambda \rightarrow \infty} T_*(\lambda) = T_0, \quad \max_{0 \leq t \leq T_*(\lambda)} |x_*(t, \lambda) - x_0(t)| = O(1/\lambda), \quad \lambda \rightarrow \infty, \quad (1.4)$$

где  $T_0 = (1+a)t_0$ ,  $t_0 = 1+1/a$ ,  $x_*(t, \lambda) = (1/\lambda) \ln(u_*(t, \lambda))$ , а  $T_0$ -периодическая функция  $x_0(t)$  задаётся равенствами

$$x_0(t) = \begin{cases} t & \text{при } 0 \leq t \leq 1, \\ 1 - a(t-1) & \text{при } 1 \leq t \leq t_0 + 1, \\ t - T_0 & \text{при } t_0 + 1 \leq t \leq T_0, \end{cases} \quad x_0(t + T_0) \equiv x_0(t). \quad (1.5)$$

Наглядное представление о релаксационных свойствах этого цикла даёт его график на плоскости  $(t, u)$ , построенный численно для случая (1.1), (1.3) при  $\lambda = 5$ ,  $a = 2$  (рис. 1). Отметим, что в дальнейшем под релаксационными колебаниями вслед за работами [2, 5] будем понимать колебательные режимы исследуемых моделей, которые сочетают быстрые и медленные движения.

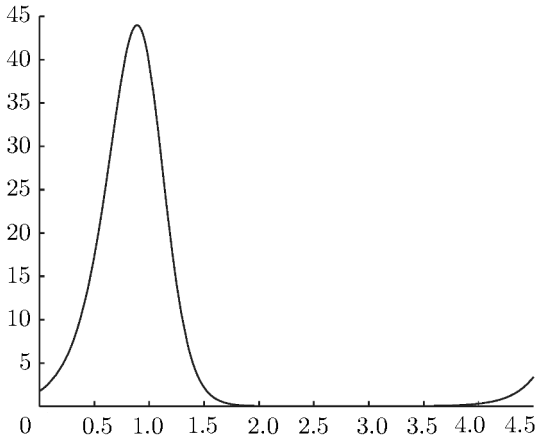


Рис. 1. График  $u_*(t, \lambda)$  при  $a = 2, \lambda = 5$ .

Предположим теперь, что имеется сеть из  $m \geq 2$  нейронов с потенциалами  $u_j = u_j(t), j = \overline{1, m}$ . Будем считать, что они электрически взаимодействуют между собой по принципу “каждый со всеми”. Тогда согласно общепринятым представлениям о природе электрических синапсов (см., например, монографии [6, с. 39–40; 7, с. 60–65] и статьи [8–10]) связь между нейронами с номерами  $j$  и  $s, j \neq s$ , осуществляется посредством соответствующего тока проводимости  $I_{j,s}^{\text{syn}}$ . Последний, в свою очередь, в силу закона Ома задаётся равенством

$$I_{j,s}^{\text{syn}} = d_{j,s}(u_s - u_j), \quad d_{j,s} = \text{const} \in \mathbb{R}. \quad (1.6)$$

Тем самым, приходим к системе вида

$$\dot{u}_j = \lambda f(u_j(t-1))u_j + \sum_{\substack{s=1 \\ s \neq j}}^m d_{j,s}(u_s - u_j), \quad j = \overline{1, m}, \quad (1.7)$$

которая, собственно, и является одной из возможных математических моделей электрических синапсов. Однако, на наш взгляд, модель (1.7) нуждается в некоторой модификации.

Суть упомянутой модификации состоит в том, чтобы заменить равенство (1.6) соотношением

$$I_{j,s}^{\text{syn}} = d_{j,s} g\left(\frac{u_s}{u_j}\right)u_j, \quad d_{j,s} = \text{const} \in \mathbb{R}, \quad (1.8)$$

где функция  $g(u) \in C^3(\mathbb{R}_+)$  удовлетворяет требованиям

$$g(u) < 0 \text{ при } u \in [0, 1), \quad g(0) = -1, \quad g(u) > 0 \text{ при } u > 1, \quad g(1) = 0, \quad g'(1) > 0, \quad (1.9)$$

$$g(u) - b, ug'(u), u^2g''(u), u^3g'''(u) = O(1/u) \text{ при } u \rightarrow +\infty, \quad b = \text{const} > 0. \quad (1.10)$$

Что же касается соответствующей модели электрических синапсов, то в данном случае она приобретает вид

$$\dot{u}_j = \left[ \lambda f(u_j(t-1)) + \sum_{\substack{s=1 \\ s \neq j}}^m d_{j,s} g\left(\frac{u_s}{u_j}\right) \right] u_j, \quad j = \overline{1, m}. \quad (1.11)$$

Следует отметить, что между прежним законом (1.6) и модифицированной омической гипотезой (1.8) имеется определённая согласованность. Действительно, в силу условий (1.9) свойства функции  $g(u)$  из (1.8) максимально близки к свойствам функции  $g(u) = u - 1$  в случае (1.6). В частности, как и система (1.7) новая модель (1.11) в силу свойства  $g(1) = 0$  допускает так называемый однородный цикл

$$u_1 = u_2 = \dots = u_m = u_*(t, \lambda), \quad (1.12)$$

где  $u_*(t, \lambda)$  – периодическое решение уравнения (1.1), о котором говорилось выше (см. (1.4), (1.5)). Что же касается отличий нового подхода к моделированию электрических синапсов от общепринятого, то их два. Во-первых, при переходе от отдельного нейрона к соответствующей сети сохраняется вольтерровская структура уравнений; во-вторых, соблюден так называемый закон насыщающих проводимостей, суть которого – выполнение условий вида (1.2), (1.10) для всех входящих в систему нелинейностей. Тем самым наши модели (как отдельного нейрона (1.1), так и нейронной сети (1.11)) базируются на одних и тех же общих принципах,

сформулированных в работе [2]. Под вольтерровской структурой уравнений (1.11) подразумевается возможность записи их правых частей в виде произведения соответствующей зависимой переменной и некоторой неособой в нуле функции. Это свойство гарантирует положительность решений с положительными начальными условиями.

Вопрос о возможных периодических режимах системы (1.11) достаточно сложен. Поэтому, не претендуя на полноту анализа, ниже ограничимся рассмотрением лишь тех из них, которые располагаются в некоторой окрестности её однородного цикла (1.12). Как оказывается, при нахождении таких режимов можно применить специальные асимптотические методы, разработанные ранее в статье [11] для аналогичных (1.7) диффузионных систем

$$\dot{u}_j = \lambda f(u_j(t-1))u_j + d(u_{j+1} - 2u_j + u_{j-1}), \quad j = \overline{1, m}, \tag{1.13}$$

где  $u_0 = u_1$ ,  $u_{m+1} = u_m$ ,  $d = \text{const} > 0$ . Описание упомянутых методов приводится ниже.

**2. Базовые теоремы.** Прежде всего перейдём в системе (1.11) к новым переменным  $x$ ,  $y_1, \dots, y_{m-1}$  по правилу

$$u_1 = \exp(x/\varepsilon), \quad u_j = \exp\left(\frac{x}{\varepsilon} + \sum_{k=1}^{j-1} y_k\right), \quad j = \overline{2, m}, \quad \varepsilon = 1/\lambda.$$

В результате получим релаксационную систему вида

$$\dot{x} = F(x(t-1), \varepsilon) + \varepsilon \sum_{s=2}^m d_{1,s} g\left(\exp\left(\sum_{r=1}^{s-1} y_r\right)\right), \tag{2.1}$$

$$\begin{aligned} \dot{y}_j = & \sum_{s=1}^j d_{j+1,s} g\left(\exp\left(-\sum_{r=s}^j y_r\right)\right) + \sum_{s=j+2}^m d_{j+1,s} g\left(\exp\left(\sum_{r=j+1}^{s-1} y_r\right)\right) - \sum_{s=1}^{j-1} d_{j,s} g\left(\exp\left(-\sum_{r=s}^{j-1} y_r\right)\right) - \\ & - \sum_{s=j+1}^m d_{j,s} g\left(\exp\left(\sum_{r=j}^{s-1} y_r\right)\right) + G_j(x(t-1), y_1(t-1), \dots, y_j(t-1), \varepsilon), \quad j = \overline{1, m-1}, \end{aligned} \tag{2.2}$$

где нелинейности  $F(x, \varepsilon)$ ,  $G_j(x, y_1, \dots, y_j, \varepsilon)$  задаются равенствами

$$F(x, \varepsilon) = f(\exp(x/\varepsilon)),$$

$$G_j(x, y_1, \dots, y_j, \varepsilon) = \frac{1}{\varepsilon} \left\{ f\left(\exp\left(x/\varepsilon + \sum_{k=1}^j y_k\right)\right) - f\left(\exp\left(x/\varepsilon + \sum_{k=1}^{j-1} y_k\right)\right) \right\}, \quad j = \overline{1, m-1}.$$

Далее, фиксируем постоянную  $\sigma_0 : 0 < \sigma_0 < \min(1, a)$  и на отрезке  $-\sigma_0 \leq t \leq T_0 + \sigma_0$  обозначим через

$$y_1^0(t, z), \dots, y_{m-1}^0(t, z), \quad z = (z_1, \dots, z_{m-1}) \in \mathbb{R}^{m-1}$$

компоненты решения импульсной задачи Коши

$$\begin{aligned} \dot{y}_j = & \sum_{s=1}^j d_{j+1,s} g\left(\exp\left(-\sum_{r=s}^j y_r\right)\right) + \sum_{s=j+2}^m d_{j+1,s} g\left(\exp\left(\sum_{r=j+1}^{s-1} y_r\right)\right) - \\ & - \sum_{s=1}^{j-1} d_{j,s} g\left(\exp\left(-\sum_{r=s}^{j-1} y_r\right)\right) - \sum_{s=j+1}^m d_{j,s} g\left(\exp\left(\sum_{r=j}^{s-1} y_r\right)\right), \quad j = \overline{1, m-1}, \end{aligned} \tag{2.3}$$

$$y_j(1+0) = y_j(1-0) - (1+a)y_j(0),$$

$$y_j(t_0+1+0) = y_j(t_0+1-0) - (1+1/a)y_j(t_0), \quad j = \overline{1, m-1}, \tag{2.4}$$

$$(y_1, \dots, y_{m-1})|_{t=0} = z, \tag{2.5}$$

где  $T_0, t_0$  – величины из (1.5). Рассмотрим затем отображение

$$z \mapsto \Phi(z) \stackrel{\text{def}}{=} (y_1^0(t, z), \dots, y_{m-1}^0(t, z))|_{t=T_0}, \tag{2.6}$$

действующее из  $\mathbb{R}^{m-1}$  в  $\mathbb{R}^{m-1}$ . Справедливо следующее утверждение.

**Теорема 1.** *Любой гиперболической неподвижной точке  $z = z_*$  отображения (2.6) в системе (2.1), (2.2) при всех достаточно малых  $\varepsilon > 0$  соответствует релаксационный цикл*

$$(x(t, \varepsilon), y_1(t, \varepsilon), \dots, y_{m-1}(t, \varepsilon)), \quad x(0, \varepsilon) \equiv 0 \tag{2.7}$$

периода  $T(\varepsilon)$ , устойчивое и неустойчивое многообразия которого определяются соответствующими многообразиями этой точки. Кроме того, справедливы соотношения

$$\lim_{\varepsilon \rightarrow 0} T(\varepsilon) = T_0, \quad \max_{-\sigma_0 \leq t \leq T_0 + \sigma_0} |x(t, \varepsilon) - x_0(t)| = O(\varepsilon), \quad \varepsilon \rightarrow 0, \tag{2.8}$$

$$\lim_{\varepsilon \rightarrow 0} \max_{t \in \Sigma(\varepsilon)} |y_j(t, \varepsilon) - y_j^0(t, z_*)| = 0, \quad \max_{-\sigma_0 \leq t \leq T_0 + \sigma_0} |y_j(t, \varepsilon)| \leq M, \quad j = \overline{1, m-1}, \tag{2.9}$$

где  $x_0(t)$  – функция (1.5),  $M = \text{const} > 0$ , а множество  $\Sigma(\varepsilon)$  представляет собой отрезок  $[-\sigma_0, T_0 + \sigma_0]$  с выброшенными интервалами

$$(1 - \varepsilon^\delta, 1 + \varepsilon^\delta), \quad (t_0 + 1 - \varepsilon^\delta, t_0 + 1 + \varepsilon^\delta), \quad \delta = \text{const} \in (0, 1).$$

Доказательство данной теоремы опустим, поскольку в аналогичных ситуациях оно подробно изложено в уже упоминавшейся работе [11], а также в статьях [12, 13].

Теорема 1 носит базовый характер, так как сводит интересующую нас проблему нахождения устойчивых периодических решений системы (2.1), (2.2) к поиску устойчивых неподвижных точек отображения (2.6). Однако достаточно очевидно, что в общем случае записать и проанализировать это отображение не представляется возможным. Поэтому сделаем одно упрощающее предположение, а именно будем считать, что коэффициенты синаптических связей малы, т.е.

$$d_{j,s} = \nu d_{j,s}^0, \quad 0 < \nu \ll 1, \quad d_{j,s}^0 = \text{const} \in \mathbb{R}. \tag{2.10}$$

Это обстоятельство в совокупности с имеющимся минимальным запасом гладкости функции  $g(u)$  (см. (1.10)) позволяет асимптотически проинтегрировать систему (2.3), (2.4) на отрезке времени  $0 \leq t \leq T_0$ .

Действительно, дополнив указанную систему начальным условием (2.5) и учтя гладкость зависимости правых частей (2.3), (2.4) от параметров, приходим к выводу, что компоненты  $y_j^0(t, z, \nu)$ ,  $j = \overline{1, m-1}$ , решения получившейся задачи Коши допускают при  $\nu \rightarrow 0$  следующие асимптотические представления:

$$y_j^0(t, z, \nu) = z_j + \nu \psi_j(z)t + O(\nu^2), \quad j = \overline{1, m-1}, \quad 0 \leq t < 1; \tag{2.11}$$

$$y_j^0(t, z, \nu) = -az_j + \nu \psi_j(z) + \nu \psi_j(-az)(t-1) + O(\nu^2), \quad j = \overline{1, m-1}, \quad 1 < t < t_0 + 1; \tag{2.12}$$

$$y_j^0(t, z, \nu) = z_j - \frac{\nu}{a} \psi_j(z) + \nu \left(1 - \frac{1}{a^2}\right) \psi_j(-az) + \nu \psi_j(z)(t - t_0 - 1) + O(\nu^2), \quad j = \overline{1, m-1}, \quad t_0 + 1 < t \leq T_0, \tag{2.13}$$

где

$$\psi_j(z) = \sum_{s=1}^j d_{j+1,s}^0 g\left(\exp\left(-\sum_{r=s}^j z_r\right)\right) + \sum_{s=j+2}^m d_{j+1,s}^0 g\left(\exp\left(\sum_{r=j+1}^{s-1} z_r\right)\right) -$$

$$-\sum_{s=1}^{j-1} d_{j,s}^0 g\left(\exp\left(-\sum_{r=s}^{j-1} z_r\right)\right) - \sum_{s=j+1}^m d_{j,s}^0 g\left(\exp\left(\sum_{r=j}^{s-1} z_r\right)\right), \quad j = \overline{1, m-1}. \quad (2.14)$$

Добавим ещё, что остатки в (2.11)–(2.13) имеют указанный порядок малости равномерно по  $t$  из соответствующих промежутков и по  $z \in \Omega$ , где  $\Omega \subset \mathbb{R}^{m-1}$  – произвольно фиксированный компакт. Кроме того, формулы (2.11)–(2.13) сохраняют силу (вместе с порядками остатков) при дифференцировании по компонентам  $z_k$ ,  $k = \overline{1, m-1}$ , вектора  $z$ .

Соотношения (2.11)–(2.14) свидетельствуют о том, что в случае (2.10) отображение (2.6) асимптотически близко (в  $C^1$ -метрике) к тождественному, а точнее говоря, имеет вид

$$z_j \mapsto z_j + \nu\left(1 - \frac{1}{a^2}\right)(a\psi_j(z) + \psi_j(-az)) + O(\nu^2), \quad j = \overline{1, m-1}. \quad (2.15)$$

В свою очередь, опираясь на асимптотику (2.15), нетрудно заметить, что в рассматриваемой ситуации интересующее нас отображение с точностью до величин порядка  $\nu^2$  совпадает с оператором сдвига за время  $t = 1$  по траекториям системы дифференциальных уравнений

$$\dot{z}_j = \nu\left(1 - \frac{1}{a^2}\right)(a\psi_j(z) + \psi_j(-az)), \quad j = \overline{1, m-1}. \quad (2.16)$$

Справедливость данного факта следует из непосредственного асимптотического интегрирования системы (2.16) и сравнения получившихся формул с соотношениями (2.15). А это означает справедливость следующего утверждения.

**Теорема 2.** *Предположим, что выполнено неравенство*

$$a > 1 \quad (2.17)$$

*и модельная система*

$$\frac{dz_j}{d\tau} = a\psi_j(z) + \psi_j(-az), \quad j = \overline{1, m-1}, \quad (2.18)$$

*получающаяся из (2.16) при замене времени  $\tau = \nu(1 - 1/a^2)t$ , допускает гиперболическое состояние равновесия*

$$z_* = (z_1^*, \dots, z_{m-1}^*). \quad (2.19)$$

*Тогда при всех достаточно малых  $\nu > 0$  исходное отображение (2.15) имеет асимптотически близкую к (2.19) гиперболическую неподвижную точку*

$$z_*(\nu) = (z_1^*(\nu), \dots, z_{m-1}^*(\nu)): \quad z_j^*(\nu) = z_j^* + O(\nu), \quad j = \overline{1, m-1}, \quad \nu \rightarrow 0 \quad (2.20)$$

*с теми же свойствами устойчивости.*

Предположим далее, что выполнено неравенство (2.17) и удалось найти требуемое состояние равновесия (2.19) модельной системы (2.18). Тогда согласно теореме 1 для любых достаточно малых  $0 < \nu_1 < \nu_2$  найдётся такое  $\varepsilon_0 = \varepsilon_0(\nu_1, \nu_2) > 0$ , что при всех  $0 < \varepsilon \leq \varepsilon_0$ ,  $\nu_1 \leq \nu \leq \nu_2$  и при условии (2.10) система (2.1), (2.2) допускает гиперболический цикл

$$(x(t, \varepsilon, \nu), y_1(t, \varepsilon, \nu), \dots, y_{m-1}(t, \varepsilon, \nu)), \quad x(0, \varepsilon, \nu) \equiv 0 \quad (2.21)$$

периода  $T(\varepsilon, \nu)$ , свойства устойчивости которого совпадают с аналогичными свойствами неподвижной точки (2.20). Кроме того, справедливы аналогичные (2.8), (2.9) соотношения

$$\lim_{\varepsilon \rightarrow 0} T(\varepsilon, \nu) = T_0, \quad \max_{-\sigma_0 \leq t \leq T_0 + \sigma_0} |x(t, \varepsilon, \nu) - x_0(t)| = O(\varepsilon), \quad \varepsilon \rightarrow 0, \quad (2.22)$$

$$\lim_{\varepsilon \rightarrow 0} \max_{t \in \Sigma(\varepsilon)} |y_j(t, \varepsilon, \nu) - y_j^0(t, z_*(\nu))| = 0, \quad \max_{-\sigma_0 \leq t \leq T_0 + \sigma_0} |y_j(t, \varepsilon, \nu)| \leq M, \quad j = \overline{1, m-1}. \quad (2.23)$$

Интересно отметить, что приведённый выше результат может быть усилен, а именно найдутся такие достаточно малые  $\varepsilon_0, \nu_0 > 0$ , что цикл (2.21) существует и гиперболичесок при

всех  $0 < \varepsilon \leq \varepsilon_0$ ,  $0 < \nu \leq \nu_0$ . Что же касается свойств (2.22), (2.23), то они выполняются равномерно по  $\nu \in (0, \nu_0]$  и, в частности, постоянная  $M$  из (2.23) не зависит от  $\varepsilon$ ,  $\nu$ .

Справедливость перечисленных фактов обусловлена тем обстоятельством, что в силу (2.10) система (2.1), (2.2) регулярна по параметру  $\nu$ . Поэтому при независимом стремлении  $\varepsilon \rightarrow 0$ ,  $\nu \rightarrow 0$  цикл (2.21) сохраняется, а его начальные условия в фазовом пространстве  $C([-1, 0]; \mathbb{R}^m)$  имеют пределы  $x = t$ ,  $y_j = z_j^*$ ,  $j = \overline{1, m-1}$ ,  $-1 \leq t \leq 0$ , где  $z_j^*$  – компоненты вектора (2.19).

Ниже применимость базовых теорем 1, 2 иллюстрируется на двух наиболее популярных нейронных сетях – диффузионной и симметричной полносвязной.

**3. Существование и устойчивость дискретных автоволн.** Аналогом диффузионной модели (1.13) в нашем случае является система

$$\dot{u}_j = \left[ \lambda f(u_j(t-1)) + d \left( g \left( \frac{u_{j+1}}{u_j} \right) + g \left( \frac{u_{j-1}}{u_j} \right) \right) \right] u_j, \quad j = \overline{1, m}, \quad (3.1)$$

где  $u_0 = u_1$ ,  $u_{m+1} = u_m$ ,  $d = \text{const} > 0$ . Исследование аттракторов этой системы начнём с вопроса об устойчивости её однородного цикла (1.12).

Заметим, что для сети (3.1), в отличие от общего случая (1.11), соответствующая импульсная система (2.3), (2.4) приобретает более простой вид

$$\dot{y}_j = d[g(\exp y_{j+1}) + g(\exp(-y_j)) - g(\exp y_j) - g(\exp(-y_{j-1}))], \quad j = \overline{1, m-1}, \quad y_0 = y_m = 0, \quad (3.2)$$

$$y_j(1+0) = y_j(1-0) - (1+a)y_j(0),$$

$$y_j(t_0+1+0) = y_j(t_0+1-0) - (1+1/a)y_j(t_0), \quad j = \overline{1, m-1}. \quad (3.3)$$

Это обстоятельство позволяет разобраться с устойчивостью нулевой неподвижной точки отображения (2.6), которая в силу свойств (1.9) функции  $g(u)$  заведомо существует.

**Теорема 3.** Нулевая неподвижная точка отображения (2.6), порождённого системой (3.2), (3.3), экспоненциально устойчива при любом  $d > 0$ .

**Доказательство.** Для нахождения матрицы Якоби  $\Phi'(0)$ , отвечающей точке  $z = 0$ , необходимо линеаризовать систему (3.2) и учесть импульсные условия (3.3). В результате получаем, что матрица  $\Phi'(0)$  представляет собой оператор сдвига по решениям импульсной системы

$$\dot{h}_j = dg'(1)(h_{j+1} - 2h_j + h_{j-1}), \quad h_j(1+0) = h_j(1-0) - (1+a)h_j(0),$$

$$h_j(t_0+1+0) = h_j(t_0+1-0) - (1+1/a)h_j(t_0), \quad j = \overline{1, m-1}, \quad h_0 = h_m = 0 \quad (3.4)$$

за время от  $t = 0$  до  $t = T_0$ . Далее, применим к (3.4) метод Фурье по собственным векторам разностного оператора Лапласа, а точнее говоря, положим

$$h_j = \sum_{k=1}^{m-1} g_k(t) \sin\left(\frac{\pi k}{m} j\right), \quad j = \overline{1, m-1}. \quad (3.5)$$

В результате убеждаемся в том, что компоненты  $g_k(t)$ ,  $k = \overline{1, m-1}$ , из (3.5) являются решениями импульсной системы

$$\dot{g} = -\sigma g, \quad g(1+0) = g(1-0) - (1+a)g(0), \quad g(t_0+1+0) = g(t_0+1-0) - (1+1/a)g(t_0) \quad (3.6)$$

при  $0 \leq t \leq T_0$ ,  $\sigma = \sigma_k$ , где в силу (1.9) имеем

$$\sigma_k = 4dg'(1) \sin^2\left(\frac{\pi k}{2m}\right) > 0, \quad k = \overline{1, m-1}. \quad (3.7)$$

На завершающем этапе дополним систему (3.6) начальным условием  $g = 1$  при  $t = 0$ , проинтегрируем получившуюся задачу Коши и положим  $\mu_k = g|_{t=T_0, \sigma=\sigma_k}$ , где  $\sigma_k$  – дискретные значения (3.7) параметра  $\sigma$ . В итоге получаем набор чисел  $\mu_k$ ,  $k = \overline{1, m-1}$ , который, как



нетрудно увидеть, образует спектр интересующей нас матрицы  $\Phi'(0)$ . Более того, справедливы соотношения  $\mu_k = \mu(\sigma)|_{\sigma=\sigma_k}$ ,  $k = \overline{1, m-1}$ , где

$$\begin{aligned} \mu(\sigma) &= ((1+a)\exp \sigma - 1)((1+1/a)\exp \sigma - 1)\exp(-\sigma T_0), \quad \mu(0) = 1, \\ \mu'(\sigma) &= ((2-T_0)\exp(2\sigma) + (T_0-1)\exp \sigma - 1)T_0 \exp(-\sigma T_0) < 0 \quad \text{при } \sigma \in (0, +\infty), \\ \mu(\sigma) &\rightarrow 0 \quad \text{при } \sigma \rightarrow +\infty. \end{aligned}$$

Отсюда очевидным образом следует, что  $\mu_k \in (0, 1)$ ,  $k = \overline{1, m-1}$ . Теорема 3 доказана.

В аналогичной (2.1), (2.2) системе, записанной для диффузионного случая, неподвижной точке  $z = 0$  отвечает цикл вида (2.7) с компонентами  $y_1 = \dots = y_{m-1} = 0$ , а в исходной диффузионной цепочке (3.1) – однородный цикл (1.12). Теоремы 1, 3 приводят к выводу, что этот цикл является экспоненциально орбитально устойчивым при любом фиксированном значении параметра  $d > 0$  и при всех достаточно больших  $\lambda$ .

Перейдём теперь к другим аттракторам диффузионной системы (3.1). По аналогии с непрерывным случаем дискретной автоволной или просто автоволной назовём любой её стационарный режим, отличный от однородного цикла (1.12).

Нахождение автоволин будем проводить при аналогичном (2.10) предположении

$$d = \nu, \quad 0 < \nu \ll 1. \tag{3.8}$$

Напомним, что в случае (2.17), (3.8) в силу базовых теорем 1, 2 интересующий нас вопрос сводится к поиску устойчивых состояний равновесия соответствующей модельной системы (2.18). В свою очередь, несложный подсчёт показывает, что для диффузионной сети (3.1) упомянутая система приобретает вид

$$\frac{dz_j}{d\tau} = \Delta(-z_j) - \Delta(z_j) + \Delta(z_{j+1}) - \Delta(-z_{j-1}), \quad j = \overline{1, m-1}, \quad z_0 = z_m = 0, \tag{3.9}$$

где

$$\Delta(z) = ag(\exp z) + g(\exp(-az)), \quad z \in \mathbb{R}. \tag{3.10}$$

Исследование автоволин начнём с простейшего случая  $m = 2$ , когда имеем дело с так называемой билोकальной моделью

$$\dot{u}_1 = \lambda f(u_1(t-1))u_1 + dg\left(\frac{u_2}{u_1}\right)u_1, \quad \dot{u}_2 = \lambda f(u_2(t-1))u_2 + dg\left(\frac{u_1}{u_2}\right)u_2. \tag{3.11}$$

Для того чтобы в случае (2.17), (3.8) применить теоремы 1, 2, обратимся к системе (3.9), которая при  $m = 2$  состоит из одного скалярного уравнения

$$\frac{dz}{d\tau} = \Psi(z), \quad \Psi(z) \stackrel{\text{def}}{=} \Delta(-z) - \Delta(z). \tag{3.12}$$

Анализ уравнения (3.12) требует некоторой информации о его правой части  $\Psi(z)$ . Для получения этой информации объединим факт нечётности  $\Psi(z)$  с формулой (3.10) и свойствами (1.9), (1.10) функции  $g(u)$ . В результате приходим к выводу, что

$$\Psi(0) = \Psi'(0) = \Psi''(0) = 0, \quad \Psi'''(0) = 2a(a^2 - 1)(g'(1) + 3g''(1) + g'''(1)), \tag{3.13}$$

$$\lim_{z \rightarrow +\infty} \Psi(z) = -(b+1)(a-1). \tag{3.14}$$

Предположим теперь, что наряду с неравенством (2.17) выполнено условие

$$g'(1) + 3g''(1) + g'''(1) > 0. \tag{3.15}$$

Тогда в силу свойств (3.13), (3.14) для функции  $\Psi(z)$  уравнение (3.12) имеет хотя бы одну пару состояний равновесия  $z = \pm z_*$ ,  $z_* > 0$ , таких, что  $\Psi'(z_*) = \Psi'(-z_*) \leq 0$ . В общем же случае справедливо строгое неравенство

$$\Psi'(z_*) < 0, \tag{3.16}$$

а значит, эти состояния равновесия экспоненциально устойчивы. Добавим, что в силу теорем 1, 2 при условии (3.16) данной паре состояний равновесия в исходной модели (3.11) при  $\lambda \gg 1$  и при надлежащем уменьшении параметра  $d$  соответствует пара устойчивых автоволновых периодических движений. Отметим ещё, что упомянутые циклы переходят друг в друга при замене переменных  $u_1 \rightarrow u_2, u_2 \rightarrow u_1$ .

Вопрос о реализуемости ограничений (3.15), (3.16) требует отдельного рассмотрения. Для этого привлечем конкретный пример функции связи  $g(u)$ , а именно всюду ниже считаем, что

$$g(u) = (u - 1)/(1 + u/b), \quad b = \text{const} > 0. \tag{3.17}$$

Тогда, как нетрудно увидеть,

$$g'(1) = \frac{b}{b+1}, \quad g''(1) = -\frac{2b}{(b+1)^2}, \quad g'''(1) = \frac{6b}{(b+1)^3},$$

а значит, неравенство (3.15) эквивалентно включению

$$b \in (0, 2 - \sqrt{3}) \cup (2 + \sqrt{3}, +\infty). \tag{3.18}$$

В случае условия устойчивости (3.16) ситуация несколько сложнее. Получить для него явные ограничения на параметры  $a, b$ , подобные (3.18), не удаётся. Однако, как показывает численный анализ, оно заведомо справедливо, например, при значениях  $a = 2.5, b = 15$  (на рис. 2 изображён соответствующий этим параметрам график функции  $\Psi(z)$ ). Что же касается пары устойчивых циклов системы (3.11), существующих у неё согласно нашей теории, то для конкретных нелинейностей (1.3), (3.17) и при  $a = 2.5, b = 15, d = 0.005, \lambda = 6$  графики на плоскости  $(t, u)$  компонент  $u = u_1(t), u = u_2(t)$  одного из них представлены на рис. 3 (сплошной линией показан график  $u_1(t)$ , а пунктирной – график  $u_2(t)$ ).

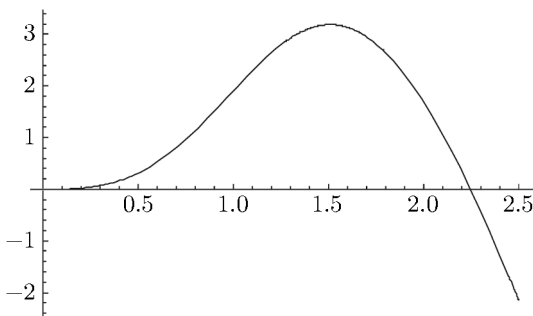


Рис. 2. График  $\Psi(z)$  при  $a = 2.5, b = 15$ .

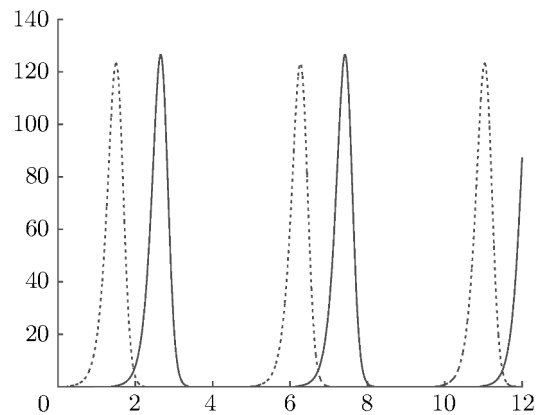


Рис. 3. Графики  $u_1(t), u_2(t)$  при  $a = 2.5, b = 15, d = 0.005, \lambda = 6$ .

Особого внимания заслуживает случай фиксированного  $a > 1$  и достаточно больших значений параметра  $b$ , когда условие (3.16) всё же удаётся проверить аналитически. Для того чтобы убедиться в этом, сделаем в уравнении (3.12) замены переменных

$$z = \ln b + \tilde{z}, \quad b\tau \rightarrow \tau \tag{3.19}$$

и учтём в его правой части вытекающие из формул (3.10), (3.17) асимптотические представления

$$\Delta(z)|_{z=\ln b+\tilde{z}} = b \left[ \frac{a \exp \tilde{z}}{1 + \exp \tilde{z}} + O(1/b) \right], \quad \Delta'(z)|_{z=\ln b+\tilde{z}} = b \left[ \frac{a \exp \tilde{z}}{(1 + \exp \tilde{z})^2} + O(1/b) \right], \quad (3.20)$$

$$\Delta(z)|_{z=-\ln b-\tilde{z}} = b[1 + O(b^{-\min(1,a-1)})],$$

$$\Delta'(z)|_{z=-\ln b-\tilde{z}} = \frac{1}{b} [a \exp(-\tilde{z}) + O(1/b)] - \frac{a \exp(a\tilde{z})}{b^{a-2}} [1 + O(b^{-\min(1,a-1)})], \quad (3.21)$$

справедливые при условии (2.17) и при  $b \rightarrow +\infty$  равномерно по  $\tilde{z}$  из любого фиксированного компакта. В результате после отбрасывания асимптотически малых добавок приходим к уравнению

$$\frac{d\tilde{z}}{d\tau} = 1 - \frac{a \exp \tilde{z}}{1 + \exp \tilde{z}}, \quad (3.22)$$

допускающему экспоненциально устойчивое состояние равновесия

$$\tilde{z} = \ln \frac{1}{a-1}. \quad (3.23)$$

Остаётся добавить, что в исходном уравнении (3.12) состоянию равновесия (3.23) отвечает состояние равновесия  $z = z_*(b)$  с асимптотикой

$$z_*(b) = \ln b + \ln \frac{1}{a-1} + O(b^{-\min(1,a-1)}), \quad b \rightarrow +\infty,$$

удовлетворяющее требуемому условию устойчивости (3.16).

Обратимся теперь к диффузионной цепочке (3.1) при  $m \geq 3$ , считая по-прежнему, что нелинейность  $g(u)$  в ней задана равенством (3.17). Далее, фиксируем произвольно индекс  $j_0$  из множества  $1 \leq j_0 \leq m-2$ , выполним во вспомогательной системе (3.9) аналогичные (3.19) замены переменных

$$z_j = \ln b + \tilde{z}_j, \quad j = \overline{1, j_0}; \quad z_j = -\ln b + \tilde{z}_j, \quad j = \overline{j_0+1, m-1}; \quad b\tau \rightarrow \tau, \quad (3.24)$$

и при условии (2.17) устремим параметр  $b$  к бесконечности. В результате, опираясь на асимптотические равенства (3.20), (3.21), для новых переменных  $\tilde{z}_j$ ,  $j = \overline{1, m-1}$ , приходим к некоторой предельной системе, которая в зависимости от индексов  $j_0$  и  $m$  записывается по-разному, а именно все возможные здесь ситуации исчерпываются следующими четырьмя случаями:

- 1) индекс  $j_0 = 1$ , при этом  $m \geq 4$ ;
- 2) средние значения индекса  $j_0$  такие, что  $2 \leq j_0 \leq m-3$ , а  $m \geq 5$ ;
- 3) максимальное значение индекса  $j_0 = m-2$ , причём  $m \geq 4$ ;
- 4) наконец,  $m = 3$ ,  $j_0 = 1$ .

Рассмотрим подробнее случай 1). Учитывая, что  $j_0 = 1$ , для перехода от переменной  $z_1$  к  $\tilde{z}_1$  используется первая формула из замены (3.24), а для всех остальных значений индекса  $j$  – вторая. Непосредственная подстановка асимптотических представлений (3.20), (3.21) в формулы (3.9), (3.10) приводит к следующей предельной системе:

$$\begin{aligned} \frac{d\tilde{z}_1}{d\tau} &= 2 - \frac{a \exp \tilde{z}_1}{1 + \exp \tilde{z}_1}, \quad \frac{d\tilde{z}_2}{d\tau} = \frac{a \exp(-\tilde{z}_2)}{1 + \exp(-\tilde{z}_2)} - 1, \\ \frac{d\tilde{z}_j}{d\tau} &= \frac{a \exp(-\tilde{z}_j)}{1 + \exp(-\tilde{z}_j)} - \frac{a \exp(-\tilde{z}_{j-1})}{1 + \exp(-\tilde{z}_{j-1})}, \quad 3 \leq j \leq m-2, \\ \frac{d\tilde{z}_{m-1}}{d\tau} &= \frac{a \exp(-\tilde{z}_{m-1})}{1 + \exp(-\tilde{z}_{m-1})} - 1 - \frac{a \exp(-\tilde{z}_{m-2})}{1 + \exp(-\tilde{z}_{m-2})}. \end{aligned} \quad (3.25)$$

В случае 2) ( $2 \leq j_0 \leq m - 3$ ,  $m \geq 5$ ) предельная система приобретает вид

$$\begin{aligned} \frac{d\tilde{z}_1}{d\tau} &= 1 - \frac{a \exp \tilde{z}_1}{1 + \exp \tilde{z}_1} + \frac{a \exp \tilde{z}_2}{1 + \exp \tilde{z}_2}, & \frac{d\tilde{z}_j}{d\tau} &= -\frac{a \exp \tilde{z}_j}{1 + \exp \tilde{z}_j} + \frac{a \exp \tilde{z}_{j+1}}{1 + \exp \tilde{z}_{j+1}}, & j &= \overline{2, j_0 - 1}, \\ \frac{d\tilde{z}_{j_0}}{d\tau} &= 1 - \frac{a \exp \tilde{z}_{j_0}}{1 + \exp \tilde{z}_{j_0}}, & \frac{d\tilde{z}_{j_0+1}}{d\tau} &= \frac{a \exp(-\tilde{z}_{j_0+1})}{1 + \exp(-\tilde{z}_{j_0+1})} - 1, \\ \frac{d\tilde{z}_j}{d\tau} &= \frac{a \exp(-\tilde{z}_j)}{1 + \exp(-\tilde{z}_j)} - \frac{a \exp(-\tilde{z}_{j-1})}{1 + \exp(-\tilde{z}_{j-1})}, & j_0 + 2 &\leq j \leq m - 2, \\ \frac{d\tilde{z}_{m-1}}{d\tau} &= \frac{a \exp(-\tilde{z}_{m-1})}{1 + \exp(-\tilde{z}_{m-1})} - \frac{a \exp(-\tilde{z}_{m-2})}{1 + \exp(-\tilde{z}_{m-2})} - 1; \end{aligned} \tag{3.26}$$

в случае 3) ( $j_0 = m - 2$ ,  $m \geq 4$ ) имеем дело с системой

$$\begin{aligned} \frac{d\tilde{z}_1}{d\tau} &= 1 - \frac{a \exp \tilde{z}_1}{1 + \exp \tilde{z}_1} + \frac{a \exp \tilde{z}_2}{1 + \exp \tilde{z}_2}, & \frac{d\tilde{z}_j}{d\tau} &= -\frac{a \exp \tilde{z}_j}{1 + \exp \tilde{z}_j} + \frac{a \exp \tilde{z}_{j+1}}{1 + \exp \tilde{z}_{j+1}}, & 2 &\leq j \leq m - 3, \\ \frac{d\tilde{z}_{m-2}}{d\tau} &= 1 - \frac{a \exp \tilde{z}_{m-2}}{1 + \exp \tilde{z}_{m-2}}, & \frac{d\tilde{z}_{m-1}}{d\tau} &= \frac{a \exp(-\tilde{z}_{m-1})}{1 + \exp(-\tilde{z}_{m-1})} - 2; \end{aligned} \tag{3.27}$$

в случае 4) ( $m = 3$ ,  $j_0 = 1$ ) получаем

$$\frac{d\tilde{z}_1}{d\tau} = 2 - \frac{a \exp \tilde{z}_1}{1 + \exp \tilde{z}_1}, \quad \frac{d\tilde{z}_2}{d\tau} = \frac{a \exp(-\tilde{z}_2)}{1 + \exp(-\tilde{z}_2)} - 2. \tag{3.28}$$

Несложный анализ каждого из вариантов (3.25)–(3.28) показывает, что при любых фиксированных  $m \geq 3$ ,  $1 \leq j_0 \leq m - 2$  и при дополнительном условии

$$a > 2 \tag{3.29}$$

соответствующая предельная система для  $\tilde{z}_j$ ,  $j = \overline{1, m - 1}$ , допускает единственное экспоненциально устойчивое состояние равновесия

$$\tilde{z}_{(j_0)} = \left( \ln \frac{1}{a - 2}, \underbrace{\ln \frac{1}{a - 1}, \dots, \ln \frac{1}{a - 1}}_{j_0 - 1}, \underbrace{-\ln \frac{1}{a - 1}, \dots, -\ln \frac{1}{a - 1}}_{m - 2 - j_0}, -\ln \frac{1}{a - 2} \right). \tag{3.30}$$

Что же касается системы (3.9), то в силу формул (3.20), (3.21), (3.24), (3.30) при условии (3.29) и при всех  $b \gg 1$  она имеет  $m - 2$  экспоненциально устойчивых состояния равновесия

$$\tilde{z}_{(j_0)}(b) = \left( \underbrace{\ln b, \dots, \ln b}_{j_0}, \underbrace{-\ln b, \dots, -\ln b}_{m - 1 - j_0} \right) + \tilde{z}_{(j_0)} + O(1/b), \quad b \rightarrow +\infty, \quad j_0 = \overline{1, m - 2}. \tag{3.31}$$

Согласно теоремам 1, 2 в диффузионной цепочке (3.1) с нелинейностью (3.17) состояниям равновесия (3.31) соответствуют экспоненциально орбитально устойчивые автоволновые периодические режимы. Отсюда в силу произвольности  $m$  заключаем, что в рамках модели (3.1), (3.17) реализуется известное явление буферности. Суть этого явления в том, что при подходящем подборе параметров  $a$ ,  $b$ ,  $\lambda$ ,  $d$  и при увеличении  $m$  можно гарантировать существование в указанной модели любого наперёд заданного конечного числа устойчивых циклов. Добавим ещё, что феномен буферности характерен именно для нейронных систем, поскольку он отражает ситуацию, когда в неокортексе человеческого мозга различные идеи и концепции соревнуются друг с другом в целях доминирования.

**4. Случай симметричной полносвязной сети.** В максимально симметричном случае, когда все связи равноправны, интересующая нас математическая модель нейронной сети приобретает вид

$$\dot{u}_j = \left[ \lambda f(u_j(t-1)) + d \sum_{\substack{s=1 \\ s \neq j}}^m g\left(\frac{u_s}{u_j}\right) \right] u_j, \quad j = \overline{1, m}, \tag{4.1}$$

где  $d = \text{const} > 0$ . Интересны вопросы существования и устойчивости специальных её периодических решений – так называемых режимов двухкластерной синхронизации.

Для описания упомянутых режимов фиксируем произвольно натуральное  $k: 1 \leq k \leq m - 1$  и предположим, что совокупность индексов  $1 \leq j \leq m$  разбита на два непересекающихся множества  $\mathcal{A}$  и  $\mathcal{B}$ , состоящих из  $k$  и  $m - k$  элементов соответственно, т.е.

$$\{1, 2, \dots, m\} = \mathcal{A} \cup \mathcal{B}. \tag{4.2}$$

Тогда, очевидно, система (4.1) допускает решения с компонентами

$$u_j = v(t) \quad \text{при} \quad j \in \mathcal{A}, \quad u_j = w(t) \quad \text{при} \quad j \in \mathcal{B}, \tag{4.3}$$

где переменные  $v, w$  удовлетворяют вспомогательной системе

$$\dot{v} = \lambda f(v(t-1))v + (m - k)dg\left(\frac{w}{v}\right)v, \quad \dot{w} = \lambda f(w(t-1))w + kdg\left(\frac{v}{w}\right)w. \tag{4.4}$$

Если же, в свою очередь, система (4.4) имеет непостоянное периодическое решение

$$v = v_{(k)}(t), \quad w = w_{(k)}(t), \quad v_{(k)}(t) \neq w_{(k)}(t), \tag{4.5}$$

то отвечающее ему решение (4.2), (4.3) исходной системы (4.1) назовём периодическим режимом двухкластерной синхронизации.

Нетрудно увидеть, что один и тот же цикл (4.5) порождает целое семейство  $\mathcal{U}_k$  периодических режимов двухкластерной синхронизации. Все циклы из этого семейства задаются равенствами (4.2), (4.3) при

$$v = v_{(k)}(t), \quad w = w_{(k)}(t),$$

а их количество очевидно равно  $C_m^k$ .

Остановимся сначала на вопросе о существовании семейства циклов  $\mathcal{U}_k$ . В связи с этим при условиях (2.17), (3.8) применим к системе (4.4) теоремы 1, 2. В результате убедимся в том, что проблема существования и устойчивости её неоднородного цикла (4.5) сводится к отысканию ненулевого устойчивого состояния равновесия аналогичного (3.12) скалярного нелинейного уравнения

$$\frac{dz}{d\tau} = k\Delta(-z) - (m - k)\Delta(z), \tag{4.6}$$

где  $\Delta(z)$  – функция (3.10).

Как и ранее, всюду ниже считаем, что нелинейность  $g(u)$  в (4.1) задана равенством (3.17). Тогда, выполнив в (4.6) замены (3.19) и устремив при фиксированном  $a > 1$  параметр  $b$  к бесконечности, для новой переменной  $\tilde{z}$  получим аналогичное (3.22) уравнение

$$\frac{d\tilde{z}}{d\tau} = k - a(m - k) \frac{\exp \tilde{z}}{1 + \exp \tilde{z}}. \tag{4.7}$$

Заметим, что при условии

$$a > \max\left(1, \frac{k}{m - k}\right) \tag{4.8}$$

уравнение (4.7) допускает экспоненциально устойчивое состояние равновесия

$$z_{(k)}^* = \ln \frac{k}{a(m-k) - k}. \tag{4.9}$$

Что же касается исходного уравнения (4.6), то при выполнении неравенства (4.8) оно имеет экспоненциально устойчивое состояние равновесия  $z_{(k)}^*(b) > 0$  с асимптотикой

$$z_{(k)}^*(b) = \ln b + z_{(k)}^* + O(b^{-\min(1, a-1)}), \quad b \rightarrow +\infty, \tag{4.10}$$

где величина  $z_{(k)}^*$  определяется в (4.9).

Итак, из наших построений следует, что при выполнении условия (3.8) при всех  $\lambda \gg 1$  и при фиксированных параметрах  $a$  и  $b$ , первый из которых удовлетворяет неравенству (4.8), а второй – достаточно большой, система (4.1) с нелинейностью (3.17) имеет искомое семейство  $\mathcal{U}_k$  периодических режимов двухкластерной синхронизации. Перейдём теперь к вопросу об устойчивости данных циклов. В связи с этим сделаем два полезных наблюдения.

Во-первых, несложная проверка показывает, что система (4.1) инвариантна относительно замен

$$(u_1, u_2, \dots, u_m) \rightarrow (u_{j_1}, u_{j_2}, \dots, u_{j_m}), \tag{4.11}$$

где  $(j_1, j_2, \dots, j_m)$  – произвольная перестановка набора индексов  $(1, 2, \dots, m)$ . Во-вторых, периодические режимы из семейства  $\mathcal{U}_k$  допускают кодирование с помощью бинарных векторов

$$(\vartheta_1, \vartheta_2, \dots, \vartheta_m) : \quad \vartheta_j = 1 \quad \text{или} \quad 0, \quad j = \overline{1, m}. \tag{4.12}$$

Точнее говоря, предполагаем, что  $j$ -я координата вектора (4.12) равна 1 или 0 при  $j \in \mathcal{A}$  или  $j \in \mathcal{B}$  соответственно. В этом случае между векторами (4.12), содержащими  $k$  единиц и  $m - k$  нулей, и циклами семейства  $\mathcal{U}_k$  имеет место взаимно-однозначное соответствие.

Суммируя сказанное выше, убеждаемся в том, что любые два цикла из  $\mathcal{U}_k$  переходят друг в друга под действием замен (4.11), а значит имеют одинаковые свойства устойчивости. Таким образом, проблема устойчивости всех режимов семейства  $\mathcal{U}_k$  сводится к исследованию устойчивости только одного цикла (4.2), (4.3), (4.5), соответствующего бинарному вектору

$$\underbrace{(1, \dots, 1)}_k, \underbrace{(0, \dots, 0)}_{m-k}. \tag{4.13}$$

Для решения поставленной проблемы обратимся к модельной системе (2.18), отвечающей системе (4.1) при условиях (2.17), (3.8), (3.17). Характерная особенность этой системы состоит в том, что нелинейности  $\psi_j(z)$  в ней теперь задаются равенствами (2.14), в которых все коэффициенты  $d_{j,s}^0$  равны единице. Учитывая данное обстоятельство, нетрудно увидеть, что при достаточно большом фиксированном значении параметра  $b$  и при дополнительном условии (4.8) циклу семейства  $\mathcal{U}_k$  с кодировкой (4.13) соответствует состояние равновесия указанной системы с компонентами

$$z_j = 0, \quad j = \overline{1, m-1}, \quad j \neq k; \quad z_k = z_{(k)}^*(b), \tag{4.14}$$

где  $z_{(k)}^*(b)$  – состояние равновесия (4.10) уравнения (4.6). Таким образом, свойства устойчивости интересующего нас цикла совпадают с аналогичными свойствами состояния равновесия (4.14).

Несложный подсчёт показывает, что система в вариациях, получающаяся при линеаризации системы (2.18) на упомянутом выше состоянии равновесия, преобразуется к виду

$$\frac{dh_j}{d\tau} = -(m-k)\Delta'(z_k)h_j, \quad 1 \leq j \leq k-1, \tag{4.15}$$

$$\frac{dh_k}{d\tau} = -\Delta'(-z_k) \sum_{s=1}^k \left( \sum_{r=s}^k h_r \right) - \Delta'(z_k) \sum_{s=k+1}^m \left( \sum_{r=k}^{s-1} h_r \right), \tag{4.16}$$

$$\frac{dh_j}{d\tau} = -k\Delta'(-z_k)h_j, \quad k+1 \leq j \leq m-1, \tag{4.17}$$

где  $z_k = z_{(k)}^*(b)$  (см. (4.10)),  $\Delta'(z)$ ,  $z \in \mathbb{R}$ , – производная функции (3.10). В свою очередь, из структуры уравнений (4.15)–(4.17) следует, что собственные значения матрицы этой системы задаются равенствами

$$\begin{aligned} \lambda_j &= -(m-k)\Delta'(z_k), \quad 1 \leq j \leq k-1; \quad \lambda_k = -k\Delta'(-z_k) - (m-k)\Delta'(z_k); \\ \lambda_j &= -k\Delta'(-z_k), \quad k+1 \leq j \leq m-1. \end{aligned} \tag{4.18}$$

При анализе знаков величин (4.18) будем считать выполненным неравенство

$$a > 3. \tag{4.19}$$

В этом случае в силу соотношений (3.20), (3.21), (4.10) имеем

$$\lambda_j = -(m-k)b \left[ \frac{a \exp z_{(k)}^*}{(1 + \exp z_{(k)}^*)^2} + O(1/b) \right] < 0, \quad 1 \leq j \leq k;$$

$$\lambda_j = -\frac{ka}{b} \exp(-z_{(k)}^*) [1 + O(b^{-\min(1, a-3)})] < 0, \quad k+1 \leq j \leq m-1,$$

отсюда и из (4.8), (4.19), в свою очередь, следует, что при выполнении неравенства

$$a > \max \left( 3, \frac{k}{m-k} \right) \tag{4.20}$$

все циклы семейства  $\mathcal{U}_k$  являются экспоненциально орбитально устойчивыми.

В итоге отметим, что, как и в предыдущем случае, в рамках модели (4.1) с нелинейностью (3.17) наблюдается феномен буферности. Действительно, из условия (4.20) следует, что при  $d \ll 1$ ,  $\lambda \gg 1$  при достаточно большом фиксированном  $b > 0$  и при  $a > \max(3, m-1)$  в указанной модели существуют все семейства  $\mathcal{U}_k$ ,  $k = \overline{1, m-1}$ , периодических режимов двухкластерной синхронизации и все режимы из этих семейств устойчивы. Отсюда в силу произвольности  $m$  заключаем, что можно добиться сосуществования любого конечного числа устойчивых циклов.

**5. Заключение.** Интересно отметить, что система (1.11) может иметь и другие аттракторы, отличные от тех, которые описывает теорема 1. Для пояснения сути дела предположим, что коэффициенты синаптических связей заданы равенствами

$$d_{j,s} = \lambda d_{j,s}^0, \quad d_{j,s}^0 = \text{const} \in \mathbb{R}. \tag{5.1}$$

Далее с учётом соотношения (5.1) произведём в (1.11) замены

$$u_j = \exp(x_j/\varepsilon), \quad j = \overline{1, m}, \quad \varepsilon = 1/\lambda \ll 1.$$

В результате приходим к системе вида

$$\dot{x}_j = F(x_j(t-1), \varepsilon) + \sum_{\substack{s=1 \\ s \neq j}}^m d_{j,s}^0 H(x_s - x_j, \varepsilon), \quad j = \overline{1, m}, \tag{5.2}$$

где  $F(x, \varepsilon)$  – функция из (2.1),  $H(x, \varepsilon) = g(\exp(x/\varepsilon))$ .

Обратим внимание, что в силу свойств (1.2), (1.9), (1.10) справедливы предельные равенства

$$\lim_{\varepsilon \rightarrow 0} F(x, \varepsilon) = R_0(x) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{при } x < 0, \\ f(1) & \text{при } x = 0, \\ -a & \text{при } x > 0, \end{cases} \quad \lim_{\varepsilon \rightarrow 0} H(x, \varepsilon) = H_0(x) \stackrel{\text{def}}{=} \begin{cases} -1 & \text{при } x < 0, \\ 0 & \text{при } x = 0, \\ b & \text{при } x > 0. \end{cases}$$

Отсюда следует, что при  $\varepsilon \rightarrow 0$  исходная релаксационная система (5.2) переходит в релейную систему

$$\dot{x}_j = R_0(x_j(t-1)) + \sum_{\substack{s=1 \\ s \neq j}}^m d_{j,s}^0 H_0(x_s - x_j), \quad j = \overline{1, m}. \tag{5.3}$$

Наличие предельного объекта (5.3) представляет собой ещё одну характерную особенность нашего подхода к математическому моделированию электрических синапсов. Опираясь на эту особенность, для отыскания аттракторов системы (5.2) можем воспользоваться общими результатами из статьи [5] о соответствии между устойчивыми циклами релейной и релаксационной систем. Однако для того чтобы применить эти результаты сначала необходимо изучить свойства предельной релейной системы. К сожалению, пока имеется весьма скудная информация об её аттракторах.

В первую очередь отметим, что система (5.3) допускает однородный цикл

$$x_1 = x_2 = \dots = x_m = x_0(t), \tag{5.4}$$

где  $x_0(t)$  – периодическая функция (1.5). Вопрос же о нахождении у неё циклов, отличных от (5.4), представляет собой отдельную и пока не решённую проблему. Добавим только, что, как показывают численные эксперименты, такие циклы могут существовать даже в простейшем случае  $m = 2$ .

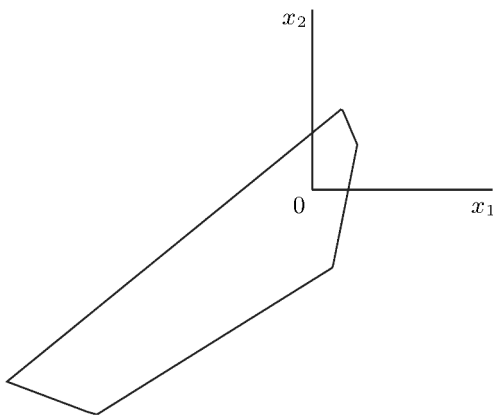
Для того чтобы убедиться в этом, обратимся к системе

$$\dot{x}_1 = R_0(x_1(t-1)) + d_1 H_0(x_2 - x_1), \quad \dot{x}_2 = R_0(x_2(t-1)) + d_2 H_0(x_1 - x_2), \tag{5.5}$$

где  $d_1, d_2 = \text{const} > 0$ . Отметим сразу, что в симметричном случае  $d_1 = d_2$  кроме однородного цикла (5.4) никаких других аттракторов в ней обнаружить не удалось. Однако в несимметричном случае численный анализ данной системы при значениях параметров  $a = 2.5$ ,  $b = 0.07$ ,  $d_1 = 1.5$ ,  $d_2 = 0.1$  и при начальных условиях  $x_1 = -1$ ,  $x_2 = t$ , заданных на отрезке  $-1 \leq t \leq 0$ , показал наличие у неё устойчивого неоднородного цикла. Характерная особенность этого цикла состоит в том, что его компоненты  $x_1(t)$ ,  $x_2(t)$  имеют по четыре переключения на промежутке времени длины периода.

Напомним (см. работы [5, 14]), что согласно общепринятой в теории релейных систем терминологии точками переключения компоненты  $x_j(t)$  системы (5.5) называются моменты времени, при прохождении через которые производная  $\dot{x}_j(t)$  претерпевает конечные скачки, а сама функция  $x_j(t)$  остаётся непрерывной. В случае найденного нами цикла у его компонент  $x_1(t)$ ,  $x_2(t)$  на отрезке длины периода имеются по две точки переключения, связанные с переменами знаков запаздывающих компонент  $x_1(t-1)$  и  $x_2(t-1)$  соответственно, и ещё две общие для обеих компонент точки переключения, обусловленные переменами знаков функции  $x_1(t) - x_2(t)$ .

Проекция упомянутого устойчивого цикла на плоскость  $(x_1, x_2)$  представлена на рис. 4. Вопрос о его аналитическом построении пока остаётся открытым.



**Рис. 4.** График кривой  $(x_1(t), x_2(t))$  при  $a = 2.5$ ,  $b = 0.07$ ,  $d_1 = 1.5$ ,  $d_2 = 0.1$ .

Исследование выполнено при поддержке Российского научного фонда (проект 22-11-00209).



## СПИСОК ЛИТЕРАТУРЫ

1. Глызин С.Д., Колесов А.Ю., Розов Н.Х. Об одном способе математического моделирования химических синапсов // Дифференц. уравнения. 2013. Т. 49. № 10. С. 1227–1244.
2. Глызин С.Д., Колесов А.Ю., Розов Н.Х. Релаксационные автоколебания в сетях импульсных нейронов // Успехи мат. наук. 2015. Т. 70. Вып. 3 (423). С. 3–76.
3. Hutchinson G.E. Circular causal systems in ecology // Ann. N.Y. Acad. of Sci. 1948. V. 50. P. 221–246.
4. Колесов А.Ю., Мищенко Е.Ф., Розов Н.Х. Об одной модификации уравнения Хатчинсона // Журн. вычислит. математики и мат. физики. 2010. Т. 50. № 12. С. 2099–2112.
5. Колесов А.Ю., Мищенко Е.Ф., Розов Н.Х. Реле с запаздыванием и его  $C^1$ -аппроксимация // Тр. Мат. ин-та им. В.А. Стеклова 1997. Т. 216. С. 126–153.
6. Scott A.C. Neuroscience: a Mathematical Primer. New York, 2002.
7. Кащенко С.А., Майоров В.В. Модели волновой памяти. М., 2013.
8. Абарбанель Г.Д.И., Рабинович М.И., Селверстон А., Баженков М.В., Хуэрта Р., Суццик М.М., Рубчинский Л.Л. Синхронизация в нейронных ансамблях // Успехи физ. наук. 1996. Т. 166. № 4. С. 363–390.
9. Леванова Т.А., Казаков А.О., Коротков А.Г., Осипов Г.В. Влияние электрической связи на динамику ансамбля нейроноподобных элементов с синаптическими тормозящими связями // Изв. вузов. Прикл. нелинейная динамика. 2018. Т. 26. № 5. С. 101–112.
10. Шапин Д.С., Некоркин В.И. Параметрически возбуждаемые хаотические спайковые последовательности и информационные аспекты в ансамбле нейронов ФитцХью–Нагумо // Письма в Журн. эксп. и теор. физики. 2021. Т. 113. Вып. 6. С. 415–420.
11. Колесов А.Ю., Розов Н.Х. Автоволновые процессы в печочках диффузионно связанных уравнений с запаздыванием // Успехи мат. наук. 2012. Т. 67. Вып. 2 (404). С. 109–156.
12. Глызин С.Д., Колесов А.Ю., Розов Н.Х. Релаксационные автоколебания в нейронных системах. II // Дифференц. уравнения. 2011. Т. 47. № 12. С. 1675–1692.
13. Колесов А.Ю., Розов Н.Х. Дискретные автоволны в системах дифференциально-разностных уравнений с запаздыванием из экологии // Тр. Мат. ин-та им. В.А. Стеклова. 2012. Т. 277. С. 101–143.
14. Глызин С.Д., Колесов А.Ю., Розов Н.Х. Релаксационные автоколебания в нейронных системах. I // Дифференц. уравнения. 2011. Т. 47. № 7. С. 919–932.

Ярославский государственный университет  
имени П.Г. Демидова

Поступила в редакцию 16.11.2021 г.  
После доработки 20.05.2022 г.  
Принята к публикации 25.05.2022 г.

===== ОБЫКНОВЕННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ =====

УДК 517.927.6

## О КРАТНОСТИ СОБСТВЕННЫХ ЗНАЧЕНИЙ ДИФФЕРЕНЦИАЛЬНОГО ОПЕРАТОРА ЧЕТВЁРТОГО ПОРЯДКА НА ГРАФЕ

© 2022 г. Р. Ч. Кулаев, А. А. Уртаева

Изучаются свойства собственных значений краевой задачи для дифференциального уравнения четвёртого порядка на геометрическом графе, моделирующей упругие деформации стержневой системы с условиями упругошарнирного соединения в узлах. Устанавливаются условия простоты точек спектра соответствующего дифференциального оператора. Выводятся оценки кратности собственных значений. Эти оценки даются в терминах топологических характеристик графа. В рамках этих понятий полученные оценки являются точными.

DOI: 10.31857/S0374064122070020, EDN: CEADRF

**Введение.** В настоящей работе рассматривается краевая задача для дифференциального уравнения четвёртого порядка на геометрическом графе (пространственной сети)  $\Gamma$ :

$$L_\lambda u \equiv \frac{d^2}{d\Gamma^2} \left( p(x) \frac{d^2 u}{d\Gamma^2} \right) = \lambda \rho(x) u, \quad u|_{\partial\Gamma} = u'|_{\partial\Gamma} = 0, \quad (1)$$

где  $\partial\Gamma$  – множество граничных вершин  $\Gamma$ . Задача (1) представляет собой модель балочной конструкции [1–7]. Во внутренних точках рёбер производная  $du/d\Gamma$  имеет классическую форму  $u'$ , а в узлах сети оператор  $L_\lambda$  задаётся наборами условий для упругого шарнирного соединения балок. Мы изучаем вопрос о кратности собственных значений задачи (1) и даём точную оценку кратности собственных значений этой задачи.

Давно замечено, что оценка кратности собственных значений краевых задач на графах существенно зависит от топологии графа. На сегодняшний день хорошо изучены спектральные свойства задачи Штурма–Лиувилля на графе с различными условиями стыковки в узловых вершинах графа (см. работы [8–12]).

Стоит сразу отметить, что изучение свойств решений дифференциальных уравнений четвёртого порядка на сети является существенно более сложной задачей, по сравнению с аналогичной задачей для уравнения Штурма–Лиувилля. Даже модели физического происхождения оказываются очень трудными для анализа (см., например, [4, 5, 13–15]). Наибольшие продвижения в изучении качественных свойств краевых задач четвёртого порядка на графах имеются для уравнения с условиями упругошарнирного сочленения стержней. Для такой краевой задачи в статьях [4, 5] получены условия разрешимости, установлен принцип максимума, на основе которого доказана положительная обратимость краевой задачи и положительность её функции Грина. В [16, 17] изучались свойства решений уравнения (1) (положительность, колеблемость, распределение нулей, неосцилляционность). Как показывают результаты этих работ, дифференциальный оператор  $L_\lambda$  задачи (1), порождаемый условиями упругошарнирного сочленения, наследует основные качественные свойства оператора Штурма–Лиувилля на сети. В связи с этим естественным становится вопрос об исследованиях осцилляционных свойств спектра оператора  $L_\lambda$ . Одним из первых шагов на этом пути является изучение вопроса о кратности собственных значений. Ниже устанавливаются условия, обеспечивающие простоту собственных значений оператора  $L_\lambda$ , а также оценки кратностей собственных значений. В частности, показывается, что для точек спектра оператора  $L_\lambda$  имеют место оценки их кратностей, аналогичные оценкам для собственных значений задачи Штурма–Лиувилля (см. [12]). При этом получаемые оценки являются наилучшими.

**1. Постановка задачи.** Далее используем терминологию и обозначения работ [2, 5]. На протяжении всей статьи  $\Gamma \subset \mathbb{R}^N$  обозначает связный и конечный *геометрический граф* без петель с множеством вершин  $V(\Gamma)$  и множеством точек рёбер графа  $E(\Gamma)$ . *Ребро* графа – это интервал конечной длины, а *вершина* графа – это концевая точка одного или нескольких ребер. Ребра графа обозначаются  $\gamma_i$ , вершины –  $a, b$  и т.д. Для любой вершины  $a \in V(\Gamma)$  через  $I(a)$  обозначим множество индексов ребер, инцидентных вершине  $a$ , через  $|I(a)|$  – количество элементов множества  $I(a)$ . Элементы множеств  $J(\Gamma) = \{a \in V(\Gamma) : |I(a)| \geq 2\}$  и  $\partial\Gamma = \{a \in V(\Gamma) : |I(a)| = 1\}$  называются *внутренними* и *граничными* вершинами соответственно. Через  $|\partial\Gamma|$  обозначаем число граничных вершин графа  $\Gamma$ . Предполагаем, что  $\Gamma = E(\Gamma) \cup J(\Gamma)$ . Обратим внимание, что граничные вершины не включены в граф. *Подграфом* графа  $\Gamma$  называется любое непустое связное подмножество  $\Gamma$ , которое попадает под определение графа. Граф, не имеющий циклов, называется *деревом*. Граф-дерево  $\Gamma$  называем *цепочкой*, если  $|I(a)| = 2$  для любой вершины  $a \in J(\Gamma)$ .

Введём функциональные пространства:

$$C[\Gamma] = \{u : \Gamma \rightarrow \mathbb{R} : u \text{ равномерно непрерывна на каждом ребре } \gamma_i \subset E(\Gamma)\};$$

$$C[E(\Gamma)] = \{u : E(\Gamma) \rightarrow \mathbb{R} : u \text{ равномерно непрерывна на каждом ребре } \gamma_i \subset E(\Gamma)\}.$$

Каждая функция  $u \in C[\Gamma]$  (или  $C[E(\Gamma)]$ ) имеет предел  $\lim_{\gamma_i \ni x \rightarrow a} u_i(x)$ ,  $i \in I(a)$ , в каждой вершине  $a \in V(\Gamma)$ ; обозначим его через  $u_i(a)$ . Обратим внимание, что  $u_k(a)$  не обязательно равны  $u_i(a)$  или  $u(a)$ , где  $k, i \in I(a)$  ( $k \neq i$ ). Пространство непрерывных на  $\Gamma$  функций определяется равенством

$$C(\Gamma) = \{u \in C[\Gamma] : u_i(a) = u(a) \text{ для всех } a \in J(\Gamma) \text{ и всех } i \in I(a)\}.$$

Теперь определим производную функции на графе. Для этого зададим функцию  $\mu(x) \in C[E(\Gamma)]$ , линейно отображающую каждое ребро  $\gamma_i \subset E(\Gamma)$  на интервал  $(0, l_i) \subset \mathbb{R}$ , где  $l_i$  – длина  $\gamma_i$ . Положим  $u'_i(x) := \lim_{\gamma_i \ni y \rightarrow x} (u_i(y) - u_i(x)) / (\mu_i(y) - \mu_i(x))$ ,  $x \in \bar{\gamma}_i$ . Аналогично определяются производные более высокого порядка.

Через  $C^n[\Gamma]$  (или  $C^n[E(\Gamma)]$ ) мы обозначаем пространство функций  $u(x) \in C[\Gamma]$  (или  $C^n[E(\Gamma)]$ ), производные которых до порядка  $n$  включительно существуют и принадлежат пространству  $C[E(\Gamma)]$ . Для функции  $u(x) \in C^n[\Gamma]$  (или  $C^n[E(\Gamma)]$ ) в любой вершине  $a \in V(\Gamma)$  определено множество производных  $u_i^{(j)}(a)$ ,  $1 \leq j \leq n$ ,  $i \in I(a)$ , вдоль ребер, смежных с  $a$ . Производные нечётного порядка зависят от ориентации ребер. В дальнейшем при записи условий связи с производными в вершинах графа нам будет удобно использовать производные по направлению “от вершины”, которые мы будем обозначать  $u_{i\nu}^{(k)}(a)$  (одномерный аналог производных по внутренней нормали). Для чётной производной ориентация не важна, и поэтому для краткости вместо  $u''_{i\nu}(a)$  пишем  $u''_i(a)$ .

Изучение вопроса о кратности точек спектра задачи (1) будет связано с исследованием свойств решений следующего уравнения:

$$Lu \equiv \frac{d^2}{d\Gamma^2} \left( p(x) \frac{d^2 u}{d\Gamma^2} \right) - r(x)u = 0, \quad x \in \Gamma. \tag{2}$$

При этом под дифференциальным уравнением (2) (совпадающим с уравнением (1) при  $r = \lambda\rho$ ) подразумевается набор обыкновенных дифференциальных уравнений на ребрах графа

$$(p_i(x)u''_i)'' - r_i(x)u = 0, \quad x \in \gamma_i \subset E(\Gamma), \tag{3}$$

и набор условий согласования в каждой внутренней вершине  $a \in J(\Gamma)$

$$u_i(a) = u(a), \quad \beta_i(a)u''_i(a) - \vartheta_i(a)u'_{i\nu}(a) = 0, \quad i \in I(a), \tag{4}$$

$$\sum_{i \in I(a)} (p_i u''_i)'_{\nu}(a) - r(a)u(a) = 0, \quad a \in J(\Gamma). \tag{5}$$

Задача (3)–(5) моделирует малые деформации стержневой системы с условиями упругошарнирного соединения (см. [2, 5] и [7, п. 4.2]). В этом случае равенства (3)–(5) можно трактовать следующим образом:  $u(x)$  обозначает смещение балки при выходе из состояния равновесия; условия (4) описывают классические локальные условия в узлах графа – перемещения всех стержней системы являются непрерывными и имеется упругошарнирное сочленение в вершине  $a$  (см. [18, раздел 5.18]). Последнее условие – это условие динамического равновесия.

Решением дифференциального уравнения (2) будем называть всякую функцию  $u(x) \in C^4[\Gamma] \cap C(\Gamma)$ , удовлетворяющую на каждом ребре графа соответствующему обыкновенному дифференциальному уравнению (3), а в каждой внутренней вершине – условиям (4), (5).

Всюду далее полагаем

$$p \in C^2[E(\Gamma)], \quad \inf_{x \in E(\Gamma)} p(x) > 0 \quad \text{и} \quad r, \rho \in C[\Gamma], \quad r(x) > 0, \quad \rho(x) > 0 \quad \text{на} \quad \Gamma;$$

$$\beta_i(a) \geq 0, \quad \vartheta_i(a) \geq 0 \quad \text{и} \quad \beta_i(a) + \vartheta_i(a) > 0 \quad \text{для любых} \quad a \in V(\Gamma), \quad i \in I(a).$$

**2. Вспомогательные утверждения.** На протяжении всей статьи будем опираться на следующие результаты.

**Лемма 1** [19]. Пусть  $u(x)$  – нетривиальное решение уравнения

$$(p(x)u''') - r(x)u = 0, \quad x \in [a, b] \subset \mathbb{R}. \tag{6}$$

Если  $u, u', u'', (pu''')$  неотрицательны в точке  $a$ , то все эти функции положительны на промежутке  $(a, b)$ .

**Следствие 1.** Пусть  $u(x)$  – нетривиальное решение уравнения (6), удовлетворяющее граничным условиям

$$u(a)(\beta(a)u''(a) - \vartheta(a)u'(a)) = 0, \quad u(b)(\beta(b)u''(b) + \vartheta(b)u'(b)) = 0. \tag{7}$$

Тогда  $u(x)$  не имеет кратных нулей на интервале  $(a, b)$ , а в конечных точках  $a$  и  $b$  не может иметь нуль кратности большей двух.

**Определение 1.**  $S$ -зоной функции  $u(x) \in C(\Gamma)$  назовём подграф  $\Gamma_0 \subset \Gamma$  такой, что  $u(x) \neq 0$  на  $\Gamma_0$ ;  $u(x) = 0$  на  $\partial\Gamma_0$ ;  $u(x)$  имеет нуль на любом подграфе  $\Gamma_1 \supset \Gamma_0$ ,  $\Gamma_1 \neq \Gamma_0$ .

**Теорема 1** [16]. Пусть  $u(x)$  – решение краевой задачи

$$Lu = 0, \quad x \in \Gamma, \quad u|_{\partial\Gamma} = (\beta u'' - \vartheta u')|_{\partial\Gamma} = 0, \tag{8}$$

имеющее  $S$ -зону  $\Gamma_0 \subset \Gamma$ . Тогда любое решение задачи (8) либо пропорционально  $u(x)$  на  $\Gamma_0$ , либо меняет знак в  $\Gamma_0$ .

**3. Условия простоты собственных значений.** Сначала опишем структуру спектра краевой задачи (1).

**Лемма 2.** При  $\lambda < 0$  краевая задача (1) имеет только тривиальное решение.

**Доказательство.** Пусть  $u$  – решение задачи (1). Умножим  $u$  на  $L_\lambda u$  и проинтегрируем дважды по частям. Используя условия согласования (4), (5), получим

$$\begin{aligned} 0 &= \int_{\Gamma} u(x)L_\lambda u(x) dx = \int_{\Gamma} p(x)(u'')^2(x) - \lambda\rho(x)u^2(x) dx - \sum_{a \in V(\Gamma)} \lambda\rho(a)u^2(a) - \\ &- \sum_{a \in V(\Gamma)} u(a) \left( \sum_{i \in I(a)} (p_i u''_i)'_\nu(a) - \lambda\rho(a)u \right) + \sum_{a \in V(\Gamma)} \sum_{\substack{i \in I(a) \\ \beta_i(a) \neq 0}} \frac{p_i(a)\vartheta_i(a)}{\beta_i(a)} (u'_i)^2(a) = \\ &= \int_{\Gamma} p(x)(u'')^2(x) - \lambda\rho(x)u^2(x) dx - \sum_{a \in V(\Gamma)} \lambda\rho(a)u^2(a) + \sum_{a \in V(\Gamma)} \sum_{\substack{i \in I(a) \\ \beta_i(a) \neq 0}} \frac{p_i(a)\vartheta_i(a)}{\beta_i(a)} (u'_i)^2(a). \end{aligned}$$

С учётом неотрицательности всех функций  $p$ ,  $\rho$  и  $\beta$ ,  $\vartheta$ , а также неравенства  $\lambda < 0$ , получим

$$0 \geq \int_{\Gamma} \rho(x)u^2(x) dx \geq 0.$$

Поскольку  $\rho(x) > 0$  на  $E(\Gamma)$ , то  $u \equiv 0$  на  $E(\Gamma)$ . Теперь для завершения доказательства леммы остаётся сослаться на непрерывность функции  $u$ .

Путём стандартных рассуждений (см., например, [6, 13, 20]) можно доказать следующий классический результат.

**Теорема 2.** Дифференциальный оператор  $L_\lambda$ , порождённый соотношениями (1), (4) и (5), является самосопряжённым. Спектр  $\Lambda$  оператора  $L_\lambda$  дискретен, т.е. существует неограниченная последовательность положительных собственных значений  $\Lambda = \{\lambda_k\}_{k=0}^\infty$  такая, что

$$0 \leq \lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_k \dots$$

Поскольку дифференциальный оператор  $L_\lambda$  является самосопряжённым, то кратность его собственных значений совпадает с размерностью соответствующего собственного подпространства. Поэтому вопрос о кратности точки спектра оператора  $L_\lambda$  удобно обсуждать в форме вопроса о размерности пространства всех решений краевой задачи (8). Обозначим это пространство через  $\mathcal{S}(L)$ . Очевидно, что  $\dim \mathcal{S}(L) \leq 4n$ , где  $n$  – число рёбер графа  $\Gamma$ . В контексте последующего изучения осцилляционных свойств спектра дифференциального оператора  $L_\lambda$  мы интересуемся условиями, обеспечивающими равенство

$$\dim \mathcal{S}(L) = 1, \tag{9}$$

что означает простоту соответствующей точки спектра  $\Lambda$ .

**Лемма 3.** Пусть  $u \in \mathcal{S}(L)$  и  $a \in V(\Gamma)$ . Если для некоторого индекса  $i \in I(a)$  выполнены равенства  $u(a) = (p_i u_i'')'(a) = 0$ , то функция  $u(x)$  равна тождественно нулю на соответствующем ребре  $\gamma_i$ .

**Доказательство.** Пусть  $\gamma_i = (a, b)$ . Тогда из условий (4) в вершине  $a$  и условий леммы следует, что значения  $u_i(a)$ ,  $u'_{i\nu}(a)$ ,  $u''_i(a)$  и  $(p_i u_i'')'_{\nu}(a)$  либо все неотрицательны, либо все неположительны, а из условий (4) в вершине  $b$  следует, что  $u'_{i\nu}(b)u''_i(b) \geq 0$ . В силу леммы 1  $u_i \equiv 0$ . Лемма доказана.

**Следствие 2.** Если  $\dim \mathcal{S}(L) \geq 2$ , то для всякой граничной вершины  $a \in \partial\Gamma$  существует нетривиальное решение задачи (8), тождественно равное нулю на примыкающем к  $a$  ребре.

**Следствие 3.** Если всякое нетривиальное решение задачи (8) имеет в  $\Gamma$  только изолированные нули, то справедливо равенство (9).

**Лемма 4.** Пусть  $u \in \mathcal{S}(L)$  и для некоторой вершины  $a \in J(\Gamma)$  выполнены равенства

$$u(a) = 0, \quad (p_i u_i'')'_{\nu}(a) = 0, \quad i \in I(a) \setminus i_0, \quad i_0 \in I(a). \tag{10}$$

Тогда  $u(x)$  равна тождественно нулю на всех ребрах графа, примыкающих к вершине  $a$ .

**Доказательство.** Из леммы 3 следует, что  $u_i \equiv 0$  для всех  $i \in I(a) \setminus i_0$ . Тогда из условия (5) в вершине  $a \in J(\Gamma)$  получаем

$$\sum_{i \in I(a)} (p_i u_i'')'_{\nu}(a) = (p_{i_0} u_{i_0}'')'_{\nu}(a) = 0.$$

Из леммы 3 имеем  $u_{i_0} \equiv 0$ . Лемма доказана.

**Следствие 4.** Пусть  $\Gamma$  является графом-цепочкой. Тогда  $\dim \mathcal{S}(L) = 1$ .

**Доказательство.** Пусть  $a \in \partial\Gamma$ . Тогда для любого нетривиального решения  $u \in \mathcal{S}(L)$  выполнено неравенство  $(pv'')'_{\nu}(a) \neq 0$ , иначе в противном случае из лемм 3 и 4 будет следовать тождество  $u \equiv 0$  на всей цепочке  $\Gamma$ .

Возьмём теперь два произвольных ненулевых решения  $u, v \in \mathcal{S}(L)$  и рассмотрим нетривиальную линейную комбинацию

$$w(x) = (pv'')'_{\nu}(a)u(x) - (pu'')'_{\nu}(a)v(x).$$

Очевидно, что  $w \in \mathcal{S}(L)$  и  $w(a) = 0$ ,  $(pw'')(a) = 0$ . Опять же, ввиду лемм 3 и 4, имеем  $w \equiv 0$  на  $\Gamma$ . Следствие доказано.

**Определение 2.** Точку  $x_0 \in \Gamma$  назовём *нетривиальным нулём* функции  $u \in C(\Gamma)$ , если  $u(x_0) = 0$ , а в любой окрестности точки  $x_0$  функция  $u$  не тождественна нулю.

**Следствие 5.** Пусть  $u \in \mathcal{S}(L)$ . Если вершина  $a \in J(\Gamma)$  является нетривиальным нулём функции  $u(x)$ , то  $u(x)$  не равна тождественно нулю по крайней мере на двух ребрах, примыкающих к вершине  $a$ .

Введём обозначение  $J_{3+}(\Gamma) = \{a \in J(\Gamma) : |I(a)| \geq 3\}$ .

**Лемма 5.** Пусть  $u(x)$  – нетривиальное решение (8), равное тождественно нулю на некотором ребре. Если  $|\partial\Gamma| \geq 3$ , то найдётся вершина  $a \in J_{3+}(\Gamma)$  такая, что  $u(a) = 0$ .

**Доказательство.** Пусть  $u \equiv 0$  на ребре  $\gamma_1$ . Так как  $u \not\equiv 0$  на  $\Gamma$ , то найдётся ребро  $\gamma_j \subset E(\Gamma)$  такое, что  $u_j(x) \not\equiv 0$ . Рассмотрим произвольную цепочку  $\Gamma(\gamma_1, \gamma_j)$  графа  $\Gamma$ , соединяющую ребра  $\gamma_1$  и  $\gamma_j$ . Без ограничения общности можно считать, что ребра и вершины графа занумерованы так, что цепочка  $\Gamma(\gamma_1, \gamma_j)$  состоит из рёбер  $\gamma_i = (a_i, a_{i+1})$ ,  $a_i \in V(\Gamma)$ , с индексами  $1 \leq i \leq j$ . Если  $a_2 \in J_{3+}(\Gamma)$ , то лемма доказана. В противном случае  $|I(a_2)| = 2$  и тождество  $u_1 \equiv 0$  вместе со следствием 5 дают  $u_2 \equiv 0$ . Поскольку  $u_j \not\equiv 0$ , то найдётся ребро  $\gamma_i \subset \Gamma(\gamma_1, \gamma_j)$  такое, что соответствующая концевая точка  $a_i$  этого ребра принадлежит множеству  $J_{3+}(\Gamma)$ . Лемма доказана.

**Определение 3.** Назовём точку  $x_0 \in \Gamma$  *простой*, если  $x_0 \notin J_{3+}(\Gamma)$  и множество  $\Gamma \setminus \{x_0\}$  не связно.

Отметим, что если  $\Gamma$  является деревом, то любая точка  $x \in \Gamma \setminus J_{3+}(\Gamma)$  является простой.

**Определение 4.** Будем говорить, что задача (8) *простая*, если у некоторого её решения все нули в  $\Gamma$  являются простыми. В частности, если существует решение задачи (8) без нулей в  $\Gamma$ , то такая задача также называется простой.

**Теорема 3.** Если задача (8) простая, то выполнено равенство (9).

**Доказательство.** С учётом следствия 4 нужно рассмотреть случай  $|\partial\Gamma| \geq 3$ . Пусть  $v \in \mathcal{S}(L)$  и  $v$  имеет нули только в простых точках  $\Gamma$ , либо вообще не имеет нулей в  $\Gamma$ . Число нулей  $v$  конечно. Действительно, иначе  $v_{i_0} \equiv 0$  на некотором ребре  $\gamma_{i_0} = (a_1, a_2)$  и, как следует из леммы 5, существует вершина  $a \in J_{3+}(\Gamma)$  такая, что  $v(a) = 0$ . Получим противоречие.

Через  $\{\Gamma_i\}_{i=1}^k$  обозначим множество всех  $S$ -зон функции  $v$ . Будем говорить, что  $S$ -зоны  $\Gamma_i$  и  $\Gamma_j$  ( $i \neq j$ ) смежны, если  $\partial\Gamma_i \cap \partial\Gamma_j \neq \emptyset$ . Поскольку все нули функции  $v$  в  $\Gamma$  простые, то для каждой пары смежных  $S$ -зон пересечение  $\partial\Gamma_i \cap \partial\Gamma_j$  является одноточечным. Отношение смежности на множестве  $\{\Gamma_i\}_{i=1}^k$  позволяет рассматривать это объединение  $S$ -зон как алгебраический граф:  $S$ -зоны являются вершинами, нули  $v$  являются рёбрами. Обозначим этот граф через  $S\Gamma$ . При этом мы допускаем, что  $S\Gamma$  является нуль-графом, т.е. не содержит рёбер (в случае, когда  $v$  не имеет нулей в  $\Gamma$ ). Легко видеть, что  $S\Gamma$  является связным и не имеет циклов.

Для дальнейшего доказательства теоремы введём следующую терминологию. Граничные вершины графа  $S\Gamma$  назовём *вершинами нулевого порядка*. Обозначим через  $S_1\Gamma$  алгебраический граф, полученный из  $S\Gamma$  удалением всех вершин степени 1 вместе с инцидентными рёбрами. Тогда граничные вершины подграфа  $S_1\Gamma$  назовём *вершинами первого порядка*. Обозначим через  $S_2\Gamma$  граф, полученный из  $S_1\Gamma$  удалением всех его вершин степени 1 вместе с инцидентными рёбрами. Граничные вершины подграфа  $S_2\Gamma$  называем *вершинами второго порядка* и т.д.

Пусть  $u(x)$  – произвольная нетривиальная функция из  $\mathcal{S}(L)$ , а  $\Gamma_{i_0}$  – произвольная вершина нулевого порядка графа  $S\Gamma$ . Тогда  $\partial\Gamma_{i_0} \setminus \partial\Gamma = x_1$ , где  $x_1$  – нуль функции  $v$ . Поэтому

$$u(x) = \beta(x)u''(x) - \vartheta(x)u'_\nu(x) = 0, \quad x \in \partial\Gamma_{i_0} \setminus x_1.$$

В силу теоремы 1 получаем  $u(x) = C_{i_0}v(x)$  при  $x \in \Gamma_{i_0}$ . Действительно, если решения  $u$  и  $v$  не пропорциональны на подграфе  $\Gamma_{i_0}$ , то согласно теореме 1 решение  $u \in \mathcal{S}(L)$  имеет в подграфе  $\Gamma_{i_0}$   $S$ -зону  $\tilde{\Gamma}_{i_0} \subset \Gamma_{i_0}$ , причём  $\tilde{\Gamma}_{i_0} \neq \Gamma_{i_0}$ . Но тогда, опять же в силу теоремы 1, функция  $v$  должна иметь нули в  $\tilde{\Gamma}_{i_0}$ . Получим противоречие.

Таким образом, если  $\Gamma_{i_0}$  – произвольная вершина нулевого порядка алгебраического графа  $S\Gamma$ , то  $u(x) = 0$  на  $\partial\Gamma_{i_0}$ .

Возьмём произвольную вершину первого порядка  $\Gamma_{i_1}$  графа  $S\Gamma$ . Легко видеть, что  $u(x) = 0$  для каждого  $x \in \partial\Gamma_{i_1} \setminus x_2$ , где  $x_2 = \partial\Gamma_{i_1} \cap E(S_1\Gamma)$ . В силу теоремы 1 получаем  $u(x) = C_{i_1}v(x)$  на  $\partial\Gamma_{i_1}$  и так далее. Таким образом, мы можем заключить, что на любой зоне знакопостоянства функции  $v$  решения  $v$  и  $u$  пропорциональны.

Поскольку  $u \not\equiv 0$ , то существует  $S$ -зона  $\Gamma_0$  функции  $v$  такая, что  $u = C_0v$  на  $\Gamma_0$  и  $C_0 \neq 0$ . Если других  $S$ -зон функция  $v$  не имеет, то теорема доказана. В противном случае рассмотрим функцию  $w(x) = u(x) - C_0v(x)$  и произвольную точку  $x_0 \in \partial\Gamma_0 \cap \Gamma$ . Очевидно, что  $w(x) \equiv 0$  на  $\Gamma_0$ , а точка  $x_0$  является простой. Из определения 3 и следствия 5 получаем  $w(x) \equiv 0$  в некоторой окрестности  $x_0$ . Отсюда следует, что  $w(x) \equiv 0$  на всех  $S$ -зонах функции  $v$ , смежных с  $\Gamma_0$ . Этот факт позволяет рассмотреть функцию  $w$  на каждой смежной с  $\Gamma_0$  зоной знакопостоянства функции  $v$  и провести те же рассуждения, что и на  $\Gamma_0$ .

Поскольку  $S\Gamma$  – конечный граф, мы заключаем, что  $w(x) \equiv 0$  на  $\Gamma$ . Теорема доказана.

**Следствие 6.** Пусть граф  $\Gamma$  является деревом. Если существует функция из  $\mathcal{S}(L)$ , не имеющая нулей в  $\Gamma$ , то  $\dim \mathcal{S}(L) = 1$ .

**4. Оценки кратности собственных значений.** Пусть  $\tilde{\Gamma}$  – подграф графа  $\Gamma$ . Множество функций из  $\mathcal{S}(L)$ , обращающихся в нуль на  $\Gamma \setminus \tilde{\Gamma}$ , обозначим через  $\mathcal{S}(L_{\tilde{\Gamma}})$ . Нетрудно видеть, что для любой функции  $u \in \mathcal{S}(L)$  её сужение на  $\tilde{\Gamma}$  является решением краевой задачи

$$Lu = 0, \quad x \in \tilde{\Gamma}, \quad u|_{\partial\tilde{\Gamma}} = (\beta u'' - \vartheta u'_\nu)|_{\partial\tilde{\Gamma}} = 0.$$

**Теорема 4.** Пусть граф  $\Gamma$  является деревом. Тогда  $\dim \mathcal{S}(L) \leq d - 1$ , где  $d$  – число граничных вершин графа  $\Gamma$ .

**Доказательство.** При  $d = 2$  доказываемое утверждение следует из следствия 4.

Пусть  $d \geq 3$  и  $a_1, a_2, \dots, a_d$  – все граничные вершины графа  $\Gamma$ . Предположив противное, можно фиксировать линейно независимый набор  $\{u^i(x)\}_{i=1}^d$  функций из  $\mathcal{S}(L)$ . Рассмотрим вершину  $a_1 \in \partial\Gamma$ . Поскольку  $d \geq 3$ , то в силу леммы 5 вершина  $a_1$  обязательно является концевой вершиной некоторой цепочки  $\Gamma(a_1, b_1)$  (возможно, состоящей из одного ребра), промежуточные вершины которой имеют кратность два в исходном графе  $\Gamma$ , соединяющей  $a_1$  с некоторой вершиной  $b_1 \in J_{3+}(\Gamma)$ . Рассуждая как и в доказательстве следствия 4, линейными комбинациями из функций  $\{u^i(x)\}_{i=1}^d$  можно построить  $d - 1$  линейно независимых решений  $\{v^i(x)\}_{i=1}^{d-1}$ , обращающихся в нуль на цепочке  $\Gamma(a_1, b_1)$ .

Положим  $\Gamma_1 = \Gamma \setminus (\Gamma(a_1, b_1) \cup a_1)$ . Очевидно, что  $\partial\Gamma_1 = \{a_2, a_3, \dots, a_d\}$ , а все функции  $v^i$  принадлежат  $\mathcal{S}(L_{\Gamma_1})$ . Теперь можно рассмотреть вершину  $a_2$  и повторить предыдущие рассуждения. Продолжая этот процесс, мы придём к графу-цепочке  $\Gamma_{d-1} \subset \Gamma$ , соединяющему  $a_{d-1}$  с  $a_d$ , и двум решениям  $w^1, w^2 \in \mathcal{S}(L_{\Gamma_{d-1}})$ . Из следствия 4 легко следует, что функции  $w^1$  и  $w^2$  линейно зависимы, что противоречит сделанному выше предположению о линейной независимости набора  $\{u^i(x)\}_{i=1}^d$ . Теорема доказана.

Следующий пример показывает, что полученная оценка является точной.

**Пример.** Рассмотрим граф  $\Gamma$ , состоящий из трёх рёбер  $\gamma_i = (a_i, b)$ ,  $i \in \{1, 2, 3\}$ , с общим концом  $b$  (внутренняя вершина графа). Предполагая, что ребра  $\gamma_i$  направлены в сторону от граничных вершин  $a_i$ , мы рассмотрим спектральную краевую задачу на  $\Gamma$ :

$$\begin{aligned} u_i^{IV} - \lambda u_i &= 0, \quad x \in \gamma_i, \\ u &\in C(\Gamma) \cap C^4[\Gamma], \quad u_i''(b) = 0, \quad u_{2\nu}'''(b) + u_{2\nu}'''(b) + u_{3\nu}'''(b) = 0; \\ u(a_i) &= u'(a_i) = 0, \quad a_i \in \partial\Gamma. \end{aligned}$$

Пусть все три ребра графа имеют единичную длину. Используя метрическую функцию  $\mu = \mu(x)$ , построим характеристический определитель краевой задачи:

$$\Delta(\lambda) = 8(1 + \cos \sqrt[4]{\lambda} \operatorname{ch} \sqrt[4]{\lambda}) \sin^2 \sqrt[4]{\lambda} \operatorname{sh}^2 \sqrt[4]{\lambda}.$$

Нули  $\Delta(\lambda)$  являются собственными значениями краевой задачи. Мы имеем две последовательности собственных значений:

$$\lambda_k^{(1)} = \left( \pi k - \frac{\pi}{2} + o(k^{-1}) \right)^4, \quad \lambda_k^{(2)} = (\pi k)^4, \quad k \in \mathbb{N}.$$

Кратность каждого собственного значения  $\lambda_k^{(1)}$  равна единице, а каждое из  $\lambda_k^{(2)}$  имеет кратность два.

**Лемма 6.** Пусть  $x_0 \in \Gamma$  и существует функция из  $\mathcal{S}(L)$ , не равная нулю в точке  $x_0$ . Тогда размерность пространства решений из  $\mathcal{S}(L)$ , которые обращаются в нуль в точке  $x_0$  равна  $\dim \mathcal{S}(L) - 1$ .

Доказательство леммы 6 следует из того факта, что множество  $\{u \in \mathcal{S}(L) : u(x_0) = 0\}$  является гиперплоскостью, порождаемой линейным функционалом  $l(u) = u(x_0)$  на конечномерном пространстве  $\mathcal{S}(L)$ .

**Лемма 7.** Для любой вершины  $a \in J(\Gamma)$  размерность пространства решений из  $\mathcal{S}(L)$ , которые обращаются в нуль на всех ребрах, примыкающих к  $a$ , не меньше  $\dim \mathcal{S}(L) - |I(a)|$ .

**Доказательство.** Содержательную часть утверждения леммы несет случай  $\dim \mathcal{S}(L) > |I(a)|$ . Обозначая  $m = \dim \mathcal{S}(L)$  и привлекая лемму 6, мы всегда можем выбрать  $m - 1$  линейно независимых решений из  $\mathcal{S}(L)$ , каждое из которых равно нулю в вершине  $a$ . Далее, переходя, как обычно, к линейным комбинациям, из этих функций можно построить  $m - |I(a)|$  линейно независимых решений, удовлетворяющих условиям (10), с последующим привлечением леммы 4. Для этого опять можно применить рассуждение из следствия 4. Лемма доказана.

Лемма 7 позволяет перенести основной результат работы [12, теорема 3], относящийся к задаче Штурма–Лиувилля на сети, на краевую задачу четвёртого порядка (1). Прежде чем сформулировать соответствующий результат, напомним некоторые определения из [12].

Пусть теперь  $\Gamma$  – произвольный граф. Вершину  $a \in J(\Gamma)$ , лежащую на каком-либо цикле, назовём *разбивающей*, если  $\Gamma \setminus a$  не связно. Пусть  $\Gamma_i(a)$  – одна из компонент связности множества  $\Gamma \setminus a$ . Множество  $\Gamma_i = \Gamma_i(a) \cup a$  является подграфом графа  $\Gamma$ , который также может иметь циклы. Он также может разбиваться на связные компоненты некоторыми вершинами, лежащими на его циклах. Во избежание недоразумений подчеркнём, что выбрасываются только вершины, лежащие на циклах. После такого разбиения снова образуем подграфы типа  $\Gamma_i$ , присоединяя к ним вершины, с помощью которых они разбивались. Продолжая этот процесс и далее, мы получаем некоторое количество подграфов. Каждый из них является либо деревом, либо связным объединением циклов, не допускающим дальнейшего разбиения описанного выше типа. Такое объединение циклов назовём *гнездом*.

Обозначим через  $\zeta(\Gamma)$  количество гнезд графа, содержащих ровно по одной разбивающей вершине, а через  $\eta(\Gamma)$  – цикломатическое число графа  $\Gamma$ .

**Теорема 5.** Кратность любого собственного значения краевой задачи (1) не превосходит величины  $|\partial\Gamma| + \zeta(\Gamma) + \eta(\Gamma) - 1$ .

Доказательство теоремы 5 почти дословно повторяет доказательство соответствующего результата из работы [12], поскольку для рассматриваемой задачи (1) лемма 7 обеспечивает выполнение условий теоремы 3 из статьи [12].

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации (соглашение 075-02-2022-890).

## СПИСОК ЛИТЕРАТУРЫ

1. Боровских А.В., Мустафокулов Р., Лазарев К.П., Покорный Ю.В. Об одном классе дифференциальных уравнений четвёртого порядка на пространственной сети // Докл. РАН. 1995. Т. 345. № 6. С. 730–732.
2. Кулаев Р.Ч. О функции Грина краевой задачи на графе-пучке // Изв. вузов. Математика. 2013. № 2. С. 56–66.
3. Кулаев Р.Ч. Неосциллятия уравнения четвертого порядка на графе // Мат. сб. 2015. Т. 206. № 12. С. 79–118.



4. *Покорный Ю.В., Мустафокулов Р.* О положительности функции Грина линейных краевых задач для уравнений четвёртого порядка на графе // Изв. вузов. Математика. 1999. № 2. С. 75–82.
5. *Borovskikh A.V., Lazarev K.P.* Fourth-order differential equations on geometric graphs // J. of Math. Sci. 2004. V. 119. № 6. P. 719–738.
6. *Mercier D., Régnier V.* Control of a network of Euler–Bernoulli beams // J. Math. Anal. Appl. 2008. V. 342. № 2. P. 874–894.
7. *Xu Gen Qi, Mastorakis N.E.* Differential Equations on Metric Graph. Boston, 2010.
8. *Lubary J.A.* On the geometric and algebraic multiplicities for eigenvalue problems on graphs // Lect. Not. in Pure and Appl. Math. V. 219. New York, 2001. P. 135–146.
9. *Покорный Ю.В., Пенкин О.М., Прядиев В.Л., Боровских А.В., Лазарев К.П., Шабров С.А.* Дифференциальные уравнения на геометрических графах. М., 2007.
10. *Kuchment P.* Quantum graphs. I. Some basic structures // Waves in Random Media. 2004. V. 14. № 1. P. 107–128.
11. *Kuchment P.* Quantum graphs. II. Some spectral properties of quantum and combinatouial graphs // J. Phys. A. 2005. V. 38. № 22. P. 4887–4900.
12. *Диаб А.Т., Калдыбекова Б.К., Пенкин О.М.* О кратности собственных значений в задаче Штурма–Лиувилля на графах // Мат. заметки. 2016. Т. 99. № 4. С. 489–501.
13. *Кулаев Р.Ч.* Необходимое и достаточное условия положительности функции Грина для уравнения четвёртого порядка на графе // Дифференц. уравнения. 2015. Т. 51. № 3. С. 302–316.
14. *Кулаев Р.Ч.* О свойстве неосцилляции уравнения на графе // Сиб. мат. журн. 2016. Т. 57. № 1. С. 85–97.
15. *Кулаев Р.Ч.* К вопросу о неосцилляции уравнения на графе // Дифференц. уравнения. 2014. Т. 50. № 11. С. 1563–1565.
16. *Кулаев Р.Ч., Уртаева А.А.* Теоремы Штурма о распределении нулей для уравнения четвёртого порядка на графе // Мат. заметки. 2022. Т. 111. № 6. С. 947–952.
17. *Kulaev R.Ch.* The qualitative theory of fourth-order differential equations on a graph // Mediterr. J. Math. 2022. V. 19. № 73. P. 1–15.
18. *Timoshenko S.P., Young D.H., Weaver Jr.W.* Vibration Problems in Engineering. Chichester, 1990.
19. *Leighton W., Nehari Z.* On the oscillation of solutions of self-adjoint linear differential equations of the fourth order // Trans. Amer. Math. Soc. 1958. V. 89. P. 325–377.
20. *Castro C., Zuazua E.* Exact boundary controllability of two Euler–Bernoulli beams connected by a point mass // Math. Comput. Model. 2000. V. 32. P. 955–969.

Южный математический институт –  
филиал Владикавказского научного центра РАН,  
Северо-Осетинский государственный университет  
имени К.Л. Хетагурова, г. Владикавказ

Поступила в редакцию 14.03.2022 г.  
После доработки 14.03.2022 г.  
Принята к публикации 25.05.2022 г.

УДК 517.95

## ОБРАТНАЯ ЗАДАЧА С ПЕРЕОПРЕДЕЛЕНИЕМ ТРЕТЬЕГО РОДА ДЛЯ АБСТРАКТНОГО ДИФФЕРЕНЦИАЛЬНОГО УРАВНЕНИЯ ВТОРОГО ПОРЯДКА

© 2022 г. И. В. Тихонов, М. Алмохамед

В банаховом пространстве рассматривается линейная обратная задача для абстрактного дифференциального уравнения второго порядка. Стационарное неоднородное слагаемое в уравнении предполагается неизвестным. В начальный момент времени заданы стандартные условия Коши. К ним добавляется дополнительное условие – финальное переопределение третьего рода. Оно содержит комбинацию значений эволюционной функции и её производной в выбранный финальный момент времени. Для изучаемой задачи установлен критерий единственности решения, выраженный в спектральных терминах – через нули характеристической целой функции. Результат носит универсальный характер и не требует ограничений на тип дифференциального уравнения. Отдельно обсуждается вопрос о распределении нулей характеристической функции. Посредством проведённого анализа получен ряд эффективных достаточных признаков единственности (и неединственности) решения. Все они просты и удобны на практике.

DOI: 10.31857/S0374064122070032, EDN: CECRGN

**1. Постановка задачи.** Пусть  $E$  – комплексное банахово пространство и  $A$  – линейный замкнутый оператор в  $E$  с областью определения  $D(A) \subset E$  (не обязательно плотной в  $E$ ). Зафиксируем вещественное число  $T > 0$ . На интервале  $(0, T)$  рассмотрим абстрактное дифференциальное уравнение второго порядка

$$u''(t) = Au(t) + g, \quad 0 < t < T, \quad (1)$$

с неизвестным элементом  $g \in E$ .

Для одновременного нахождения функции  $u: [0, T] \rightarrow E$  и элемента  $g$  добавим условия Коши

$$u(0) = u_0, \quad u'(0) = u_1 \quad (2)$$

и специальное переопределение

$$\alpha u(T) + \beta u'(T) = u_2. \quad (3)$$

Элементы  $u_0, u_1, u_2 \in E$  и числовые значения  $\alpha, \beta \in \mathbb{C} \setminus \{0\}$  считаем заданными.

Условие (3) будем называть *финальным переопределением третьего рода*. Задача (1)–(3) относится к классу *обратных задач* (см. [1, 2]). Модельная форма (1)–(3), простая по виду, но общая по характеру, позволяет охватить единым методом широкий круг примеров из математической физики. Главное, чтобы изучаемое дифференциальное уравнение имело второй порядок по выделенной переменной  $t$ , неформально полагаемой *временем*. Цель нашей работы – получить критерий единственности решения обратной задачи (1)–(3) без ограничений на оператор  $A$ , кроме отмеченных *линейности* и *замкнутости*. Поясним понятие решения.

Пару  $(u(t), g)$  назовём *ослабленным решением* обратной задачи (1)–(3), если

$$u \in C^2((0, T), E) \cap C^1([0, T], E), \quad u(t) \in D(A) \quad \text{при} \quad 0 < t < T, \quad g \in E, \quad (4)$$

и все соотношения (1)–(3) выполнены. При этом автоматически  $Au \in C((0, T), E)$ .

Повышая требования, будем считать ослабленное решение  $(u(t), g)$  *классическим*, если

$$u \in C^2([0, T], E), \quad u(t) \in D(A) \quad \text{при} \quad 0 \leq t \leq T, \quad g \in E. \quad (5)$$

При этом уравнение (1) должно выполняться на отрезке  $[0, T]$ , а в соотношениях (2) для согласования условий естественно выбирать элемент  $u_0$  из  $D(A)$ .

В основном мы будем рассматривать именно ослабленные решения (4), называя их далее просто *решениями*. Тогда, установив критерий единственности для ослабленных решений (4), получим результат, пригодный и для классических решений (5).

По-видимому, поставленная обратная задача ещё не изучена с должной подробностью. Нам известны лишь прежние результаты [2; п. 8.3], где для уравнений “эллиптического типа” с позитивным оператором  $A$  (в смысле П.Е. Соболевского, см. [3; с. 275]) обсуждался аналог обратной задачи (1)–(3) со значениями  $\alpha, \beta \in \mathbb{R}$ ,  $\alpha^2 + \beta^2 \neq 0$ .

Отметим ещё, что специальные случаи  $\alpha \neq 0$ ,  $\beta = 0$  (*финальное переопределение первого рода*) и  $\alpha = 0$ ,  $\beta \neq 0$  (*финальное переопределение второго рода*) более известны, и общие критерии единственности для них получены в работах [4] и [5] соответственно. Поэтому далее, не снижая общности, ограничимся предположением  $\alpha \neq 0$ ,  $\beta \neq 0$ , как и было указано при постановке задачи. Точнее, наша постановка задачи (1)–(3) зависит от трёх числовых параметров  $\alpha, \beta \in \mathbb{C} \setminus \{0\}$  и  $T > 0$ . Как мы увидим в дальнейшем, картина с единственностью решения по-сути обусловлена одним “безразмерным” параметром  $p \equiv 2\beta/(\alpha T)$ .

Завершая этот вводный раздел, отметим роль А.И. Прилепко, интерес и настойчивость которого во многом сформировали подходы к изучению подобных обратных задач. Первоначальный импульс развитию дал, возможно, обзор [6], где был намечен переход от обратных задач теории потенциала к некоторым идейно близким обратным задачам для эволюционных уравнений (см. также [7]). Кроме того, выделим работы [8–12], идейно повлиявшие на наше исследование и посвящённые различным неклассическим задачам с теми или иными финальными условиями (см. также [13–15]).

Теперь, используя методику работы [4] и некоторые факты из теории целых функций, установим универсальный критерий единственности решения для обратной задачи (1)–(3), действующий без ограничений на тип дифференциального уравнения (1).

**2. Однородная обратная задача.** Допустим, что обратная задача (1)–(3) с некоторыми элементами  $u_0, u_1, u_2 \in E$  разрешима. Поставим вопрос о единственности её ослабленного решения. Он очевидно сводится к вопросу об отсутствии нетривиальных решений у однородной обратной задачи для уравнения (1) с условиями

$$u(0) = 0, \quad u'(0) = 0, \quad \alpha u(T) + \beta u'(T) = 0. \quad (6)$$

Как обычно, *тривиальным решением* однородной задачи (1), (6) считаем пару

$$u(t) \equiv 0, \quad g = 0. \quad (7)$$

Все другие решения задачи (1), (6) (если они есть) называем тогда *нетривиальными\**.

Укажем простой способ, позволяющий строить нетривиальные решения однородной обратной задачи (1), (6). Рассмотрим её “операционный аналог”, состоящий из скалярной задачи Коши

$$y''(t) = \lambda y(t) + 1, \quad 0 \leq t \leq T, \quad (8)$$

$$y(0) = 0, \quad y'(0) = 0 \quad (9)$$

и финального условия

$$\alpha y(T) + \beta y'(T) = 0. \quad (10)$$

Здесь  $\lambda$  – спектральный параметр. Требуется найти значения  $\lambda \in \mathbb{C}$ , при которых в скалярной задаче (8)–(10) существуют решения  $y \in C^2[0, T]$ .

\*) Очевидно, что у таких нетривиальных решений  $(u(t), g)$  функция  $u(t)$  не должна обращаться в тождественный нуль на  $[0, T]$ , в то время как возможность  $g = 0$  изначально исключать нельзя. Эта возможность действительно реализуется в некоторых специальных ситуациях, когда собственные значения оператора  $A$  заполняют всю комплексную плоскость, и не реализуется, если хотя бы одна точка из  $\mathbb{C}$  не является собственным значением оператора  $A$  (подробнее см. [16]).

Нетрудно понять, что каждому такому  $\lambda$  отвечает лишь одно возможное решение  $y(t)$  (причём последнее не будет тождественным нулём). Этим спектральная задача (8)–(10) отличается от привычных спектральных задач типа Штурма–Лиувилля, где есть естественное свойство линейности, и собственные функции можно умножать на ненулевые константы.

Анализ спектральной задачи (8)–(10) не сложен. Решение задачи Коши (8), (9) может быть записано в виде

$$y(t) = y(t, \lambda) = \frac{\operatorname{ch}(\sqrt{\lambda}t) - 1}{\lambda} \quad (y(t) = t^2/2 \text{ при } \lambda = 0). \tag{11}$$

При этом

$$y'(t) = y'_t(t, \lambda) = \frac{\operatorname{sh}(\sqrt{\lambda}t)}{\sqrt{\lambda}} \quad (y'(t) = t \text{ при } \lambda = 0). \tag{12}$$

Последующая подстановка выражений (11), (12) в финальное условие (10) даёт уравнение

$$\alpha \frac{\operatorname{ch}(\sqrt{\lambda}T) - 1}{\lambda} + \beta \frac{\operatorname{sh}(\sqrt{\lambda}T)}{\sqrt{\lambda}} = 0. \tag{13}$$

Именно корни  $\lambda \in \mathbb{C}$  уравнения (13) служат спектральными значениями задачи (8)–(10).

Поясним одно обстоятельство. Функции (11), (12) представимы степенными рядами

$$\frac{\operatorname{ch}(\sqrt{\lambda}t) - 1}{\lambda} = \frac{t^2}{2!} + \lambda \frac{t^4}{4!} + \lambda^2 \frac{t^6}{6!} + \dots + \lambda^n \frac{t^{2n+2}}{(2n+2)!} + \dots, \tag{14}$$

$$\frac{\operatorname{sh}(\sqrt{\lambda}t)}{\sqrt{\lambda}} = \frac{t}{1!} + \lambda \frac{t^3}{3!} + \lambda^2 \frac{t^5}{5!} + \dots + \lambda^n \frac{t^{2n+1}}{(2n+1)!} + \dots, \tag{15}$$

т.е. они являются *целыми* как по переменной  $t$ , так и по параметру  $\lambda$ . Тем самым никакой “многозначности”, связанной с присутствием  $\sqrt{\lambda}$ , в формулах (11)–(13) не возникает. Отметим также связь разложений (14), (15) с теорией так называемых *обобщённых экспонент* и теорией *функций типа Миттаг-Лёффлера* (подробнее см. [17]; см. также [18; с. 33]).

Для последующего использования удобно ввести целые функции<sup>\*)</sup>

$$L_1(\lambda) = L_1(\lambda; T) \equiv \frac{\operatorname{ch}(\sqrt{\lambda}T) - 1}{\lambda}, \quad L_2(\lambda) = L_2(\lambda; T) \equiv \frac{\operatorname{sh}(\sqrt{\lambda}T)}{\sqrt{\lambda}} \tag{16}$$

переменной  $\lambda \in \mathbb{C}$ . Составим из функций (16) следующую *характеристическую функцию* изучаемой обратной задачи:

$$L(\lambda) = L(\lambda; T, \alpha, \beta) \equiv \alpha L_1(\lambda) + \beta L_2(\lambda), \quad \lambda \in \mathbb{C}. \tag{17}$$

Здесь числа  $\alpha, \beta$  те же, что в условии (3) (а также в последнем из соотношений (6) и в формуле (10)). Нули функции (17) совпадают с корнями уравнения (13) и со спектральными значениями  $\lambda$  скалярной задачи (8)–(10).

Поскольку  $L(\lambda)$  есть целая функция нецелого порядка  $\rho = 1/2$ , то множество нулей

$$\Lambda = \Lambda(T, \alpha, \beta) \equiv \{\lambda \in \mathbb{C} : L(\lambda) = 0\} \tag{18}$$

должно быть бесконечным (см. [19; с. 38–41]). Ясно, что это счётное множество с единственной предельной точкой на бесконечности. Структура множества  $\Lambda$  и вопрос о правильной индексации нулей будут рассмотрены отдельно, в пп. 8, 9 ниже. Там, в частности, отмечено, что

---

<sup>\*)</sup> Это элементарные целые функции порядка  $\rho = 1/2$  по переменной  $\lambda \in \mathbb{C}$ . Все необходимые сведения из теории целых функций см. в [19, 20].

нули функции (17) являются, как правило, простыми, и особых проблем с кратностью точек в множестве (18) не возникает.

Выберем фиксированное значение  $\lambda$  из множества (18), обозначив его временно через  $\lambda_0$ . Поскольку это некоторый нуль функции  $L(\lambda)$  и, автоматически, спектральное значение скалярной задачи (8)–(10), то при  $\lambda = \lambda_0$  задача (8)–(10) имеет решение

$$y(t) = y(t, \lambda_0) = \frac{\operatorname{ch}(\sqrt{\lambda_0}t) - 1}{\lambda_0} \equiv \frac{t^2}{2!} + \lambda_0 \frac{t^4}{4!} + \dots + \lambda_0^n \frac{t^{2n+2}}{(2n+2)!} + \dots, \quad 0 \leq t \leq T.$$

Допустим, что то же число  $\lambda_0 \in \Lambda$  оказалось собственным значением оператора  $A$  из уравнения (1), т.е.  $Af_0 = \lambda_0 f_0$  с собственным вектором  $f_0 \in D(A)$ ,  $f_0 \neq 0$ . Тогда пара

$$u(t) = y(t, \lambda_0)f_0 = \frac{\operatorname{ch}(\sqrt{\lambda_0}t) - 1}{\lambda_0} f_0, \quad g = f_0 \quad (19)$$

удовлетворяет всем соотношениям (1), (6), образуя нетривиальное решение этой задачи. Такие решения (если они есть) будем называть *элементарными решениями* однородной обратной задачи (1), (6). Зафиксируем точное утверждение.

**Лемма 1.** Пусть некоторый нуль  $\lambda_0$  характеристической функции (17) является собственным значением оператора  $A$  с собственным вектором  $f_0 \in D(A)$ ,  $f_0 \neq 0$ . Тогда однородная обратная задача (1), (6) имеет нетривиальное элементарное решение вида (19). Точнее, подобных решений будет бесконечно много, так как собственный вектор  $f_0$  можно умножить на любую ненулевую константу.

Лемма проверяется непосредственно. Понятно, что в её условиях единственность решения обратной задачи будет нарушаться, поскольку в однородной версии (1), (6) появятся нетривиальные решения вида (19). Эти решения оказываются даже классическими, так как для них требования (5) очевидно выполнены и уравнение (1) действует на всём отрезке  $[0, T]$ .

Итак, для единственности решения в обратной задаче необходимо, чтобы ни один нуль характеристической функции  $L(\lambda)$  из формулы (17) не совпадал с собственным значением оператора  $A$ . Нарботанный прежде опыт (см. [4, 5]) подсказывает, что такое условие может быть ещё и достаточным, и результат – критерий единственности решения – не исключён в максимальной общности. Покажем, что эта надежда вполне обоснована.

**3. Формулировка главного результата.** Рассматриваем задачу (1)–(3) в исходных предположениях из п. 1. Главный результат работы состоит в следующем.

**Теорема 1.** Пусть  $A$  – линейный замкнутый оператор в  $E$ . Предположим, что при некоторых  $u_0, u_1, u_2 \in E$  обратная задача (1)–(3) имеет ослабленное решение  $(u(t), g)$ . Для того чтобы это решение было единственным необходимо и достаточно, чтобы ни один нуль характеристической функции (17) не являлся собственным значением оператора  $A$ .

При выборе  $u_0 = u_1 = u_2 = 0$  теорема 1 приобретает следующий специальный вид.

**Теорема 2.** Пусть  $A$  – линейный замкнутый оператор в  $E$ . Для того чтобы однородная обратная задача (1), (6) имела только тривиальное решение (7) и не имела других ослабленных решений необходимо и достаточно, чтобы ни один нуль характеристической функции (17) не являлся собственным значением оператора  $A$ .

Ввиду очевидной линейности изучаемой обратной задачи теоремы 1 и 2 эквивалентны, и можно ограничиться доказательством только теоремы 2. При этом необходимость в ней получается ссылкой на лемму 1 и элементарные решения вида (19). Итак, зафиксируем логику: нужный критерий будет установлен, если обосновать достаточность условия из теоремы 2 для единственности решения однородной обратной задачи (1), (6).

Последующее довольно объёмное доказательство существенно использует соображения из теории целых функций. Чтобы не прерывать изложение, удобно сразу обсудить несколько технических фактов, связанных с функциями  $L_1(\lambda)$ ,  $L_2(\lambda)$  из формулы (16) и с образованной из них характеристической функцией  $L(\lambda)$  вида (17).

**4. Элементарные целые функции и их поведение на мнимой оси.** Рассматриваем целые функции  $L_1(\lambda)$ ,  $L_2(\lambda)$  из формулы (16). Нам нужно охарактеризовать их поведение

на мнимой оси. Заметим, что обе эти функции являются *вещественными* в том смысле, что их тейлоровские разложения по переменной  $\lambda$  содержат только вещественные коэффициенты (см. формулы (14), (15), взятые при  $t = T > 0$ ). Поэтому для каждой из функций значения в комплексно сопряжённых точках плоскости будут комплексно сопряжены. Как следствие имеем  $|L_1(+i\tau)| = |L_1(-i\tau)|$  и  $|L_2(+i\tau)| = |L_2(-i\tau)|$  в симметричных точках  $\pm i\tau$  на мнимой оси. Здесь и всюду далее обозначаем через  $i$  мнимую единицу.

Введём характеристики

$$V_1(\tau) \equiv |L_1(i\tau)|, \quad V_2(\tau) \equiv |L_2(i\tau)|, \tag{20}$$

зависящие от переменной  $\tau \in \mathbb{R}$  или, точнее, от  $|\tau| \geq 0$ .

Прямой подсчёт показывает, что

$$V_1(\tau) \equiv \left| \frac{\operatorname{ch}(\sqrt{i\tau}T) - 1}{i\tau} \right| = \frac{\operatorname{ch}(\sqrt{|\tau|/2T}) - \cos(\sqrt{|\tau|/2T})}{|\tau|}, \tag{21}$$

$$V_2(\tau) \equiv \left| \frac{\operatorname{sh}(\sqrt{i\tau}T)}{\sqrt{i\tau}} \right| = \sqrt{\frac{\operatorname{ch}^2(\sqrt{|\tau|/2T}) - \cos^2(\sqrt{|\tau|/2T})}{|\tau|}}. \tag{22}$$

Затем, используя формулы  $\operatorname{ch}^2 a = (1 + \operatorname{ch} 2a)/2$  и  $\cos^2 a = (1 + \cos 2a)/2$ , заметим связь

$$V_2(\tau) = \sqrt{\frac{\operatorname{ch}(\sqrt{2|\tau|}T) - \cos(\sqrt{2|\tau|}T)}{2|\tau|}} = \sqrt{2V_1(4\tau)}. \tag{23}$$

Отметим также степенное разложение

$$V_1(\tau) = \frac{T^2}{2!} + \frac{\tau^2 T^6}{2^2 6!} + \frac{\tau^4 T^{10}}{2^4 10!} + \dots = \sum_{n=0}^{\infty} \frac{\tau^{2n} T^{4n+2}}{2^{2n} (4n+2)!}, \quad \tau^{2n} = |\tau|^{2n}, \tag{24}$$

элементарно получаемое из (21). Перечисленные соотношения дают такой результат.

**Лемма 2.** Пусть функции  $L_1(\lambda)$ ,  $L_2(\lambda)$  определены по формулам (16). Тогда их характеристики  $V_1(\tau)$ ,  $V_2(\tau)$  вида (20) являются строго положительными при  $\tau \in \mathbb{R}$  и строго возрастающими при возрастании  $|\tau| \geq 0$ . При  $|\tau| \rightarrow \infty$  верны асимптотики

$$V_1(\tau) \sim \frac{\operatorname{ch}(\sqrt{|\tau|/2T})}{|\tau|}, \quad V_2(\tau) \sim \frac{\operatorname{ch}(\sqrt{|\tau|/2T})}{|\tau|^{1/2}}, \tag{25}$$

откуда  $V_1(\tau) = o(V_2(\tau))$  при  $|\tau| \rightarrow \infty$ .

**Доказательство.** То, что функция  $V_1(\tau)$  строго положительна при  $\tau \in \mathbb{R}$  и строго возрастает при возрастании  $|\tau|$ , очевидно следует из формулы (24). Затем, глядя на формулу (23), делаем аналогичные выводы относительно  $V_2(\tau)$ . И, наконец, асимптотики (25), действующие при  $|\tau| \rightarrow \infty$  вместе с соотношением  $V_1(\tau) = o(V_2(\tau))$ , без труда выводятся из явных представлений (21), (22). Лемма доказана.

Обратимся теперь к характеристической функции  $L(\lambda) = \alpha L_1(\lambda) + \beta L_2(\lambda)$ , построенной по правилу (17) с фиксированными значениями  $\alpha, \beta \in \mathbb{C} \setminus \{0\}$ . По аналогии с (20) введём величину

$$V(\tau) \equiv |L(i\tau)| = |\alpha L_1(i\tau) + \beta L_2(i\tau)|, \quad \tau \in \mathbb{R}. \tag{26}$$

Так как коэффициенты  $\alpha$ ,  $\beta$  могут быть не вещественными, то утверждать, что функция  $V(\tau)$  зависит лишь от  $|\tau|$  здесь уже, конечно, нельзя. Кроме того, неотрицательная функция  $V(\tau)$  уже не обязательно строго положительна при  $\tau \in \mathbb{R}$  (ибо какие-то нули функции  $L(\lambda)$  могут оказаться на мнимой оси). Но поведение  $V(\tau)$  при  $|\tau| \rightarrow \infty$  легко охарактеризовать.

Действительно, имеем оценку сверху

$$V(\tau) = |\alpha L_1(i\tau) + \beta L_2(i\tau)| \leq |\alpha| |L_1(i\tau)| + |\beta| |L_2(i\tau)| = |\alpha| V_1(\tau) + |\beta| V_2(\tau), \quad \tau \in \mathbb{R},$$

и соответствующую оценку снизу

$$V(\tau) = |\alpha L_1(i\tau) + \beta L_2(i\tau)| \geq |\beta| |L_2(i\tau)| - |\alpha| |L_1(i\tau)| = |\beta| V_2(\tau) - |\alpha| V_1(\tau), \quad \tau \in \mathbb{R}.$$

Отсюда получаем, что

$$1 - \frac{|\alpha| V_1(\tau)}{|\beta| V_2(\tau)} \leq \frac{V(\tau)}{|\beta| V_2(\tau)} \leq 1 + \frac{|\alpha| V_1(\tau)}{|\beta| V_2(\tau)}, \quad \tau \in \mathbb{R}, \tag{27}$$

где  $V_1(\tau) = o(V_2(\tau))$  при  $|\tau| \rightarrow \infty$  согласно лемме 2. Поэтому, переходя к пределу в (27) и используя вторую асимптотику (25), замечаем, что

$$V(\tau) \equiv |L(i\tau)| = |\alpha L_1(i\tau) + \beta L_2(i\tau)| \sim |\beta| V_2(\tau) = |\beta| |L_2(i\tau)| \sim |\beta| \frac{\text{ch}(\sqrt{|\tau|/2T})}{|\tau|^{1/2}} \tag{28}$$

при  $|\tau| \rightarrow \infty$ . Зафиксируем нужный результат в следующей форме.

**Лемма 3.** Пусть функции  $L_1(\lambda)$ ,  $L_2(\lambda)$  определены по формуле (16) и неотрицательная величина  $V(\tau)$  определена по формуле (26) с фиксированными значениями  $\alpha, \beta \in \mathbb{C} \setminus \{0\}$ . Тогда  $V(\tau)$  строго положительна при достаточно больших  $|\tau| > 0$ , и верно соотношение

$$\lim_{|\tau| \rightarrow \infty} \frac{|\alpha| V_1(\tau) + |\beta| V_2(\tau)}{V(\tau)} = \lim_{|\tau| \rightarrow \infty} \frac{|\alpha| |L_1(i\tau)| + |\beta| |L_2(i\tau)|}{|\alpha L_1(i\tau) + \beta L_2(i\tau)|} = 1. \tag{29}$$

**Доказательство.** Строгая положительность  $V(\tau)$  при достаточно больших  $|\tau| > 0$  следует из асимптотики (28). Для вычисления предела (29) достаточно заметить, что числитель и знаменатель фигурирующей там дроби образуют величины, эквивалентные при  $|\tau| \rightarrow \infty$  (см. асимптотические формулы (25) и (28)). Лемма доказана.

Соотношение (29) играет важную роль в дальнейшем. Вернёмся теперь к обратной задаче.

**5. Начало доказательства главного результата: операторное тождество.** Критерий единственности решения доказываем в форме теоремы 2. Напомним, что в силу леммы 1 осталось установить *достаточность* условия из этого критерия.

Итак, предположим, что ни один нуль характеристической функции (17) не является собственным значением оператора  $A$ . Возьмём произвольное ослабленное решение  $(u(t), g)$  однородной обратной задачи (1), (6) и покажем, что  $u(t) \equiv 0$  на  $[0, T]$  и  $g = 0$ , т.е. что решение может быть только тривиальным.

Для реализации нашего плана потребуются ряд вспомогательных функций. Их построение должно быть согласовано с изучаемой задачей (1), (6). Это означает, что функции должны зависеть от выбора значений  $\alpha, \beta \in \mathbb{C} \setminus \{0\}$ . Поскольку в конкретной ситуации такие значения можно считать фиксированными, мы не указываем на  $\alpha$  и  $\beta$  в используемых далее обозначениях. Другими словами, вместо  $Y(t, \lambda; \alpha, \beta)$  или  $Z(t, \lambda; \alpha, \beta)$  пишем просто  $Y(t, \lambda)$ ,  $Z(t, \lambda)$  и т.п. Аналогично всюду в доказательстве считаем фиксированным значение  $T > 0$ .

Начнём с того, что, комбинируя выражения (11) и (12), составим функцию

$$Y(t, \lambda) = \alpha \frac{\text{ch}(\sqrt{\lambda}t) - 1}{\lambda} + \beta \frac{\text{sh}(\sqrt{\lambda}t)}{\sqrt{\lambda}} \tag{30}$$

с аргументами  $t \in [0, T]$  и  $\lambda \in \mathbb{C}$ . Отметим, что

$$Y(0, \lambda) \equiv 0, \quad Y(T, \lambda) = L(\lambda), \quad \lambda \in \mathbb{C}. \tag{31}$$

Первое соотношение в (31) очевидно следует из определения (30); второе получается сравнением с формулами (16), (17), задающими характеристическую функцию  $L(\lambda)$ .

Центральную роль в этом разделе играет даже не сама функция  $Y(t, \lambda)$ , а её производная по переменной  $t$ . Введём обозначение

$$Z(t, \lambda) \equiv Y'_t(t, \lambda) = \alpha \frac{\text{sh}(\sqrt{\lambda}t)}{\sqrt{\lambda}} + \beta \text{ch}(\sqrt{\lambda}t). \tag{32}$$

Продифференцировав дальше, получим

$$Z'_t(t, \lambda) = \alpha \operatorname{ch}(\sqrt{\lambda}t) + \beta\sqrt{\lambda} \operatorname{sh}(\sqrt{\lambda}t), \tag{33}$$

$$Z''_{tt}(t, \lambda) = \alpha\sqrt{\lambda} \operatorname{sh}(\sqrt{\lambda}t) + \beta\lambda \operatorname{ch}(\sqrt{\lambda}t) = \lambda Z(t, \lambda). \tag{34}$$

Выделим значения

$$Z(0, \lambda) = \beta, \quad Z'_t(0, \lambda) = \alpha, \quad \lambda \in \mathbb{C}. \tag{35}$$

Отметим также то, что все функции  $Y(t, \lambda)$ ,  $Z(t, \lambda)$ ,  $Z'_t(t, \lambda)$ ,  $Z''_{tt}(t, \lambda)$  являются целыми по переменным  $t$  и  $\lambda$ , разложимыми в степенные ряды наподобие (14), (15).

Теперь, сочетая функцию  $Z(t, \lambda)$  из формулы (32) и функцию  $u(t)$  – первый компонент выбранного решения  $(u(t), g)$ , определим векторную целую функцию

$$f(\lambda) = \int_0^T Z(t, \lambda)u(T-t) dt, \quad \lambda \in \mathbb{C}. \tag{36}$$

Поскольку поведение  $u(t)$  вблизи границ отрезка  $[0, T]$  может в определённом смысле “портиться” (см. условия (4)), возьмём малое  $\varepsilon > 0$  и рассмотрим аппроксимацию

$$f_\varepsilon(\lambda) = \int_\varepsilon^{T-\varepsilon} Z(t, \lambda)u(T-t) dt, \quad \lambda \in \mathbb{C}. \tag{37}$$

Тогда

$$f_\varepsilon(\lambda) = \int_\varepsilon^{T-\varepsilon} Z(t, \lambda)u(T-t) dt \rightarrow \int_0^T Z(t, \lambda)u(T-t) dt = f(\lambda) \tag{38}$$

при  $\varepsilon \rightarrow 0 + 0$  для любого фиксированного  $\lambda \in \mathbb{C}$ . В предельном переходе (38) используется то, что  $u \in C([0, T], E)$  по определению ослабленного решения (4).

Напомним также, что  $u(t) \in D(A)$  при  $0 < t < T$  и  $Au \in C((0, T), E)$  с линейным замкнутым оператором  $A$ . Поэтому, согласно известным свойствам векторного интеграла Римана, получаем, что  $f_\varepsilon(\lambda) \in D(A)$  для любого  $\lambda \in \mathbb{C}$ , и вычисление  $Af_\varepsilon(\lambda)$  можно проводить внесением оператора  $A$  под знак интеграла в (37) (см. [21, теорема 3.3.2]).

С учётом уравнения (1) имеем

$$Af_\varepsilon(\lambda) = \int_\varepsilon^{T-\varepsilon} Z(t, \lambda)Au(T-t) dt = \int_\varepsilon^{T-\varepsilon} Z(t, \lambda)u''(T-t) dt - g \int_\varepsilon^{T-\varepsilon} Z(t, \lambda) dt. \tag{39}$$

Последний “скалярный” интеграл в (39) допускает предельный переход

$$\int_\varepsilon^{T-\varepsilon} Z(t, \lambda) dt = \int_\varepsilon^{T-\varepsilon} Y'_t(t, \lambda) dt = Y(T-\varepsilon, \lambda) - Y(\varepsilon, \lambda) \rightarrow L(\lambda) \tag{40}$$

при  $\varepsilon \rightarrow 0 + 0$  для любого фиксированного  $\lambda \in \mathbb{C}$ . В выкладке (40) использована связь (32) с соотношениями (31).

Предшествующий “нескалярный” интеграл в (39) обрабатывается чуть сложнее. Проинтегрируем два раза по частям

$$\int_\varepsilon^{T-\varepsilon} Z(t, \lambda)u''(T-t) dt = (Z(t, \lambda)u'(T-t) + Z'_t(t, \lambda)u(T-t)) \Big|_{T-\varepsilon}^\varepsilon + \int_\varepsilon^{T-\varepsilon} Z''_{tt}(t, \lambda)u(T-t) dt.$$



При  $\varepsilon \rightarrow 0 + 0$  внеинтегральное слагаемое обращается в нуль. Действительно, с учётом действующих условий (6) и (35), а также того, что  $u \in C^1([0, T], E)$ , имеем

$$Z(\varepsilon, \lambda)u'(T - \varepsilon) + Z'_t(\varepsilon, \lambda)u(T - \varepsilon) - Z(T - \varepsilon, \lambda)u'(\varepsilon) - Z'_t(T - \varepsilon, \lambda)u(\varepsilon) \rightarrow \beta u'(T) + \alpha u(T) - 0 - 0 = 0.$$

Затем, принимая во внимание связь  $Z''_{tt}(t, \lambda) = \lambda Z(t, \lambda)$  (см. (34)), замечаем, что

$$\int_{\varepsilon}^{T-\varepsilon} Z''_{tt}(t, \lambda)u(T - t) dt = \lambda \int_{\varepsilon}^{T-\varepsilon} Z(t, \lambda)u(T - t) dt = \lambda f_{\varepsilon}(\lambda) \rightarrow \lambda f(\lambda) = \lambda \int_0^T Z(t, \lambda)u(T - t) dt$$

при  $\varepsilon \rightarrow 0 + 0$  для любого фиксированного  $\lambda \in \mathbb{C}$  (здесь  $f(\lambda)$ ,  $f_{\varepsilon}(\lambda)$  – векторные целые функции из формул (36), (37) соответственно). В результате имеем соотношение

$$\int_{\varepsilon}^{T-\varepsilon} Z(t, \lambda)u''(T - t) dt \rightarrow 0 + \lambda \int_0^T Z(t, \lambda)u(T - t) dt = \lambda f(\lambda) \tag{41}$$

при  $\varepsilon \rightarrow 0 + 0$  для любого фиксированного  $\lambda \in \mathbb{C}$ .

Теперь можем положить  $\varepsilon \rightarrow 0 + 0$  в формуле (39). Учитывая замкнутость оператора  $A$ , а также соотношения (38), (40), (41), выводим из (39), что  $f(\lambda) \in D(A)$  при всех  $\lambda \in \mathbb{C}$  и

$$Af(\lambda) = \lambda f(\lambda) - L(\lambda)g, \quad \lambda \in \mathbb{C}. \tag{42}$$

Это *основное операторное тождество* для функции  $f(\lambda)$  из формулы (36). Напомним, что векторная целая функция  $f(\lambda)$  определяется первым компонентом  $u(t)$  решения  $(u(t), g)$  однородной обратной задачи (1), (6). Второй компонент  $g \in E$ , умноженный на характеристическую функцию  $L(\lambda)$  из формулы (17), входит в правую часть (42).

Равенство (42) можно дифференцировать по  $\lambda$ , внося производные под знак линейного замкнутого оператора  $A$ . Последовательное дифференцирование приводит к тождествам

$$Af^{(m)}(\lambda) = \lambda f^{(m)}(\lambda) + m f^{(m-1)}(\lambda) - L^{(m)}(\lambda)g, \quad \lambda \in \mathbb{C}, \tag{43}$$

верным при всех  $m \in \mathbb{N}$ . При  $m = 0$  формула (43) вновь обращается в тождество (42).

Воспользуемся теперь нашим базовым предположением о том, что ни один нуль характеристической функции  $L(\lambda)$  не является собственным значением оператора  $A$ .

Допустим, например, что  $L(\lambda_0) = 0$  в точке  $\lambda_0 \in \mathbb{C}$ . Тогда подстановка  $\lambda = \lambda_0$  в тождество (42) даёт результат  $Af(\lambda_0) = \lambda_0 f(\lambda_0)$ . Отсюда, в силу упомянутого базового предположения, выводим, что  $f(\lambda_0) = 0$ . Это означает, что в действующих условиях всякий нуль характеристической функции  $L(\lambda)$  из формулы (17) будет нулём векторной целой функции  $f(\lambda)$ , определённой по формуле (36). Более того, если  $\lambda = \lambda_0$  – нуль кратности  $k \geq 2$  для  $L(\lambda)$ , то  $L^{(m)}(\lambda_0) = 0$  при  $m = 0, \dots, k - 1$ , и последовательное применение формулы (43) показывает, что  $f^{(m)}(\lambda_0) = 0$  при  $m = 0, \dots, k - 1$ .

Зафиксируем установленный факт: при сделанном предположении всякий нуль характеристической функции  $L(\lambda)$  из формулы (17) является нулём не меньшей кратности\*) для векторной целой функции  $f(\lambda)$  из формулы (36). Другими словами, отношение

$$\frac{f(\lambda)}{L(\lambda)} = \frac{1}{L(\lambda)} \int_0^T Z(t, \lambda)u(T - t) dt \tag{44}$$

определяет векторную целую функцию переменной  $\lambda \in \mathbb{C}$ .

\*) Как станет ясно в дальнейшем (см. п. 8 ниже), обычно все нули функции  $L(\lambda)$  являются простыми, и лишь при некоторых специальных сочетаниях параметров  $\alpha, \beta \in \mathbb{C} \setminus \{0\}$  и  $T > 0$  характеристическая функция (17) имеет также один нуль кратности два. Иначе говоря, хотя кратных нулей у  $L(\lambda)$  почти никогда не бывает, их потенциальную возможность всё же приходится учитывать.

Выведем отсюда, что  $u(t) \equiv 0$  при  $0 \leq t \leq T$ . Неформальную идею последующих рассуждений дал нам пример 200 из известного задачника [22; отд. 4, гл. 3, § 5]. Этот пример был разобран в [22] со ссылкой на Карлемана. Для применения идеи в нашей ситуации схему Карлемана пришлось видоизменить.

**6. Продолжение доказательства: принцип Фрагмена–Линделёфа.** Поскольку предстоит активная работа с целыми функциями, удобно перейти полностью к скалярному случаю. Возьмём линейный непрерывный функционал  $f^*$  из сопряжённого банахова пространства  $E^*$  и подействуем им на векторные компоненты в (44).

Обозначим  $F(\lambda) \equiv f^*(f(\lambda))$  при  $\lambda \in \mathbb{C}$  и  $\psi(t) \equiv f^*(u(t))$  при  $0 \leq t \leq T$ . Тогда, согласно формуле (36), имеем

$$F(\lambda) = \int_0^T Z(t, \lambda) \psi(T-t) dt, \quad \lambda \in \mathbb{C}. \quad (45)$$

При взятии функционала  $f^*$  нули функции  $f(\lambda)$  не исчезают, а их кратности не уменьшаются. Отсюда заключаем, что отношение

$$Q(\lambda) \equiv \frac{F(\lambda)}{L(\lambda)} = \frac{f^*(f(\lambda))}{L(\lambda)} \quad (46)$$

определяет целую функцию переменной  $\lambda \in \mathbb{C}$  (ср. (46) с прежним отношением (44)). Для нужных оценок дроби (46) перепишем её числитель  $F(\lambda)$  в виде, отличном от (45).

Вспомним сначала, что  $Z(t, \lambda) = Y'_t(t, \lambda)$ , где  $Y(0, \lambda) \equiv 0$  (см. формулу (32), а затем (31)). Для функции  $\psi(t) \equiv f^*(u(t))$  отметим соотношения

$$\psi \in C^1[0, T], \quad \psi(0) = 0, \quad (47)$$

выполненные согласно (4) и (6). (Из (6) также следует, что  $\psi'(0) = 0$  и  $\alpha\psi(T) + \beta\psi'(T) = 0$ , но эти равенства уже не понадобятся.) Применив сказанное к интегралу (45), получаем, что

$$F(\lambda) = \int_0^T Y'_t(t, \lambda) \psi(T-t) dt = \int_0^T Y(t, \lambda) \psi'(T-t) dt, \quad \lambda \in \mathbb{C}, \quad (48)$$

где  $Y(T, \lambda)\psi(0) - Y(0, \lambda)\psi(T) \equiv 0$  из-за значений  $\psi(0) = 0$  и  $Y(0, \lambda) \equiv 0$ .

Теперь преобразуем запись функции  $Y(t, \lambda)$ . Исходя из определения (30), имеем

$$Y(t, \lambda) = \alpha \left( \frac{t}{T} \right)^2 \frac{\text{ch}(\sqrt{(t/T)^2 \lambda T}) - 1}{(t/T)^2 \lambda} + \beta \left( \frac{t}{T} \right) \frac{\text{sh}(\sqrt{(t/T)^2 \lambda T})}{\sqrt{(t/T)^2 \lambda}}.$$

Последующее сравнение с элементарными функциями (16) даёт выражение

$$Y(t, \lambda) = \alpha(t/T)^2 L_1((t/T)^2 \lambda) + \beta(t/T) L_2((t/T)^2 \lambda), \quad 0 \leq t \leq T, \quad \lambda \in \mathbb{C}. \quad (49)$$

Подставим (49) в интеграл (48) и сделаем там замену  $s = t/T$ . Получим представление

$$F(\lambda) = \int_0^1 (\alpha s^2 L_1(\lambda s^2) + \beta s L_2(\lambda s^2)) \eta'(1-s) ds, \quad \lambda \in \mathbb{C}, \quad (50)$$

где  $\eta(s) \equiv \psi(Ts)$ ,  $\eta'(s) = T\psi'(Ts)$  и  $\eta'(1-s) = T\psi'(T-Ts)$  при  $0 \leq s \leq 1$ . Из (47) следует, что

$$\eta \in C^1[0, 1], \quad \eta(0) = 0. \quad (51)$$

Представление (50) с функцией  $\eta(s)$  типа (51) используем при анализе отношения (46), составленного из функций  $F(\lambda)$  и  $L(\lambda) = \alpha L_1(\lambda) + \beta L_2(\lambda)$ .

Напомним (см. [19, 20]), что монотонные при  $r > 0$  характеристики

$$M(F; r) \equiv \max_{|\lambda|=r} |F(\lambda)|, \quad M(L_1; r) \equiv \max_{|\lambda|=r} |L_1(\lambda)|, \quad M(L_2; r) \equiv \max_{|\lambda|=r} |L_2(\lambda)|$$

отражают рост целых функций  $F(\lambda)$ ,  $L_1(\lambda)$ ,  $L_2(\lambda)$  на плоскости  $\mathbb{C}$ . Основываясь на (50), запишем оценку

$$M(F; r) \leq C_\eta (|\alpha| M(L_1; r) + |\beta| M(L_2; r)), \quad r > 0, \tag{52}$$

с константой

$$C_\eta \equiv \int_0^1 |\eta'(s)| ds = \text{Var} \{ \eta(s) \}_0^1. \tag{53}$$

Элементарные функции (16), т.е. функции  $L_1(\lambda)$ ,  $L_2(\lambda)$ , имеют порядок  $\rho = 1/2$ , а их характеристики  $M(L_1; r)$ ,  $M(L_2; r)$  при  $r \geq 1$  оцениваются сверху через  $\exp(\sqrt{r}T)$ . Подставив данную мажоранту в (52), получим, что и  $F(\lambda)$  не может расти быстрее, чем целая функция порядка  $\rho = 1/2$ . Тот же характер роста имеет целая функция  $L(\lambda) = \alpha L_1(\lambda) + \beta L_2(\lambda)$ .

Но тогда по теореме о категориях (см. [19; с. 37]) рост целой функции  $Q(\lambda)$ , задаваемой отношением (46), не может быть выше, чем у функции порядка  $\rho = 1/2$ . В частности, заведомо можем утверждать, что  $Q(\lambda)$  есть целая функция нулевого экспоненциального типа, т.е. такая, что  $\ln M(Q; r) = o(r)$  при  $r \rightarrow +\infty$ .

Оценим  $Q(\lambda) \equiv F(\lambda)/L(\lambda)$  на мнимой оси. Обозначим  $\lambda = i\tau$ , где  $\tau \in \mathbb{R}$ , и используем характеристики  $V_1(\tau)$ ,  $V_2(\tau)$  из формулы (20). Учитывая их строгую монотонность, действующую при возрастании  $|\tau| \geq 0$  (см. лемму 2), выводим из (50), что

$$|F(i\tau)| \leq C_\eta (|\alpha| V_1(\tau) + |\beta| V_2(\tau)), \quad \tau \in \mathbb{R}, \tag{54}$$

с той же константой  $C_\eta$  вида (53). Согласно лемме 3 величина  $V(\tau) \equiv |L(i\tau)|$  строго положительна при достаточно больших  $|\tau| > 0$ . Поэтому с учётом (54) можем корректно оценить

$$|Q(i\tau)| = \frac{|F(i\tau)|}{|L(i\tau)|} \leq C_\eta \frac{|\alpha| V_1(\tau) + |\beta| V_2(\tau)}{V(\tau)} \tag{55}$$

при тех же достаточно больших  $|\tau| > 0$ . Снова по лемме 3, если  $|\tau| \rightarrow \infty$ , то дробь в правой части (55) стремится к единице. Следовательно, непрерывная величина  $|Q(i\tau)|$  является ограниченной при  $\tau \in \mathbb{R}$ , или, другими словами, целая функция нулевого экспоненциального типа  $Q(\lambda)$  оказывается ограниченной на мнимой оси. В силу известного варианта теоремы Фрагмена–Линделёфа (см. [19; с. 71]) такая функция может быть только константой.

Итак, получили, что  $Q(\lambda) \equiv C$  в плоскости  $\mathbb{C}$ . Но тогда  $F(\lambda) = CL(\lambda)$  при всех  $\lambda \in \mathbb{C}$  с некоторой константой  $C \in \mathbb{C}$ . Используя представление (50), запишем

$$\int_0^1 (\alpha s^2 L_1(\lambda s^2) + \beta s L_2(\lambda s^2)) \eta'(1-s) ds = C(\alpha L_1(\lambda) + \beta L_2(\lambda)), \quad \lambda \in \mathbb{C}. \tag{56}$$

Покажем, что тождество (56) с функцией  $\eta(s)$ , удовлетворяющей требованиям (51), возможно, только если  $C = 0$  и  $\eta(s) \equiv 0$  при  $0 \leq s \leq 1$ .

**7. Завершение доказательства: теорема Мюнца.** Проанализируем тождество (56), выполненное при всех  $\lambda \in \mathbb{C}$  с некоторой константой  $C \in \mathbb{C}$ . Используя явные формулы (16) (см. также (14) и (15)), разложим в степенные ряды

$$\alpha L_1(\lambda) + \beta L_2(\lambda) = \sum_{n=0}^{\infty} \left( \alpha \frac{T^{2n+2}}{(2n+2)!} + \beta \frac{T^{2n+1}}{(2n+1)!} \right) \lambda^n = \sum_{n=0}^{\infty} \left( \frac{\alpha T}{2n+2} + \beta \right) \frac{T^{2n+1}}{(2n+1)!} \lambda^n,$$

$$\begin{aligned} \alpha s^2 L_1(\lambda s^2) + \beta s L_2(\lambda s^2) &= \sum_{n=0}^{\infty} \left( \alpha \frac{s^{2n+2} T^{2n+2}}{(2n+2)!} + \beta \frac{s^{2n+1} T^{2n+1}}{(2n+1)!} \right) \lambda^n = \\ &= \sum_{n=0}^{\infty} \left( \alpha T \frac{s^{2n+2}}{2n+2} + \beta s^{2n+1} \right) \frac{T^{2n+1}}{(2n+1)!} \lambda^n. \end{aligned}$$

Подставим эти разложения в (56) и приравняем коэффициенты при одинаковых степенях  $\lambda$ . С учётом очевидных сокращений на  $T^{2n+1}/(2n+1)!$  получим счётный набор соотношений

$$\int_0^1 \left( \alpha T \frac{s^{2n+2}}{2n+2} + \beta s^{2n+1} \right) \eta'(1-s) ds = C \left( \frac{\alpha T}{2n+2} + \beta \right), \quad n = 0, 1, 2, \dots \quad (57)$$

Модуль интеграла в равенстве (57) мажорируется величиной

$$\left( |\alpha| T \frac{1}{(2n+2)(2n+3)} + |\beta| \frac{1}{2n+2} \right) \max_{0 \leq s \leq 1} |\eta'(s)|,$$

стремящейся к нулю при  $n \rightarrow \infty$ . Но тогда и сам интеграл стремится к нулю при  $n \rightarrow \infty$ . Тем самым, переходя к пределу в (57), получаем при  $n \rightarrow \infty$  что  $0 = C\beta$ . Здесь  $\beta \neq 0$  по условию. Поэтому  $C = 0$ .

Подставив найденное значение  $C$  в (57), получим соотношения

$$\int_0^1 \left( \alpha T \frac{s^{2n+2}}{2n+2} + \beta s^{2n+1} \right) \eta'(1-s) ds = 0, \quad n = 0, 1, 2, \dots \quad (58)$$

Заметим, что

$$\int_0^1 \frac{s^{2n+2}}{2n+2} \eta'(1-s) ds = \frac{s^{2n+2}}{2n+2} \eta(1-s) \Big|_1^0 + \int_0^1 s^{2n+1} \eta(1-s) ds = \int_0^1 s^{2n+1} \eta(1-s) ds$$

с учётом значения  $\eta(0) = 0$  (см. (51)). Такое преобразование в (58) даёт результат

$$\int_0^1 (\alpha T \eta(1-s) + \beta \eta'(1-s)) s^{2n+1} ds = 0, \quad n = 0, 1, 2, \dots \quad (59)$$

Полученное возможно, только если

$$\alpha T \eta(1-s) + \beta \eta'(1-s) \equiv 0, \quad 0 \leq s \leq 1. \quad (60)$$

Факт (60) выводится из (59) разными способами, но проще сразу воспользоваться теоремой Мюнца, утверждающей в своей слабой редакции<sup>\*)</sup>, что если непрерывная функция ортогональна на  $[0, 1]$  системе степеней  $\{s^{a_n}\}_{n=0}^{\infty}$ , где  $0 < a_0 < a_1 < \dots < +\infty$  и  $\sum_{n=0}^{\infty} a_n^{-1} = +\infty$ , то функция тождественно равна нулю на  $[0, 1]$  (см. номер 198 в [22; с. 44] или [23; с. 304–305]; см. также [24; с. 109] и [25; с. 116, 122–124]).

Итак, установлено соотношение (60). Оно очевидно эквивалентно обыкновенному дифференциальному уравнению  $\eta'(s) + (\alpha T/\beta)\eta(s) = 0$  при  $0 \leq s \leq 1$ , к которому, согласно (51),

<sup>\*)</sup> Более известна сильная версия теоремы Мюнца о плотности линейных комбинаций степеней в пространствах типа  $C[0, 1]$  или  $L^p[0, 1]$ . Отметим для точности, что в оригинальной работе Мюнца [23] рассматривался лишь классический вариант пространства  $C[0, 1]$ . Но случай  $L^p$  оказался идейно близок, и перенесённый туда результат тоже стали называть *теоремой Мюнца* (см., например, [24; гл. 3, § 6]).

надо добавить начальное условие  $\eta(0) = 0$ . Отсюда следует, что  $\eta(s) \equiv 0$  при  $0 \leq s \leq 1$ . Вспомнив связь  $\eta(s) \equiv \psi(Ts)$  при  $0 \leq s \leq 1$ , заключаем, что функция  $\psi(t) \equiv f^*(u(t))$  тождественно равна нулю всюду на  $[0, T]$ . Выбор функционала  $f^* \in E^*$  был произвольным, поэтому на основании известного следствия теоремы Хана–Банаха (см. [21; теорема 2.7.4]) можем утверждать, что  $u(t) \equiv 0$  при  $0 \leq t \leq T$ . Подстановка такой функции в уравнение (1) показывает, что  $g = 0$ .

Подведём итог: при сделанном предположении (о том, что ни один нуль характеристической функции (17) не является собственным значением оператора  $A$ ) неизбежно получаем, что решение  $(u(t), g)$  однородной обратной задачи (1), (6) может быть только тривиальным. Критерий единственности решения в форме теоремы 2 полностью доказан. Отсюда автоматически следует теорема 1. Обоснование главного результата завершено.

**8. Нули характеристической функции: общие сведения.** Теперь для применения установленного критерия требуется информация о распределении нулей характеристической функции  $L(\lambda) = L(\lambda; T, \alpha, \beta)$  из формулы (17). Уделим этому вопросу особое внимание.

Как и в формуле (18), рассматриваем множество  $\Lambda = \Lambda(T, \alpha, \beta) \equiv \{\lambda \in \mathbb{C} : L(\lambda) = 0\}$ . Опишем его состав в зависимости от выбора параметров  $T > 0$  и  $\alpha, \beta \in \mathbb{C} \setminus \{0\}$ . Специально подчеркнём, что соглашение  $\alpha \neq 0, \beta \neq 0$  действует далее без оговорок – при обращении в нуль одного из коэффициентов ситуация сразу упрощается и конкретный вид множества  $\Lambda$  находится элементарным образом.

Действительно, если взять  $\alpha \neq 0, \beta = 0$  (как для финального переопределения первого рода), то  $L(\lambda) = \alpha L_1(\lambda)$  с функцией  $L_1(\lambda) = L_1(\lambda; T)$  из формулы (16) и

$$\Lambda(T, \alpha, 0) = \Lambda(T, 1, 0) \equiv \{\lambda \in \mathbb{C} : L_1(\lambda) = 0\} = \{-4k^2\pi^2/T^2\}_{k \in \mathbb{N}}, \tag{61}$$

причём все нули в множестве (61) имеют кратность два (см. также [4]). В свою очередь если взять  $\alpha = 0, \beta \neq 0$  (как для финального переопределения второго рода), то  $L(\lambda) = \beta L_2(\lambda)$  с функцией  $L_2(\lambda) = L_2(\lambda; T)$  из формулы (16) и

$$\Lambda(T, 0, \beta) = \Lambda(T, 0, 1) \equiv \{\lambda \in \mathbb{C} : L_2(\lambda) = 0\} = \{-m^2\pi^2/T^2\}_{m \in \mathbb{N}}, \tag{62}$$

причём все нули в множестве (62) являются простыми (см. также [5]). Любопытная особенность – присутствие в обоих множествах (61) и (62) общей серии нулей  $(-4k^2\pi^2/T^2)$ , взятых при  $k \in \mathbb{N}$ , сохранится и при рассмотрении финального переопределения третьего рода со значениями  $\alpha \neq 0, \beta \neq 0$ . Сосредоточимся сейчас именно на этом случае.

Используем элементарные формулы  $\operatorname{ch} a - 1 = 2 \operatorname{sh}^2(a/2)$  и  $\operatorname{sh} a = 2 \operatorname{sh}(a/2) \operatorname{ch}(a/2)$ . С их помощью для функций (16) запишем

$$L_1(\lambda) = (2/\lambda) \operatorname{sh}^2(\sqrt{\lambda}T/2), \quad L_2(\lambda) = (2/\sqrt{\lambda}) \operatorname{sh}(\sqrt{\lambda}T/2) \operatorname{ch}(\sqrt{\lambda}T/2).$$

Отсюда видно, что характеристическая функция (17) допускает альтернативное представление

$$L(\lambda) = \frac{2}{\sqrt{\lambda}} \operatorname{sh} \frac{\sqrt{\lambda}T}{2} \left( \frac{\alpha}{\sqrt{\lambda}} \operatorname{sh} \frac{\sqrt{\lambda}T}{2} + \beta \operatorname{ch} \frac{\sqrt{\lambda}T}{2} \right) = G_1(\lambda)G_2(\lambda), \tag{63}$$

т.е. функция  $L(\lambda) = \alpha L_1(\lambda) + \beta L_2(\lambda)$  распадается в произведение двух целых функций

$$G_1(\lambda) = G_1(\lambda; T) \equiv \frac{2}{\sqrt{\lambda}} \operatorname{sh} \frac{\sqrt{\lambda}T}{2}, \tag{64}$$

$$G_2(\lambda) = G_2(\lambda; T, \alpha, \beta) \equiv \frac{\alpha}{\sqrt{\lambda}} \operatorname{sh} \frac{\sqrt{\lambda}T}{2} + \beta \operatorname{ch} \frac{\sqrt{\lambda}T}{2} \tag{65}$$

переменной  $\lambda \in \mathbb{C}$ . Тем самым множество нулей характеристической функции  $L(\lambda)$  оказывается составленным из множества нулей функции (64) и множества нулей функции (65).

Удобно сократить число параметров и ввести стандартизированные функции

$$H_1(\zeta) \equiv \frac{2}{\sqrt{\zeta}} \operatorname{sh} \frac{\sqrt{\zeta}}{2}, \tag{66}$$

$$H_2(\zeta) = H_2(\zeta; p) \equiv \frac{2}{\sqrt{\zeta}} \operatorname{sh} \frac{\sqrt{\zeta}}{2} + p \operatorname{ch} \frac{\sqrt{\zeta}}{2} \tag{67}$$

переменной  $\zeta \in \mathbb{C}$ . В записи (67) остался единственный параметр  $p \in \mathbb{C} \setminus \{0\}$ . Очевидно, что

$$G_1(\lambda; T) = TH_1(\lambda T^2), \quad G_2(\lambda; T, \alpha, \beta) = (\alpha T/2)H_2(\lambda T^2; 2\beta/(\alpha T)). \tag{68}$$

Поэтому характеристическая функция (17) представима в виде

$$L(\lambda) = L(\lambda; T, \alpha, \beta) = cH_1(\lambda T^2)H_2(\lambda T^2; p), \tag{69}$$

где

$$c = \alpha T^2/2, \quad p = 2\beta/(\alpha T). \tag{70}$$

Условия (70) обеспечивают эквивалентность представлений (63) и (69) для  $L(\lambda)$ .

Формулы перехода (68) показывают, что нули функций  $G_1(\lambda)$ ,  $G_2(\lambda)$  связаны с нулями соответствующих функций  $H_1(\zeta)$ ,  $H_2(\zeta)$  преобразованием подобия  $\lambda = \zeta/T^2$ , сохраняющим не только все геометрические соотношения нулей, но и их кратности.

Завершим этот первичный разбор ситуации следующим базовым утверждением.

**Теорема 3.** *Множество нулей характеристической функции  $L(\lambda)$  вида (17) или, что эквивалентно, вида (63) состоит из двух счётных серий. Первая универсальная серия нулей*

$$\lambda_k^{(1)} = -4k^2\pi^2/T^2, \quad k \in \mathbb{N}, \tag{71}$$

*отвечает множителю (64) и не зависит от выбора значений  $\alpha, \beta \in \mathbb{C} \setminus \{0\}$ . Вторая серия нулей отвечает множителю (65) и может быть записана в виде*

$$\lambda_k^{(2)} = \zeta_k/T^2, \quad k \in J, \tag{72}$$

где  $\zeta_k = \zeta_k(p)$  – нули элементарной целой функции (67) с параметром  $p = 2\beta/(\alpha T)$ . Нумерация  $k \in J = J(p) \subset \mathbb{Z}$  в (72) может зависеть от значения  $p \in \mathbb{C} \setminus \{0\}$ . Множества (71) и (72) не пересекаются, т.е. ни один нуль функции (64) не является нулём функции (65).

**Доказательство.** Все корни уравнения  $\operatorname{sh}(a/2) = 0$  на плоскости  $\mathbb{C}$  выражаются формулой  $a_k = 2k\pi i$ , где  $k \in \mathbb{Z}$ . При этом  $\operatorname{ch}(a_k/2) = \cos k\pi \neq 0$ . Отсюда, во-первых, следует, что все нули функции (64) имеют вид (71) с указанной там нумерацией  $k \in \mathbb{N}$ , а, во-вторых, что ни один такой нуль  $\lambda_k^{(1)}$  не может быть нулём функции (65).

В то же время функция  $G_2(\lambda) = G_2(\lambda; T, \alpha, \beta)$  вида (65) как целая функция нецелого порядка  $\rho = 1/2$  имеет бесконечное (счётное) множество нулей. Эти нули, учитывая вторую связь в (68), можно формально записать в виде (72) через нули стандартизированной функции (67). Вопрос о наиболее удобной нумерации  $k \in J$  в множестве таких нулей должен уточняться отдельно. Теорема доказана.

Как видим, наличие общего множителя  $(2/\sqrt{\lambda}) \operatorname{sh}(\sqrt{\lambda}T/2)$  приводит к тому, что серия нулей (71) оказывается общей для всех трёх функций  $L_1(\lambda)$ ,  $L_2(\lambda)$  и  $L(\lambda) = \alpha L_1(\lambda) + \beta L_2(\lambda)$ . Тем самым (см. теорему 1) отсутствие среди чисел (71) собственных значений оператора  $A$  есть *необходимое условие* единственности решения изучаемой обратной задачи (1)–(3) с фиксированным  $T > 0$  при любом выборе параметров  $\alpha, \beta \in \mathbb{C} \setminus \{0\}$  (см. также работы [4] и [5] про случаи  $\alpha \neq 0, \beta = 0$  и  $\alpha = 0, \beta \neq 0$  соответственно). Эта особенность изучаемой обратной задачи представляется весьма оригинальной.

Обсудим теперь вопрос о кратности нулей характеристической функции  $L(\lambda)$ . Используем представление  $L(\lambda) = G_1(\lambda)G_2(\lambda)$  с функциями (64), (65). Поскольку, согласно теореме 3,

множества нулей указанных сомножителей не пересекаются, вопрос о кратных нулях достаточно рассмотреть отдельно для  $G_1(\lambda)$  и  $G_2(\lambda)$ , точнее даже, для их регуляризированных аналогов – функций  $H_1(\zeta)$  и  $H_2(\zeta)$  из формул (66) и (67) соответственно.

Вопрос об отсутствии кратных нулей у функции  $H_1(\zeta)$  решается просто. Имеем

$$H_1(\zeta) \equiv \frac{2}{\sqrt{\zeta}} \operatorname{sh} \frac{\sqrt{\zeta}}{2}, \quad H_1'(\zeta) \equiv \frac{1}{2\zeta} \left( \operatorname{ch} \frac{\sqrt{\zeta}}{2} - \frac{2}{\sqrt{\zeta}} \operatorname{sh} \frac{\sqrt{\zeta}}{2} \right).$$

Как уже отмечалось при доказательстве теоремы 3, величины  $\operatorname{sh}(a/2)$  и  $\operatorname{ch}(a/2)$  никогда одновременно не обращаются в нуль. Отсюда следует, что соотношение  $H_1(\zeta) = H_1'(\zeta) = 0$  невозможно ни в какой точке  $\zeta \in \mathbb{C}$ .

Рассмотрим теперь функцию  $H_2(\zeta) = H_2(\zeta; p)$  с параметром  $p \in \mathbb{C} \setminus \{0\}$ . Картина с её кратными нулями оказывается довольно неожиданной. В решении поставленной задачи важную роль играют комплексно сопряжённые корни уравнения

$$\operatorname{sh} z = z, \tag{73}$$

расположенные в полуплоскости  $\operatorname{Re} z > 0$ . Множество таких корней запишем в виде

$$z = z_n, \quad n \in \mathbb{Z} \setminus \{0\}, \quad z_{-n} = \bar{z}_n. \tag{74}$$

Считаем, что основная серия корней  $z = z_n$  при  $n \in \mathbb{N}$  расположена в первом координатном угле  $\operatorname{Re} z > 0, \operatorname{Im} z > 0$  и занумерована в порядке возрастания модулей.

Следующий результат получен авторами совместно с В.Б. Шерстюковым.

**Теорема 4.** *Рассматриваем функцию  $H_2(\zeta; p)$  вида (67) с параметром  $p \in \mathbb{C} \setminus \{0\}$ . Определим счётное множество значений*

$$p_n = -\frac{2}{1 + \operatorname{ch} z_n}, \quad n \in \mathbb{Z} \setminus \{0\}, \quad p_{-n} = \bar{p}_n, \tag{75}$$

где  $z_n$  – корни (74) трансцендентного уравнения (73), попадающие в полуплоскость  $\operatorname{Re} z > 0$ . Тогда справедливы утверждения:

1. При каждом  $p \in \mathbb{C} \setminus \{0\}$ , не входящем в множество (75), функция  $H_2(\zeta; p)$  имеет только простые нули.

2. При  $p = p_n$  из множества (75) функция  $H_2(\zeta; p_n)$  помимо бесконечного числа простых нулей имеет ровно один кратный нуль кратности два, который находится по формуле

$$\zeta = z_n^2 \tag{76}$$

с тем же  $z_n$  из множества (74), что и при выборе параметра  $p_n = -2/(1 + \operatorname{ch} z_n)$ .

Добавим ещё, что при  $n \in \mathbb{N}$  значения  $p_n$  из формулы (75) попадают в область

$$-2 < \operatorname{Re} p < 0, \quad 0 < \operatorname{Im} p < 1, \tag{77}$$

причём  $p_n \rightarrow 0$  при  $n \rightarrow \infty$ . Как следствие, при всех номерах  $n \in \mathbb{Z} \setminus \{0\}$  значения  $p_n$  не могут быть вещественными.

Подробное доказательство теоремы 4 и обсуждение сопутствующих обстоятельств будет дано отдельно (см. также [26, 27]). Сейчас отметим лишь, что все величины из теоремы 4 допускают эффективное исследование с выводом необходимых оценок, представляющих интерес как для математической физики, так и для теории функций. В частности, уравнение (73) посредством замены  $z = i\xi$  с переменной  $\xi \in \mathbb{C}$  приводится к виду

$$\sin \xi = \xi. \tag{78}$$

Последнее встречается в спектральной теории при рассмотрении ряда задач механики сплошной среды (см., например, [28–30]). С аналитической точки зрения уравнение (78) рассматривал ещё Харди [31]. Простой и понятный разбор ситуации, связанной с корнями подобных

уравнений, дал Маркушевич [32; с. 64–68]. Основываясь на его изложении, получаем приближённую формулу

$$z_n \approx \ln(4n\pi) + \frac{1}{4n} + i \left( 2n\pi + \frac{\pi}{2} - \frac{1}{2n\pi} \ln(4n\pi) \right) \tag{79}$$

для основной серии корней “нашего” уравнения (73). Родственный вариант, восходящий к работе Харди [31], имеет вид (см. [29; с. 392])

$$z_n = \ln((4n + 1)\pi) + \delta_n + i \left( 2n\pi + \frac{\pi}{2} - \frac{2}{(4n + 1)\pi} \ln((4n + 1)\pi) + \varepsilon_n \right). \tag{80}$$

Отдельно показывается, что приближения (79) и (80) согласованы с нумерацией  $n \in \mathbb{N}$  (см. [31]). Высокую точность таких соотношений подтверждают следующие оценки В.Б. Шерстюкова:

$$0 < \delta_n < \frac{2}{(4n + 1)^2\pi^2} \ln^2((4n + 1)\pi), \quad |\varepsilon_n| < \frac{4}{(4n + 1)^3\pi^3} \ln^3((4n + 1)\pi),$$

действующие при всех номерах  $n \in \mathbb{N}$  для остатков  $\delta_n, \varepsilon_n \in \mathbb{R}$  в формуле (80). Используя информацию о корнях уравнения (73), можно весьма точно оценить величины (75), (76) из теоремы 4 и указать для них ряд полезных аналитических соотношений. Здесь применимы также методы компьютерной математики.

Упомянем, к примеру, что область (77), где локализуются значения  $p_n$  при  $n \in \mathbb{N}$ , получена из общих соображений, связанных с действием отображения  $p_n = -2/(1 + \operatorname{ch} z_n)$  на корни уравнения (73), расположенные в первом координатном угле. Однако, как показывает численный расчёт, можно, не погрешив против истины, заменить множество (77) гораздо более точным прямоугольником  $-0.11 < \operatorname{Re} p < 0, \quad 0 < \operatorname{Im} p < 0.22$ .

Вернёмся к рассмотрению характеристической функции  $L(\lambda) = L(\lambda; T, \alpha, \beta)$ . Используем представление (69) с параметром  $p = 2\beta/(\alpha T)$ , заданным согласно (70). Проведённый анализ позволяет утверждать, что в большинстве ситуаций, т.е. в случае общего положения, все нули функции  $L(\lambda) = L(\lambda; T, \alpha, \beta)$  являются простыми. Исключения составляют редкие примеры, когда параметры  $\alpha, \beta \in \mathbb{C} \setminus \{0\}$  и  $T > 0$  связаны специальным соотношением

$$\alpha T + (1 + \operatorname{ch} z_n)\beta = 0 \tag{81}$$

с некоторым корнем  $z_n$  из множества (74) (ср. (81) с формулой (75) из теоремы 4). При таком условии функция  $L(\lambda) = L(\lambda; T, \alpha, \beta)$  помимо бесконечного числа простых нулей имеет единственный нуль кратности два, и он находится по формуле  $\lambda = z_n^2/T^2$  с тем же значением  $z_n$ , что и в (81) (ср. формулу (72) с формулой (76) из теоремы 4).

Согласно теореме 4 значения  $1 + \operatorname{ch} z_n = -2/p_n$  не могут быть вещественными. Поэтому соотношение (81) невозможно, если  $\alpha, \beta \in \mathbb{R} \setminus \{0\}$ . Отсюда получаем такое утверждение.

**Теорема 5.** *При любом выборе параметров  $\alpha, \beta \in \mathbb{R} \setminus \{0\}$  и  $T > 0$  все нули характеристической функции  $L(\lambda) = L(\lambda; T, \alpha, \beta)$  из формулы (17) являются простыми.*

Выбор вещественных параметров  $\alpha$  и  $\beta$  в условии (3) является, безусловно, приоритетным. Сосредоточимся именно на этом случае. Дополним теорему 5 следующим результатом.

**Теорема 6.** *При любом выборе параметров  $\alpha, \beta \in \mathbb{R} \setminus \{0\}$  и  $T > 0$  все нули характеристической функции  $L(\lambda) = L(\lambda; T, \alpha, \beta)$  из формулы (17) являются вещественными.*

**Доказательство.** С учётом основной теоремы 3 достаточно установить, что все нули  $\zeta_k$  стандартизированной функции (67) при любом выборе параметра  $p \in \mathbb{R} \setminus \{0\}$  являются вещественными. Зафиксируем значение  $p \in \mathbb{R} \setminus \{0\}$  и рассмотрим скалярную задачу

$$4y''(\tau) - \zeta y(\tau) = 0, \quad 0 \leq \tau \leq 1, \quad y(0) = 0, \quad y(1) + py'(1) = 0 \tag{82}$$

со спектральным параметром  $\zeta \in \mathbb{C}$ . С точностью до умножения на константу решение дифференциального уравнения из системы (82), удовлетворяющее первому условию  $y(0) = 0$ ,



имеет вид  $y(\tau) = (2/\sqrt{\zeta}) \operatorname{sh}(\sqrt{\zeta}\tau/2)$ . После подстановки данной функции во второе краевое условие  $y(1) + py'(1) = 0$  приходим к уравнению

$$\frac{2}{\sqrt{\zeta}} \operatorname{sh} \frac{\sqrt{\zeta}}{2} + p \operatorname{ch} \frac{\sqrt{\zeta}}{2} = 0 \tag{83}$$

относительно неизвестной  $\zeta \in \mathbb{C}$ . Корни уравнения (83) те же, что нули целой функции (67) – они совпадают со спектральными значениями задачи (82). Но спектр в задаче (82) может быть только вещественным.

Действительно, исходя из (82), введём оператор  $B = 4d^2/d\tau^2$  в линейном вещественном пространстве  $C[0, 1]$  на области определения  $D_p \equiv \{y \in C^2[0, 1] : y(0) = 0, y(1) + py'(1) = 0\}$ . Используем в  $C[0, 1]$  скалярное произведение

$$(y_1, y_2) = \int_0^1 y_1(\tau)y_2(\tau) d\tau, \quad (y, y) \equiv \|y\|^2.$$

Стандартное интегрирование по частям показывает симметричность оператора  $B$  в том смысле, что  $(By_1, y_2) = (y_1, By_2)$  для любых  $y_1, y_2 \in D_p$ . Как известно, такой оператор не может иметь комплексных собственных значений  $\zeta = \gamma_1 + i\gamma_2$ ,  $\gamma_1, \gamma_2 \in \mathbb{R}$ ,  $\gamma_2 \neq 0$ .

В самом деле, предположим, что  $B(y_1 + iy_2) = (\gamma_1 + i\gamma_2)(y_1 + iy_2)$  с некоторыми  $y_1, y_2 \in D_p$ , т.е.  $By_1 = \gamma_1 y_1 - \gamma_2 y_2$  и  $By_2 = \gamma_1 y_2 + \gamma_2 y_1$ . Тогда вычисление числа  $(By_1, y_2) = (y_1, By_2)$  даёт равенство

$$\gamma_1(y_1, y_2) - \gamma_2(y_2, y_2) = \gamma_1(y_1, y_2) + \gamma_2(y_1, y_1) \sim \gamma_2(\|y_1\|^2 + \|y_2\|^2) = 0.$$

Но если  $\gamma_2 \neq 0$ , то  $y_1(\tau) \equiv 0$  и  $y_2(\tau) \equiv 0$  всюду на  $[0, 1]$ .

Итак, ни сам оператор  $B$ , ни ассоциированная с ним спектральная задача (82) не могут иметь комплексных (невещественных) собственных значений. Следовательно, корни уравнения (83) и совпадающие с ними нули целой функции (67) будут исключительно вещественными. Возвращаясь к исходной характеристической функции (17) (и учитывая теорему 3), получаем утверждение теоремы 6. Доказательство завершено.

Подытожим: при любом выборе значений  $\alpha, \beta \in \mathbb{R} \setminus \{0\}$  все без исключения нули характеристической функции  $L(\lambda) = L(\lambda; T, \alpha, \beta)$  являются вещественными и простыми. Для того чтобы завершить описание таких нулей, осталось, согласно теореме 3, в зависимости от параметра  $p = 2\beta/(\alpha T) \in \mathbb{R} \setminus \{0\}$  указать зоны локализации чисел  $\zeta_k$  из формулы (72), или, что то же самое, зоны локализации нулей  $\zeta_k = \zeta_k(p) \in \mathbb{R}$  стандартизированной функции  $H_2(\zeta; p)$  из формулы (67).

**9. Подробно про зоны локализации.** Разбор начнём в эквивалентных терминах – на языке корней уравнения (83). При этом для лучшего понимания динамики корней в зависимости от параметра  $p$  рассмотрим все значения  $p \in \mathbb{R}$ , включая  $p = 0$ .

Случай  $p = 0$  является особым, так как тогда корни уравнения (83) совпадают с нулями (простыми нулями) функции  $H_1(\zeta)$  из формулы (66) и имеют вид  $\zeta_k = -4k^2\pi^2$  при  $k \in \mathbb{N}$ .

Ещё один особый случай связан с возможным корнем  $\zeta = 0$ . Подстановка такого числа в уравнение (83) даёт соотношение  $1 + p \cdot 1 = 0$ , откуда очевидно, что  $\zeta = 0$  является корнем для (83) тогда и только тогда, когда  $p = -1$ .

Приступим к изучению случая “общего положения”. Поскольку при  $p \in \mathbb{R}$  корни уравнения (83) могут быть только вещественными, всюду в данном пункте считаем, что  $\zeta \in \mathbb{R}$ . Используем замену

$$\zeta = \begin{cases} 4\mu^2, & \mu \geq 0, \\ -4\mu^2, & \mu < 0. \end{cases} \tag{84}$$

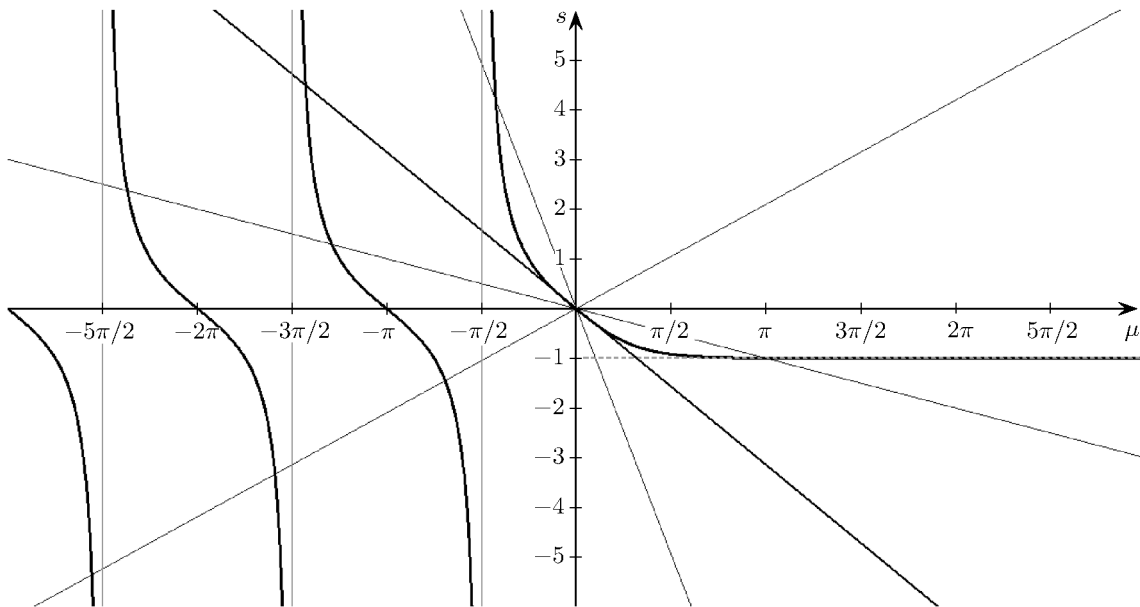
Формула (84) задаёт биекцию, точнее, сохраняющий порядок изотонный гомеоморфизм между прямой  $\mu \in \mathbb{R}$  и прямой  $\zeta \in \mathbb{R}$ . Уравнение (83) перейдёт в эквивалентную совокупность

$$\begin{cases} (1/\mu) \operatorname{sh} \mu + p \operatorname{ch} \mu = 0, & \mu \geq 0, \\ (1/\mu) \sin \mu + p \cos \mu = 0, & \mu < 0. \end{cases} \tag{85}$$

Вновь видим, что корень  $\mu = 0$ , соответствующий прежнему  $\zeta = 0$ , возникает в (85) лишь при  $p = -1$ . Временно исключим  $\mu = 0$  из рассмотрения. Тогда при  $\mu \in (-\infty, 0) \cup (0, +\infty)$  совокупность (85) приводится к следующему эквивалентному виду:

$$\begin{cases} p\mu = -\operatorname{th} \mu, & \mu > 0, \\ p\mu = -\operatorname{tg} \mu, & \mu < 0. \end{cases} \tag{86}$$

Полученные уравнения проанализируем графически (рисунок).



**Рисунок.** Геометрический анализ совокупности уравнений (86).

Введём вспомогательную функцию  $\Phi(\mu)$ , составленную из счётного числа ветвей  $\varphi_k(\mu)$  следующим образом. Особая нулевая ветвь  $\varphi_0(\mu)$  определяется на множестве  $(-\pi/2, +\infty)$  по формулам  $\varphi_0(\mu) = -\operatorname{tg} \mu$  при  $\mu \in (-\pi/2, 0)$  и  $\varphi_0(\mu) = -\operatorname{th} \mu$  при  $\mu \in [0, +\infty)$ . Ясно, что это непрерывная (точнее, непрерывно дифференцируемая), строго убывающая, выпуклая вниз функция на  $(-\pi/2, +\infty)$ . Затем, при  $k \in \mathbb{N}$ , положим  $\varphi_k(\mu) = -\operatorname{tg} \mu$  на соответствующих промежутках  $\mu \in (-(2k+1)\pi/2, -(2k-1)\pi/2)$ .

Рассмотрим графики функций  $s = \Phi(\mu)$  и  $s = p\mu$ . Абсциссы их точек пересечения на множестве  $\mu \in (-\infty, 0) \cup (0, +\infty)$  как раз дадут все корни, возникающие в (86). Для того чтобы вернуться к исходной совокупности (85), добавим корень  $\mu = 0$ , отвечающий единственному значению параметра  $p = -1$ . При таком  $p$  секущая  $s = p\mu$  к нулевой ветви  $s = \varphi_0(\mu)$  перейдет в прямую  $s = -\mu$ , касающуюся этой ветви в точке  $\mu = 0$ .

В результате получаем счётное множество корней  $\mu_k = \mu_k(p)$ , непрерывно зависящих от параметра  $p \in \mathbb{R}$ . Точнее, особый корень  $\mu_0 = \mu_0(p)$  определён лишь при  $p \in (-\infty, 0)$  — он соответствует ветви  $\varphi_0(\mu)$  и непрерывно возрастает с ростом  $p$ . Так, при  $p \in (-\infty, -1)$  этот корень  $\mu_0 = \mu_0(p)$  изменяется в промежутке  $(-\pi/2, 0)$  как корень уравнения  $p\mu = -\operatorname{tg} \mu$ . Затем при  $p = -1$  возникает промежуточное значение  $\mu_0(0) = 0$ . Наконец, при  $p \in (-1, 0)$  корень  $\mu_0 = \mu_0(p)$  изменяется в промежутке  $(0, +\infty)$  уже как корень уравнения  $p\mu = -\operatorname{th} \mu$  с асимптотикой

$$\mu_0(p) \sim -1/p, \quad p \rightarrow 0 - 0. \tag{87}$$

Асимптотика (87) устанавливается из простых соображений и даёт отличное приближение уже при  $p \geq -1/\pi$ . Действительно, при таких  $p$  точка пересечения графиков  $s = p\mu$  и  $s = -\operatorname{th} \mu$  будет расположена очень близко к асимптоте  $s = -1$  графика  $s = -\operatorname{th} \mu$ . Следовательно, уравнение  $p\mu = -\operatorname{th} \mu$  с высокой точностью выражается как  $p\mu = -1$ .

Остальные “регулярные” корни  $\mu_k = \mu_k(p)$  с индексом  $k \in \mathbb{N}$  определены при всех  $p \in \mathbb{R}$ . Точнее, при фиксированном  $k \in \mathbb{N}$  каждый такой корень соответствует ветви  $\varphi_k(\mu) = -\operatorname{tg} \mu$  и непрерывно возрастает с ростом  $p$  в пределах от  $-(2k+1)\pi/2$  до  $-(2k-1)\pi/2$ . В частности, при  $p = 0$  получаем элементарные значения  $\mu_k(0) = -k\pi$ , разделяющие основные нужные нам случаи  $\mu_k(p) \in (-(2k+1)\pi/2, -k\pi)$  при  $-\infty < p < 0$  и  $\mu_k(p) \in (-k\pi, -(2k-1)\pi/2)$  при  $0 < p < +\infty$  (см. рисунок).

Таким образом, описание корней для совокупности (85) получено. Теперь, чтобы вернуться к исходному уравнению (83), используем замену (84). Следующий результат есть прямое следствие проведённых выше рассуждений.

**Теорема 7.** Пусть  $p \in \mathbb{R}$ . Тогда все корни уравнения (83) принадлежат  $\mathbb{R}$  и при анализе их расположения на прямой надо различать три случая.

1. Если  $p < 0$ , то корни уравнения (83) представимы в виде

$$-\infty < \dots < \zeta_{k+1} < \zeta_k < \dots < \zeta_1 < \zeta_0. \tag{88}$$

При любом фиксированном  $k \in \mathbb{N}$  корень  $\zeta_k = \zeta_k(p)$  строго возрастает с ростом  $p \in (-\infty, 0)$ , непрерывно двигаясь от левой до правой границы в интервале  $(-(2k+1)^2\pi^2, -4k^2\pi^2)$ . Кроме того, имеется особый корень  $\zeta_0 = \zeta_0(p)$ , который строго возрастает в интервале  $(-\pi^2, +\infty)$ , непрерывно двигаясь от левой до правой границы так, что

$$-\pi^2 < \zeta_0(p) < 0 \quad \text{при } p \in (-\infty, -1), \quad \zeta_0(p) = 0 \quad \text{при } p = -1, \quad \zeta_0(p) > 0 \quad \text{при } p \in (-1, 0)$$

с асимптотикой  $\zeta_0(p) \sim 4/p^2$  при  $p \rightarrow 0 - 0$ .

2. Если  $p = 0$ , то корни уравнения (83) принимают вид  $\zeta_k = -4k^2\pi^2$  при  $k \in \mathbb{N}$ .

3. Если  $p > 0$ , то корни уравнения (83) представимы в виде

$$-\infty < \dots < \zeta_{k+1} < \zeta_k < \dots < \zeta_1. \tag{89}$$

При любом фиксированном  $k \in \mathbb{N}$  корень  $\zeta_k = \zeta_k(p)$  строго возрастает с ростом  $p \in (0, +\infty)$ , непрерывно двигаясь от левой до правой границы в интервале  $(-4k^2\pi^2, -(2k-1)^2\pi^2)$ . Как следствие отсюда получаем, что при всех  $p > 0$  корни (89) будут отрицательными, строго меньшими числа  $(-\pi^2)$ .

Для того чтобы завершить картину и получить нули серии (72) для характеристической функции (17) остаётся разделить на  $T^2$  числа (88) при  $p < 0$  или числа (89) при  $p > 0$  с их описанием, указанным в теореме 7. Нетрудно убедиться при этом, что нули (71) и нули (72) в теореме 3 при фиксированном  $p \in (-\infty, 0) \cup (0, +\infty)$  обладают следующим свойством перемежаемости: между двумя любыми последовательными нулями серии (72) находится ровно один нуль серии (71). Точнее, при  $p < 0$  имеем

$$\lambda_k^{(2)} = \zeta_k/T^2 < \lambda_k^{(1)} = -4k^2\pi^2/T^2 < \lambda_{k-1}^{(2)} = \zeta_{k-1}/T^2, \quad k \in \mathbb{N},$$

а при  $p > 0$  имеем

$$\lambda_{k+1}^{(2)} = \zeta_{k+1}/T^2 < \lambda_k^{(1)} = -4k^2\pi^2/T^2 < \lambda_k^{(2)} = \zeta_k/T^2, \quad k \in \mathbb{N},$$

с учётом нумерации, введённой в теореме 7.

Отметим ещё несколько следствий, важных с практической точки зрения. Напомним, что переход от параметра  $p$  к изначальным  $\alpha$ ,  $\beta$  и  $T$  происходит по формуле  $p = 2\beta/(\alpha T)$  (см. (70)). Утверждая, что нули функции (17) локализуются в каких-то интервалах, мы имеем в виду, что вне этих интервалов нулей нет совсем, но в каждый из интервалов попадает хотя бы один или даже два таких нуля. Слово “интервал” понимаем здесь в расширенном смысле – как любой связный промежуток вещественной оси.

**Теорема 8.** Для характеристической функции  $L(\lambda) = L(\lambda; T, \alpha, \beta)$  из формулы (17) выделим следующие типичные возможности.

1. При любом выборе параметров  $\alpha > 0$ ,  $\beta > 0$  и  $T > 0$  все нули функции (17) являются вещественными и отрицательными, строго меньшими числа  $(-\pi^2/T^2)$ . Точнее, эти нули локализируются в интервалах  $[-4k^2\pi^2/T^2, -(2k-1)^2\pi^2/T^2]$  при  $k \in \mathbb{N}$ .

2. Также отрицательными будут все нули функции (17), если параметры  $\alpha, \beta \in \mathbb{R} \setminus \{0\}$  и  $T > 0$  связаны условием

$$-\infty < \beta/\alpha < -T/2. \quad (90)$$

Точнее, тогда нули локализируются в интервалах  $(-\pi^2/T^2, 0)$  и  $(-(2k+1)^2\pi^2/T^2, -4k^2\pi^2/T^2]$  при  $k \in \mathbb{N}$ .

3. Далее, если параметры  $\alpha, \beta \in \mathbb{R} \setminus \{0\}$  и  $T > 0$  связаны специальным условием

$$\beta/\alpha = -T/2 \quad \sim \quad \alpha T + 2\beta = 0, \quad (91)$$

то характеристическая функция (17) имеет простой нуль в точке  $\lambda = 0$  и остальные нули, локализованные в интервалах  $(-(2k+1)^2\pi^2/T^2, -4k^2\pi^2/T^2]$  при  $k \in \mathbb{N}$ .

4. Если, наконец, параметры  $\alpha, \beta \in \mathbb{R} \setminus \{0\}$  и  $T > 0$  связаны условием

$$-T/2 < \beta/\alpha < 0, \quad (92)$$

то характеристическая функция (17) имеет один простой нуль на луче  $(0, +\infty)$  и остальные нули, локализованные в интервалах  $(-(2k+1)^2\pi^2/T^2, -4k^2\pi^2/T^2]$  при  $k \in \mathbb{N}$ .

**Доказательство.** Перечисленные результаты получаются простым сравнением базовой теоремы 3 и теоремы 7 (с учётом также теорем 5 и 6). Условия (90)–(92) соответствуют случаям  $p \in (-\infty, -1)$ ,  $p = -1$ ,  $p \in (-1, 0)$ , разобранным в теореме 7 и отнесённым сейчас к параметру  $p = 2\beta/(\alpha T)$ . Других обоснований не требуется.

Информацию, заложенную в теорему 8, можно уточнять *ad infinitum* путём дополнительно изучения корней трансцендентных уравнений (86). Подобные уравнения известны в анализе, начиная с Эйлера [33; с. 300–301] (см. также [34; с. 170–171]); ряд результатов представлен даже в учебной литературе (см. [35; с. 218, 453]). Полученного, впрочем, достаточно для многих содержательных выводов, связанных с исходной обратной задачей (1)–(3).

**10. Признаки единственности и неединственности решения.** Для сокращения формулировок рассматриваем обратную задачу (1)–(3) в её однородной версии (1), (6). Оператор  $A$  по-прежнему считаем линейным и замкнутым. Начнём с результатов абсолютного характера, когда единственность ослабленного решения обратной задачи гарантирована независимо от выбора значения  $T > 0$ .

**Теорема 9.** Пусть у оператора  $A$  нет вещественных собственных значений. Тогда при любом выборе параметров  $\alpha, \beta \in \mathbb{R} \setminus \{0\}$  и  $T > 0$  однородная обратная задача (1), (6) имеет только тривиальное решение  $u(t) \equiv 0$ ,  $g = 0$ .

**Доказательство** непосредственно следует из теорем 2 и 6.

**Теорема 10.** Пусть у оператора  $A$  нет собственных значений на луче  $(-\infty, 0) \subset \mathbb{R}$ . Тогда при любом выборе параметров  $\alpha > 0$ ,  $\beta > 0$  и  $T > 0$  однородная обратная задача (1), (6) имеет только тривиальное решение  $u(t) \equiv 0$ ,  $g = 0$ .

**Доказательство** следует из теоремы 2 и части 1 теоремы 8.

Теоремы 9 и 10 охватывают множество ситуаций. Так, под действие теоремы 10 можно подвести линейные обратные задачи о нахождении источника для эллиптических уравнений в цилиндрических областях, как в работах [7, 36, 37].

Укажем теперь два результата условного характера – когда единственность решения обратной задачи зависит от выбора финального значения  $T > 0$ .

**Теорема 11.** Пусть у оператора  $A$  нет собственных значений на луче  $(-\infty, -b) \subset \mathbb{R}$  с некоторым  $b > 0$ . Тогда при любом выборе параметров  $\alpha > 0$ ,  $\beta > 0$  и  $T \in (0, \pi/\sqrt{b}]$  однородная обратная задача (1), (6) имеет только тривиальное решение  $u(t) \equiv 0$ ,  $g = 0$ .

**Доказательство.** Так как  $T \in (0, \pi/\sqrt{b}]$ , то  $-\pi^2/T^2 \leq -b$ . Тем самым, по теореме 8, часть 1, все нули характеристической функции (17) попадают на луч  $(-\infty, -b)$ , где нет собственных значений оператора  $A$ . Применяя теорему 2, получаем нужный результат.

**Теорема 12.** Пусть у оператора  $A$  нет собственных значений на луче  $(-\infty, 0) \subset \mathbb{R}$ . Пусть параметры  $\alpha, \beta \in \mathbb{R} \setminus \{0\}$  взяты разных знаков. Тогда при любом  $T$  из интервала

$$0 < T < -2\beta/\alpha \quad (93)$$

однородная обратная задача (1), (6) имеет только тривиальное решение  $u(t) \equiv 0$ ,  $g = 0$ .

**Доказательство.** Из условия (93) следует условие (90). Поэтому, применяя часть 2 теоремы 8, заключаем, что нули характеристической функции (17) попадают на луч  $(-\infty, 0)$ , где нет собственных значений оператора  $A$ . В итоге по теореме 2 получаем нужный результат.

Можно дать много других, более специальных, утверждений, связанных с непопаданием собственных значений оператора  $A$  на интервалы локализации нулей из теоремы 8. Не будем перегружать работу такими формулировками. Отметим ещё два результата противоположного характера, когда единственность решения обратной задачи заведомо нарушается.

**Теорема 13.** Пусть оператор  $A$  имеет отрицательное собственное значение  $\lambda_0 = -\gamma^2$ , где  $\gamma > 0$ . Пусть  $T = 2k\pi/\gamma > 0$  с некоторым  $k \in \mathbb{N}$ . Тогда при всех  $\alpha, \beta \in \mathbb{C} \setminus \{0\}$  однородная обратная задача (1), (6) имеет нетривиальное решение

$$u(t) = \frac{T^2}{4k^2\pi^2} \left( 1 - \cos \frac{2k\pi t}{T} \right) f_0, \quad g = f_0 \quad (94)$$

с собственным вектором  $f_0 \neq 0$  таким, что  $f_0 \in D(A)$  и  $Af_0 = -\gamma^2 f_0 = -(4k^2\pi^2/T^2)f_0$ .

**Доказательство.** По теореме 3 собственное значение  $\lambda_0 = -\gamma^2 = -4k^2\pi^2/T^2$  попадает в серию (71) и является нулём характеристической функции (17). Но тогда, согласно лемме 1, нетривиальное решение задачи (1), (6) можно построить по формуле (19). Конкретная реализация этой формулы даёт пару (94). Непосредственная подстановка в задачу подтверждает, что соотношения (1), (6) выполнены при всех  $\alpha, \beta \in \mathbb{C} \setminus \{0\}$ . Теорема доказана.

В простых “гиперболических” примерах для уравнения колебаний струны оператор  $A$  обладает серией собственных значений  $\lambda_m = -(m\pi/l)^2$  при  $m \in \mathbb{N}$ , где  $l > 0$  – длина струны. Тогда, как следует из теоремы 13, обратная задача (1), (6) имеет нетривиальные решения вида (94) в любой момент времени  $T = T_{k,m} \equiv 2l(k/m)$ , где  $k, m \in \mathbb{N}$ . Указанные моменты образуют счётное всюду плотное множество на луче  $0 < T < \infty$ , т.е. этот “гиперболический” случай в обратных задачах подобен по своей некорректности так называемой задаче Дирихле для уравнения колебаний струны (см. [38]; ср. с [1; с. 140–143] и [15; с. 1619–1620]).

Завершим следующим обособленным результатом.

**Теорема 14.** Пусть оператор  $A$  имеет собственное значение  $\lambda_0 = 0$ , и  $Af_0 = 0$  с элементом  $f_0 \in D(A)$ ,  $f_0 \neq 0$ . Предположим, что параметры  $\alpha, \beta \in \mathbb{R} \setminus \{0\}$  и  $T > 0$  связаны условием  $\alpha T + 2\beta = 0$ . Тогда однородная обратная задача (1), (6) имеет нетривиальное решение вида

$$u(t) = (t^2/2)f_0, \quad g = f_0. \quad (95)$$

**Доказательство.** По теореме 8 с условием (91) собственное значение  $\lambda_0 = 0$  является нулём характеристической функции (17). Воспользуемся формулой (19) вместе с уточнением, указанным в (11) для  $\lambda = 0$ . Получим пару (95). Проверка соотношений (1), (6) здесь не представляет труда. Теорема доказана.

Ясно, что когда много собственных значений оператора  $A$  совпадает с нулями характеристической функции (17), мы можем комбинировать возникающие пары вида (19), осуществляя синтез все более сложных решений однородной обратной задачи (1), (6). В совсем уже специальных случаях, когда характеристическая функция (17) имеет кратный нуль кратности два (см. теорему 4), возможны примеры, где кроме элементарных решений вида (19) появляются присоединённые элементарные решения  $(u(t), g)$ , учитывающие наличие у оператора  $A$  не только собственных векторов из  $D(A)$ , но и присоединённых векторов из  $D(A^2)$  (см. [26]).

Авторы признательны А.Б. Костину, А.И. Прилепко и Ю.С. Эйдельману за проявленный интерес и поддержку наших исследований. Особые благодарности А.Ю. Попову – за ценные консультации по теории целых функций и В.Б. Шерстюкову – за научные контакты.

## СПИСОК ЛИТЕРАТУРЫ

1. *Денисов А.М.* Введение в теорию обратных задач. М., 1994.
2. *Prilepko A.I., Orlovsky D.G., Vasin I.A.* Methods for Solving Inverse Problems in Mathematical Physics. New York; Basel, 2000.
3. *Красносельский М.А., Забрейко П.П., Пустыльник Е.И., Соболевский П.Е.* Интегральные операторы в пространствах суммируемых функций. М., 1966.
4. *Тихонов И.В., Эйдельман Ю.С.* Обратная задача для дифференциального уравнения в банаховом пространстве и распределение нулей целой функции типа Миттаг-Лёффлера // Дифференц. уравнения. 2002. Т. 38. № 5. С. 637–644.
5. *Алмохамед М.* Критерий единственности решения в линейной обратной задаче с финальным переопределением второго рода // Вестн. Воронежского гос. ун-та. Сер. Физика. Математика. 2019. № 3. С. 50–58.
6. *Прилепко А.И.* Обратные задачи теории потенциала (эллиптические, параболические, гиперболические уравнения и уравнение переноса) // Мат. заметки. 1973. Т. 14. № 5. С. 755–767.
7. *Прилепко А.И.* Избранные вопросы в обратных задачах математической физики // Условно-корректные задачи мат. физики и анализа. Новосибирск, 1992. С. 151–162.
8. *Искендеров А.Д., Тагиев Р.Г.* Обратная задача об определении правых частей эволюционных уравнений в банаховом пространстве // Вопросы прикл. мат. и киберн. Науч. тр. Азербайджанского ун-та. 1979. № 1. С. 51–56.
9. *Эйдельман Ю.С.* Единственность решения обратной задачи для дифференциального уравнения в банаховом пространстве // Дифференц. уравнения. 1987. Т. 23. № 9. С. 1647–1649.
10. *Тихонов И.В., Эйдельман Ю.С.* Единственность решения двухточечной обратной задачи для абстрактного дифференциального уравнения с неизвестным параметром // Дифференц. уравнения. 2000. Т. 36. № 8. С. 1132–1133.
11. *Тихонов И.В.* Обобщённая задача Уорда для абстрактных дифференциальных уравнений // Дифференц. уравнения. 2005. Т. 41. № 3. С. 325–336.
12. *Тихонов И.В., Алмохамед М.* Об одной обратной задаче для дифференциального уравнения высокого порядка в банаховом пространстве // Некоторые актуальные проблемы современной математики и математического образования. СПб., 2019. С. 91–95.
13. *Амиров А.Х.* О разрешимости обратных задач для уравнения второго порядка // Функц. анализ и его прил. 1986. Т. 20. Вып. 3. С. 80–81.
14. *Орловский Д.Г.* Об одной обратной задаче для дифференциального уравнения второго порядка в банаховом пространстве // Дифференц. уравнения. 1989. Т. 25. № 6. С. 1000–1009.
15. *Орловский Д.Г.* К задаче определения параметра эволюционного уравнения // Дифференц. уравнения. 1990. Т. 26. № 9. С. 1614–1621.
16. *Тихонов И.В.* Структурные свойства нуль-решений абстрактной задачи Коши // Интегральные преобразования и специальные функции. Информ. бюллетень. 2002. Т. 3. № 1. С. 22–38.
17. *Тихонов И.В., Алмохамед М.* Обобщённые экспоненты и их применение в теории дифференциальных уравнений // Системы компьютерной математики и их приложения. Вып. 21. Смоленск, 2020. С. 345–353.
18. *Попов А.Ю., Седлецкий А.М.* Распределение корней функций Миттаг-Лёффлера // Совр. математика. Фунд. направления. 2011. Т. 40. С. 3–171.
19. *Левин Б.Я.* Распределение корней целых функций. М., 1956.
20. *Леонтьев А.Ф.* Целые функции. Ряды экспонент. М., 1983.
21. *Хилле Э., Филлипс Р.* Функциональный анализ и полугруппы. М., 1962.
22. *Полиа Г., Сеге Г.* Задачи и теоремы из анализа. Часть 2. Теория функций. Распределение нулей. Полиномы. Определители. Теория чисел. М., 1978.
23. *Müntz Ch.H.* Über den Approximationssatz von Weierstraß // Mathematische Abhandlungen Hermann Amandus Schwarz zu seinem Fünfzigjährigen Doctorjubiläum. Berlin; Heidelberg, 1914. S. 303–312.
24. *Качмаж С., Штейнгауз Г.* Теория ортогональных рядов. М., 1958.
25. *Masayoshi Hata.* Problems and Solutions in Real Analysis. Series on Number Theory and Its Applications. V. 4. New Jersey; London; Singapore, 2007.
26. *Алмохамед М., Тихонов И.В.* О некоторых спектральных исследованиях, связанных с теорией обратных задач // Совр. проблемы теории функций и их приложения: материалы 21-й междунар. Саратовской зимней школы. Саратов, 2022. С. 20–26.
27. *Тихонов И.В., Шерстюков В.Б., Алмохамед М.* О некоторых трансцендентных уравнениях, важных для математической физики // Совр. проблемы теории функций и их приложения: материалы 21-й междунар. Саратовской зимней школы. Саратов, 2022. С. 294–299.

28. *Fadle J.* Die Selbstspannungs-Eigenwertfunktionen der quadratischen Scheibe // Ingenieur-Archiv ( $\equiv$  Archive of Applied Mechanics). 1940. Bd. 11. S. 125–149.
29. *Buchwald V.T.* Eigenfunctions of plane elastostatics. I. The strip // Proceedings of the Royal Society of London. Ser. A. Math. and Phys. Sci. 1964. V. 277. № 1370. P. 385–400.
30. *Katopodes F.V., Davis A.M.J., Stone H.A.* Piston flow in two-dimensional channel // Physics of Fluids. 2000. V. 12. № 5. P. 1240–1243.
31. *Hardy G.H.* On the zeroes of integral function  $x - \sin x = \sum_1^{\infty} (-1)^{n-1} \frac{x^{2n+1}}{2n+1!}$  // The Messenger of Math. 1902. V. 31. № 11. P. 161–165.
32. *Маркушевич А.И.* Целые функции. Элементарный очерк. М., 1975.
33. *Эйлер Л.* Введение в анализ бесконечных. Т. 2. М., 1961.
34. *Comtet L.* Advanced Combinatorics. The Art of Finite and Infinite Expansions. Dordrecht; Holland, 1974.
35. *Кудрявцев Л.Д., Кутасов А.Д., Чехлов В.И., Шабунин М.И.* Сборник задач по математическому анализу. Т. 1. Предел. Непрерывность. Дифференцируемость. М., 2003.
36. *Соловьев В.В.* Разрешимость обратных задач для эллиптического уравнения в цилиндре // Вестн. Московского гос. обл. ун-та. Сер. Физика. Математика. 2012. № 1. С. 27–38.
37. *Альмохамед М.* Восстановление правой части в уравнении Пуассона при помощи специальных краевых условий // Совр. проблемы теории функций и их приложения: материалы 20-й междунар. Саратовской зимней школы. Саратов, 2020. С. 30–32.
38. *Bourgoin D.G., Duffin R.* The Dirichlet problem for the vibrating string equation // Bull. of the Amer. Math. Soc. 1939. V. 45. № 12. Part 1. P. 851–858.

Московский государственный университет  
имени М.В. Ломоносова,  
Московский педагогический государственный  
университет,  
Университет Алеппо,  
г. Алеппо, Сирия

Поступила в редакцию 15.02.2022 г.  
После доработки 15.02.2022 г.  
Принята к публикации 25.05.2022 г.

УДК 519.633.2

## ЧИСЛЕННОЕ РЕШЕНИЕ ЗАДАЧИ КОШИ ДЛЯ ИНТЕГРО-ДИФФЕРЕНЦИАЛЬНОГО УРАВНЕНИЯ ВТОРОГО ПОРЯДКА

© 2022 г. П. Н. Вабищевич

В конечномерном гильбертовом пространстве рассматривается задача Коши для эволюционного интегро-дифференциального уравнения второго порядка с памятью, подынтегральное выражение в котором представляет собой произведение разностного ядра и линейного оператора от производной решения по времени. Основные трудности в нахождении приближённого значения решения таких нелокальных задач в данный момент времени обусловлены необходимостью работы с приближёнными значениями решения для всех предыдущих моментов времени. Предложена трансформация рассматриваемого интегро-дифференциального уравнения к системе эволюционных слабо связанных между собой локальных уравнений. Она основывается на аппроксимации разностного ядра суммой экспонент. Формулируется локальная задача для слабо связанной системы уравнений с дополнительными обыкновенными дифференциальными уравнениями. Для решения соответствующей задачи Коши приведены оценки устойчивости решения по начальным данным и правой части. Основное внимание уделяется построению и исследованию устойчивости трёхслойных разностных схем и их вычислительной реализации.

DOI: 10.31857/S0374064122070044, EDN: CEFRSC

**Введение.** При математическом моделировании нестационарных процессов наиболее часто используются параболические и гиперболические уравнения второго порядка [1, гл. I; 2, гл. 2]. В последнее время всё возрастающее внимание уделяется уравнениям, которые частично наследуют свойства как параболических, так и гиперболических уравнений. Характерным примером могут служить эволюционные интегро-дифференциальные уравнения [3, гл. 3; 4, гл. 1]. Наиболее важной особенностью таких уравнений является их нелокальность: решение в текущий момент времени зависит от всей предыстории процесса.

Мы выделяем два класса эволюционных интегро-дифференциальных уравнений с памятью. Первый из них характеризуется нелокальностью решения, а второй – нелокальностью производной решения по времени, т.е. когда подынтегральное выражение в уравнении зависит только соответственно от решения или от производной решения по времени. Эти классы уравнений называются *уравнениями с памятью решений* и *уравнениями с памятью производной решения* по времени соответственно. При рассмотрении динамических процессов вязкоупругости [5, гл. I; 6, гл. 2] типичными являются эволюционные интегро-дифференциальные уравнения второго порядка с памятью производной решения по времени. Отметим также, что для задач с разностным ядром мы имеем возможность перейти от одного из указанных классов к другому при помощи введения другого разностного ядра. В силу этой возможности те же уравнения вязкоупругости могут (см., например, [1, гл. I]) записываться как интегро-дифференциальные уравнения с памятью решения, а не с памятью производной решения по времени.

Приближённое решение краевых задач для уравнений с памятью проводится с использованием стандартных конечно-элементных или конечно-объёмных аппроксимаций по пространству [7, гл. 3, 6; 8, гл. 3]. Мы приходим к задаче Коши для операторных уравнений с памятью в соответствующем конечномерном гильбертовом пространстве. Основное внимание необходимо уделить проблемам выбора аппроксимаций по времени. При численном решении задач для эволюционных уравнений первого порядка с памятью решения естественно ориентироваться [9, гл. 4] на использование тех или иных квадратур для интегрального члена и обычных двухслойных аппроксимаций производной по времени. Такие исследования неявной схемы Эйлера и схемы Кранка–Николсон выполнены, например, в работах [10, 11]. Соответствующие



квадратурные формулы для интегрального члена используются и при рассмотрении интегро-дифференциальных уравнений с памятью производной решения по времени.

Стандартные вычислительные алгоритмы приближённого решения задачи Коши для интегро-дифференциальных уравнений с памятью связаны с необходимостью работать со значениями решения для предшествующих моментов времени. Наиболее интересные возможности уменьшения вычислительной работы представят подходы с переходом от нелокальной задачи к локальной, тогда, в частности, требования к объёму используемой памяти существенно уменьшаются.

Для интегральных уравнений Вольтерры хорошо известный (см., например, [12, гл. 6]) переход к более простым в плане вычислительной реализации задачам обеспечивается выбором специальных аппроксимаций разностного ядра. Выделим как наиболее перспективную аппроксимацию разностного ядра суммой экспонент. При такой аппроксимации мы приходим к системе локальных слабо связанных между собой эволюционных уравнений. Возможности данного подхода при приближённом решении задачи Коши для интегро-дифференциального уравнения первого порядка с памятью решения рассмотрены в работе [13]. Подобное исследование для задач с памятью производной решения по времени выполнено в работе [14].

В настоящей работе в вещественном конечномерном гильбертовом пространстве рассмотрена задача Коши для эволюционного интегро-дифференциального уравнения Вольтерры второго порядка с памятью производной с положительно определённым самосопряжённым оператором. При аппроксимации разностного ядра суммой экспонент выполнена трансформация нелокальной задачи для уравнения с памятью в локальную систему уравнений. Получены соответствующие априорные оценки для решения задачи Коши, гарантирующие устойчивость решения по начальным данным и правой части. Предложены и исследованы на устойчивость трёхслойные разностные схемы для системы уравнений, которые имеют удобную вычислительную реализацию.

**1. Постановка задачи.** Рассматривается задача Коши для эволюционного уравнения второго порядка с памятью производной решения в вещественном конечномерном гильбертовом пространстве  $H$ . Функция  $u(t)$  удовлетворяет интегро-дифференциальному уравнению с разностным ядром

$$\frac{d^2 u}{dt^2} + B \frac{du}{dt} + \int_0^t k(t-s)C \frac{du}{ds}(s) ds + Au = f(t), \quad t > 0, \tag{1}$$

и начальным условиям

$$u(0) = u_0, \quad \frac{du}{dt}(0) = w_0. \tag{2}$$

Линейные постоянные (не зависящие от  $t$ ) операторы  $A$ ,  $B$ ,  $C$  являются самосопряжёнными и положительно определёнными, т.е.

$$A = A^* \geq \nu_A I, \quad \nu_A > 0, \quad B = B^* \geq \nu_B I, \quad \nu_B > 0, \quad C = C^* \geq \nu_C I, \quad \nu_C > 0, \tag{3}$$

где  $I$  – единичный оператор в  $H$ . Будем использовать обычные обозначения  $(\cdot, \cdot)$  и  $\|\cdot\|$  для скалярного произведения и нормы в  $H$ . Для самосопряжённого положительно определённого оператора  $D$  определим гильбертово пространство  $H_D$ , задав в векторном пространстве  $H$  скалярное произведение и норму  $(u, v)_D = (Du, v)$  и  $\|u\|_D = (u, v)_D^{1/2}$ .

Как и при рассмотрении эволюционных уравнений первого порядка с памятью решения [10, 11, 13] и памятью производной решения [14], считаем, что ядро  $k(t)$  положительно определено. В этом случае для всех  $T > 0$  ядро  $k(t)$  принадлежит пространству  $L_1(0, T)$  и имеет место неравенство

$$\int_0^T \psi(t) \int_0^t k(t-s)\psi(s) ds dt \geq 0 \quad \text{для любой } \psi \in C[0, T]. \tag{4}$$

Отметим также [15] достаточное условие положительной определённости ядра  $k(t)$ :

$$k(t) \geq 0, \quad \frac{dk}{dt}(t) \leq 0, \quad \frac{d^2k}{dt^2}(t) \geq 0, \quad t > 0. \tag{5}$$

В своем исследовании основное внимание мы уделяем получению системы локальных эволюционных уравнений, задача Коши для которой даёт приближённое решение задачи (1), (2). При исследовании устойчивости разностных аппроксимаций по времени для нас ориентиром является следующее утверждение.

**Теорема 1.** Пусть операторы  $A, B, C$  удовлетворяют условиям (3) и  $k(t)$  – положительно определённое ядро. Тогда для решения задачи (1), (2) справедлива следующая оценка устойчивости по начальным данным и правой части:

$$\left\| \frac{du}{dt}(t) \right\|^2 + \|u(t)\|_A^2 \leq \|w_0\|^2 + \|u_0\|_A^2 + \frac{1}{2} \int_0^t \|f(s)\|_{B^{-1}}^2 ds, \quad t > 0. \tag{6}$$

**Доказательство.** Домножим уравнение (1) скалярно в  $H$  на  $du(t)/dt$  и с учётом положительной определённости операторов  $A, B, C$  придём к неравенству

$$\frac{1}{2} \frac{d}{dt} \left( \left\| \frac{du}{dt}(t) \right\|^2 + \|u(t)\|_A^2 \right) + \int_0^t k(t-s) \left( C^{1/2} \frac{du}{ds}(s), C^{1/2} \frac{du}{dt}(t) \right) ds \leq \frac{1}{4} \|f(t)\|_{B^{-1}}^2.$$

Интегрируя его по  $(0, T)$ , получаем

$$\begin{aligned} & \frac{1}{2} \left( \left\| \frac{du}{dt}(T) \right\|^2 + \|u(T)\|_A^2 - \left\| \frac{du}{dt}(0) \right\|^2 - \|u(0)\|_A^2 \right) + \\ & + \int_0^T \int_0^t k(t-s) \left( C^{1/2} \frac{du}{ds}(s), C^{1/2} \frac{du}{dt}(t) \right) ds dt \leq \frac{1}{4} \int_0^T \|f(t)\|_{B^{-1}}^2 dt. \end{aligned}$$

Принимая во внимание неравенство (4) и начальное условие (2), убеждаемся в справедливости оценки (6). Теорема доказана.

Нелокальный член в уравнении (1) становится локальным в двух важных случаях: когда ядро  $k(t)$  постоянно и когда ядро является  $\delta$ -функцией. Мы можем выделить такие условия отдельно, заменяя  $k(t)$  выражением

$$k(t) \rightarrow \gamma_1 + \gamma_2 \delta(t) + k(t), \quad \gamma_1 > 0, \quad \gamma_2 > 0. \tag{7}$$

Это соответствует заменам

$$B \rightarrow B + \gamma_2 C, \quad A \rightarrow A + \gamma_1 C, \quad f(t) \rightarrow f(t) + \gamma_1 C u_0 \tag{8}$$

в уравнении (1). Таким образом, мы остаёмся в классе рассматриваемых задач (1)–(3).

**2. Локальная система уравнений.** В вычислительном плане наибольший интерес вызывают подходы к построению приближённого решения нелокальной задачи (1), (2) с памятью производной решения по времени с помощью решений некоторых локальных эволюционных задач. Они могут строиться на основе аппроксимации ядра суммой экспонент.

Ядро  $k(t)$  аппроксимируется функцией  $\tilde{k}(t)$ , которая имеет вид

$$\tilde{k}(t) = \sum_{i=1}^m a_i \exp(-b_i t), \quad t \geq 0. \tag{9}$$

Будем считать, что для коэффициентов  $a_i, b_i, i = 1, 2, \dots, m$ , выполнены предположения об их положительности:

$$a_i > 0, \quad b_i > 0, \quad i = 1, 2, \dots, m. \tag{10}$$

С учётом условий (5) при ограничениях (10) ядро  $\tilde{k}(t)$  положительно определено.

Приближённое решение задачи (1), (2) обозначим через  $v(t)$ . Оно определяется как решение задачи Коши

$$\frac{d^2v}{dt^2} + B \frac{dv}{dt} + \int_0^t \tilde{k}(t-s)C \frac{dv}{ds}(s) ds + Av = f(t), \quad t > 0, \tag{11}$$

$$v(0) = u_0, \quad \frac{dv}{dt}(0) = w_0. \tag{12}$$

Для перехода от нелокальной задачи (11), (12) к локальной введём [12–14] функции

$$v_i(t) = \int_0^t \exp(-b_i(t-s)) \frac{dv}{ds}(s) ds, \quad i = 1, 2, \dots, m.$$

С учётом представления (9) уравнение (11) записывается в виде

$$\frac{d^2v}{dt^2} + B \frac{dv}{dt} + \sum_{i=1}^m a_i C v_i + Av = f(t). \tag{13}$$

Для вспомогательных функций  $v_i(t), i = 1, 2, \dots, m$ , имеем уравнения

$$\frac{dv_i}{dt} + b_i v_i - \frac{dv}{dt} = 0, \quad i = 1, 2, \dots, m. \tag{14}$$

Система уравнений (13), (14) дополняется начальными условиями

$$v(0) = u_0, \quad \frac{dv}{dt}(0) = w_0, \quad v_i(0) = 0, \quad i = 1, 2, \dots, m. \tag{15}$$

Уравнение (13) удобно записать в несколько другой форме. Подстановка

$$v_i = \frac{1}{b_i} \frac{dv}{dt} - \frac{1}{b_i} \frac{dv_i}{dt}, \quad i = 1, 2, \dots, m,$$

в (13) даёт

$$\frac{d^2v}{dt^2} + \left( B + \sum_{i=1}^m \frac{a_i}{b_i} C \right) \frac{dv}{dt} - \sum_{i=1}^m \frac{a_i}{b_i} C \frac{dv_i}{dt} + Av = f(t). \tag{16}$$

**Теорема 2.** Пусть операторы  $A, B, C$  удовлетворяют условиям (3). Тогда для решения задачи (14)–(16) при выполнении условия (10) справедлива следующая оценка устойчивости по начальным данным и правой части:

$$\left\| \frac{dv}{dt}(t) \right\|^2 + \|v(t)\|_A^2 + \sum_{i=1}^m a_i \|v_i(t)\|_C^2 \leq \|w_0\|^2 + \|u_0\|_A^2 + \frac{1}{2} \int_0^t \|f(s)\|_{B^{-1}}^2 ds, \quad t > 0. \tag{17}$$

**Доказательство.** Домножим уравнение (16) на  $dv(t)/dt$ , а отдельное уравнение (14) для  $i = 1, 2, \dots, m$  на  $a_i b_i^{-1} C dv_i(t)/dt$ . Это даёт равенства

$$\begin{aligned} \left\| \frac{dv}{dt} \right\|_B^2 + \frac{1}{2} \frac{d}{dt} \left( \left\| \frac{dv}{dt}(t) \right\|^2 + \|v\|_A^2 \right) + \sum_{i=1}^m \frac{a_i}{b_i} \left( C \frac{dv}{dt}, \frac{dv}{dt} \right) - \sum_{i=1}^m \frac{a_i}{b_i} \left( C \frac{dv_i}{dt}, \frac{dv}{dt} \right) &= \left( f, \frac{dv}{dt} \right), \\ \frac{a_i}{b_i} \left( C \frac{dv_i}{dt}, \frac{dv_i}{dt} \right) - \frac{a_i}{b_i} \left( C \frac{dv}{dt}, \frac{dv_i}{dt} \right) + a_i \frac{1}{2} \frac{d}{dt} \|v_i\|_C^2 &= 0, \quad i = 1, 2, \dots, m. \end{aligned}$$

Складывая их, получаем

$$\left\| \frac{dv}{dt} \right\|_B^2 + \frac{1}{2} \frac{d}{dt} \left( \left\| \frac{dv}{dt}(t) \right\|^2 + \|v\|_A^2 + \sum_{i=1}^m a_i \|v_i\|_C^2 \right) + \sum_{i=1}^m \frac{a_i}{b_i} \left\| \frac{dv}{dt} - \frac{dv_i}{dt} \right\|_C^2 = \left( f, \frac{dv}{dt} \right),$$

откуда вытекает неравенство

$$\frac{d}{dt} \left( \left\| \frac{dv}{dt}(t) \right\|^2 + \|v\|_A^2 + \sum_{i=1}^m a_i \|v_i\|_C^2 \right) \leq \frac{1}{2} \|f(t)\|_{B^{-1}}^2$$

и доказываемая оценка (17). Теорема доказана.

Нам удобно записать систему (14), (16) в форме одного векторного эволюционного уравнения второго порядка. Определим векторы  $\mathbf{v} = \{v, v_1, \dots, v_m\}$  и  $\mathbf{f} = \{f, 0, \dots, 0\}$  и запишем задачу (14)–(16) в следующем виде:

$$C \frac{d^2 \mathbf{v}}{dt^2} + B \frac{d\mathbf{v}}{dt} + A\mathbf{v} = \mathbf{f}, \tag{18}$$

$$\mathbf{v}(0) = \mathbf{v}_0, \quad C \frac{d\mathbf{v}}{dt}(0) = \mathbf{w}_0, \tag{19}$$

где  $\mathbf{v}_0 = \{u_0, 0, \dots, 0\}$  и  $\mathbf{w}_0 = \{w_0, 0, \dots, 0\}$ , а операторные матрицы  $C$ ,  $B$  и  $A$  имеют представления

$$C = \text{diag}(I, 0, \dots, 0), \quad A = \text{diag}(A, a_1 C, \dots, a_m C),$$

$$B = \begin{pmatrix} B + \sum_{i=1}^m \frac{a_i}{b_i} C & -\frac{a_1}{b_1} C & \dots & -\frac{a_m}{b_m} C \\ -\frac{a_1}{b_1} C & \frac{a_1}{b_1} C & \dots & 0 \\ \dots & \dots & \dots & \dots \\ -\frac{a_m}{b_m} C & 0 & \dots & \frac{a_m}{b_m} C \end{pmatrix}. \tag{20}$$

Задачу Коши (18), (19) мы рассматриваем на прямой сумме пространств  $\mathbf{H} = H \oplus \dots \oplus H$ , на которой скалярное произведение и норма определяются выражениями

$$(\mathbf{v}, \mathbf{w}) = (v, w) + \sum_{i=1}^m (v_i, w_i), \quad \|\mathbf{v}\| = (v, v)^{1/2},$$

где  $\mathbf{v}, \mathbf{w} \in \mathbf{H}$ . Принимая во внимание условия (3) и (10), получаем

$$C = C^* \geq 0, \quad B = B^* \geq 0, \quad A = A^* > 0. \tag{21}$$

Для доказательства оценки (17) домножим скалярно в  $\mathbf{H}$  уравнение (18) на  $d\mathbf{v}/dt$ . С учётом условий (21) и того, что  $C^2 = C$ , это даёт соотношение

$$\left( B \frac{d\mathbf{v}}{dt}, \frac{d\mathbf{v}}{dt} \right) + \frac{1}{2} \frac{d}{dt} \left( \left\| C \frac{d\mathbf{v}}{dt} \right\|^2 + \|\mathbf{v}\|_A^2 \right) = \left( \mathbf{f}, \frac{d\mathbf{v}}{dt} \right).$$

Принимая во внимание неравенства

$$\left( B \frac{d\mathbf{v}}{dt}, \frac{d\mathbf{v}}{dt} \right) \geq \left( B \frac{dv}{dt}, \frac{dv}{dt} \right), \quad \left( \mathbf{f}, \frac{d\mathbf{v}}{dt} \right) \leq \left( B \frac{dv}{dt}, \frac{dv}{dt} \right) + \frac{1}{4} (B^{-1} f, f),$$

имеем

$$\left\| C \frac{dv}{dt}(t) \right\|^2 + \|v(t)\|_A^2 \leq \|Cw_0\|^2 + \|v_0\|_A^2 + \frac{1}{2} \int_0^t \|f(s)\|_{B^{-1}} ds. \tag{22}$$

В нашем случае

$$\begin{aligned} \left\| C \frac{dv}{dt}(t) \right\|^2 &= \left\| \frac{dv}{dt}(t) \right\|^2, \quad \|v(t)\|_A^2 = \|v(t)\|_A^2 + \sum_{i=1}^m a_i \|v_i(t)\|_C^2, \\ \|Cw_0\|^2 &= \|w_0\|^2, \quad \|v_0\|_A^2 = \|u_0\|_A^2, \end{aligned}$$

поэтому из неравенства (22) вытекает оценка (17).

Вместо аппроксимации (9) может рассматриваться несколько более общий случай, как следует из (7),

$$\tilde{k}(t) = \gamma_1 + \gamma_2 \delta(t) + \sum_{i=1}^m a_i \exp(-b_i t), \quad t \geq 0.$$

Переход к рассмотренному случаю обеспечивается согласно (8).

**3. Аппроксимация по времени.** При приближённом решении задачи Коши (18), (19) будем ориентироваться на безусловно устойчивые разностные схемы. Используем равномерную сетку по времени с шагом  $\tau$  и пусть  $y^n = y(t^n)$ ,  $t^n = n\tau$ ,  $n = 0, 1, \dots$ . Трёхслойная разностная схема с весом  $\sigma = \text{const} > 0$  имеет вид

$$C \frac{y^{n+1} - 2y^n + y^{n-1}}{\tau^2} + B \frac{y^{n+1} - y^{n-1}}{2\tau} + Ay^{n+\sigma} = f^n, \quad n = 1, 2, \dots, \tag{23}$$

$$y^0 = v^0, \quad y^1 = \tilde{v}^1 \tag{24}$$

при использовании обозначений

$$y^{n+\sigma} = \sigma y^{n+1} + (1 - 2\sigma)y^n + \sigma y^{n-1}, \quad y^n = \{y^n, y_1^n, \dots, y_m^n\}.$$

Для правой части и начального условия имеем

$$f^n = \{f^n, 0, \dots, 0\}, \quad v^0 = \{u_0, 0, \dots, 0\}.$$

В схеме (23), (24) второе начальное условие (24) должно быть предварительно рассчитано. Для сохранения второго порядка на решениях задачи (14)–(16) положим, например,

$$\begin{aligned} \frac{y_i^1 - y_i^0}{\tau} + b_i \frac{y_i^1 + y_i^0}{2} - \frac{y^1 - y^0}{\tau} &= 0, \quad i = 1, 2, \dots, m, \\ \frac{y^1 - y^0}{\tau} &= w_0 + \frac{\tau}{2} \left( \frac{f^1 + f^0}{2} - \left( B + \sum_{i=1}^m \frac{a_i}{b_i} C \right) \frac{y^1 - y^0}{\tau} + \sum_{i=1}^m \frac{a_i}{b_i} C \frac{y_i^1 - y_i^0}{\tau} - A \frac{y^1 - y^0}{2} \right). \end{aligned}$$

Разностная схема (23), (24) аппроксимирует задачу (18), (19) со вторым порядком по  $\tau$  при достаточной гладкости решения  $v(t)$ . При изучении устойчивости рассматриваемых трёхслойных разностных схем используем результаты теории устойчивости (корректности) операторно-разностных схем [16, гл. 6; 17, гл. 4].

**Теорема 3.** Трёхслойная разностная схема (3), (20), (23), (24) является безусловно устойчивой при  $\sigma \geq 0.25$ . При этих ограничениях априорная оценка для приближённого решения задачи (18), (19) имеет вид

$$\begin{aligned} &\left\| \frac{y^{n+1} - y^n}{\tau} \right\|^2 + \left( \sigma - \frac{1}{4} \right) \tau^2 \left\| \frac{y^{n+1} - y^n}{\tau} \right\|_A^2 + \left\| \frac{y^{n+1} + y^n}{2} \right\|_A^2 \leq \\ &\leq \left\| \frac{\tilde{v}^1 - v^0}{\tau} \right\|^2 + \left( \sigma - \frac{1}{4} \right) \tau^2 \left\| \frac{\tilde{v}^1 - v^0}{\tau} \right\|_A^2 + \left\| \frac{\tilde{v}^1 + v^0}{2} \right\|_A^2 + \frac{1}{2} \sum_{k=1}^n \tau \|f^k\|_{B^{-1}}^2, \quad n = 1, 2, \dots \end{aligned} \tag{25}$$

**Доказательство.** С учётом тождества

$$\mathbf{y}^n = \frac{1}{4}(\mathbf{y}^{n+1} + 2\mathbf{y}^n + \mathbf{y}^{n-1}) - \frac{1}{4}(\mathbf{y}^{n+1} - 2\mathbf{y}^n + \mathbf{y}^{n-1})$$

запишем схему (23) в виде

$$\left( \mathbf{C} + \left( \sigma - \frac{1}{4} \right) \tau^2 \mathbf{A} \right) \frac{\mathbf{y}^{n+1} - 2\mathbf{y}^n + \mathbf{y}^{n-1}}{\tau^2} + \mathbf{B} \frac{\mathbf{y}^{n+1} - \mathbf{y}^{n-1}}{2\tau} + \mathbf{A} \frac{\mathbf{y}^{n+1} + 2\mathbf{y}^n + \mathbf{y}^{n-1}}{4} = \mathbf{f}^n. \quad (26)$$

При использовании новых переменных

$$\mathbf{z}^n = \frac{1}{2}(\mathbf{y}^n + \mathbf{y}^{n-1}), \quad \mathbf{w}^n = \frac{\mathbf{y}^n - \mathbf{y}^{n-1}}{\tau} \quad (27)$$

из (26) следует равенство

$$\left( \mathbf{C} + \left( \sigma - \frac{1}{4} \right) \tau^2 \mathbf{A} \right) \frac{\mathbf{w}^{n+1} - \mathbf{w}^n}{\tau} + \mathbf{B} \frac{\mathbf{w}^{n+1} + \mathbf{w}^n}{2} + \mathbf{A} \frac{\mathbf{z}^{n+1} + \mathbf{z}^n}{2} = \mathbf{f}^n,$$

умножая которое на

$$2(\mathbf{z}^{n+1} - \mathbf{z}^n) = \tau(\mathbf{w}^{n+1} + \mathbf{w}^n),$$

получаем

$$\begin{aligned} & \left( \left( \mathbf{C} + \left( \sigma - \frac{1}{4} \right) \tau^2 \mathbf{A} \right) \mathbf{w}^{n+1}, \mathbf{w}^{n+1} \right) - \left( \left( \mathbf{C} + \left( \sigma - \frac{1}{4} \right) \tau^2 \mathbf{A} \right) \mathbf{w}^n, \mathbf{w}^n \right) + \\ & + \frac{\tau}{2} (\mathbf{B}(\mathbf{w}^{n+1} + \mathbf{w}^n), \mathbf{w}^{n+1} + \mathbf{w}^n) + \|\mathbf{z}^{n+1}\|_{\mathbf{A}}^2 - \|\mathbf{z}^n\|_{\mathbf{A}}^2 = \tau(\mathbf{f}^n, (\mathbf{w}^{n+1} + \mathbf{w}^n)). \end{aligned} \quad (28)$$

Так как

$$\begin{aligned} & (\mathbf{C}\mathbf{w}^{n+1}, \mathbf{w}^{n+1}) = \|\mathbf{w}^{n+1}\|^2, \\ & (\mathbf{B}(\mathbf{w}^{n+1} + \mathbf{w}^n), \mathbf{w}^{n+1} + \mathbf{w}^n) \geq (\mathbf{B}(\mathbf{w}^{n+1} + \mathbf{w}^n), \mathbf{w}^{n+1} + \mathbf{w}^n), \\ & (\mathbf{f}^n, (\mathbf{w}^{n+1} + \mathbf{w}^n)) \leq \frac{1}{2}(\mathbf{B}(\mathbf{w}^{n+1} + \mathbf{w}^n), \mathbf{w}^{n+1} + \mathbf{w}^n) + \frac{1}{2}(\mathbf{B}^{-1}\mathbf{f}^n, \mathbf{f}^n), \end{aligned}$$

то из (28) следует оценка

$$\|\mathbf{w}^{n+1}\|^2 - \|\mathbf{w}^n\|^2 + \left( \sigma - \frac{1}{4} \right) \tau^2 (\|\mathbf{w}^{n+1}\|_{\mathbf{A}}^2 - \|\mathbf{w}^n\|_{\mathbf{A}}^2) + \|\mathbf{z}^{n+1}\|_{\mathbf{A}}^2 - \|\mathbf{z}^n\|_{\mathbf{A}}^2 \leq \frac{1}{2} \tau \|\mathbf{f}^n\|_{\mathbf{B}^{-1}}^2.$$

С учётом введённых обозначений (27) это неравенство приводит нас к оценке (25). Теорема доказана.

Запишем рассматриваемую трёхслойную векторную схему с весами (23), (24) для отдельных компонент. В наиболее интересном случае  $\sigma = 0.25$  для приближённого решения задачи (14)–(16) имеем

$$\begin{aligned} & \frac{y_i^{n+1} - 2y_i^n + y_i^{n-1}}{\tau^2} + \left( \mathbf{B} + \sum_{i=1}^m \frac{a_i}{b_i} \mathbf{C} \right) \frac{y_i^{n+1} - y_i^{n-1}}{2\tau} - \\ & - \sum_{i=1}^m \frac{a_i}{b_i} \mathbf{C} \frac{y_i^{n+1} - y_i^{n-1}}{2\tau} + \mathbf{A} \frac{y_i^{n+1} + 2y_i^n + y_i^{n-1}}{4} = f_i^n, \end{aligned} \quad (29)$$

$$\frac{y_i^{n+1} - y_i^{n-1}}{2\tau} + b_i \frac{y_i^{n+1} + 2y_i^n + y_i^{n-1}}{4} - \frac{y_i^{n+1} - y_i^{n-1}}{2\tau} = 0, \quad i = 1, 2, \dots, m, \quad n = 1, 2, \dots, \quad (30)$$

при заданных начальных условиях

$$y^0 = u_0, \quad y^1 = \tilde{v}^1, \quad y_i^0 = 0, \quad y_i^1 = \tilde{v}_i^1, \quad i = 1, 2, \dots, m. \quad (31)$$

Для решения задачи (29)–(31) из априорной оценки (25) следует

$$\begin{aligned} & \left\| \frac{y^{n+1} - y^n}{\tau} \right\|^2 + \left\| \frac{y^{n+1} + y^n}{2} \right\|_A^2 + \sum_{i=1}^m a_i \left\| \frac{y_i^{n+1} + y_i^n}{2} \right\|_C^2 \leq \\ & \leq \left\| \frac{\tilde{v}^1 - v^0}{\tau} \right\|^2 + \left\| \frac{\tilde{v}^1 + v^0}{2} \right\|_A^2 + \frac{1}{2} \sum_{i=1}^m a_i \|\tilde{v}_i^1\|_C^2 + \frac{1}{2} \sum_{k=1}^n \tau \|f^k\|_{B^{-1}}^2, \quad n = 1, 2, \dots \end{aligned} \quad (32)$$

Оценка (32) является разностным аналогом оценки (17) для решения дифференциальной задачи (14)–(16).

Отдельного внимания при решении нелокальных задач заслуживает проблема вычислительной реализации. В случае (30) имеем равенства

$$\begin{aligned} y_i^{n+1} &= \frac{1}{1 + 0.5b_i\tau} y^{n+1} + \chi_i^n, \\ \chi_i^n &= \frac{1}{1 + 0.5b_i\tau} (y_i^{n-1} - y^{n-1} - 0.5b_i\tau(2y_i^n + y_i^{n-1})), \quad i = 1, 2, \dots, m. \end{aligned} \quad (33)$$

Их подстановка в (29) даёт для нахождения  $y^{n+1}$  уравнение

$$\left( \frac{2}{\tau} I + B + \frac{\tau}{2} (\mu C + A) \right) y^{n+1} = \chi^n. \quad (34)$$

Для коэффициента  $\mu$  и правой части имеем

$$\begin{aligned} \mu &= \sum_{i=1}^m \frac{a_i}{1 + 0.5b_i\tau}, \\ \chi^n &= 2\tau f^n + \frac{2}{\tau} (2y^n - y^{n-1}) + B y^{n-1} + \sum_{i=1}^m \frac{a_i}{b_i} C (y^{n-1} - y_i^{n-1} + \chi_i^n) - \frac{\tau}{2} A (2y^n + y^{n-1}). \end{aligned}$$

Тем самым переход на новый  $n + 1$  слой по времени обеспечивается решением стандартной задачи (34) для  $y^{n+1}$  и расчётом вспомогательных величин  $y_i^{n+1}$ ,  $i = 1, 2, \dots, m$ , согласно равенствам (33). Вычислительная сложность приближённого решения рассматриваемой нелокальной задачи (1), (2), (5) не намного выше, чем при решении локальной задачи. Необходимо дополнительно оперировать с решениями  $m$  простых вспомогательных локальных эволюционных задач при явных вычислениях их решения на новом слое по времени.

### СПИСОК ЛИТЕРАТУРЫ

1. *Dautray R., Lions J.-L.* Mathematical Analysis and Numerical Methods for Science and Technology. V. 1. Berlin, 2000.
2. *Evans L.C.* Partial Differential Equations. Providence, 2010.
3. *Gripenberg G., Londen S.-O., Staffans O.* Volterra Integral and Functional Equations. Cambridge, 1990.
4. *Prüss J.* Evolutionary Integral Equations and Applications. Basel, 1993.
5. *Christensen R.M.* Theory of Viscoelasticity: an Introduction. New York, 1982.
6. *Marques S.P., Creus G.J.* Computational Viscoelasticity. Berlin, 2012.
7. *Knabner P., Angermann L.* Numerical Methods for Elliptic and Parabolic Partial Differential Equations. New York, 2003.

8. *Quarteroni A., Valli A.* Numerical Approximation of Partial Differential Equations. Berlin, 1994.
9. *Chen C., Shih T.* Finite Element Methods for Integrodifferential Equations. Singapore, 1998.
10. *McLean W., Thomee V.* Numerical solution of an evolution equation with a positive-type memory term // The ANZIAM J. 1993. V. 35. № 1. P. 23–70.
11. *McLean W., Thomee V., Wahlbin L. B.* Discretization with variable time steps of an evolution equation with a positive-type memory term // J. of Comput. and Appl. Math. 1996. V. 69. № 1. P. 49–69.
12. *Linz P.* Analytical and Numerical Methods for Volterra Equations. Philadelphia, 1985.
13. *Vabishchevich P.N.* Numerical solution of the Cauchy problem for Volterra integrodifferential equations with difference kernels // Appl. Numer. Math. 2022. V. 174. P. 177–190.
14. *Vabishchevich P.N.* Approximate solution of the Cauchy problem for a first-order integrodifferential equation with solution derivative memory // arXiv. 2021. № 2111.05121. P. 1–13.
15. *Halanay A.* On the asymptotic behavior of the solutions of an integro-differential equation // J. of Math. Anal. and Appl. 1965. V. 10. № 2. P. 319–324.
16. *Samarskii A.A.* The Theory of Difference Schemes. New York, 2001.
17. *Samarskii A.A., Matus P.P., Vabishchevich P.N.* Difference Schemes with Operator Factors. Dordrecht, 2002.

Институт проблем безопасного развития  
атомной энергетики РАН, г. Москва,  
Северо-Кавказский центр математических исследований,  
Северо-Кавказский федеральный университет,  
г. Ставрополь

Поступила в редакцию 10.01.2022 г.  
После доработки 10.01.2022 г.  
Принята к публикации 25.05.2022 г.



УДК 519.632+519.633

## О РЕШЕНИИ УРАВНЕНИЯ СМЕШАННОГО ТИПА В НЕОГРАНИЧЕННОЙ ОБЛАСТИ

© 2022 г. М. П. Галанин, Д. Л. Сорокин, А. Р. Ухова

Разработаны способы численного решения задачи для уравнения смешанного типа в неограниченной области. Предложен способ задания искусственных граничных условий, позволяющий проводить расчёты в ограниченной области такой, что численное решение сходится к проекции точного решения на эту область. Подробно рассмотрен случай, когда решение удовлетворяет волновому уравнению в ограниченной области и уравнению Лапласа в оставшейся части пространства. Приведён алгоритм для случая, когда в ограниченной области процесс описывается уравнением теплопроводности. Приведены примеры решений в двумерном случае. Исследована скорость сходимости итерационного алгоритма.

DOI: 10.31857/S0374064122070056, EDN: CEFVXE

**Введение.** В работе рассмотрены задачи, описываемые гиперболическими или параболическими уравнениями во внутренней подобласти и уравнением Лапласа – во внешней, причём внутренняя подобласть конечных размеров со всех сторон окружена внешней бесконечной подобластью. Построены алгоритмы численного решения таких задач.

Алгоритмы базируются на основной интегральной формуле Грина для оператора Лапласа. Суть методов состоит во введении расчётной области конечных размеров, построении искусственных граничных условий и итерационного алгоритма для реализации этих условий. Подобные алгоритмы использованы в работе [1] для решения внешних задач для оператора Лапласа.

В п. 1 приведена постановка задачи, построен вычислительный алгоритм, исследована скорость его сходимости. С помощью интегро-интерполяционного метода в п. 2 построена разностная схема. В п. 3 приведены примеры решения задачи с помощью алгоритма, проведено сравнение теоретической и экспериментальной оценок скоростей сходимости.

**1. Постановка задачи. Построение итерационного алгоритма.** Пусть в ограниченной области  $D$  (рис. 1) некоторый процесс описывается волновым уравнением, вне этой области – уравнением Лапласа. На границе  $\partial D$  решение непрерывно вместе с нормальной производной, а также является ограниченным на бесконечности и в нуле:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \Delta u, \quad \mathbf{r} \in D, \quad t > 0;$$

$$\Delta u = 0, \quad \mathbf{r} \in \mathbb{R}^n \setminus D, \quad n = 1, 2, 3;$$

$$u|_{t=0} = \alpha(\mathbf{r}), \quad \mathbf{r} \in D;$$

$$u_t|_{t=0} = \beta(\mathbf{r}), \quad \mathbf{r} \in D;$$

$$[u] = 0, \quad \left[ \frac{\partial u}{\partial \mathbf{n}} \right] = 0 \quad \text{на } \partial D;$$

$$|u| < +\infty. \tag{1}$$

Здесь  $[\cdot]$  – знак скачка величины, указанной в скобках,  $c$  – константа.

Целью работы является построение алгоритма численного решения задачи в ограниченной области  $D_2 : D \subset D_2$  (рис. 1) такого, что проекция решения (1) на указанную ограниченную область будет совпадать с решением задачи в ограниченной области.

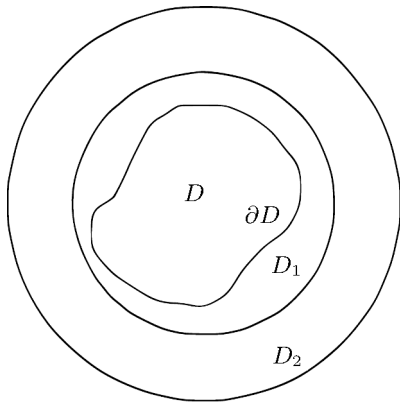


Рис. 1. Рассматриваемая область.

Без потери общности под  $D_2$  будем понимать шар в пространстве соответствующей размерности радиуса  $R_2$ . Введём обозначение  $R$  – радиус шара, полностью включающего в себя область  $D$  ( $R < R_2$ ). Форма области  $D$  может быть сложной, что не позволяет задать для соответствующей задачи функцию Грина. Поэтому для такого случая в статье [1] рассмотрена дополнительная граница  $\partial D_1$  в области  $D_2 \setminus D$ , представляющая собой сферу радиуса  $R_1 : R < R_1 < R_2$  (рис. 1).

Для нахождения решения в ограниченной области будем использовать условие

$$u|_{|\mathbf{r}|=R_2} = Su|_{|\mathbf{r}|=R_1}. \tag{2}$$

Здесь  $S$  – линейный оператор. В двумерном или трёхмерном случаях  $S$  – это интеграл Пуассона (приведён далее).

В одномерном плоском (декартовом,  $n = 1$ ) случае потребуем выполнения условия

$$\left. \frac{\partial u}{\partial r} \right|_{|\mathbf{r}|=0} = 0. \tag{3}$$

В одномерном сферически симметричном и трёхмерном случаях от решения (1) будем требовать регулярности на бесконечности:

$$|u| \rightarrow 0, \quad |\mathbf{r}| \rightarrow \infty.$$

Введённых условий достаточно для существования и единственности решений рассматриваемых задач. Не возникает сомнений, что при должной аппроксимации задачи (1)–(3) в ограниченной области  $D_2$  будет получено искомое (приближённое к точному) решение.

Рассмотрим итерационный вариант решения, при котором значение функции  $u$  на каждой итерации находится при заданном значении приближения граничного условия на границе  $|\mathbf{r}| = R_2$ , а затем по полученному решению пересчитывается по (2), и решение в области ищется снова.

В одномерных плоском (декартовом) и цилиндрически симметричном случаях справедливо равенство

$$Su|_{|\mathbf{r}|=R_1} = u|_{|\mathbf{r}|=R_1},$$

а в одномерном сферически симметричном случае

$$Su|_{|\mathbf{r}|=R_1} = \frac{R_1}{R_2} u|_{|\mathbf{r}|=R_1}.$$

Исследуем сходимость итерационного процесса в одномерном случае.

Ограничимся исследованием дифференциально-разностной задачи. Для этого заменим производные по времени на конечно-разностные с шагом  $\tau$ . Пусть разыскивается решение на новом временном слое в соответствии с задачей

$$\begin{aligned} \frac{1}{\tau^2} y &= c^2 \Delta y + f, \quad \mathbf{r} \in D; \\ \Delta y &= 0, \quad \mathbf{r} \in \mathbb{R}^n \setminus D, \quad n = 1, 2, 3; \\ [y] &= 0, \quad \left[ \frac{\partial y}{\partial \mathbf{n}} \right] = 0 \quad \text{на } \partial D; \\ |u| &< +\infty; \\ y|_{|\mathbf{r}|=R_2} &= Sy|_{|\mathbf{r}|=R_1}. \end{aligned} \tag{4}$$

Здесь правая часть  $f$  определяется решением с предыдущих временных слоёв.

Устроим итерационный процесс: ищем решение на  $k$ -й итерации путём нахождения решения (4) с граничным условием

$$y^k|_{|r|=R_2} = \mu^{k-1},$$

после нахождения которого пересчитаем условие по правилу

$$\mu^k = Sy^k|_{|r|=R_1}.$$

Сходимость итераций будет иметь место, если  $|\delta\mu^k| \leq q|\delta\mu^{k-1}|$  с  $0 < q < 1$ . Здесь  $\delta\mu^k = \mu - \mu^k$ ,  $\mu = y|_{|r|=R_2}$ .

**Утверждение 1.** В одномерных случаях описанный итерационный процесс сходится со скоростью геометрической прогрессии с показателем  $q$ , равным

$$q = \frac{c\tau \operatorname{cth}(R/(c\tau)) + (R_1 - R)}{c\tau \operatorname{cth}(R/(c\tau)) + (R_2 - R_1)}$$

в плоском (декартовом) случае,

$$q = \frac{c\tau I_0(R/(c\tau))/I_1(R/(c\tau)) + R(\ln R_1 - \ln R)}{c\tau I_0(R/(c\tau))/I_1(R/(c\tau)) + R(\ln R_2 - \ln R)}$$

в цилиндрически симметричном случае,

$$q = \frac{c\tau \operatorname{th}(R/(c\tau)) + (R_1 - R)}{c\tau \operatorname{th}(R/(c\tau)) + (R_2 - R)}$$

в сферически симметричном случае.

В данном утверждении  $I_0$  и  $I_1$  – функции Инфельда (Бесселя комплексного аргумента).

Доказательство предложения состоит в прямом решении задачи (4) для  $\delta y^k = y - y^k$  и анализе граничного условия.

Как видно, во всех случаях показатели  $q < 1$  и в основном при малых  $\tau$  определяются слагаемыми, которые не зависят от  $\tau$ . В значительной степени данные показатели соответствуют аналогичным коэффициентам задачи для уравнения Лапласа в пространстве (см. [1]).

Рассмотрим также задачу, аналогичную (1), но с заменой волнового уравнения на уравнение теплопроводности с постоянной температуропроводностью  $a^2$ :

$$\frac{\partial u}{\partial t} = a^2 \Delta u, \quad \mathbf{r} \in D, \quad t > 0, \quad (5)$$

с очевидными изменениями начальных условий (1).

Будем решать эту задачу также итерационным способом, который исследуем, ограничившись дифференциально-разностным приближением.

**Утверждение 2.** Для задачи с уравнением теплопроводности (5) итерационный процесс сходится с показателем геометрической прогрессии  $q$  из утверждения 1, в котором коэффициент  $c$  заменён на  $a$ , а  $\tau$  – на  $\sqrt{\tau}$ .

Сходимость подобного итерационного процесса в одномерном случае исследована в работе [2]. В статье [3] приведены результаты численного моделирования в двумерном случае.

**Замечание.** В общем трёхмерном случае для решения задачи справедлив принцип максимума, поэтому сходимость итераций также имеет место, как и для уравнения Лапласа [1, 2, 4].

**2. Построение разностной схемы.** Положим далее значение скорости  $c = 1$ . Рассмотрим пространственно двумерный случай. Задачу будем решать в цилиндрических координатах.

Решение внешней задачи для уравнения Лапласа для круга известно и задаётся интегралом Пуассона [5, с. 314, 6, с. 157]. В двумерном случае он имеет вид

$$u(r, \varphi) = Su|_{|\tilde{r}|=R_1} = \frac{1}{2\pi} \int_0^{2\pi} u(R_1, \psi) \frac{r^2 - R_1^2}{r^2 - 2rR_1 \cos(\varphi - \psi) + R_1^2} d\psi, \quad r > R_1. \quad (6)$$

Вычислив интеграл Пуассона от значений искомой функции на границе  $\partial D_1$ , можно получить значения функции на границе области  $D_2$ .

Тогда исходная задача (1), (2) в области уравнения Лапласа примет вид

$$\Delta u = 0, \quad \vec{r} \in D_2 \setminus D;$$

$$u|_{\partial D_2} = \frac{1}{2\pi} \int_0^{2\pi} u(R_1, \psi) \frac{R_2^2 - R_1^2}{R_2^2 - 2R_1R_2 \cos(\varphi - \psi) + R_1^2} d\psi.$$

Построим итерационный алгоритм для учёта интегрального граничного условия (6). Значение  $u^s$  на границе области  $D_2$  на каждой новой итерации будет вычисляться через уже известные значения  $u^{s-1}$  на предыдущей итерации. Для нахождения  $u^s$  имеем задачу с условиями Дирихле на внешней границе ограниченной области:

$$u^s|_{\partial D_2} = Su^{s-1}|_{\partial D_1}.$$

Схожий алгоритм применён в работах [1–3].

Для численного расчёта введём равномерную сетку в

$$\Omega = \{0 \leq r \leq R_2, \quad 0 \leq \varphi \leq 2\pi, \quad 0 \leq t \leq T\}$$

с шагом  $h_r$  по координате  $r$ , шагом  $h_\varphi$  по координате  $\varphi$  и шагом  $\tau$  по времени. Выберем шаг так, чтобы узлы сетки  $N$ ,  $N_1$ ,  $N_2$  попадали на границы областей радиусов  $R$ ,  $R_1$  и  $R_2$  соответственно.

Построим разностную схему для волнового уравнения интегро-интерполяционным методом (см. [7, с. 215]). Она имеет вид

$$y_{\bar{t}t} = \frac{1}{r_i h^r} (r_{i+1/2} y_r - r_{i-1/2} y_{\bar{r}})^{(\sigma)} + \frac{1}{r_i^2} y_{\bar{\varphi}\varphi}^{(\sigma)}. \tag{7}$$

Здесь  $y^{(\sigma)} = \sigma \hat{y} + (1 - 2\sigma)y + \sigma \check{y}$ , где  $y$  – сеточная функция, приближающая решение  $u$ ,  $\sigma$  – вес. Неизвестной является функция  $\hat{y}$ . Обозначения разностных производных и сеточных функций на разных временных слоях соответствуют обозначениям в работах [7, с. 210; 8, гл. 1].

Полученная схема является трёхслойной, для нахождения неизвестного значения  $\hat{y}_{ij}$  необходимо знать значения решения с двух предыдущих слоёв. Значения  $u$  на нулевом слое известны из начального условия  $u|_{t=0} = \alpha(r, \varphi)$ , на первом слое можно вычислить решение из условия  $u_t|_{t=0} = \beta(r, \varphi)$ . Схема (7) имеет второй порядок по пространству и времени, поэтому аппроксимируем второе начальное условие тоже со вторым порядком. Для этого воспользуемся формулой Тейлора.

Отсюда значение решения на первом временном слое имеет вид

$$y_{ij}^1 = \alpha + \tau\beta + \frac{\tau^2}{2} \left( \frac{1}{r_i h^r} (r_{i+1/2} \alpha_r - r_{i-1/2} \alpha_{\bar{r}}) + \frac{1}{r_i^2} \alpha_{\bar{\varphi}\varphi} \right).$$

Аналогично (7) построим разностную схему для уравнения Лапласа:

$$\frac{1}{r_i h^r} (r_{i+1/2} y_r - r_{i-1/2} y_{\bar{r}})^{(\sigma)} + \frac{1}{r_i^2} y_{\bar{\varphi}\varphi}^{(\sigma)} = 0.$$

Для приближённого вычисления интеграла Пуассона (6) воспользуемся квадратурной формулой трапеций

$$y_{N_2,j} = \frac{h_\varphi (R_2^2 - R_1^2)}{2\pi} \left( \frac{0.5 y_{N_1,0}}{R_1^2 - 2R_1R_2 \cos(\varphi_j - \varphi_0) + R_2^2} + \frac{0.5 y_{N_1,N_\varphi}}{R_1^2 - 2R_1R_2 \cos(\varphi_j - \varphi_{N_\varphi}) + R_2^2} + \sum_{m=1}^{N_\varphi-1} \frac{y_{N_1,m}}{R_1^2 - 2R_1R_2 \cos(\varphi_j - \varphi_m) + R_2^2} \right).$$

Здесь  $N_1, N_2$  – номера узлов сетки по радиусу, соответствующие  $r = R_1$  и  $r = R_2$ . Число узлов сетки по углу  $\varphi$  равно  $N_\varphi + 1$ . По углу аппроксимированы периодические граничные условия.

Решение задачи должно быть ограничено в нуле по радиусу. Аппроксимируем уравнение (1) в нуле так, как это описано в работах [7, с. 313; 8, с. 434].

На границе области  $D$  поставлены условия сопряжения. Для аппроксимации условия равенства потоков рассмотрим ячейки  $[r_{N-1/2}, r_N] \times [\varphi_{j-1/2}, \varphi_{j+1/2}]$  и  $[r_N, r_{N+1/2}] \times [\varphi_{j-1/2}, \varphi_{j+1/2}]$ . В первой ячейке проинтегрируем волновое уравнение, а во второй – уравнение Лапласа. Из данных равенств выразим поток  $u$  в точке  $r_N$ , приравняем и получим аппроксимацию условия сопряжения

$$y_{\bar{t}t} = \frac{2}{h_i^r r_{N_1}} (r_{N_1+1/2} y_r - r_{N_1-1/2} y_{\bar{r}})^{(\sigma)} + \frac{2}{r_{N_1}^2} y_{\bar{\varphi}\varphi}^{(\sigma)}.$$

Если границей области  $D$  является не окружность, то возникает необходимость постановки условий сопряжения на границе, например, вида  $\varphi = \text{const}$ , которое построим также путём интегрирования уравнений по ячейке  $[r_{i-1/2}, r_{i+1/2}] \times [\varphi_{N_\varphi-1/2}, \varphi_{N_\varphi+1/2}]$ , где  $N_\varphi$  – номер узла по  $\varphi$ , попадающего на границу  $D$ :

$$y_{\bar{t}t} = \frac{2}{h_i^r r_i} (r_{i+1/2} y_r - r_{i-1/2} y_{\bar{r}})^{(\sigma)} + \frac{2}{r_i^2} y_{\bar{\varphi}\varphi}^{(\sigma)}.$$

Для проверки итерационного алгоритма (получения эталонного решения в отсутствии аналитического) рассмотрим вспомогательную задачу. Вместо условия (6) на внешней границе поставим условие равенства нулю производной решения по  $r$ , а радиус внешней области  $R_2$  увеличим, чтобы уменьшить влияние данного граничного условия на решение в исходной области. Таким образом, необходимо решить задачу в области  $D'_2$  радиуса  $R'_2 = QR_2$ ,  $Q > 0$  – некоторый коэффициент, с другим граничным условием, которое аппроксимируем интегроинтерполяционным методом, проинтегрировав уравнение Лапласа в ячейке  $[r_{N_2-1/2}, r_{N_2}] \times [\varphi_{j-1/2}, \varphi_{j+1/2}]$ :

$$-\frac{2r_{N_2-1/2}}{r_{N_2} h_r} y_{\bar{r}}^{(\sigma)} + \frac{1}{r_{N_2}^2} y_{\bar{\varphi}\varphi}^{(\sigma)} = 0,$$

где  $N_2$  – узел, попадающий на границу  $\partial D'_2$ .

**3. Построение разностной схемы.** Рассмотрим четыре варианта задачи: пример 1 – одномерная цилиндрически симметричная задача, примеры 2 и 3 – двумерная задача с областью  $D$  в виде круга, при этом в примере 2 аналитическое решение известно, а для примера 3 особым способом найдено “эталонное” решение, в примере 4 область  $D$  представляет собой половину круга. Во всех расчётах будем использовать  $\sigma = 0.5$ . Для решения системы линейных алгебраических уравнений используем решатель BiCGStab – стабилизированный метод бисопряжённых градиентов для задач с разреженными матрицами, реализованный в библиотеке Eigen. Для оценки норм ошибок используется равномерная норма.

**Пример 1.** Пусть  $R$  – корень уравнений  $J_1(R) = 0$ ,  $J_1$  – функция Бесселя первого рода, граница области  $D$  совпадает с кругом радиуса  $R$ ,  $T = 0.5$  – время протекания процесса. Начальные условия при  $r < R$ :

$$u(r, \varphi, 0) = J_0(r), \quad u_t(r, \varphi, 0) = 0.$$

Точное решение задаётся выражением

$$u(r, \varphi, t) = \begin{cases} J_0(r) \cos t, & 0 \leq r \leq R; \\ J_0(R) \cos t, & r > R. \end{cases}$$

Для нахождения решения итерационным способом на каждом временном слое воспользуемся разностными схемами, построенными ранее.

Согласно результатам вычислительных экспериментов (табл. 1) разностная схема позволяет получить решение одномерной задачи, при этом наблюдается второй порядок сходимости.

**Таблица 1.** Сравнение численного решения одномерной задачи (пример 1) с точным.  $R_1 = 2R$ ,  $R_2 = 4R$ . Начальный шаг  $h = (h_r, h_\varphi, \tau) = (0.383171, \pi/4, 0.05)$

Шаг	Относительная ошибка	Отношение ошибок
$h$	$4.92 \cdot 10^{-3}$	
$h/2$	$7.53 \cdot 10^{-4}$	6.54
$h/4$	$1.88 \cdot 10^{-4}$	3.99
$h/8$	$3.94 \cdot 10^{-5}$	4.78

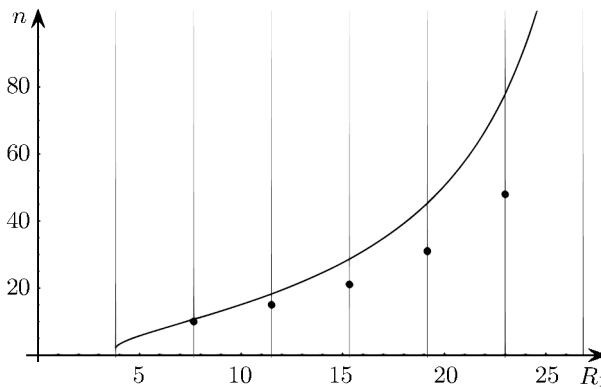
Рассмотрим влияние положения границы области  $D_1$  на сходимость. Зафиксируем границы областей  $D$ ,  $D_2$ , а положение  $\partial D_1$  будем менять.

Согласно утверждению 1 должна наблюдаться сходимость со скоростью геометрической прогрессии с коэффициентом  $q$ . Тогда число итераций можно оценить по формуле

$$n > \frac{\ln \varepsilon}{\ln q}, \tag{8}$$

где  $\varepsilon$  – убывание ошибки,  $\|y^s - y^*\|/\|y^0 - y^*\| < \varepsilon$ .

На рис. 2 приведены два графика: сплошная линия – график зависимости априорной оценки необходимого числа итераций до сходимости по формуле (8), а точками отмечен результат вычислительных экспериментов. Видим, что качественно и количественно графики близки. График оценки числа итераций мажорирует график результатов вычислительного эксперимента.



**Рис. 2.** Зависимость числа итераций от положения границы  $R_1$  при решении одномерной задачи (пример 1)  $R_2 = 8R$ .

Зависимость коэффициента геометрической прогрессии от  $R_1$  в расчётах вычислим как

$$\log_q 10^{-5} = n(R_1). \tag{9}$$

По результатам, представленным в табл. 2, можно сделать вывод, что погрешность в определении коэффициента  $q$  составляет около 10 % и связана с тем, что разностная задача лишь аппроксимирует исходную.

**Таблица 2.** Сравнение коэффициентов геометрической прогрессии, предсказанных в утверждении 1 и полученных в ходе вычислительных экспериментов

$R_1$	$R$	$2R$	$3R$	$4R$	$5R$	$6R$
$q$ согласно утверждению 1	0.338	0.531	0.669	0.775	0.863	0.936
$q$ по (9)	0.316	0.464	0.578	0.690	0.787	0.940

**Пример 2.** Пусть внутренняя область  $D$  имеет форму круга радиуса  $R = 2.40483$ . Необходимо решить задачу (1) с начальными условиями

$$u|_{t=0} = 0, \quad u_t|_{t=0} = J_1(r) \sin \varphi.$$

Точное решение задаётся выражением

$$u(r, \varphi, t) = \begin{cases} J_1(r) \sin t \sin \varphi, & 0 \leq r \leq R; \\ 1.24846r^{-1} \sin t \sin \varphi, & r > R. \end{cases}$$

Результаты серии расчётов (табл. 3) показывают, что построенный метод сходится к точному решению в области  $D_2$  со вторым порядком.

Можно предположить, что наихудшая скорость сходимости в двумерном случае должна наблюдаться на задачах типа примера 1, решения которых построены на основной гармонике. Результаты решения примеров 1 и 2 с разным положением границы  $R_1$  это подтверждают (табл. 4).

**Таблица 3.** Сравнение численного решения двумерной задачи (пример 2) с точным  $R_1 = 2R$ ,  $R_2 = 4R$ . Начальный шаг  $h = (h_r, h_\varphi, \tau) = (0.229031, \pi/16, 0.1)$

Шаг	Относительная ошибка	Отношение ошибок
$h$	$3.03 \cdot 10^{-3}$	
$h/2$	$7.63 \cdot 10^{-4}$	3.97
$h/4$	$1.92 \cdot 10^{-4}$	3.98
$h/8$	$4.79 \cdot 10^{-5}$	4.00

**Таблица 4.** Зависимость числа итераций до сходимости от положения границы  $\partial D_1$

$R_1$	$R$	$2R$	$3R$	$4R$	$5R$	$6R$
$n$ (пример 1)	6	8	11	14	19	29
$n$ (пример 2)	5	7	8	11	16	29

**Пример 3.** Пусть внутренняя область  $D$  имеет форму круга радиуса  $R = 1$ . Необходимо решить задачу (1) с начальными условиями

$$u|_{t=0} = 0, \quad u_t|_{t=0} = r \sin \varphi. \tag{10}$$

В силу того, что аналитическое решение достаточно сложно, воспользуемся методом расширения расчётной области для получения эталонного решения.

Результаты серии расчётов (табл. 6) показывают, что норма отклонения (для двух последовательных значений  $Q$ ) численного решения в области  $D_2$  уменьшается с увеличением  $Q$ , что соответствует ожиданиям. Для вычисления нормы отклонения численного решения использована равномерная норма.

Также проведена серия расчётов (табл. 6), которая демонстрирует, что измельчение сетки позволяет получить более точное решение, но измельчение сетки более чем в 4 раза при неизменном  $Q$  не позволяет получить более точного решения.

**Таблица 5.** Зависимость отклонения решения примера 3 в  $D_2$  от размера расчётной области

$Q$	Норма отклонения
5	$3.05 \cdot 10^{-3}$
7	$1.69 \cdot 10^{-3}$
8	$1.37 \cdot 10^{-3}$
12	$1.12 \cdot 10^{-3}$

**Таблица 6.** Зависимость нормы ошибки численного решения примера 3 от шага сетки.  $h = (h_r, h_\varphi, \tau) = (0.1, 0.25\pi, 0.1)$  – шаги по пространству и времени,  $Q = 5$

Шаг	Норма ошибки
$h$	$3.26 \cdot 10^{-3}$
$h/2$	$3.05 \cdot 10^{-3}$
$h/4$	$2.78 \cdot 10^{-3}$

**Пример 4.** Рассмотрим задачу, в которой внутренняя область  $D$  имеет форму полукруга радиуса  $R = 1$  (рис. 3). В этой области будем решать задачу (1) с начальными условиями (10).

Для данного примера также проблематично записать аналитическое решение и применён метод расширения расчётной области. Результаты расчётов свидетельствуют о том, что ошибка уменьшается с увеличением параметра  $Q$  и с уменьшением шагов по времени и пространству (табл. 7 и 8).

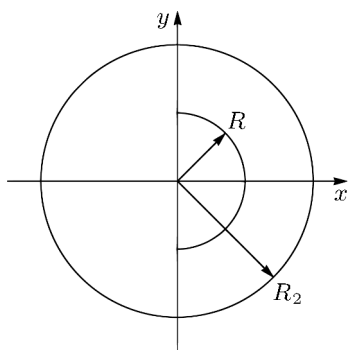
**Таблица 7.** Зависимость отклонения решения в  $D_2$  примера 4 от размера расчётной области

$q$	Норма отклонения
4	$4.77 \cdot 10^{-3}$
5	$1.92 \cdot 10^{-3}$
7	$1.12 \cdot 10^{-3}$
8	$1.13 \cdot 10^{-3}$

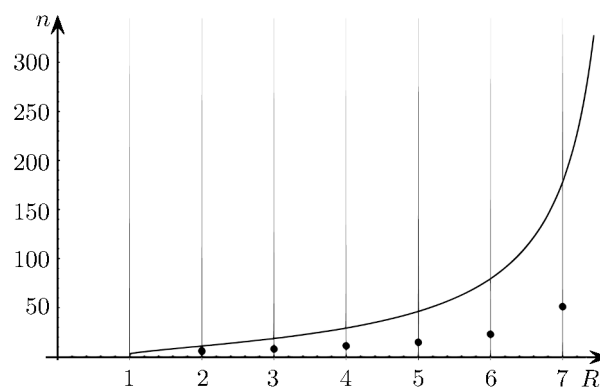
**Таблица 8.** Зависимость нормы ошибки численного решения примера 4 от шага сетки.  $h = (h_r, h_\varphi, \tau) = (0.1, 0.25\pi, 0.1)$  – шаги по пространству и времени,  $Q = 5$

Шаг	Норма ошибки
$h$	$4.77 \cdot 10^{-3}$
$h/2$	$1.92 \cdot 10^{-3}$
$h/4$	$1.77 \cdot 10^{-3}$

Для примеров 3 и 4 получены оценки зависимости числа итераций до сходимости от размера  $R_1$ . Результаты вычислительных экспериментов представлены на рис. 4. Графики для примеров 3 и 4 идентичны. В этих случаях  $R = 1$  – радиус внутренней области  $D$ ,  $R_1$  – радиус дополнительной области  $D_1$  (варьировался),  $R_2 = 3$  – радиус внешней области  $D_2$ . С увеличением радиуса дополнительной области количество итераций увеличивается, значит чем ближе эта граница к границе внутренней области, тем выше скорость сходимости.



**Рис. 3.** Структура расчётной области в примере 4.



**Рис. 4.** Графики зависимости числа итераций от положения границы  $R_1$  (для примеров 3 и 4).

**Заключение.** Разработан алгоритм решения задач для уравнений смешанного типа в неограниченной области в случае, когда в некоторой финитной области процесс описывается уравнением гиперболического (или параболического) типа, а вне – эллиптического. Исследована скорость сходимости построенного вычислительного алгоритма. Проведено сравнение оценки скорости сходимости алгоритма с результатами вычислительных экспериментов в двумерном случае. Данный алгоритм может быть легко обобщён на трёхмерный случай. Даже в двумерном случае вычислительные затраты при применении метода меньше, чем при использовании расширения расчётной области, но могут уступать методу задания интегрального граничного условия. В трёхмерном же случае ожидается, что итерационный метод будет наиболее эффективен среди перечисленных.

Исследование выполнено за счёт Российского научного фонда (грант 22-21-00260).

#### СПИСОК ЛИТЕРАТУРЫ

1. Галанин М.П., Сорокин Д.Л. О решении внешних краевых задач для уравнения Лапласа // Дифференц. уравнения. 2020. Т. 56. № 7. С. 918–926.
2. Галанин М.П., Сорокин Д.Л., Ухова А.Р. Методы численного решения дифференциального уравнения смешанного типа в неограниченной области // Мат. моделирование и численные методы. 2021. № 1. С. 91–109.



3. *Galanin M.P., Sorokin D.L.* Numerical solution tasks with mixed operator in unlimited area // CEUR Workshop Proceedings. 2020. V. 2543. P. 370–376.
4. *Савченко А.О., Ильин В.П., Бутюгин Д.С.* Метод решения внешней трёхмерной краевой задачи для уравнения Лапласа // Сиб. журн. индустр. мат. 2016. Т. 19. № 2. С. 88–99.
5. *Тихонов А.Н., Самарский А.А.* Уравнения математической физики. М., 1972.
6. *Свешников А.Г., Боголюбов А.Н., Кравцов В.В.* Лекции по математической физике. М., 1993.
7. *Галанин М.П., Савенков Е.Б.* Методы численного анализа математических моделей. М., 2010.
8. *Самарский А.А.* Введение в теорию разностных схем. М., 1971.

Институт прикладной математики  
имени М.В. Келдыша РАН, г. Москва,  
Московский государственный технический  
университет имени Н.Э. Баумана

Поступила в редакцию 18.02.2022 г.  
После доработки 18.02.2022 г.  
Принята к публикации 25.05.2022 г.

УДК 519.633+517.962.2+517.958

## О СВОЙСТВАХ ОДНОГО РАЗНОСТНОГО МЕТОДА РЕШЕНИЯ ДВУХФАЗНОЙ ЗАДАЧИ СТЕФАНА

© 2022 г. А. О. Гусев, О. В. Щерица, О. С. Мажорова

Исследуются свойства разностного метода решения двухфазной задачи Стефана на сетке, согласованной с формой границы раздела фаз. Математическая модель учитывает конвективное движение, возникающее в жидкости, теплоперенос в твёрдой и жидкой фазах, на границе раздела фаз – условие баланса внутренней энергии. Фазовый переход происходит при постоянной температуре. Положение межфазной границы определяется с помощью метода выпрямления фронта. Показано, что предложенная вычислительная схема является консервативной и энергетически нейтральной.

DOI: 10.31857/S0374064122070068, EDN: CEIGQC

**Введение.** Накопленный опыт применения метода конечных разностей для решения сложных прикладных задач показывает, что при конструировании вычислительного алгоритма недостаточно опираться только на такие свойства разностных схем как аппроксимация, устойчивость и сходимость. Необходимо также учитывать конкретные физические свойства исследуемого явления. Одним из основных критериев качества вычислительной процедуры является выполнение разностных аналогов законов сохранения, т.е. консервативность. Для построения консервативных разностных схем используется интегро-интерполяционный метод [1–3]. Применение неконсервативных разностных схем может существенно понизить точность результатов расчётов, а в ряде случаев привести к неверным результатам. А.Н. Тихонов и А.А. Самарский, изучив способы аппроксимации уравнений с разрывными коэффициентами, показали, что разностная схема, обеспечивающая второй порядок точности в классе достаточно гладких коэффициентов, не сходится в классе разрывных коэффициентов [1, 4–7].

В работе А.А. Самарского и Ю.П. Попова [3] сформулирован принцип полной консервативности разностных схем. Он состоит в выполнении на разностном уровне не только аналогов основных законов сохранения, присущих тому или иному явлению, но и дополнительных соотношений, свойственных дифференциальной задаче (наличие дивергентной и недивергентной формы записи законов сохранения, выполнение балансов отдельных видов энергии и т.п.). Таким образом, в надёжной дискретной модели различные формы записи законов сохранения должны переходить друг в друга с помощью тождественных преобразований, аналогичных тем, которые имеют место в дифференциальном случае.

Принцип полной консервативности был первоначально сформулирован для уравнений газовой динамики и магнитной гидродинамики, а затем применён к другим классам задач математической физики. В работе [8] в одномерном приближении были рассмотрены некоторые особенности применения принципа полной консервативности при решении эволюционных задач с подвижными границами. Для задачи о кристаллизации многокомпонентного раствора на подвижной и фиксированной сетках были построены дивергентные и недивергентные разностные схемы, гарантирующие выполнение балансов внутренней энергии и массы. Как и в дифференциальном случае, разностные уравнения переходят друг в друга с помощью замены переменных.

В статье [9] рассмотрена задача Стефана в цилиндрической системе координат в осесимметричном приближении. Основу математической модели составляют уравнения Навье–Стокса в переменных функции тока–вихрь, уравнения тепло- и массопереноса в твёрдой и жидкой фазах, условия термодинамического равновесия на границе раздела фаз. Для определения положения границы раздела фаз используется метод выпрямления фронта [10]. Задача решается в системе координат, полученной в результате замены переменных специального вида. В этой системе координат положение границы раздела фиксировано. Разностная схема построена в

новой системе координат на неподвижной сетке. Получены дивергентная и недивергентная формы записи, аппроксимирующие соответствующие формы записи исходной дифференциальной задачи, при этом на разностном уровне выполняются законы сохранения кинетической и внутренней энергии, балансы масс для компонентов. В работе [11] показано, что в разностной схеме, построенной с помощью метода выпрямления фронта, можно выполнить дискретный аналог преобразования системы координат и получить алгебраически эквивалентную вычислительную схему на подвижной сетке, согласованной с формой границы раздела фаз.

В данной работе продолжены исследования свойств разностной схемы, предложенной в статьях [9, 11]. Доказано, что разностный оператор, аппроксимирующий диссипативные члены в уравнениях движения жидкости и теплопереноса в расчётной системе координат, самосопряжён и отрицательно определён. Показано, что вычислительная схема, полученная с помощью метода выпрямления фронта, является энергетически нейтральной. Рассмотрены граничные условия для вихря, обеспечивающие выполнение закона сохранения завихрённости на разностном уровне. Результаты, полученные ниже и в работах [9, 11], позволяют утверждать, что построенный класс схем наследует основные свойства исходной дифференциальной задачи о фазовом переходе.

**1. Постановка задачи.** Задача о фазовом переходе рассматривается в рамках следующих предположений. Твёрдая и жидкая фазы имеют одинаковые плотности  $\rho^s = \rho^{lq}$  и удельные теплоёмкости  $c_p^s = c_p^{lq}$ , но различные коэффициенты теплопроводности –  $k^s$  и  $k^{lq}$  соответственно. Жидкая фаза является вязкой несжимаемой жидкостью с кинематической вязкостью  $\nu$ . Поля температуры  $T(t, r, z)$  и скоростей  $\mathbf{V}(t, r, z)$  жидкости предполагаются осесимметричными. Моделирование осуществляется в цилиндрической системе координат  $Orz$  в области  $\Omega(r, z) = [0, R] \times [0, Z]$ . В подобласти  $\Omega^s(t, r, z) = \{(r, z) : r \in [0, R], z \in [0, \zeta(t, r)]\}$  находится твёрдая фаза, жидкая фаза располагается в подобласти  $\Omega^{lq}(t, r, z) = \{(r, z) : r \in [0, R], z \in [\zeta(t, r), Z]\}$ . Здесь  $z = \zeta(t, r)$  – граница раздела фаз, положение которой изменяется в ходе процесса.

Движение жидкости описывается уравнениями Навье–Стокса в переменных функции тока–завихрённости. Компоненты вектора скорости  $\mathbf{V} = (V^r, 0, V^z)$  записаны в виде

$$V^r = \frac{1}{r} \frac{\partial \psi}{\partial z}, \quad V^z = -\frac{1}{r} \frac{\partial \psi}{\partial r}, \quad (1)$$

где  $\psi$  – функция тока. Завихрённость  $\boldsymbol{\omega} = \text{rot } \mathbf{V}$  определяется следующим образом:

$$\boldsymbol{\omega} = (0, \omega_\varphi, 0), \quad \omega_\varphi = \partial_z V^r - \partial_r V^z,$$

где  $\partial_r = \partial/\partial r$ ,  $\partial_z = \partial/\partial z$ . Для удобства в качестве неизвестной функции в уравнениях движения будем использовать  $\omega = -\omega_\varphi/r$ . Конвективный перенос в жидкости описывается с помощью уравнения переноса завихрённости

$$\frac{\partial \omega}{\partial t} + \frac{1}{r} \mathcal{K}^{(r,z)}(\omega, \psi) = \frac{1}{r} \mathcal{D}^{(r,z)}(r^2 \omega) + \beta_T g \frac{1}{r} \frac{\partial T}{\partial r},$$

уравнения, связывающего завихрённость и функцию тока

$$-\omega = \frac{1}{r} \mathcal{D}^{(r,z)}(\psi),$$

и уравнения конвективного теплопереноса

$$\frac{\partial T}{\partial t} + \frac{1}{r} \mathcal{K}^{(r,z)}(T, \psi) = \frac{1}{r} \mathcal{D}^{(r,z)}(T).$$

Здесь  $\beta_T$  – коэффициент температурного расширения,  $g$  – ускорение свободного падения,  $\mathcal{K}^{(r,z)}/r$  – оператор, описывающий конвективный перенос

$$\mathcal{K}^{(r,z)}(f, \psi) = \partial_r \mathcal{H}^r + \partial_z \mathcal{H}^z, \quad f = \omega, T, \quad (2)$$

где компоненты конвективного потока  $\mathcal{H}^{(r,z)} = (\mathcal{H}^r, \mathcal{H}^z)$  имеют вид

$$\mathcal{H}^r(f) = rV^r f = [\partial_z \psi]f, \quad \mathcal{H}^z(f) = rV^z f = -[\partial_r \psi]f,$$

$\mathcal{D}^{(r,z)}/r$  – эллиптический оператор, описывающий диссипативные процессы в уравнениях переноса завихрённости и конвективной теплопроводности:

$$\mathcal{D}^{(r,z)}(r^\beta f) = \partial_r \mathcal{W}^r + \partial_z \mathcal{W}^z, \quad f = \omega, \psi, T, \tag{3}$$

**Таблица.** Параметры диссипативного потока

$f$	$\varkappa$	$\alpha$	$\beta$
$\omega$	$\nu$	-1	2
$\psi$	1	-1	0
$T$	$k/(c_p \rho)$	1	0

компоненты потока  $\mathcal{W}^{(r,z)} = (\mathcal{W}^r, \mathcal{W}^z)$  вычисляются следующим образом:

$$\mathcal{W}^r(r^\beta f) = \varkappa r^\alpha \partial_r(r^\beta f), \quad \mathcal{W}^z(r^\beta f) = \varkappa r^\alpha \partial_z(r^\beta f),$$

значения параметров  $\varkappa, \alpha, \beta$  приведены в таблице.

Распределение температуры в твёрдой фазе описывается уравнением теплопроводности

$$\frac{\partial T}{\partial t} = \frac{1}{r} \mathcal{D}^{(r,z)}(T).$$

На межфазной границе температура равна температуре фазового перехода

$$T|_{z=\zeta(t,r)} = T_{\text{ph}}$$

и выполняется закон сохранения энергии (условие Стефана):

$$[[k \nabla T \cdot \mathbf{N}]] = \lambda \rho v_n. \tag{4}$$

Здесь  $[[q]] = q^s - q^l$  – скачок величины  $q$  на фронте кристаллизации,  $\nabla = (\partial_r, \partial_z)$ ,  $\lambda$  – скрытая теплота плавления,  $v_n = v_{\text{ph}}(\mathbf{e}_z \cdot \mathbf{N})$ ,  $v_{\text{ph}} = v_{\text{ph}}(t, r) = d\zeta/dt$  – скорость движения границы раздела фаз,  $\mathbf{N}$  – единичная нормаль к межфазной границе, направленная в жидкую фазу,  $\mathbf{e}_z$  – единичный вектор в направлении оси  $z$ .

Граничные условия для температуры имеют вид

$$T|_{z=0} = T_{\text{bot}}(t, r), \quad T|_{z=Z} = T_{\text{top}}(t, r), \quad T|_{r=R} = T_{\text{wall}}(t, z).$$

Условия для функции тока и температуры на оси ампулы вытекают из условий симметрии

$$\psi = 0, \quad \omega_\varphi = 0, \quad \partial_r T = 0.$$

Краевые условия для функции тока на фронте кристаллизации, верхней и боковой границах жидкой фазы:

$$\psi = 0, \quad \partial_n \psi = 0.$$

Здесь  $\partial_n$  – производная по направлению внешней нормали к соответствующей границе.

**2. Метод выпрямления фронта.** Для решения задачи с внутренней подвижной границей, положение которой необходимо определять в ходе решения задачи, применяется метод выпрямления фронта (см. [10]). Физическая область  $\Omega(r, z)$  отображается в расчётную область  $S(\xi, \eta)$  так, что в новой системе координат положение границы раздела фаз фиксировано и совпадает с координатной линией  $\eta = \text{const}$ .

**2.1. Преобразование системы координат.** В данной работе используется замена переменных, при которой области  $\Omega^s, \Omega^l$  в новой системе координат становятся прямоугольниками  $S^s$  и  $S^l$  (рис. 1), а границы областей  $z = 0, z = \zeta(t, r), z = Z$  переходят в прямые  $\eta = 0, \eta = 1, \eta = 2$  соответственно. Связь между системами координат имеет вид

$$\tilde{t} = t, \quad \xi = r, \quad \eta = \begin{cases} z/l^s, & z \in [0, \zeta(t, r)], \\ 1 + (z - \zeta)/l^l, & z \in [\zeta(t, r), Z], \end{cases} \tag{5}$$

где  $l^s = l^s(t, \xi) = \zeta(t, \xi)$ ,  $l^{lq} = l^{lq}(t, \xi) = Z - \zeta(t, \xi)$  – толщины зон  $\Omega^s$  и  $\Omega^{lq}$  соответственно. Якобиан преобразования (5):  $J_m = \partial(\tilde{t}, \xi, \eta) / \partial(t, r, z) = 1/l^m$ ,  $m = s, lq$ .

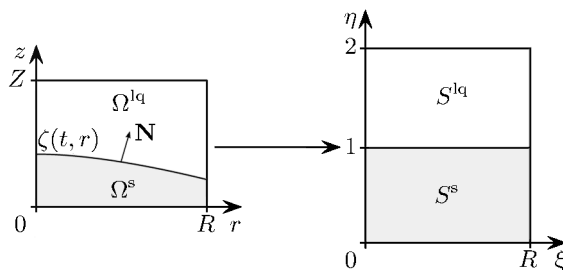


Рис. 1. Преобразование системы координат.

Запишем задачу в новой системе координат  $(\tilde{t}, \xi, \eta)$ . Частные производные неизвестных функций  $f = T, \omega, \psi$  преобразуются следующим образом:

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial \tilde{t}} + \frac{\partial \eta}{\partial t} \frac{\partial f}{\partial \eta}, \quad \frac{\partial f}{\partial r} = \frac{\partial f}{\partial \xi} + \frac{\partial \eta}{\partial r} \frac{\partial f}{\partial \eta}, \quad \frac{\partial f}{\partial z} = \frac{\partial \eta}{\partial z} \frac{\partial f}{\partial \eta}.$$

Для вычисления производных  $\partial \eta / \partial t$ ,  $\partial \eta / \partial r$ ,  $\partial \eta / \partial z$  воспользуемся обратным для (5) преобразованием координат

$$t = \tilde{t}, \quad r = \xi, \quad z = \begin{cases} \varphi^s(\tilde{t}, \xi, \eta) = Z_1 + l^s(\eta - 1), & \eta \in [0, 1], \\ \varphi^{lq}(\tilde{t}, \xi, \eta) = Z_2 + l^{lq}(\eta - 2), & \eta \in [1, 2]. \end{cases} \quad (6)$$

Якобиан этого преобразования равен  $J_m^{-1} = \partial(t, r, z) / \partial(\tilde{t}, \xi, \eta) = l^m$ . Нетрудно показать, что элементы матрицы Якоби вычисляются по формулам [12, 13]

$$\frac{\partial \eta}{\partial t} = -\frac{1}{J_m^{-1}} \frac{\partial z}{\partial \tilde{t}} = -\frac{1}{l^m} \frac{\partial \varphi^m}{\partial \tilde{t}}, \quad \frac{\partial \eta}{\partial r} = -\frac{1}{J_m^{-1}} \frac{\partial z}{\partial \xi} = -\frac{1}{l^m} \frac{\partial \varphi^m}{\partial \xi}, \quad \frac{\partial \eta}{\partial z} = \frac{1}{J_m^{-1}} \frac{\partial r}{\partial \xi} = \frac{1}{l^m}.$$

Для производной по времени получим

$$\frac{\partial f}{\partial t} = \frac{1}{\xi l^m} \left[ \frac{\partial}{\partial \tilde{t}} (\xi l^m f) - \frac{\partial}{\partial \eta} (\xi \varphi_t^m f) \right] = \frac{1}{\xi l^m} \mathcal{T}^{(\xi, \eta)}(f).$$

В дальнейшем для краткости будем опускать волну над  $\tilde{t}$  и верхний индекс  $m$  у функций  $l^m$ ,  $\varphi^m$ ,  $\varphi_t^m = \partial \varphi^m / \partial \tilde{t}$ ,  $\varphi_\xi^m = \partial \varphi^m / \partial \xi$ .

Формулы для вычисления скорости движения жидкости (1) преобразуются следующим образом:

$$V^r = \frac{1}{\xi l} \frac{\partial \psi}{\partial \eta}, \quad V^z = -\frac{1}{\xi l} \left[ l \frac{\partial \psi}{\partial \xi} - \varphi_\xi \frac{\partial \psi}{\partial \eta} \right]. \quad (7)$$

Тогда для конвективных членов (2) имеем

$$\frac{1}{r} \mathcal{K}^{(r, z)}(f, \psi) = \frac{1}{\xi l} \mathcal{K}^{(\xi, \eta)}(f, \psi), \quad \mathcal{K}^{(\xi, \eta)}(f, \psi) = \partial_\xi \mathcal{H}^\xi + \partial_\eta \mathcal{H}^\eta. \quad (8)$$

Компоненты конвективного потока  $\mathcal{H}^{(\xi, \eta)} = (\mathcal{H}^\xi, \mathcal{H}^\eta)$  в расчётной области имеют вид

$$\mathcal{H}^\xi = \xi U^\xi f = [\partial_\eta \psi] f, \quad \mathcal{H}^\eta = \xi U^\eta f = -[\partial_\xi \psi] f. \quad (9)$$

Здесь величины  $U^\xi$  и  $U^\eta$  равны скоростям переноса  $f$  вдоль соответствующих координатных направлений и вычисляются по формулам

$$U^\xi = \partial_\eta \psi / (\xi l), \quad U^\eta = -\partial_\xi \psi / (\xi l).$$

Важно отметить, что выражения (8), (9), полученные в расчётной системе координат  $(\tilde{t}, \xi, \eta)$  для конвективных членов, с точностью до якобиана совпадают с записью соответствующего оператора в физической области  $\Omega^{lq}(t, r, z)$ .

Оператор (3) преобразуется следующим образом:

$$\frac{1}{r} \mathcal{D}^{(r,z)}(r^\beta f) = \frac{1}{\xi l} \mathcal{D}^{(\xi,\eta)}(\xi^\beta f), \quad \mathcal{D}^{(\xi,\eta)}(\xi^\beta f) = \partial_\xi \mathcal{W}^\xi + \partial_\eta \mathcal{W}^\eta,$$

где компоненты потока  $\mathcal{W}^{(\xi,\eta)} = (\mathcal{W}^\xi, \mathcal{W}^\eta)$  в расчётной системе координат имеют вид

$$\mathcal{W}^\xi(\xi^\beta f) = \varkappa \xi^\alpha [L^{\xi\xi} \partial_\xi(\xi^\beta f) + L^{\xi\eta} \partial_\eta(\xi^\beta f)], \tag{10}$$

$$\mathcal{W}^\eta(\xi^\beta f) = \varkappa \xi^\alpha [L^{\eta\xi} \partial_\xi(\xi^\beta f) + L^{\eta\eta} \partial_\eta(\xi^\beta f)]. \tag{11}$$

Коэффициенты  $L^{\xi\xi}$ ,  $L^{\xi\eta}$ ,  $L^{\eta\xi}$ ,  $L^{\eta\eta}$  вычисляются по формулам

$$L^{\xi\xi} = l, \quad L^{\xi\eta} = L^{\eta\xi} = -\varphi_\xi, \quad L^{\eta\eta} = (1 + \varphi_\xi^2)/l.$$

Таким образом, уравнения Навье–Стокса в расчётной системе координат имеют вид

$$\frac{1}{\xi l} \mathcal{T}^{(\xi,\eta)}(\omega) + \frac{1}{\xi l} \mathcal{K}^{(\xi,\eta)}(\omega, \psi) = \frac{1}{\xi l} \mathcal{D}^{(\xi,\eta)}(\xi^2 \omega) + \beta_T g \frac{1}{\xi l} \left[ \frac{\partial}{\partial \xi}(\xi l T) - \frac{\partial}{\partial \eta}(\xi \varphi_\xi T) \right], \tag{12}$$

$$-\omega = \frac{1}{\xi l} \mathcal{D}^{(\xi,\eta)}(\psi). \tag{13}$$

Для уравнения теплопереноса получим

$$\frac{1}{\xi l} \mathcal{T}^{(\xi,\eta)}(T) + \frac{1}{\xi l} \mathcal{K}^{(\xi,\eta)}(T, \psi) = \frac{1}{\xi l} \mathcal{D}^{(\xi,\eta)}(T). \tag{14}$$

Здесь и далее предполагается, что в твёрдой фазе конвективные члены тождественно равны нулю.

Граничные условия и условия на фронте преобразуются аналогично. Для условия Стефана (4) имеем (см. [9])

$$\left[ \left[ k \left( L^{\eta\xi} \frac{\partial T}{\partial \xi} + L^{\eta\eta} \frac{\partial T}{\partial \eta} \right) \right] \right] = \lambda \rho v_{ph}. \tag{15}$$

Определим скалярное произведение функций  $f$ ,  $g$ , интегрируемых с квадратом в области  $S(\xi, \eta)$ , следующим образом:

$$(f, g)_S = \int_S f g \xi l dS.$$

Здесь и далее  $dS = d\xi d\eta$ .

**2.2. Свойства модели в расчётной системе координат.** Система уравнений (12)–(15) обладает следующими свойствами.

**Закон сохранения завихрённости.** Запишем уравнение (13), связывающее функцию тока и завихрённость в неподвижной системе координат:

$$-\omega = [\partial_\xi(lV^z) - \partial_\eta(V^r + \varphi_\xi V^z)]/(\xi l). \tag{16}$$

Здесь мы воспользовались выражением (7) для скорости жидкости и равенством  $l_\xi = \varphi_\xi \eta$ . Применив теорему Остроградского–Гаусса, из уравнения (16) в силу условий прилипания получим закон сохранения завихрённости

$$-\int_{S^{lq}} \omega \xi l dS = \int_{S^{lq}} [\partial_\xi(lV^z) - \partial_\eta(V^r + \varphi_\xi V^z)] dS = \int_1^2 (lV^z)|_{\xi=0} d\eta. \tag{17}$$

**Закон сохранения кинетической энергии.** Конвективные члены в уравнении переноса завихренности (12) не вносят вклад в баланс кинетической энергии [14], т.е.

$$\left( \frac{\mathcal{K}^{(\xi, \eta)}(\omega, \psi)}{\xi l}, \psi \right)_{S^{1q}} = \int_{S^{1q}} \mathcal{K}^{(\xi, \eta)}(\omega, \psi) \psi \, dS = 0. \quad (18)$$

**Закон сохранения внутренней энергии.** В рассматриваемой системе выполняется закон сохранения теплоты. При этом для конвективных членов в уравнениях теплопереноса (14) имеем

$$\left( \frac{\mathcal{K}^{(\xi, \eta)}(T, \psi)}{\xi l}, 1 \right)_{S^{1q}} = 0, \quad \left( \frac{\mathcal{K}^{(\xi, \eta)}(T, \psi)}{\xi l}, T \right)_{S^{1q}} = 0. \quad (19)$$

**Свойства эллиптического оператора.** Оператор  $\frac{\mathcal{D}^{(\xi, \eta)}(f)}{\xi l}$ ,  $f = \xi^2 \omega, \psi, T$ , является самосопряжённым и отрицательно определённым в пространстве функций  $D(S') = \{f \in C^2(S') : f|_{\partial S'} = 0\}$ .

Покажем самосопряжённость оператора. Воспользуемся тем, что на границе области функции  $f$  и  $g$  обращаются в нуль; проинтегрировав дважды, получим

$$\begin{aligned} \left( \frac{\mathcal{D}^{(\xi, \eta)}(f)}{\xi l}, g \right)_{S'} &= \int_{S'} [\partial_\xi (\varkappa \xi^\alpha L^{\xi\xi} \partial_\xi f + \varkappa \xi^\alpha L^{\xi\eta} \partial_\eta f) + \partial_\eta (\varkappa \xi^\alpha L^{\eta\eta} \partial_\eta f + \varkappa \xi^\alpha L^{\eta\xi} \partial_\xi f)] g \, dS = \\ &= \int_{S'} [\partial_\xi (\varkappa \xi^\alpha L^{\xi\xi} \partial_\xi g + \varkappa \xi^\alpha L^{\eta\xi} \partial_\eta g) + \partial_\eta (\varkappa \xi^\alpha L^{\eta\eta} \partial_\eta g + \varkappa \xi^\alpha L^{\xi\eta} \partial_\xi g)] f \, dS = \left( f, \frac{\mathcal{D}^{(\xi, \eta)}(g)}{\xi l} \right)_{S'}. \end{aligned} \quad (20)$$

Самосопряжённость оператора имеет место в силу того, что  $L^{\xi\eta} = L^{\eta\xi}$ .

Покажем отрицательную определённость оператора:

$$\begin{aligned} \left( \frac{\mathcal{D}^{(\xi, \eta)}(f)}{\xi l}, f \right)_{S'} &= \int_{S'} \mathcal{D}^{(\xi, \eta)}(f) f \, dS = \\ &= - \int_{S'} [\varkappa \xi^\alpha L^{\xi\xi} (\partial_\xi f)^2 + 2\varkappa \xi^\alpha L^{\xi\eta} \partial_\xi f \partial_\eta f + \varkappa \xi^\alpha L^{\eta\eta} (\partial_\eta f)^2] \, dS. \end{aligned} \quad (21)$$

В силу критерия Сильвестра подынтегральное выражение является положительно определённой квадратичной формой, поэтому оператор  $\frac{\mathcal{D}^{(\xi, \eta)}(f)}{\xi l}$  – отрицательно определённый.

Численный метод решения задачи будем строить таким образом, чтобы на разностном уровне выполнялись дискретные аналоги свойств (17)–(21).

### 3. Разностная задача.

**3.1. Сетка и сеточные функции.** В расчётной области  $S(\xi, \eta)$  введём прямоугольную сетку  $\omega_h^{(\xi, \eta)} = \omega_h^\xi \times \omega_h^\eta$ :  $\omega_h^\xi = \{\xi_i, i = \overline{0, M}, \xi_0 = 0, \xi_M = R\}$ ,  $\omega_h^\eta = \{\eta_j, j = \overline{0, N}, \eta_0 = 0, \eta_{j^*} = 1, \eta_N = 2\}$  таким образом, чтобы фронт кристаллизации располагался в узлах сетки с координатами  $(\xi_i, \eta_{j^*})$ ,  $i = \overline{0, M}$ ; шаги сетки  $\omega_h^{(\xi, \eta)}$ :  $h_{i+1/2}^\xi = \xi_{i+1} - \xi_i$ ,  $h_{j+1/2}^\eta = \eta_{j+1} - \eta_j$ . Введём также потоковые узлы:  $\xi_{i+1/2} = (\xi_{i+1} + \xi_i)/2$ ,  $\eta_{j+1/2} = (\eta_{j+1} + \eta_j)/2$ . Расчётную область  $S(\xi, \eta)$  разобьём на прямоугольные ячейки  $S_{ij}^{(\xi, \eta)} = [\xi_{i-1/2}, \xi_{i+1/2}] \times [\eta_{j-1/2}, \eta_{j+1/2}]$  с границами  $\partial S_{i+1/2, j}^{(\xi, \eta)}$ ,  $\partial S_{i, j-1/2}^{(\xi, \eta)}$ ,  $\partial S_{i-1/2, j}^{(\xi, \eta)}$ ,  $\partial S_{i, j+1/2}^{(\xi, \eta)}$ ; длины границ ячейки  $S_{ij}^{(\xi, \eta)}$  равны  $\tilde{h}_i^\xi = 0.5(h_{i+1/2}^\xi + h_{i-1/2}^\xi)$  и  $\tilde{h}_j^\eta = 0.5(h_{j+1/2}^\eta + h_{j-1/2}^\eta)$ , её площадь  $dS_{ij}^{(\xi, \eta)} = \tilde{h}_i^\xi \tilde{h}_j^\eta$ . Также рассмотрим ячейки

$S_{i+1/2 j+1/2}^{(\xi, \eta)}$  с центрами в точках  $(\xi_{i+1/2}, \eta_{j+1/2})$  (рис. 2). Сетка по времени  $\omega_\tau = \{t_0 = 0, t_{k+1} = t_k + \tau, k = 0, 1, \dots\}$ , где  $\tau$  – шаг по времени.

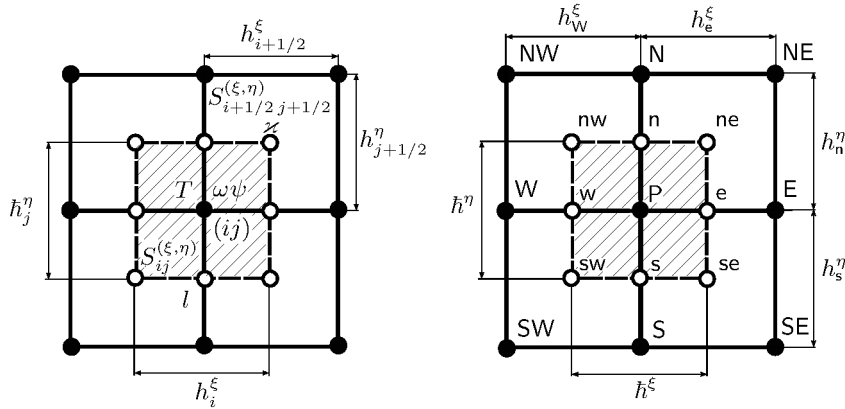


Рис. 2. Сетка и сеточные функции.

Сеточные функции  $f_{ij}^k = f(t_k, \xi_i, \eta_j)$ ,  $f = T, \psi, \omega$ , будем относить к узлам сетки. Определим эти функции внутри расчётных ячеек:  $f(t_k, \xi, \eta) = f_{ij}^k$  при  $(\xi, \eta) \in S_{ij}^{(\xi, \eta)}$ . Длины фаз  $l_{ij+1/2}$  задаются в серединах горизонтальных границ ячеек  $S_{ij}^{(\xi, \eta)}$ . Длина фазы в центре ячейки вычисляется с помощью линейной интерполяции

$$l_{ij} = (l_{ij+1/2} h_{j+1/2}^\eta + l_{ij-1/2} h_{j-1/2}^\eta) / (2h_j^\eta).$$

Таким образом,  $l_{ij} = l_i^s$  в твёрдой фазе,  $l_{ij} = l_i^{lq}$  в жидкой фазе,

$$l_{ij}^* = (l_i^{lq} h_{j^*+1/2}^\eta + l_i^s h_{j^*-1/2}^\eta) / (2h_{j^*}^\eta)$$

на фронте кристаллизации. Физические параметры областей  $\varkappa$  и метрические коэффициенты  $L^{\xi\xi}, L^{\xi\eta}, L^{\eta\eta}$  относятся к центрам ячеек  $S_{i+1/2 j+1/2}^{(\xi, \eta)}$  (см. рис. 2).

Подмножество узлов сетки  $\omega_h^{(\xi, \eta)}$ , принадлежащих жидкой фазе, обозначим через  $\omega_{lq}^{(\xi, \eta)}$ . Введём множество индексов узлов сетки  $\omega_\gamma^{(\xi, \eta)} : (\overline{I}_\gamma \times \overline{J}_\gamma)$ ,  $\gamma = h, lq$ , а также множество индексов внутренних узлов сетки  $\omega_\gamma^{(\xi, \eta)} : (I_\gamma \times J_\gamma)$ . Для проверки выполнения интегральных свойств сеточного решения определим скалярное произведение для сеточных функций, заданных в центрах ячеек  $S_{ij}^{(\xi, \eta)}$ , следующим образом:

$$(f, g)_\gamma = \sum_{(i,j) \in (\overline{I}_\gamma \times \overline{J}_\gamma)} f_{ij} g_{ij} \xi_i l_{ij} dS_{ij}^{(\xi, \eta)}, \quad \gamma = h, lq. \tag{22}$$

Для обозначения узлов сетки введём также локальную географическую нотацию [2] (см. рис. 2). Определим сеточные операторы разностного дифференцирования. Для пространственных производных в направлении оси  $\xi$  имеем  $f_{P, \xi} = (f_E - f_P) / h_e^\xi$ ,  $f_{P, \bar{\xi}} = f_{W, \xi}$ ; разностные аппроксимации пространственных производных в направлении оси  $\eta$ :  $f_{P, \eta} = (f_N - f_P) / h_n^\eta$ ,  $f_{P, \bar{\eta}} = f_{S, \eta}$ ,  $f_{P, \bar{\eta}} = (f_n - f_P) / (0.5h_n^\eta)$ ,  $f_{P, \bar{\eta}} = f_{s, \bar{\eta}}$ . Разностную производную по времени обозначим через  $f_t = (\hat{f}_P - f_P) / \tau$ , где  $\hat{f}_P = f(t_k + \tau, \xi_P, \eta_P)$ .

**3.2. Разностная схема.** С помощью интегро-интерполяционного метода построим консервативную разностную схему. Проинтегрируем уравнения (12)–(14) по ячейке  $S_P^{(\xi, \eta)}$ . В ре-



гулярных точках сетки для эволюционных уравнений (12), (14) получим

$$\int_{S_p^{(\xi,\eta)}} \mathcal{T}^{(\xi,\eta)}(f) dS + \int_{S_p^{(\xi,\eta)}} \mathcal{K}^{(\xi,\eta)}(f, \psi) dS - \int_{S_p^{(\xi,\eta)}} \mathcal{D}^{(\xi,\eta)}(\xi^\beta f) dS = 0. \quad (23)$$

(Аппроксимация подъемной силы в уравнении переноса завихрённости строится стандартным образом.)

Построим сначала аппроксимации для второго и третьего интегралов в уравнении (23). Перейдём с помощью теоремы Остроградского–Гаусса к интегралам по границе ячейки и получим

$$\int_{S_p^{(\xi,\eta)}} \mathcal{K}^{(\xi,\eta)}(f, \psi) dS \approx \mathcal{K}_h^{(\xi,\eta)}(f, \psi) \bar{h}^\xi \bar{h}^\eta = \int_{\partial S^{(\xi,\eta)}} [\mathcal{H}^\xi N^\xi + \mathcal{H}^\eta N^\eta] d\gamma, \quad (24)$$

$$\int_{S_p^{(\xi,\eta)}} \mathcal{D}^{(\xi,\eta)}(\xi^\beta f) dS \approx \mathcal{D}_h^{(\xi,\eta)}(\xi^\beta f) \bar{h}^\xi \bar{h}^\eta = \int_{\partial S^{(\xi,\eta)}} [\mathcal{W}^\xi N^\xi + \mathcal{W}^\eta N^\eta] d\gamma. \quad (25)$$

Здесь  $N^\xi$ ,  $N^\eta$  – компоненты вектора внешней нормали к границе ячейки  $S_p^{(\xi,\eta)}$ .

Процесс построения аппроксимаций для операторов  $\mathcal{K}^{(\xi,\eta)}$  и  $\mathcal{D}^{(\xi,\eta)}$  сводится к вычислению значений соответствующих потоков в центрах границ ячейки  $S_p^{(\xi,\eta)}$ . При этом требуется вычислять значения сеточных функций, а также их производных в узлах, в которых они не определены. Переинтерполяцию соответствующих величин можно осуществлять различными способами, однако предпочтение следует отдавать аппроксимациям, обеспечивающим на разностном уровне выполнение дискретных аналогов свойств (17)–(21) дифференциальной задачи (12)–(15). В дальнейшем при выборе интерполяционных формул будем руководствоваться данным принципом.

**3.2.1. Аппроксимация конвективных членов.** Для конвективных членов имеем

$$\int_{\partial S^{(\xi,\eta)}} [\mathcal{H}^\xi N^\xi + \mathcal{H}^\eta N^\eta] d\gamma = \int_{\partial S_e^{(\xi,\eta)}} \mathcal{H}^\xi d\eta - \int_{\partial S_w^{(\xi,\eta)}} \mathcal{H}^\xi d\eta + \int_{\partial S_n^{(\xi,\eta)}} \mathcal{H}^\eta d\xi - \int_{\partial S_s^{(\xi,\eta)}} \mathcal{H}^\eta d\xi.$$

Отсюда следует, что выражение (24) можно представить в виде

$$\mathcal{K}_h^{(\xi,\eta)}(f, \psi) \bar{h}^\xi \bar{h}^\eta = (\mathcal{H}_e^\xi - \mathcal{H}_w^\xi) \bar{h}^\eta + (\mathcal{H}_n^\eta - \mathcal{H}_s^\eta) \bar{h}^\xi.$$

В уравнении переноса завихрённости (12) конвективный поток в центрах границ ячейки будем аппроксимировать следующим образом:

$$\mathcal{H}_e^\xi = 0.5[\psi_{E,\eta} \circ \omega_E + \psi_{P,\eta} \circ \omega_P], \quad \mathcal{H}_n^\eta = 0.5[\psi_{N,\xi} \circ \omega_N + \psi_{P,\xi} \circ \omega_P].$$

Потоки через “западную” и “южную” границы вычисляются аналогично. Таким образом, аппроксимация конвективных членов в уравнениях Навье–Стокса имеет вид

$$\mathcal{K}_h^{(\xi,\eta)}(\omega, \psi) = (\psi_{P,\eta} \circ \omega_P)_\xi - (\psi_{P,\xi} \circ \omega_P)_\eta. \quad (26)$$

Выражение (26) совпадает с соотношением, предложенным А. Аракавой в работе [14], для аппроксимации конвективных членов в уравнении Навье–Стокса в декартовой системе координат. В работе [14] доказано следующее утверждение.

**Утверждение 1.** *Конвективные члены, записанные в форме (26), не вносят вклад в баланс кинетической энергии, т.е.*

$$\left( \frac{\mathcal{K}_h^{(\xi,\eta)}(\omega, \psi)}{\xi l}, \psi \right)_{lq} = 0.$$

В уравнении теплопереноса (14) компоненты конвективного потока в центрах границ ячеек вычисляются следующим образом:

$$\mathcal{H}_e^\xi(T) = \psi_{e,\eta} \circ T_e, \quad \mathcal{H}_n^\eta(T) = \psi_{n,\xi} \circ T_n,$$

а конвективные члены записываются в виде

$$\mathcal{K}_h^{(\xi,\eta)}(T, \psi) = (\psi_{e,\eta} \circ T_e)_\xi - (\psi_{n,\xi} \circ T_n)_\eta. \tag{27}$$

Выражение (27) также предложено в работе [14] для аппроксимации конвективных членов в уравнении теплопереноса в декартовой системе координат. Там же доказано ещё одно утверждение.

**Утверждение 2.** *Конвективные члены, записанные в форме (27), не влияют на баланс теплоты, а также на среднее по области значение квадрата температуры, т.е.*

$$\left( \frac{\mathcal{K}_h^{(\xi,\eta)}(T, \psi)}{\xi l}, 1 \right)_{lq} = 0, \quad \left( \frac{\mathcal{K}_h^{(\xi,\eta)}(T, \psi)}{\xi l}, T \right)_{lq} = 0.$$

Таким образом, использование выражений (26), (27) для аппроксимации конвективных членов в уравнениях переноса завихрённости и теплопереноса позволяет построить энергетически нейтральную вычислительную схему.

**3.2.2. Аппроксимация эллиптического оператора.** Для оператора  $\mathcal{D}^{(\xi,\eta)}$  из соотношения (25) получим

$$\mathcal{D}_h^{(\xi,\eta)}(\xi^\beta f) \hbar^\xi \hbar^\eta = (\mathcal{W}_e^\xi - \mathcal{W}_w^\xi) \hbar^\eta + (\mathcal{W}_n^\eta - \mathcal{W}_s^\eta) \hbar^\xi. \tag{28}$$

Аппроксимируем потоки в выражении (28) на границе ячейки  $S_p^{(\xi,\eta)}$ . Рассмотрим для определённости “восточную” границу

$$\mathcal{W}_e^\xi \hbar^\eta = \int_{\partial S_e^{(\xi,\eta)}} \mathcal{W}^\xi(\xi^\beta f) d\eta = \int_{\partial S_e^{(\xi,\eta)}} [\mathcal{L}^{\xi\xi} \partial_\xi(\xi^\beta f) + \mathcal{L}^{\xi\eta} \partial_\eta(\xi^\beta f)] d\eta. \tag{29}$$

Здесь  $\mathcal{L}^{\xi\xi} = \xi^\alpha \varkappa L^{\xi\xi}$ ,  $\mathcal{L}^{\xi\eta} = \xi^\alpha \varkappa L^{\xi\eta}$  и  $\mathcal{L}^{\eta\eta} = \xi^\alpha \varkappa L^{\eta\eta}$ ,  $\mathcal{L}^{\eta\xi} = \xi^\alpha \varkappa L^{\eta\xi}$ . Для интегралов по границе ячейки в (29) имеем

$$\begin{aligned} \int_{\partial S_e^{(\xi,\eta)}} \mathcal{L}^{\xi\xi} \partial_\xi(\xi^\beta f) d\eta &\approx \mathcal{L}_e^{\xi\xi}(\xi^\beta f)_\xi \hbar^\eta, \\ \int_{\partial S_e^{(\xi,\eta)}} \mathcal{L}^{\xi\eta} \partial_\eta(\xi^\beta f) d\eta &\approx 0.5[h_n^\eta \mathcal{L}_{ne}^{\xi\eta}(\xi^\beta f)_{e,\eta} + h_s^\eta \mathcal{L}_{se}^{\xi\eta}(\xi^\beta f)_{e,\bar{\eta}}]. \end{aligned} \tag{30}$$

Для вычисления метрических коэффициентов и разностных производных в центре границы  $\partial S_e^{(\xi,\eta)}$  будем использовать следующие интерполяционные формулы:

$$\begin{aligned} \mathcal{L}_e^{\xi\xi} &= [h_n^\eta \mathcal{L}_{ne}^{\xi\xi} + h_s^\eta \mathcal{L}_{se}^{\xi\xi}] / (2\hbar^\eta), \quad (\xi^\beta f)_{e,\eta} = 0.5[(\xi^\beta f)_{P,\eta} + (\xi^\beta f)_{E,\eta}], \\ (\xi^\beta f)_{e,\bar{\eta}} &= 0.5[(\xi^\beta f)_{P,\bar{\eta}} + (\xi^\beta f)_{E,\bar{\eta}}]. \end{aligned}$$

Поток через “северную” границу  $\partial S_n^{(\xi,\eta)}$  вычисляется по формуле

$$\mathcal{W}_n^\eta \hbar^\xi = \int_{\partial S_n^{(\xi,\eta)}} \mathcal{W}^\eta(\xi^\beta f) d\xi \approx \mathcal{L}_n^{\eta\eta}(\xi^\beta f)_\eta \hbar^\xi + 0.5[h_e^\xi \mathcal{L}_{ne}^{\eta\xi}(\xi^\beta f)_{n,\xi} + h_w^\xi \mathcal{L}_{nw}^{\eta\xi}(\xi^\beta f)_{n,\bar{\xi}}],$$

а в центре границы имеем

$$\mathcal{L}_n^{\eta\eta} = [h_e^\xi \mathcal{L}_{ne}^{\eta\eta} + h_w^\xi \mathcal{L}_{nw}^{\eta\eta}] / (2h^\xi), \quad (\xi^\beta f)_{n,\xi} = 0.5[(\xi^\beta f)_{P,\xi} + (\xi^\beta f)_{N,\xi}],$$

$$(\xi^\beta f)_{n,\bar{\xi}} = 0.5[(\xi^\beta f)_{P,\bar{\xi}} + (\xi^\beta f)_{N,\bar{\xi}}].$$

Выражения для потоков через “западную” и “южную” границы вычисляются аналогичным образом.

Рассмотрим пространство сеточных функций  $D_h(\omega_\gamma^{(\xi,\eta)}) = \{f_{ij}, (i, j) \in (\bar{I}_\gamma \times \bar{J}_\gamma) : f_{ij} = 0, (i, j) \in (\bar{I}_\gamma \times \bar{J}_\gamma) \setminus (I_\gamma \times J_\gamma)\}$ . Докажем следующие утверждения.

**Утверждение 3.** Для сеточных функций из пространства  $D_h(\omega_\gamma^{(\xi,\eta)})$  разностный оператор  $\mathcal{D}_h^{(\xi,\eta)}(f_{ij}) / (\xi_i l_{ij})$  является отрицательно определённым относительно сеточного скалярного произведения (22):

$$\left( \frac{\mathcal{D}_h^{(\xi,\eta)}(f)}{\xi l}, f \right)_\gamma < 0, \quad \gamma = h, lq. \tag{31}$$

**Доказательство.** Левую часть неравенства (31) запишем в виде

$$\left( \frac{\mathcal{D}_h^{(\xi,\eta)}(f)}{\xi l}, f \right)_\gamma = \sum_{(i,j) \in (I_\gamma \times J_\gamma)} (\mathcal{W}_{i+j/2}^\xi - \mathcal{W}_{i-j/2}^\xi) f_{ij} h_j^\eta + \sum_{(i,j) \in (I_\gamma \times J_\gamma)} (\mathcal{W}_{ij+1/2}^\eta - \mathcal{W}_{ij-1/2}^\eta) f_{ij} h_i^\xi. \tag{32}$$

Рассмотрим первую сумму в правой части (32). Преобразуем в ней члены, содержащие коэффициент  $\mathcal{L}^{\xi\xi}$ . Представим эти члены в виде полусуммы разностных производных назад и вперед в направлении  $\xi$ :

$$\sum_{(i,j) \in (I_\gamma \times J_\gamma)} [\mathcal{L}_{i+j/2}^{\xi\xi} f_\xi - \mathcal{L}_{i-j/2}^{\xi\xi} f_{\bar{\xi}}] f_{ij} h_j^\eta =$$

$$= \frac{1}{2} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} [h_{i-1/2}^\xi (\mathcal{L}_{i+j/2}^{\xi\xi} f_\xi)_{\bar{\xi}} + h_{i+1/2}^\xi (\mathcal{L}_{i-j/2}^{\xi\xi} f_{\bar{\xi}})_\xi] f_{ij} h_j^\eta. \tag{33}$$

Используем в правой части равенства (33) формулу суммирования по частям [15, с. 46] по индексу  $i$ . С учётом условий на границе области имеем

$$\frac{1}{2} \sum_{j \in J_\gamma} \sum_{i \in I_\gamma} [h_{i-1/2}^\xi (\mathcal{L}_{i+j/2}^{\xi\xi} f_\xi)_{\bar{\xi}} + h_{i+1/2}^\xi (\mathcal{L}_{i-j/2}^{\xi\xi} f_{\bar{\xi}})_\xi] f_{ij} h_j^\eta =$$

$$= -\frac{1}{2} \sum_{j \in J_\gamma} \sum_{i \in I_\gamma} [h_{i+1/2}^\xi \mathcal{L}_{i+j/2}^{\xi\xi} (f_\xi)^2 + h_{i-1/2}^\xi \mathcal{L}_{i-j/2}^{\xi\xi} (f_{\bar{\xi}})^2] h_j^\eta. \tag{34}$$

Запишем члены, содержащие смешанные производные, в виде

$$\mathcal{S}_1 = \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i-1/2}^\xi h_{j+1/2}^\eta (\mathcal{L}_{i+j/2+1/2}^{\xi\eta} f_\eta)_{\bar{\xi}} f_{ij} + \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i+1/2}^\xi h_{j+1/2}^\eta (\mathcal{L}_{i-j/2+1/2}^{\xi\eta} f_\eta)_\xi f_{ij} +$$

$$+ \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i-1/2}^\xi h_{j-1/2}^\eta (\mathcal{L}_{i+j/2-1/2}^{\xi\eta} f_\eta)_{\bar{\xi}} f_{ij} + \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i+1/2}^\xi h_{j-1/2}^\eta (\mathcal{L}_{i-j/2-1/2}^{\xi\eta} f_\eta)_\xi f_{ij}. \tag{35}$$

Применим к первому слагаемому в (35) формулу суммирования по частям по индексу  $i$ :

$$\frac{1}{4} \sum_{j \in J_\gamma} \sum_{i \in I_\gamma} h_{i-1/2}^\xi h_{j+1/2}^\eta (\mathcal{L}_{i+j/2+1/2}^{\xi\eta} f_\eta)_{\bar{\xi}} f_{ij} = -\frac{1}{4} \sum_{j \in J_\gamma} \sum_{i \in I_\gamma} h_{i+1/2}^\xi h_{j+1/2}^\eta \mathcal{L}_{i+j/2+1/2}^{\xi\eta} f_\xi f_\eta.$$

Преобразуем второе, третье и четвертое слагаемые в (35) аналогичным образом и получим

$$\begin{aligned} \mathcal{S}_1 = & -\frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} [h_{i+1/2}^\xi h_{j+1/2}^\eta \mathcal{L}_{i+j/2+1/2}^{\xi\eta} f_\xi f_\eta + h_{i-1/2}^\xi h_{j+1/2}^\eta \mathcal{L}_{i-j/2+1/2}^{\xi\eta} f_{\bar{\xi}} f_\eta] - \\ & -\frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} [h_{i-1/2}^\xi h_{j-1/2}^\eta \mathcal{L}_{i-j/2-1/2}^{\xi\eta} f_{\bar{\xi}} f_{\bar{\eta}} + h_{i+1/2}^\xi h_{j-1/2}^\eta \mathcal{L}_{i+j/2-1/2}^{\xi\eta} f_\xi f_{\bar{\eta}}]. \end{aligned} \quad (36)$$

Рассмотрим вторую сумму в правой части выражения (32). Для членов, содержащих коэффициент  $\mathcal{L}^{\eta\eta}$ , имеем

$$\begin{aligned} & \sum_{(i,j) \in (I_\gamma \times J_\gamma)} [\mathcal{L}_{ij+1/2}^{\eta\eta} f_\eta - \mathcal{L}_{ij-1/2}^{\eta\eta} f_{\bar{\eta}}] f_{ij} h_i^\xi = \\ & = -\frac{1}{2} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} [h_{j+1/2}^\eta \mathcal{L}_{ij+1/2}^{\eta\eta} (f_\eta)^2 + h_{j-1/2}^\eta \mathcal{L}_{ij-1/2}^{\eta\eta} (f_{\bar{\eta}})^2] h_i^\xi. \end{aligned} \quad (37)$$

Члены, содержащие смешанные производные, запишем в виде

$$\begin{aligned} \mathcal{S}_2 = & -\frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} [h_{i+1/2}^\xi h_{j+1/2}^\eta \mathcal{L}_{i+j/2+1/2}^{\xi\eta} f_\xi f_\eta + h_{i+1/2}^\xi h_{j-1/2}^\eta \mathcal{L}_{i+j/2-1/2}^{\xi\eta} f_\xi f_{\bar{\eta}}] - \\ & -\frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} [h_{i-1/2}^\xi h_{j+1/2}^\eta \mathcal{L}_{i-j/2+1/2}^{\xi\eta} f_{\bar{\xi}} f_\eta + h_{i-1/2}^\xi h_{j-1/2}^\eta \mathcal{L}_{i-j/2-1/2}^{\xi\eta} f_{\bar{\xi}} f_{\bar{\eta}}]. \end{aligned} \quad (38)$$

В силу того, что  $\mathcal{L}^{\xi\eta} = \mathcal{L}^{\eta\xi}$ , выражение (38) совпадает с (36), т.е.  $\mathcal{S}_1 = \mathcal{S}_2$ .

Заменив в правой части соотношения (32) потоки их представлениями (34), (36)–(38), получим

$$\begin{aligned} \left( \frac{\mathcal{D}_h^{(\xi,\eta)}(f)}{\xi l}, f \right)_\gamma = & -\frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i+1/2}^\xi h_{j+1/2}^\eta [\mathcal{L}_{i+j/2+1/2}^{\xi\xi} (f_\xi)^2 + 2\mathcal{L}_{i+j/2+1/2}^{\xi\eta} f_\xi f_\eta + \mathcal{L}_{i+j/2+1/2}^{\eta\eta} (f_\eta)^2] - \\ & -\frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i-1/2}^\xi h_{j+1/2}^\eta [\mathcal{L}_{i-j/2+1/2}^{\xi\xi} (f_{\bar{\xi}})^2 + 2\mathcal{L}_{i-j/2+1/2}^{\xi\eta} f_{\bar{\xi}} f_\eta + \mathcal{L}_{i-j/2+1/2}^{\eta\eta} (f_\eta)^2] - \\ & -\frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i+1/2}^\xi h_{j-1/2}^\eta [\mathcal{L}_{i+j/2-1/2}^{\xi\xi} (f_\xi)^2 + 2\mathcal{L}_{i+j/2-1/2}^{\xi\eta} f_\xi f_{\bar{\eta}} + \mathcal{L}_{i+j/2+1/2}^{\eta\eta} (f_{\bar{\eta}})^2] - \\ & -\frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i-1/2}^\xi h_{j-1/2}^\eta [\mathcal{L}_{i-j/2-1/2}^{\xi\xi} (f_{\bar{\xi}})^2 + 2\mathcal{L}_{i-j/2-1/2}^{\xi\eta} f_{\bar{\xi}} f_{\bar{\eta}} + \mathcal{L}_{i-j/2-1/2}^{\eta\eta} (f_{\bar{\eta}})^2]. \end{aligned}$$

В силу критерия Сильвестра квадратичные формы, находящиеся под знаками сумм, являются положительно определёнными:

$$\mathcal{L}_{pq}^{\xi\xi} = l_{pq} > 0, \quad \mathcal{L}_{pq}^{\xi\xi} \mathcal{L}_{pq}^{\eta\eta} - (\mathcal{L}_{pq}^{\xi\eta})^2 = 1 > 0, \quad p = i \pm 1/2, \quad q = j \pm 1/2,$$

откуда

$$\left( \frac{\mathcal{D}_h^{(\xi,\eta)}(f)}{\xi l}, f \right)_\gamma < 0, \quad \gamma = h, lq.$$

Утверждение доказано.

**Утверждение 4.** Для сеточных функций из пространства  $D_h(\omega_\gamma^{(\xi,\eta)})$  разностный оператор  $\mathcal{D}_h^{(\xi,\eta)}(f_{ij})/(\xi_i l_{ij})$  является самосопряжённым относительно сеточного скалярного произведения (22):

$$\left(\frac{\mathcal{D}_h^{(\xi,\eta)}(f)}{\xi l}, g\right)_\gamma = \left(f, \frac{\mathcal{D}_h^{(\xi,\eta)}(g)}{\xi l}\right)_\gamma, \quad \gamma = h, lq. \tag{39}$$

**Доказательство.** Запишем левую часть равенства (39) следующим образом:

$$\begin{aligned} \left(\frac{\mathcal{D}_h^{(\xi,\eta)}(f)}{\xi l}, g\right)_\gamma &= \sum_{(i,j) \in (I_\gamma \times J_\gamma)} \mathcal{D}_h^{(\xi,\eta)}(f_{ij}) g_{ij} \bar{h}_i^\xi \bar{h}_j^\eta = \\ &= \sum_{(i,j) \in (I_\gamma \times J_\gamma)} (\mathcal{W}_{i+j/2}^\xi - \mathcal{W}_{i-j/2}^\xi) g_{ij} \bar{h}_j^\eta + \sum_{(i,j) \in (I_\gamma \times J_\gamma)} (\mathcal{W}_{ij+1/2}^\eta - \mathcal{W}_{ij-1/2}^\eta) g_{ij} \bar{h}_i^\xi. \end{aligned} \tag{40}$$

Рассмотрим первую сумму в правой части (40). Преобразуем в ней члены, содержащие коэффициент  $\mathcal{L}^{\xi\xi}$ . Для этого дважды воспользуемся формулой суммирования по частям по индексу  $i$ :

$$\begin{aligned} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i-1/2}^\xi (\mathcal{L}_{i+j/2}^{\xi\xi} f_\xi)_{\bar{\xi}} g_{ij} \bar{h}_j^\eta &= - \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i+1/2}^\xi \mathcal{L}_{i+j/2}^{\xi\xi} f_\xi g_\xi \bar{h}_j^\eta = \\ &= \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i-1/2}^\xi (\mathcal{L}_{i+j/2}^{\xi\xi} g_\xi)_{\bar{\xi}} f_{ij} \bar{h}_j^\eta. \end{aligned} \tag{41}$$

Применив последовательно формулу суммирования по частям по индексам  $i$  и  $j$ , запишем члены, содержащие смешанные производные, в виде

$$\begin{aligned} &\frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i-1/2}^\xi h_{j+1/2}^\eta (\mathcal{L}_{i+j/2+1/2}^{\xi\eta} f_\eta)_{\bar{\xi}} g_{ij} + \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i+1/2}^\xi h_{j+1/2}^\eta (\mathcal{L}_{i-j/2+1/2}^{\xi\eta} f_\eta)_\xi g_{ij} + \\ &+ \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i-1/2}^\xi h_{j-1/2}^\eta (\mathcal{L}_{i+j/2-1/2}^{\xi\eta} f_\eta)_{\bar{\xi}} g_{ij} + \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i+1/2}^\xi h_{j-1/2}^\eta (\mathcal{L}_{i-j/2-1/2}^{\xi\eta} f_\eta)_\xi g_{ij} = \\ &= \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i+1/2}^\xi h_{j-1/2}^\eta (\mathcal{L}_{i+j/2+1/2}^{\xi\eta} g_\xi)_{\bar{\eta}} f_{ij} + \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i-1/2}^\xi h_{j-1/2}^\eta (\mathcal{L}_{i-j/2+1/2}^{\xi\eta} g_\xi)_{\bar{\eta}} f_{ij} + \\ &+ \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i+1/2}^\xi h_{j+1/2}^\eta (\mathcal{L}_{i+j/2-1/2}^{\xi\eta} g_\xi)_\eta f_{ij} + \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i-1/2}^\xi h_{j+1/2}^\eta (\mathcal{L}_{i-j/2-1/2}^{\xi\eta} g_\xi)_\eta f_{ij}. \end{aligned} \tag{42}$$

Здесь воспользовались тем, что

$$\begin{aligned} &\frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i-1/2}^\xi h_{j+1/2}^\eta (\mathcal{L}_{i+j/2+1/2}^{\xi\eta} f_\eta)_{\bar{\xi}} g_{ij} = \\ &= -\frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i+1/2}^\xi h_{j+1/2}^\eta \mathcal{L}_{i+j/2+1/2}^{\xi\eta} f_\eta g_\xi = \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i+1/2}^\xi h_{j-1/2}^\eta (\mathcal{L}_{i+j/2+1/2}^{\xi\eta} g_\xi)_{\bar{\eta}} f_{ij}, \end{aligned}$$

оставшиеся слагаемые в левой части равенства (42) преобразованы аналогично.

Рассмотрим вторую сумму в правой части (40). Для членов, содержащих коэффициент  $\mathcal{L}^{\eta\eta}$ , получим

$$\sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{j-1/2}^\eta (\mathcal{L}_{j+1/2}^{\eta\eta} f_\eta)_{\bar{\eta}} g_{ij} \bar{h}_i^\xi = \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{j-1/2}^\eta (\mathcal{L}_{ij+1/2}^{\eta\eta} g_\eta)_{\bar{\eta}} f_{ij} \bar{h}_i^\xi. \tag{43}$$

Аналогично (42) запишем

$$\begin{aligned} & \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i+1/2}^\xi h_{j-1/2}^\eta (\mathcal{L}_{i+j/2+1/2}^{\eta\xi} f_\xi)_{\bar{\eta}} g_{ij} + \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i+1/2}^\xi h_{j+1/2}^\eta (\mathcal{L}_{i+j/2-1/2}^{\eta\xi} f_\xi)_\eta g_{ij} + \\ & + \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i-1/2}^\xi h_{j-1/2}^\eta (\mathcal{L}_{i-j/2+1/2}^{\eta\xi} f_\xi)_{\bar{\eta}} g_{ij} + \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i-1/2}^\xi h_{j+1/2}^\eta (\mathcal{L}_{i-j/2-1/2}^{\eta\xi} f_\xi)_\eta g_{ij} = \\ & = \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i-1/2}^\xi h_{j+1/2}^\eta (\mathcal{L}_{i+j/2+1/2}^{\eta\xi} g_\eta)_{\bar{\xi}} f_{ij} + \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i-1/2}^\xi h_{j-1/2}^\eta (\mathcal{L}_{i+j/2-1/2}^{\eta\xi} g_\eta)_{\bar{\xi}} f_{ij} + \\ & + \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i+1/2}^\xi h_{j+1/2}^\eta (\mathcal{L}_{i-j/2+1/2}^{\eta\xi} g_\eta)_\xi f_{ij} + \frac{1}{4} \sum_{(i,j) \in (I_\gamma \times J_\gamma)} h_{i+1/2}^\xi h_{j-1/2}^\eta (\mathcal{L}_{i-j/2-1/2}^{\eta\xi} g_\eta)_\xi f_{ij}. \end{aligned} \quad (44)$$

Подставим правые части выражений (41)–(44) в правую часть равенства (40). В силу того, что  $\mathcal{L}^{\xi\eta} = \mathcal{L}^{\eta\xi}$ , имеем

$$\left( \frac{\mathcal{D}_h^{(\xi,\eta)}(f)}{\xi l}, g \right)_\gamma = \left( f, \frac{\mathcal{D}_h^{(\xi,\eta)}(g)}{\xi l} \right)_\gamma, \quad \gamma = h, lq.$$

Утверждение доказано.

Возможны и другие способы переинтерполяции искомых функций в потоковые точки. Соответствующие разностные выражения могут иметь тот же порядок аппроксимации, что и операторы (26)–(28). Однако построенные на их основе разностные уравнения не будут обладать свойствами, аналогичными свойствам исходной дифференциальной задачи. В работе [16] предложена следующая аппроксимация оператора  $\mathcal{D}^{(\xi,\eta)}$ :

$$\widetilde{\mathcal{D}}_h^{(\xi,\eta)}(\xi^\beta f) \hbar^\xi \hbar^\eta = (\widetilde{\mathcal{W}}_e^\xi - \widetilde{\mathcal{W}}_w^\xi) \hbar^\eta + (\widetilde{\mathcal{W}}_n^\eta - \widetilde{\mathcal{W}}_s^\eta) \hbar^\xi,$$

$$\widetilde{\mathcal{W}}_e^\xi = \mathcal{L}_e^{\xi\xi}(\xi^\beta f)_\xi + \mathcal{L}_e^{\xi\eta}[(\xi^\beta f)_{ne} - (\xi^\beta f)_{se}]/\hbar^\eta, \quad \widetilde{\mathcal{W}}_n^\eta = \mathcal{L}_n^{\eta\eta}(\xi^\beta f)_\eta + \mathcal{L}_n^{\eta\xi}[(\xi^\beta f)_{ne} - (\xi^\beta f)_{nw}]/\hbar^\xi.$$

Нетрудно показать, что разностный оператор  $\widetilde{\mathcal{D}}_h^{(\xi,\eta)}(f_{ij})/(\xi_i l_{ij})$  не является самосопряжённым относительно сеточного скалярного произведения (22). При этом выражения, используемые в статье [16] для аппроксимации конвективных членов, не обеспечивают выполнение разностных аналогов законов сохранения кинетической энергии, массы и теплоты.

**3.2.3. Аппроксимация производной по времени.** Для оператора  $\mathcal{T}^{(\xi,\eta)}(f)$  получим

$$\int_{S_P^{(\xi,\eta)}} \mathcal{T}^{(\xi,\eta)}(f) d\xi d\eta \approx \mathcal{T}_h^{(\xi,\eta)}(f) \hbar^\xi \hbar^\eta = \partial_t(l_P f_P)_{\xi_P} \hbar^\xi \hbar^\eta - [(\varphi_t f)_n - (\varphi_t f)_s]_{\xi_P} \hbar^\xi. \quad (45)$$

Здесь  $(\varphi_t f)_n = \varphi_{n,t} f_n$ ,  $(\varphi_t f)_s = \varphi_{s,t} f_s$ . Значения функции  $f$  в центрах горизонтальных граней ячейки вычисляется по формулам  $f_n = (f_P + f_N)/2$ ,  $f_s = (f_P + f_S)/2$ . Проинтегрируем (45) по отрезку  $[t_k, t_{k+1}]$ , полученное выражение разделим на шаг  $\tau$  и площадь ячейки  $dS_P^{(\xi,\eta)}$ :

$$\frac{1}{\xi_P} \mathcal{T}_{h,\tau}^{(\xi,\eta)}(f) = (l_P f_P)_t - [(\varphi_t \widehat{f})_n - (\varphi_t \widehat{f})_s]/\hbar^\eta. \quad (46)$$

Аппроксимация (46) содержит нелинейный член  $\widehat{l_P f_P}$ , для удобства запишем её в недивергентной форме:

$$\frac{1}{\xi_P} \mathcal{T}_{h,\tau}^{(\xi,\eta)}(f) = (l_P f_P)_t - [(\varphi_t \widehat{f})_{\bar{\eta}} h_n^\eta + (\varphi_t \widehat{f})_{\bar{\eta}} h_s^\eta]/(2\hbar^\eta).$$

Воспользуемся формулами разностного дифференцирования произведения [15, с. 51]:

$$(l_P f_P)_t = l_{P,t} \widehat{f}_P + l_P f_{P,t} = [l_{n,t} h_n^\eta + l_{s,t} h_s^\eta] / (2\hbar^\eta) \widehat{f}_P + [l_n h_n^\eta + l_s h_s^\eta] / (2\hbar^\eta) f_{P,t},$$

$$[(\varphi_t \widehat{f})_{\bar{\eta}} h_n^\eta + (\varphi_t \widehat{f})_{\bar{\eta}} h_s^\eta] / (2\hbar^\eta) = [(\varphi_{t\bar{\eta}} \widehat{f}_P + \varphi_{t,n} \widehat{f}_{\bar{\eta}}) h_n^\eta + (\varphi_{t\bar{\eta}} \widehat{f}_P + \varphi_{t,s} \widehat{f}_{\bar{\eta}}) h_s^\eta] / (2\hbar^\eta).$$

Отсюда с учётом того, что  $l_{n,t} = \varphi_{t\bar{\eta}}$ ,  $l_{s,t} = \varphi_{t\bar{\eta}}$ , получаем недивергентную форму записи производной по времени:

$$\frac{1}{\xi_P} \mathcal{T}_{h,\tau}^{(\xi,\eta)}(f) = l_P f_{P,t} - [\varphi_{t,n} \widehat{f}_{\bar{\eta}} h_n^\eta + \varphi_{t,s} \widehat{f}_{\bar{\eta}} h_s^\eta] / (2\hbar^\eta). \tag{47}$$

**3.2.4. Аппроксимация уравнения, связывающего значения функции тока и завихрённости.** Аппроксимация уравнения (13) во внутренних точках области имеет вид

$$-\omega_P l_P \xi_P dS_P^{(\xi,\eta)} = \mathcal{D}_h^{(\xi,\eta)}(\psi) dS_P^{(\xi,\eta)}. \tag{48}$$

Граничные условия для завихрённости строятся следующим образом: уравнение (13) интегрируется по приграничным ячейкам с учётом условий прилипания и непротекания. Рассмотрим для определённости ячейку на вертикальной границе области. Получим

$$-\omega_P r_P l_P \frac{h_W^\xi}{2} \hbar^\eta = [\mathcal{W}_P^\xi(\psi) - \mathcal{W}_W^\xi(\psi)] \hbar^\eta + [\mathcal{W}_n^\eta(\psi) - \mathcal{W}_s^\eta(\psi)] \frac{h_W^\xi}{2}.$$

В силу условия прилипания на границе жидкой фазы имеем  $\psi = 0$ ,  $\mathcal{W}_P^\xi(\psi) = 0$ ,  $\psi_{P,\xi} = \psi_{N,\xi} = \psi_{S,\xi} = 0$ , откуда

$$-\omega_P r_P l_P \frac{h_W^\xi}{2} \hbar^\eta = \frac{\mathcal{L}_{nW}^{\xi\xi} h_n^\eta + \mathcal{L}_{sW}^{\xi\xi} h_s^\eta}{2h_W^\xi} \psi_W - \frac{\mathcal{L}_{nW}^{\xi\eta}}{2} \psi_{NW} + \frac{\mathcal{L}_{sW}^{\xi\eta}}{2} \psi_{SW}. \tag{49}$$

В декартовой системе координат  $\mathcal{L}^{\xi\xi} = 1$ ,  $\mathcal{L}^{\xi\eta} = \mathcal{L}^{\eta\xi} = 0$  и условие (49) преобразуется в известное условие А. Тома [17]. Уравнения для завихрённости на других участках границы жидкой фазы аппроксимируются аналогичным образом.

**Утверждение 5.** Пусть сеточная функция  $\omega_{ij}$  удовлетворяет уравнению (48). Тогда выполнен разностный аналог закона сохранения завихрённости в форме

$$-\sum_{(i,j) \in (\bar{I}_{1q} \times \bar{J}_{1q})} \omega_{ij} l_{ij} \xi_i dS_{ij}^{(\xi,\eta)} = \sum_{j \in \bar{J}_{1q}} (V^z l)_{1/2j} \hbar_j^\eta.$$

**Доказательство.** Из уравнения (48) следует, что

$$-\sum_{(i,j) \in (\bar{I}_{1q} \times \bar{J}_{1q})} \omega_{ij} l_{ij} \xi_i dS_{ij}^{(\xi,\eta)} = \sum_{(i,j) \in (I_{1q} \times J_{1q})} [(\mathcal{W}_{i+j/2}^\xi - \mathcal{W}_{i-j/2}^\xi) \hbar_j^\eta + (\mathcal{W}_{ij+1/2}^\eta - \mathcal{W}_{ij-1/2}^\eta) \hbar_i^\xi] - \sum_{j \in J_{1q}} [\omega_{0j} l_{0j} \xi_0 dS_{0j}^{(\xi,\eta)} + \omega_{Nj} l_{Nj} \xi_N dS_{Nj}^{(\xi,\eta)}] - \sum_{i \in I_{1q}} [\omega_{ij^*} l_{0j^*+0} \xi_i dS_{ij^*}^{(\xi,\eta)} + \omega_{iM} l_{iM} \xi_i dS_{iM}^{(\xi,\eta)}]. \tag{50}$$

Сумму потоков, протекающих через грани внутренних ячеек сетки, можно свести к следующей сумме по граничным контрольным объёмам:

$$\sum_{(i,j) \in (I_{1q} \times J_{1q})} [(\mathcal{W}_{i+j/2}^\xi - \mathcal{W}_{i-j/2}^\xi) \hbar_j^\eta + (\mathcal{W}_{ij+1/2}^\eta - \mathcal{W}_{ij-1/2}^\eta) \hbar_i^\xi] =$$

$$= \sum_{j \in J_{1q}} [\mathcal{W}_{N-j/2}^\xi - \mathcal{W}_{j/2}^\xi] \hbar_j^\eta + \sum_{i \in I_{1q}} [\mathcal{W}_{iM-1/2}^\eta - \mathcal{W}_{ij^*+1/2}^\eta] \hbar_i^\xi. \tag{51}$$

Подставив (50) в (51), получим

$$\begin{aligned} - \sum_{(i,j) \in (\bar{I}_{1q} \times \bar{J}_{1q})} \omega_{ij} l_{ij} \xi_i dS_{ij}^{(\xi,\eta)} &= - \sum_{j \in J_{1q}} [\omega_{0j} l_{0j} \xi_0 dS_{0j}^{(\xi,\eta)} + \mathcal{W}_{j/2}^\xi \hbar_j^\eta] - \sum_{j \in J_{1q}} [\omega_{Nj} l_{Nj} \xi_N dS_{Nj}^{(\xi,\eta)} - \mathcal{W}_{N-j/2}^\xi \hbar_j^\eta] - \\ &- \sum_{i \in I_{1q}} [\omega_{ij^*} l_{ij^*+0} \xi_i dS_{ij^*}^{(\xi,\eta)} + \mathcal{W}_{ij^*+1/2}^\eta \hbar_i^\xi] - \sum_{i \in I_{1q}} [\omega_{iM} l_{iM} \xi_i dS_{iM}^{(\xi,\eta)} - \mathcal{W}_{iM-1/2}^\eta \hbar_i^\xi]. \end{aligned} \tag{52}$$

Завихрённость на оси  $\omega_{0j} = 0$ , и в силу граничных условий (49) вторая, третья и четвёртая суммы в правой части (52) также обращаются в нуль. Тогда

$$- \sum_{(i,j) \in (\bar{I}_{1q} \times \bar{J}_{1q})} \omega_{ij} l_{ij} \xi_i dS_{ij}^{(\xi,\eta)} = \sum_{j \in \bar{J}_{1q}} -\mathcal{W}_{1/2j}^\xi \hbar_j^\eta.$$

Сравнив выражение (7) для скорости  $V^z$  и соотношение (10) для потока  $\mathcal{W}^\xi$ , получим, что  $V^z l = -\mathcal{W}^\xi(\psi)$ , откуда

$$- \sum_{(i,j) \in (\bar{I}_{1q} \times \bar{J}_{1q})} \omega_{ij} l_{ij} \xi_i dS_{ij}^{(\xi,\eta)} = \sum_{j \in \bar{J}_{1q}} (V^z l)_{1/2j} \hbar_j^\eta.$$

Утверждение доказано.

Заменим в уравнениях (12), (13) дифференциальные операторы их разностными аппроксимациями (26)–(28) и (47). Таким образом, неявная разностная схема для уравнений Навье–Стокса имеет вид

$$\begin{aligned} \mathcal{T}_{h,\tau}^{(\xi,\eta)}(\omega) dS_P^{(\xi,\eta)} + \mathcal{K}_h^{(\xi,\eta)}(\hat{\omega}, \hat{\psi}) dS_P^{(\xi,\eta)} &= \mathcal{D}_h^{(\xi,\eta)}(\hat{\omega}) dS_P^{(\xi,\eta)} - \mathcal{F}_h^{(\xi,\eta)}(T) dS_P^{(\xi,\eta)}, \\ - \hat{\omega}_{\bar{l}P} \xi_P dS_P^{(\xi,\eta)} &= \mathcal{D}_h^{(\xi,\eta)}(\hat{\psi}) dS_P^{(\xi,\eta)}. \end{aligned}$$

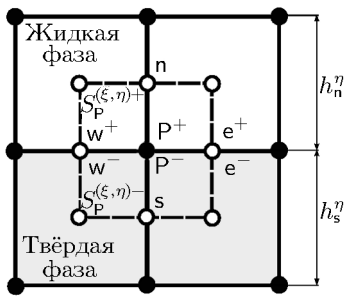


Рис. 3. Ячейка на фронте кристаллизации.

Функция тока и завихрённость в граничных условиях (49) также отнесены к верхнему временному слою.

**3.2.5. Аппроксимация условия Стефана.** Для аппроксимации условия Стефана (15) проинтегрируем уравнение теплопереноса (14) по ячейке  $S_P^{(\xi,\eta)}$ , содержащей границу раздела фаз. Область интегрирования разобьём на две подобласти:  $S_P^{(\xi,\eta)+}$  – расположенную в жидкости и  $S_P^{(\xi,\eta)-}$  – лежащую в твёрдой фазе (рис. 3). Аппроксимация  $\mathcal{T}^{(\xi,\eta)}(T)$  на межфазной границе задаётся формулой (45). Однако стоит помнить, что  $l_n$ ,  $\varphi_{t,n}$  и  $l_s$ ,  $\varphi_{t,s}$  относятся к жидкой и твёрдой фазам соответственно. Проинтегрировав конвективные члены по области  $S_P^{(\xi,\eta)+}$ , получим

$$\mathcal{K}_h^{(\xi,\eta)}(f, \psi) \hbar^\xi \frac{h_n^\eta}{2} = [\mathcal{H}_e^\xi(T) - \mathcal{H}_w^\xi(T)] \frac{h_n^\eta}{2} + [\mathcal{H}_n^\eta(T) - \mathcal{H}_p^\eta(T)] \hbar^\xi.$$

Для вычисления  $\mathcal{H}_w^\xi(T)$ ,  $\mathcal{H}_e^\xi(T)$  используются следующие выражения:  $\mathcal{H}_w^\xi(T) = \psi_{w,\eta} T_w$ ,  $\mathcal{H}_e^\xi(T) = \psi_{e,\eta} T_e$ ; при этом компонента  $\mathcal{H}_p^\eta(T) = 0$ , так как  $\psi_{p,\xi} = 0$ . Проинтегрировав диссипативные члены по областям  $S_P^{(\xi,\eta)+}$  и  $S_P^{(\xi,\eta)-}$ , имеем

$$\mathcal{D}_h^{(\xi,\eta)+}(T) \hbar^\xi \frac{h_n^\eta}{2} \approx \int_{S_P^{(\xi,\eta)+}} \mathcal{D}^{(\xi,\eta)}(T) d\Omega = [\mathcal{W}_{e^+}^\xi - \mathcal{W}_{w^+}^\xi] \frac{h_n^\eta}{2} + [\mathcal{W}_n^\eta - \mathcal{W}_{p^+}^\eta] \hbar^\xi, \tag{53}$$



$$\mathcal{D}_h^{(\xi, \eta)-}(T) \hbar^\xi \frac{h_s^\eta}{2} \approx \int_{S_P^{(\xi, \eta)-}} \mathcal{D}^{(\xi, \eta)}(T) d\Omega = [\mathcal{W}_{e^-}^\xi - \mathcal{W}_{w^-}^\xi] \frac{h_s^\eta}{2} + [\mathcal{W}_{P^-}^\eta - \mathcal{W}_s^\eta] \hbar^\xi. \quad (54)$$

Здесь величины с индексами  $P^+$ ,  $e^+$ ,  $w^+$  принадлежат  $S_P^{(\xi, \eta)+}$ , а с индексами  $P^-$ ,  $e^-$ ,  $w^-$  располагаются в  $S_P^{(\xi, \eta)-}$ . Просуммировав (53) и (54), получим следующую аппроксимацию:

$$\mathcal{D}_h^{(\xi, \eta)}(T) \hbar^\xi \hbar^\eta = [\mathcal{W}_e^\xi - \mathcal{W}_w^\xi] \hbar^\eta + [\mathcal{W}_n^\eta - \mathcal{W}_s^\eta] \hbar^\xi + [\mathcal{W}_{P^-}^\eta - \mathcal{W}_{P^+}^\eta] \hbar^\xi.$$

Из условия Стефана (15) следует, что

$$\mathcal{W}_{P^-}^\eta - \mathcal{W}_{P^+}^\eta = \lambda \rho \xi_P v_{ph}.$$

Таким образом на границе раздела фаз аппроксимация диссипативных членов (28) дополняется слагаемым, описывающим выделение/поглощение теплоты при смене агрегатного состояния вещества.

Во внутренних точках сетки неявная разностная схема для уравнений теплопереноса имеет вид

$$\mathcal{T}_{h, \tau}^{(\xi, \eta)}(T) dS_P^{(\xi, \eta)} + \mathcal{K}_h^{(\xi, \eta)}(\widehat{\psi}, \widehat{T}) dS_P^{(\xi, \eta)} = \mathcal{D}_h^{(\xi, \eta)}(\widehat{T}) dS_P^{(\xi, \eta)}.$$

Благодаря тому, что аппроксимации условий на границе раздела фаз и границах областей согласованы с аппроксимациями соответствующих уравнений во внутренних узлах сетки, на разностном уровне выполнен закон сохранения энергии. Из выполнения дискретного аналога закона сохранения для квадрата искомой величины следует, что температура на протяжении всего расчёта ограничена в любой точке сетки. Данное обстоятельство исключает возможность возникновения в расчёте нелинейной вычислительной неустойчивости (nonlinear computational instability [14, 18]).

**Заключение.** Для двухфазной задачи Стефана в цилиндрической системе координат построена энергетически нейтральная консервативная разностная схема, в которой на разностном уровне выполняются дискретные аналоги законов сохранения завихрённости, кинетической и внутренней энергии. Показано, что разностный оператор, аппроксимирующий диссипативные члены в уравнениях Навье–Стокса и теплопереноса, является самосопряжённым и отрицательно определённым. В работе [11] показано, что преобразование системы координат, выполненное на дискретном уровне в построенной разностной схеме, позволяет получить консервативную разностную схему для задачи Стефана на подвижной сетке, согласованной с формой фронта кристаллизации. Таким образом, предложенная разностная схема наследует основные свойства исходной дифференциальной задачи как в неподвижной, так и в физической системе координат. Важно отметить, что рассмотренная разностная схема естественным образом обобщается на случай численного моделирования процесса кристаллизации многокомпонентного соединения [9].

## СПИСОК ЛИТЕРАТУРЫ

1. Самарский А.А. Введение в теорию разностных схем. М., 1971.
2. Patankar S. Numerical Heat Transfer and Fluid Flow. London, 1981.
3. Самарский А.А., Попов Ю.П. Разностные методы решения задач газовой динамики. М., 1971.
4. Тихонов А.Н., Самарский А.А. О разностных схемах для уравнений с разрывными коэффициентами // Докл. АН СССР. 1956. Т. 108. № 3. С. 393–396.
5. Тихонов А.Н., Самарский А.А. О сходимости разностных схем в классе разрывных коэффициентов // Докл. АН СССР. 1959. Т. 124. № 3. С. 529–532.
6. Тихонов А.Н., Самарский А.А. Об однородных разностных схемах // Журн. вычислит. математики и мат. физики. 1961. Т. 1. № 1. С. 5–63.
7. Тихонов А.Н., Самарский А.А. Об однородных разностных схемах // Докл. АН СССР. 1958. Т. 122. № 4. С. 562–566.

8. *Мажорова О.С., Попов Ю.П., Щерица О.В.* Консервативные разностные схемы для термодиффузионной задачи Стефана // Дифференц. уравнения. 2013. Т. 49. № 7. С. 897–905.
9. *Gusev A.O., Shcheritsa O.V., Mazhorova O.S.* Conservative finite volume strategy for investigation of solution crystal growth techniques // Computers & Fluids. 2020. V. 202. P. 104501.
10. *Landau H.G.* Heat conduction in a melting solid // J. of Appl. Math. 1950. V. 8. P. 81–94.
11. *Гусев А.О., Щерица О.В., Мажорова О.С.* К вопросу об эквивалентности разностных методов решения задачи Стефана на подвижных и фиксированных сетках // Дифференц. уравнения. 2021. Т. 57. № 7. С. 907–921.
12. *Флетчер К.* Вычислительные методы в динамике жидкости. Т. 2. М., 1991.
13. *Steger J.* Implicit finite-difference simulation of flow about arbitrary two-dimensional geometries // Amer. Inst. of Aeronaut. Astronaut. J. 1978. V. 16. № 7. P. 679–686.
14. *Arakawa A.* Computational design for long-term numerical integration of the equation of fluid motion: two dimensional incompressible flow // J. Comput. Phys. 1966. V. 1. P. 119–143.
15. *Самарский А.А.* Теория разностных схем. М., 1989.
16. *Lan C.W.* Newton's method for solving heat transfer, fluid flow and interface shapes in a floating molten zone // Int. J. Numer. Methods Fluids. 1994. V. 19. P. 41–65.
17. *Thom A.* The flow past circular cylinders at low speeds // Proc. Roy. Soc. London. Ser. A. 1933. V. 141. № 4. P. 651–669.
18. *Phillips N.* An Example of Non-linear Computational Instability. The Atmosphere and the Sea in Motion. New York, 1959. P. 501–504.

Институт прикладной математики  
имени М.В. Келдыша РАН, г. Москва,

Поступила в редакцию 10.03.2022 г.  
После доработки 10.03.2022 г.  
Принята к публикации 25.05.2022 г.

УДК 519.63+517.957

## КОМПАКТНЫЕ И МОНОТОННЫЕ РАЗНОСТНЫЕ СХЕМЫ ДЛЯ ОБОБЩЁННОГО УРАВНЕНИЯ ФИШЕРА

© 2022 г. П. П. Матус, Б. Д. Утебаев

Для обобщённого уравнения Фишера с нелинейной конвекцией строятся и изучаются на стандартных шаблонах монотонные и компактные разностные схемы  $4+1$  и  $4+2$  порядков аппроксимации при нестандартных соотношениях на временной и пространственной шаг. Получены априорные оценки разностного решения в нелинейном случае на основе установленных двусторонних оценок сеточного решения. Эти результаты обобщаются на двумерное обобщённое уравнение Фишера с нелинейной конвекцией. Проведённый вычислительный эксперимент иллюстрирует эффективность рассматриваемых методов.

DOI: 10.31857/S037406412207007X, EDN: CELASR

**Введение.** Разработка численных методов исследования математических моделей реакционно-диффузионных систем открывает новые возможности для изучения нелинейных явлений в сложных физических, химических, биологических и других системах. К числу таких моделей принадлежит популяционная модель Фишера или Колмогорова–Петровского–Пискунова (КПП) (см. работы [1, 2])

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u(1 - u), \quad (1)$$

в которой учитываются основные механизмы, характерные для популяционной динамики. Уравнение Фишера (или КПП) и его модификации встречаются в различных задачах, например, в теории горения, в теории фазовых переходов, в физике плазмы и др. [3]. В книге [4, с. 454] было предложено обобщение уравнения Фишера (обобщённое уравнение Бюргерса–Фишера с нелинейной конвекцией)

$$\frac{\partial u}{\partial t} + \frac{\partial h(u)}{\partial x} = \frac{\partial^2 u}{\partial x^2} + u(1 - u)$$

для учёта различных видов конвективного движения.

В последнее время много усилий было направлено на разработку компактных схем высокого порядка точности, которые используют только узлы сетки, непосредственно примыкающие к центральному узлу. Под *компактными* разностными схемами понимаются разностные схемы повышенных порядков аппроксимации и/или точности, записываемые на стандартных для данного уравнения шаблонах [5, с. 11; 6]. В работе [7] построены компактные разностные схемы для уравнений конвекции–диффузии с дивергентными и недивергентными конвективными слагаемыми, которые базируются на использовании экспоненциальных схем. Следует отметить также статью [8], в которой рассматриваются компактные и монотонные разностные схемы  $(4+1)$ , т.е. четвёртого порядка по пространственной переменной и первого по временной для уравнения Фишера (1).

В работе А.А. Самарского [9] предложены безусловно монотонные схемы (нет соотношений между сеточными шагами и коэффициентами уравнения), равномерно сходящиеся со скоростью  $O(h^2)$  для эллиптического уравнения второго порядка типа конвекции–диффузии. Идея построения монотонных разностных схем высокого порядка аппроксимации на стандартных шаблонах была развита в статьях В.К. Полевикова [10, 11], в которых построены компактные разностные схемы четвёртого порядка точности при соотношениях на шаги сетки, удовлетворяющих неравенствам  $1/\sqrt{5} \leq h_1/h_2 \leq \sqrt{5}$ .

Монотонные разностные схемы играют важную роль при математическом моделировании прикладных задач, так как они позволяют получить численные решения без нефизических

осцилляций (см. [12–14]). Существует достаточно много определений монотонности [15, с. 169; 16; 17, с. 229]. Наиболее часто используют определение, связанное с выполнением принципа максимума на сеточном уровне [15, с. 169; 18, с. 27; 19]. В линейном случае оно достаточно близко к определению монотонности схемы по Фридрихсу [20], т.е. в канонической форме записи разностного уравнения все коэффициенты неотрицательны, а их сумма равна единице (требование аппроксимации). В нелинейном случае такие схемы позволяют получить не только априорные оценки, но и двусторонние оценки разностного решения для начально-краевых задач для параболических уравнений [21], которые согласованы с аналогичными оценками в дифференциальном случае, полученными О.А. Ладыженской [22, с. 22].

Настоящая работа посвящена построению и исследованию компактных и монотонных разностных схем на стандартных шаблонах для обобщённого уравнения Фишера с нелинейной конвекцией. В п. 1 приводятся вспомогательные результаты, необходимые для дальнейшего исследования разностных схем, в частности, лемма о двусторонних оценках разностного решения и компактные безусловно монотонные схемы четвёртого порядка точности для одномерного стационарного уравнения конвекции–диффузии.

В п. 2 построены компактные и монотонные разностные схемы с  $(4 + 1)$  порядком аппроксимации для обобщённого уравнения Фишера с квазилинейной конвекцией. Доказана монотонность таких схем при нестандартных соотношениях на временной и пространственной шаг  $\tau \geq c_1 h^2$ . Получены априорные оценки разностного решения в нелинейном случае на основе установленных двусторонних оценок сеточного решения. Также строятся компактные схемы  $(4 + 2)$ . Для реализации нелинейной разностной схемы построен итерационный метод, сохраняющий второй порядок аппроксимации по временной переменной при специальном выборе начального приближения. Приведены результаты тестовых расчётов, подтверждающие повышенный порядок точности разностных схем на стандартных шаблонах.

В п. 3 полученные результаты обобщаются на двумерное обобщённое уравнение Фишера с нелинейной конвекцией.

## 1. Вспомогательные результаты.

**1.1. Двусторонние оценки.** Пусть задано начальное число точек – сетка  $\bar{\omega}_h$ . Окрестностью точки  $x$  называется множество  $M'(x) = M(x) \setminus x$ ,  $M(x)$  – шаблон. Пусть заданы функции  $A(x)$ ,  $B(x, \xi)$ ,  $F(x)$ , определённые при  $x \in \bar{\omega}_h$  и принимающие вещественные значения. Далее каждой точке  $x \in \bar{\omega}_h$  сопоставим одно и только одно уравнение вида [15, с. 226]

$$A(x)y(x) = \sum_{\xi \in M'(x)} B(x, \xi)y(\xi) + F(x), \quad x \in \bar{\omega}_h, \quad (2)$$

называемое канонической формой записи разностной схемы. В соответствии с [15, с. 226] точка  $x$  называется граничным узлом сетки, если в ней задано условие Дирихле:

$$y(x) = \mu(x), \quad x \in \gamma_h, \quad (3)$$

где  $\gamma_h$  – множество граничных узлов. Отметим, что при аппроксимации граничных условий второго или третьего рода сетка может не содержать граничных узлов, т.е. все точки сетки будут являться только внутренними узлами. Будем предполагать, что выполняются обычные условия положительности коэффициентов:

$$A(x) > 0, \quad B(x, \xi) > 0 \quad \text{для всех } \xi \in M'(x), \quad (4)$$

$$D(x) = A(x) - \sum_{\xi \in M'(x)} B(x, \xi) > 0 \quad \text{для всех } \xi \in M'(x). \quad (5)$$

В соответствии с монографией А.А. Самарского [15, с. 169] разностные схемы (2), (3), удовлетворяющие условиям (4), (5), будем называть *монотонными*.

В дальнейшем будем пользоваться следующими сеточными нормами:

$$\|\cdot\|_{\bar{C}} = \max_{x \in \bar{\omega}_h} |\cdot|, \quad \|\cdot\|_C = \max_{x \in \omega_h} |\cdot|.$$

Сформулируем основные результаты, позволяющие установить двусторонние оценки сеточного решения через входные данные задачи при незнакоопределённых входных данных задачи  $F(x)$ .

**Лемма** (см. [16]). Пусть выполнены условия положительности коэффициентов (4), (5). Тогда максимальное и минимальное значения решения разностной схемы (2) принадлежит интервалу изменения входных данных

$$\min_{x \in \bar{\omega}_h} \frac{F(x)}{D(x)} \leq y(x) \leq \max_{x \in \bar{\omega}_h} \frac{F(x)}{D(x)}. \tag{6}$$

**Следствие 1** (см. [15, с. 230]). Пусть выполнены условия леммы 1. Тогда для решения разностной схемы (2) справедлива оценка в сеточном аналоге нормы  $C$

$$\|y\|_C = \max_{x \in \bar{\omega}_h} |y(x)| \leq \|F/D\|_C.$$

**1.2. Компактные и безусловно монотонные схемы четвёртого порядка аппроксимации для одномерного уравнения конвекции–диффузии.** Отметим, что рассматриваемые ниже конечно-разностные методы являются следствием разностных схем, построенных В.К. Полевиковым для двумерного стационарного уравнения конвекции–диффузии в статье [11].

Рассмотрим краевую задачу

$$\frac{d}{dx} \left( k(x) \frac{du}{dx} \right) - r(x) \frac{du}{dx} - q(x)u = -f(x), \quad 0 < x < l, \tag{7}$$

$$u(0) = \mu_1, \quad u(l) = \mu_2, \tag{8}$$

где

$$0 < k_1 \leq k(x) \leq k_2, \quad |r(x)| \leq k_3, \quad q(x) \geq 0.$$

Для упрощения дальнейших исследований ниже мы будем использовать следующие обозначения:

$$L^{(k,r)}u = Lu - r(x) \frac{du}{dx}, \quad Lu = \frac{d}{dx} \left( k(x) \frac{du}{dx} \right).$$

Здесь и далее предполагаем, что решение рассматриваемых дифференциальных задач существует, единственно и обладает всеми непрерывными производными, необходимыми по ходу изложения.

При построении монотонных схем четвёртого порядка точности на трёхточечном шаблоне  $(x_{i-1}, x_i, x_{i+1})$  с использованием равномерной сетки  $\bar{\omega}_h = \omega_h \cup \gamma_h = \{x_i = ih, i = \overline{0, N}, h = 1/N\}$ ,  $\omega_h = \{x_i = ih, i = \overline{1, N-1}\}$ ,  $\gamma_h = \{x_0 = 0, x_N = l\}$  будем ориентироваться на соотношение [15, с. 170]

$$L^{(k,r)}u \sim \Lambda^{(a,b)}u = \kappa(au_{\bar{x}})_x - b^+au_{\bar{x}} - b^-a^{(+1)}u_x$$

с коэффициентами

$$a = 6 \left( \frac{1}{k^{(-1)}} + \frac{4}{k^{(-0.5)}} + \frac{1}{k} \right) > 0, \quad b^\pm = \frac{r^\pm}{k} + O(h^4),$$

$$r^+ = 0.5(r + |r|) \geq 0, \quad r^- = 0.5(r - |r|) \leq 0, \quad \kappa = \frac{1}{1 + R + R^2 + R^3} > 0, \quad R = \frac{h|r|}{2k} > 0. \tag{9}$$

Для аппроксимации коэффициента  $k(x)$  используется шаблонный функционал, предложенный А.А. Самарским в работе [23]. Полученная ниже схема не обобщается на квазилинейные уравнения параболического типа ввиду неопределённости коэффициента в точке  $k^{(0.5)}$ .

Здесь используются стандартные обозначения теории разностных схем (см. [15]):

$$u_{\bar{x}} = \frac{u - u^{(-1)}}{h}, \quad u_x = \frac{u^{(+1)} - u}{h}, \quad u = u(x), \quad u^{(+1)} = u(x + h), \quad u^{(-1)} = u(x - h),$$

$$k = k(x), \quad k^{(0.5)} = k(x + 0.5h), \quad k^{(-0.5)} = k(x - 0.5h).$$

На равномерной сетке  $\bar{\omega}_h$  дифференциальную задачу (7), (8) заменим неконсервативной разностной схемой

$$\tilde{a}\Lambda^{(a,\tilde{b})}y - \tilde{q}y + \tilde{f} = 0, \tag{10}$$

$$y_0 = \mu_1, \quad y_N = \mu_2, \tag{11}$$

где

$$\Lambda^{(a,\tilde{b})}y = \tilde{\kappa}(ay_{\bar{x}})_x - \tilde{b}^+ay_{\bar{x}} - \tilde{b}^-a^{(+1)}y_x, \quad \tilde{\kappa} = \frac{1}{1 + \tilde{R} + \tilde{R}^2 + \tilde{R}^3}, \quad \tilde{R} = \frac{h|\tilde{r}|}{2k},$$

$$\tilde{a} = 1 + \frac{h^2}{12}\tilde{p}_1, \quad \tilde{p}_1 = p_1 + \sigma_1, \quad \tilde{b} = \frac{\tilde{r}}{k} + O(h^4), \quad p_1 = \left(\frac{r}{k}\right)^2 + \frac{r}{k^2}\frac{dk}{dx} - 2\frac{d}{dx}\left(\frac{r}{k}\right),$$

$$\tilde{r} = \frac{1}{\tilde{k}}\left(\frac{r}{k} + \frac{h^2}{12}\tilde{p}_2\right), \quad \tilde{p}_2 = \left(p_2 + 2\frac{d}{dx}\left(\frac{q}{k}\right) + \sigma_1\frac{r}{k}\right), \quad p_2 = L^{(k,r)}\left(\frac{r}{k^2}\right),$$

$$\tilde{q} = q + \frac{h^2}{12}\left[L^{(k,r)}\left(\frac{q}{k}\right) + \left(\frac{q}{k} + \sigma_1\right)q\right], \quad \tilde{f} = f + \frac{h^2}{12}\left[L^{(k,r)}\left(\frac{f}{k}\right) + \left(\frac{q}{k} + \sigma_1\right)f\right]. \tag{12}$$

В соответствии с работами [23; 24, с. 70] покажем, что разностная схема (10), (11) имеет порядок аппроксимации  $O(h^4)$ , т.е. для её невязки

$$\psi = \tilde{a}\Lambda^{(a,\tilde{b})}u - L^{(k,r)}u - \tilde{q}u + \tilde{f} + qu - f$$

имеет место априорная оценка

$$\|\psi\|_C \leq Mh^4, \quad M = \text{const} > 0.$$

Разностный оператор  $\Lambda^{(a,b)}$  аппроксимирует соответствующий дифференциальный оператор  $L^{(k,r)}u$  со вторым порядком. С учётом предположений (9) в отношении коэффициента  $a$  справедливы следующие асимптотические разложения:

$$au_{\bar{x}} = k\frac{du}{dx} - \frac{h}{2}Lu + \frac{h^2\sqrt{k}}{6}\frac{d}{dx}\left(\frac{1}{\sqrt{k}}Lu\right) - \frac{h^3}{24}L\left(\frac{1}{k}Lu\right) + O(h^4),$$

$$a^{(+1)}u_x = k\frac{du}{dx} + \frac{h}{2}Lu + \frac{h^2\sqrt{k}}{6}\frac{d}{dx}\left(\frac{1}{\sqrt{k}}Lu\right) + \frac{h^3}{24}L\left(\frac{1}{k}Lu\right) + O(h^4). \tag{13}$$

С помощью разложений (13) приходим к соотношению

$$\Lambda^{(a,b)}u = L^{(k,r)}u + \frac{h^2}{12}L_1u + O(h^4), \tag{14}$$

где

$$L_1u = \left(L\left(\frac{1}{k}Lu\right) - 2\frac{r}{k}\sqrt{k}\frac{d}{dx}\left(\frac{1}{\sqrt{k}}Lu\right)\right).$$

Далее, чтобы избавиться от производных высоких порядков, не поддающихся разностной аппроксимации на минимальном шаблоне, упростим выражение  $L_1u$ . Так как

$$Lu = L^{(k,r)}u + \frac{r}{k}k\frac{du}{dx},$$

то получим

$$\begin{aligned} L_1 u &= L^{(k,r)} \left( \frac{1}{k} L^{(k,r)} u \right) + L^{(k,r)} \left( \frac{r}{k^2} k \frac{du}{dx} \right) + \\ &+ \frac{r}{k} k \frac{d}{dx} \left( \frac{1}{k} L^{(k,r)} u + \frac{r}{k^2} k \frac{du}{dx} \right) - 2 \frac{r}{k} \sqrt{k} \frac{d}{dx} \left( \frac{1}{\sqrt{k}} L u \right) = \\ &= L^{(k,r)} \left( \frac{1}{k} L^{(k,r)} u \right) + p_2 k \frac{du}{dx} - p_1 L u, \end{aligned} \tag{15}$$

где  $p_1, p_2$  определяются из формул (12).

Подставив в (15)  $L^{(k,r)} u = qu - f$  из уравнения (7) и добавив параметр регуляризации  $-\sigma_1$ , равный тождественно нулю на точном решении  $u(x)$ , приходим к соотношению

$$\begin{aligned} L_1 u &= L^{(k,r)} \left( \frac{1}{k} (qu - f) \right) + p_2 k \frac{du}{dx} - p_1 L u - \sigma_1 (L^{(k,r)} u - qu + f) = \\ &= L^{(k,r)} \left( \frac{q}{k} u \right) - p_1 L u + p_2 k \frac{du}{dx} - \sigma_1 L^{(k,r)} u - (f - qu) \sigma_1 - L^{(k,r)} \left( \frac{f}{k} \right). \end{aligned} \tag{16}$$

Так как

$$\begin{aligned} L^{(k,r)} \left( \frac{q}{k} u \right) &= L \left( \frac{q}{k} u \right) - \frac{r}{k} k \frac{d}{dx} \left( \frac{q}{k} u \right) = L^{(k,r)} \left( \frac{q}{k} \right) u + \frac{q}{k} L^{(k,r)} u + 2k \frac{du}{dx} \frac{d}{dx} \left( \frac{q}{k} \right) = \\ &= L^{(k,r)} \left( \frac{q}{k} \right) u + \frac{q}{k} (q - f) + 2k \frac{du}{dx} \frac{d}{dx} \left( \frac{q}{k} \right), \end{aligned}$$

то

$$\begin{aligned} L_1 u &= L^{(k,r)} \left( \frac{q}{k} \right) u + \frac{q}{k} (q - f) + 2k \frac{du}{dx} \frac{d}{dx} \left( \frac{q}{k} \right) - \\ &- p_1 L u + p_2 k \frac{du}{dx} - \sigma_1 L^{(k,r)} u - (f - qu) \sigma_1 - L^{(k,r)} \left( \frac{f}{k} \right) = \\ &= -p_1 L u + \left( p_2 + 2 \frac{d}{dx} \left( \frac{q}{k} \right) \right) k \frac{du}{dx} - \sigma_1 L^{(k,r)} u + \left[ L^{(k,r)} \left( \frac{q}{k} \right) + \left( \frac{q}{k} + \sigma_1 \right) q \right] u - \\ &- \left[ L^{(k,r)} \left( \frac{f}{k} \right) + \left( \frac{q}{k} + \sigma_1 \right) f \right]. \end{aligned} \tag{17}$$

С учётом выражений (9), (12), (14) и (17) получим

$$\begin{aligned} \psi &= \tilde{a} (\Lambda^{(k,\tilde{r})} u - L^{(k,\tilde{r})} u) + \frac{h^2}{12} \tilde{p}_1 L u - \frac{h^2}{12} \tilde{p}_2 k \frac{du}{dx} - (\tilde{q} - q) u + (\tilde{f} - f) + O(h^4) = \\ &= \tilde{a} (\Lambda^{(k,\tilde{r})} u - L^{(k,\tilde{r})} u) + \frac{h^2}{12} L_1 u + O(h^4) = \frac{h^2}{12} \tilde{p}_1 (\Lambda^{(k,\tilde{r})} u - L^{(k,\tilde{r})} u) + O(h^4) = O(h^4). \end{aligned}$$

Параметр регуляризации  $\sigma_1 = O(1)$  в выражении (16) будем выбирать из условий  $\tilde{a} \geq 1$  и  $\tilde{q} \geq 0$ , т.е.

$$\sigma_1 = \sigma_1(x) = - \min \left\{ p_1, \frac{1}{q} L^{(k,r)} \left( \frac{q}{k} \right) \right\}.$$

Для исследования монотонности разностной схемы (10), (11) запишем её в каноническом виде:

$$\begin{aligned} C_i y_i &= A_i y_{i-1} + B_i y_{i+1} + F_i, \quad i = \overline{1, N-1}, \\ y_0 &= \mu_1, \quad y_N = \mu_1, \end{aligned}$$

в котором

$$A_i = a_i \left( \frac{\tilde{a}}{h^2} (\tilde{\kappa} + h\tilde{b}^+) \right) > 0, \quad B_i = a_{i+1} \left( \frac{\tilde{a}}{h^2} (\tilde{\kappa} - h\tilde{b}^-) \right) > 0,$$

$$C_i = (a_{i+1} + a_i) \frac{\tilde{a}\tilde{\kappa}}{h^2} + \frac{\tilde{a}}{h} (\tilde{b}^+ a_i - \tilde{b}^- a_{i+1}) + \tilde{q} > 0, \quad F_i = \tilde{f}, \quad D_i = C_i - A_i - B_i > 0. \quad (18)$$

Так как выполнены все условия положительности коэффициентов (18), то разностная схема (10), (11) является безусловно монотонной.

Итак, согласно сформулированной выше лемме получаем двустороннюю оценку решения разностной схемы (10) при произвольных знакопостоянных входных данных задачи

$$\min \left\{ \mu_1, \mu_2, \min_{1 \leq i \leq N-1} \frac{F_i}{D_i} \right\} \leq y_i \leq \max \left\{ \mu_1, \mu_2, \max_{1 \leq i \leq N-1} \frac{F_i}{D_i} \right\}.$$

Также из следствия 1 получаем, что разностная схема (10) устойчива относительно правой части и граничных условий, и для её решения справедлива априорная оценка

$$\|y\|_{\bar{C}} \leq \max \left\{ |\mu_1|, |\mu_2|, \left\| \frac{F_i}{D_i} \right\|_C \right\}.$$

**2. Компактные разностные схемы (4+1) для обобщённого уравнения Фишера с нелинейной конвекцией.** Целью данной работы является исследование компактных и монотонных разностных схем (в смысле отсутствия ограничений типа Куранта на временной шаг), для реализации которых не требуется применение итерационных методов [8]. При построении таких схем на стандартном трёхточечном шаблоне на верхнем слое будем пользоваться разностной схемой (10).

В области  $\bar{Q}_T = \bar{\Omega} \times [0, T]$ ,  $\bar{\Omega} = \{0 \leq x \leq l\}$ ,  $\bar{\Omega} = \Omega \cup \Gamma$ ,  $\Gamma$  – граница области, рассмотрим начально краевую задачу для обобщённого уравнения Фишера с нелинейной конвекцией

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} - h'(u) \frac{\partial u}{\partial x} + u(1 - u), \quad 0 < x < l, \quad 0 < t \leq T, \quad (19)$$

с начальными и граничными условиями

$$u(x, 0) = u_0(x), \quad u_0(x) \geq 0, \quad x \in \bar{\Omega}, \quad (20)$$

$$u(0, t) = \mu_1(t) \geq 0, \quad u(l, t) = \mu_2(t) \geq 0, \quad t \in (0, T], \quad (21)$$

т.е. априори предполагаем неотрицательность входных данных.

Введём стандартные обозначения теории разностных схем [15, с. 260]:

$$v = v_i^n = v(x_i, t_n), \quad \hat{v} = v_i^{n+1} = v(x_i, t_{n+1}), \quad v_t = \frac{\hat{v} - v}{\tau}.$$

На равномерной сетке  $\bar{\omega} = \bar{\omega}_h \times \bar{\omega}_\tau = \{(x_i, t_n) \in \bar{Q}_T\}$ ,  $\bar{\omega}_h = \{x_i = ih, \quad i = \overline{0, N}, \quad hN = l\}$ ,  $\bar{\omega}_\tau = \{t_n = n\tau, \quad n = \overline{0, N_0}, \quad \tau N_0 = T\}$ ,  $\bar{\omega}_\tau = \omega_\tau \cup \{t_{N_0} = T\}$  дифференциальную задачу (19)–(21) заменим разностной схемой

$$(1 + \lambda)y_t = \sigma(\kappa\hat{y}_{\bar{x}x} - b^+\hat{y}_{\bar{x}} - b^-\hat{y}_x) + (1 - \sigma)(\kappa y_{\bar{x}x} - b^+y_{\bar{x}} - b^-y_x) +$$

$$+ \frac{h^2}{12}(\kappa(y - y\hat{y})_{\bar{x}x} - b^+(y - y\hat{y})_{\bar{x}} - b^-(y - y\hat{y})_x) +$$

$$+ \tilde{a}(\tilde{\kappa}\hat{y}_{\bar{x}x} - \tilde{b}^+\hat{y}_{\bar{x}} - \tilde{b}^-\hat{y}_x) + (y - y\hat{y})(1 + \lambda), \quad (22)$$

$$y(x, 0) = u_0(x), \quad x \in \bar{\omega}_h \quad \hat{y}_0 = \mu_1, \quad \hat{y}_N = \mu_2, \quad (23)$$



где

$$\begin{aligned} \lambda &= \frac{h^2}{12}\sigma_1 \geq 0, \quad \sigma = 1 - \frac{h^2}{12\tau}, \\ \kappa &= \frac{1}{1 + R + R^2 + R^3}, \quad R = \frac{h|b|}{2}, \quad \tilde{\kappa} = \frac{1}{1 + \tilde{R} + \tilde{R}^2 + \tilde{R}^3}, \quad \tilde{R} = \frac{h|\tilde{b}|}{2}, \\ b &= b(y) = h'(y_i^n) + O(h^4), \quad b^+ = 0.5(b + |b|) \geq 0, \quad b^- = 0.5(b - |b|) \leq 0, \\ \tilde{b} &= \frac{r_2 + \sigma_1 b}{r_1 + \sigma_1}, \quad \tilde{b}^+ = 0.5(\tilde{b} + |\tilde{b}|) \geq 0, \quad \tilde{b}^- = 0.5(\tilde{b} - |\tilde{b}|) \leq 0, \\ r_2 &= (b(y))_{\bar{x}x} - b(y)(b(y))_x, \quad \tilde{a} = \frac{h^2}{12}(r_1 + \sigma_1), \quad r_1 = (b(y))^2 - 2(b(y))_x. \end{aligned} \tag{24}$$

Параметр регуляризации  $\sigma_1$  в выражении (24) будем выбирать из условия  $\tilde{a} > 0$ , т.е.

$$\sigma_1 = \sigma_1(y_i^n) = \begin{cases} -\min_{\substack{0 \leq i \leq N \\ 1 \leq n \leq N_0}} \{r_1(y(x_i, t_{n-1}))\}, & \text{если } r_1 < 0, \\ 0, & \text{если } r_1 > 0. \end{cases}$$

Покажем, что разностная схема (22), (23) имеет четвёртый порядок аппроксимации по пространству и первый по времени, т.е. для её невязки

$$\begin{aligned} \psi &= -(1 + \lambda)u_t + \sigma(\kappa\hat{u}_{\bar{x}x} - b^+\hat{u}_{\bar{x}} - b^-\hat{u}_x) + (1 - \sigma)(\kappa u_{\bar{x}x} - b^+u_{\bar{x}} - b^-u_x) + \\ &+ \frac{h^2}{12}(\kappa(u - u\hat{u})_{\bar{x}x} - b^+(u - u\hat{u})_{\bar{x}} - b^-(u - u\hat{u})_x) + \\ &+ \tilde{a}(\tilde{\kappa}\hat{u}_{\bar{x}x} - \tilde{b}^+\hat{u}_{\bar{x}} - \tilde{b}^-\hat{u}_x) + (u - u\hat{u})(1 + \lambda) \end{aligned}$$

имеет место оценка

$$\|\psi\|_C \leq M_1(h^4 + \tau), \quad M_1 = \text{const} > 0. \tag{25}$$

С учётом того, что

$$\kappa(u)u_{\bar{x}x} - b^+(u)u_{\bar{x}} - b^-(u)u_x \sim \frac{\partial^2 u}{\partial x^2} - h'(u)\frac{\partial u}{\partial x} + \frac{h^2}{12}Eu + O(h^4),$$

где

$$\begin{aligned} Eu &= \left( \frac{\partial^3 u}{\partial x^2 \partial t} - h'(u)\frac{\partial^2 u}{\partial x \partial t} \right) - \left( \frac{\partial^2 u(1 - \hat{u})}{\partial x^2} - h'(u)\frac{\partial}{\partial x}(u(1 - \hat{u})) \right) - \\ &- \left( \left( (h'(u))^2 - 2\frac{\partial}{\partial x}h'(u) \right) \frac{\partial^2 \hat{u}}{\partial x^2} - \left( \frac{\partial^2}{\partial x^2}h'(u) + h'(u)\frac{\partial}{\partial x}h'(u) \right) \frac{\partial \hat{u}}{\partial x} \right) - \\ &- \sigma_1 \left( \frac{\partial^2 \hat{u}}{\partial x^2} - h'(u)\frac{\partial \hat{u}}{\partial x} + u(1 - \hat{u}) - \frac{\partial u}{\partial t} \right) \end{aligned}$$

и

$$\begin{aligned} \frac{\partial^3 u}{\partial x^2 \partial t} - h'(u)\frac{\partial^2 u}{\partial x \partial t} &\sim \kappa u_{\bar{x}xt} - b^+u_{\bar{x}t} - b^-u_{xt}, \\ \frac{\partial^2 u(1 - \hat{u})}{\partial x^2} - h'(u)\frac{\partial}{\partial x}(u(1 - \hat{u})) &\sim \kappa(u(1 - \hat{u}))_{\bar{x}x} - b^+(u(1 - \hat{u}))_{\bar{x}} - b^-(u(1 - \hat{u}))_x, \\ \left( (h'(u))^2 - 2\frac{\partial}{\partial x}h'(u) \right) \frac{\partial^2 \hat{u}}{\partial x^2} - \left( \frac{\partial^2}{\partial x^2}h'(u) + h'(u)\frac{\partial}{\partial x}h'(u) \right) \frac{\partial \hat{u}}{\partial x} &+ \\ + \sigma_1 \left( \frac{\partial^2 \hat{u}}{\partial x^2} - h'(u)\frac{\partial \hat{u}}{\partial x} \right) &\sim \tilde{a}(u)(\tilde{\kappa}(u)\hat{u}_{\bar{x}x} - \tilde{b}^+(u)\hat{u}_{\bar{x}} - \tilde{b}^-(u)\hat{u}_x), \end{aligned}$$

приходим к оценке (25).

**2.1. Двусторонняя оценка.** Для получения априорных оценок сеточного решения запишем разностную схему (22) в каноническом виде:

$$C_i y_i^{n+1} = A_i y_{i-1}^{n+1} + B_i y_{i+1}^{n+1} + F_i^n, \quad i = \overline{1, N-1},$$

в котором

$$\begin{aligned} A_i &= \frac{\tau}{h^2} \left[ \left( 1 - \frac{h^2}{12\tau} - \frac{h^2}{12} y_{i-1}^n \right) (\kappa + hb^+) + \tilde{a}_i (\tilde{\kappa} + h\tilde{b}^+) \right], \\ B_i &= \frac{\tau}{h^2} \left[ \left( 1 - \frac{h^2}{12\tau} - \frac{h^2}{12} y_{i+1}^n \right) (\kappa - hb^-) + \tilde{a}_i (\tilde{\kappa} - h\tilde{b}^-) \right], \\ C_i &= \frac{\tau}{h^2} \left[ \left( 1 - \frac{h^2}{12\tau} - \frac{h^2}{12} y_i^n \right) (\kappa + hb^+) + \tilde{a}_i (\tilde{\kappa} + h\tilde{b}^+) \right] + \\ &+ \frac{\tau}{h^2} \left[ \left( 1 - \frac{h^2}{12\tau} - \frac{h^2}{12} y_i^n \right) (\kappa - hb^-) + \tilde{a}_i (\tilde{\kappa} - h\tilde{b}^-) \right] + (1 + \lambda)(1 + \tau y_i^n), \\ F_i &= \frac{1}{12} (\kappa + hb^+) (1 + \tau) y_{i-1}^n + \frac{1}{12} (\kappa - hb^-) (1 + \tau) y_{i+1}^n + \\ &+ (1 + \tau) \left[ (1 + \lambda) - \frac{1}{12} (\kappa + hb^+) - \frac{1}{12} (\kappa - hb^-) \right] y_i^n. \end{aligned}$$

Пусть

$$m_1 = \min_{(x,t) \in \bar{Q}_T} \{ \min\{\mu_1(t), \mu_2(t)\}, u_0(x) \}, \quad m_2 = \max_{(x,t) \in \bar{Q}_T} \{ \max\{\mu_1(t), \mu_2(t)\}, u_0(x) \}.$$

Имеет место следующее утверждение.

**Теорема 1.** Пусть выполнены следующие условия:

$$h < h_0, \quad h_0^2 = \frac{12}{m_2} e^T, \quad \tau > \tau_0, \quad \tau_0 = \frac{h^2}{12 - h^2 e^T m_2}. \tag{26}$$

Тогда разностное решение  $y(x, t)$ ,  $(x, t) \in \bar{\omega}$ , неотрицательно и для всех  $i = \overline{0, N}$ ,  $k = \overline{0, N_0}$  имеет место двусторонняя оценка

$$0 \leq y_i^k \leq e^{tk} m_2. \tag{27}$$

**Доказательство.** Очевидно, что  $0 \leq m_1 \leq y_i^0 = u_0(x_i) \leq m_2$ . По индукции предполагаем, что оценка (27) имеет место для всех  $k = \overline{1, n}$ . Докажем теперь справедливость этих неравенств и для  $k = n + 1$ . Действительно, в силу условий теоремы (26) и предположения индукции

$$A_i^n \geq r - \frac{\tau}{12} \max_{1 \leq i \leq N-1} y_i^n \geq r - \frac{\tau m_2}{12} e^{\lambda t_n} > 0, \quad r = \frac{\tau}{h^2} - \frac{1}{12},$$

$$B_i^n > 0, \quad C_i^n > 0, \quad D_i^n > 0, \quad i = \overline{1, N-1}.$$

В соответствии с двусторонней оценкой (6) имеет место неравенство

$$m_1^{n+1} \leq y_i^{n+1} \leq m_2^{n+1},$$

где

$$m_1^{n+1} = \min_{0 < i < N} \frac{F_i^n}{D_i^n} \geq 0, \quad m_2^{n+1} = \max_{0 < i < N} \frac{F_i^n}{D_i^n} \geq 0.$$

Заметим, что

$$\max_i \frac{F_i^n}{D_i^n} \leq \frac{(1 + \tau) \max y_i^n}{1 + \tau \min y_i^n} \leq e^\tau \max y_i^n.$$

Следовательно,

$$y_i^{n+1} \leq \max\{m_2^{n+1}, e^\tau \max y_i^n\} \leq m_2 e^{t_{n+1}}.$$

Теорема доказана.

**Следствие 2.** Для решения разностной схемы (22), (23) имеет место априорная оценка

$$\|y^n\|_{\overline{C}} \leq m_2 e^{t^n}, \quad n = \overline{0, N_0}.$$

**2.2. Вычислительный эксперимент.** Чтобы продемонстрировать точность настоящего метода, рассмотрим уравнение

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \frac{\partial^2 u}{\partial x^2} + u(1 - u), \tag{28}$$

сочетающее механизмы реакции, конвекции и диффузии, с начальным и граничными условиями

$$u(x, 0) = \frac{1}{2} - \frac{1}{2} \operatorname{th} \left( \frac{x}{4} \right),$$

$$u(0, t) = \frac{1}{2} + \frac{1}{2} \operatorname{th} \left( \frac{5t}{8} \right), \quad u(l, t) = \frac{1}{2} + \frac{1}{2} \operatorname{th} \left( \frac{5t}{8} - \frac{l}{4} \right). \tag{29}$$

Это уравнение известно как уравнение Бюргерса–Фишера из-за свойств конвективного явления от уравнения Бюргерса и характеристики диффузионного переноса и реакции типа уравнения Фишера. Соответствующее точное решение этого уравнения имеет следующий вид [25]:

$$u(x, t) = \frac{1}{2} + \frac{1}{2} \operatorname{th} \left( \frac{5t}{8} - \frac{x}{4} \right).$$

Для определения порядка скорости сходимости по временной и пространственной переменным (4+1) в нормах  $L_\infty = C$  и  $L_2$  воспользуемся правилом Рунге

$$p_{L_\infty(L_2)}^h = \log_2 \frac{\|z(2h, \tau)\|_{L_\infty(L_2)}}{\|z(h, \tau)\|_{L_\infty(L_2)}}, \quad p_{L_\infty(L_2)}^\tau = \log_2 \frac{\|z(h, 2\tau)\|_{L_\infty(L_2)}}{\|z(h, \tau)\|_{L_\infty(L_2)}}. \tag{30}$$

Для проверки скорости сходимости вдоль пространственного и временного направлений выберем шаги  $h$  и  $\tau$  так, чтобы выполнялись условия  $h^4 \geq \tau$  для определения порядка скорости сходимости разностного решения по пространственной переменной  $p_\infty^h$  и  $\tau \geq h^4$  для определения порядка скорости сходимости сеточного решения по временной переменной  $p_\infty^\tau$  (см. [8]).

В табл. 1, 2 приведены порядки скорости сходимости по пространственному и временному направлениям в нормах  $L_\infty$  и  $L_2$ . Величины, представленные в таблицах, соответствуют моменту времени  $T = 1$ .

**Таблица 1.** Скорость сходимости по пространственному направлению в нормах  $L_\infty$  и  $L_2$

$h = 0.1$	$\tau$	$\ z\ _{L_\infty}$	$p_\infty^h$	$\ z\ _{L_2}$	$p_{L_2}^h$
$h$	$\tau$	0.00202609	–	0.00148301	–
$h/2$	$\tau/2^4$	0.00012658	3.99714	9.21E-05	4.00917
$h/4$	$\tau/4^4$	7.70E-06	4.03757	5.60E-06	4.03892
$h/8$	$\tau/8^4$	2.75E-07	4.80735	1.97E-07	4.82915

**Таблица 2.** Скорость сходимости по временному направлению в нормах  $L_\infty$  и  $L_2$  ( $h = 0.001$ )

$\tau = 0.1$	$\ z\ _{L_\infty}$	$p_\infty^{\tilde{}}$	$\ z\ _{L_2}$	$p_{L_2}^{\tilde{}}$
$\tau$	0.00203821	–	0.00148323	–
$\tau/2$	0.00101756	0.99716	0.00074039	1.00233
$\tau/4$	0.00050808	0.99145	0.00036966	1.00214
$\tau/8$	0.00025383	1.00569	0.00018467	1.00125

Из результатов, представленных в табл. 1, 2, видно, что указанная разностная схема имеет четвёртый порядок аппроксимации по пространству и первый по времени.

**2.3. Компактные разностные схемы (4 + 2) для обобщённого уравнения Фишера с нелинейной конвекцией.** В данном пункте строятся компактные схемы (4 + 2). Для реализации нелинейной разностной схемы построен итерационный метод, сохраняющий второй порядок аппроксимации по временной переменной при специальном выборе начального приближения.

В области  $\bar{Q}_T$  рассмотрим начально краевую задачу для обобщённого уравнения Фишера с нелинейной конвекцией (19)–(21).

На равномерной сетке  $\bar{\omega}$  дифференциальную задачу (19)–(21) заменим разностной схемой

$$\begin{aligned}
 (1 + \lambda)y_t &= \sigma(\kappa\hat{y}_{\bar{x}x} - b^+\hat{y}_{\bar{x}} - b^-\hat{y}_x) + (1 - \sigma)(\kappa y_{\bar{x}x} - b^+y_{\bar{x}} - b^-y_x) + \\
 &+ \frac{h^2}{12}(\kappa(y^{(0.5)} - y\hat{y})_{\bar{x}x} - b^+(y^{(0.5)} - y\hat{y})_{\bar{x}} - b^-(y^{(0.5)} - y\hat{y})_x) + \\
 &+ \tilde{a}(\tilde{\kappa}(y^{(0.5)})_{\bar{x}x} - \tilde{b}^+(y^{(0.5)})_{\bar{x}} - \tilde{b}^-(y^{(0.5)})_x) + (y^{(0.5)} - y\hat{y})(1 + \lambda),
 \end{aligned} \tag{31}$$

где

$$\begin{aligned}
 \lambda &= \frac{h^2}{12}\sigma_2 \geq 0, \quad \sigma = \frac{1}{2} - \frac{h^2}{12\tau}, \quad y^{(0.5)} = \frac{\hat{y} + y}{2}, \\
 \kappa &= \frac{1}{1 + R + R^2 + R^3}, \quad R = \frac{h|b|}{2}, \quad \tilde{\kappa} = \frac{1}{1 + \tilde{R} + \tilde{R}^2 + \tilde{R}^3}, \quad \tilde{R} = \frac{h|\tilde{b}|}{2}, \\
 b &= b(y) = \frac{h'(\hat{y}) + h'(y)}{2}, \quad b^+ = 0.5(b + |b|) \geq 0, \quad b^- = 0.5(b - |b|) \leq 0, \\
 \tilde{b} &= \frac{r_2 + \sigma_2 b}{r_1 + \sigma_2}, \quad \tilde{b}^+ = 0.5(\tilde{b} + |\tilde{b}|) \geq 0, \quad \tilde{b}^- = 0.5(\tilde{b} - |\tilde{b}|) \leq 0, \\
 r_2 &= (b(y))_{\bar{x}x} - b(y)(b(y))_x, \quad \tilde{a} = \frac{h^2}{12}(r_1 + \sigma_2), \quad r_1 = b(y)^2 - 2(b(y))_x.
 \end{aligned}$$

Здесь параметр регуляризации  $\sigma_2 = \sigma_2(y)$  определяется из условия  $\tilde{a} > 0$ .

Разностная схема (31) является нелинейной. Для её реализации можно использовать следующий итерационный метод:

$$\begin{aligned}
 (1 + \lambda)\frac{{}^k\hat{y} - y}{}^\tau &= \sigma(\kappa {}^k\hat{y}_{\bar{x}x} - b^+ {}^k\hat{y}_{\bar{x}} - b^- {}^k\hat{y}_x) + (1 - \sigma)(\kappa y_{\bar{x}x} - b^+y_{\bar{x}} - b^-y_x) + \\
 &+ \frac{h^2}{12}\left(\kappa\left(\frac{{}^k\hat{y} + y}{}^2 - y {}^k\hat{y}\right)_{\bar{x}x} - b^+\left(\frac{{}^k\hat{y} + y}{}^2 - y {}^k\hat{y}\right)_{\bar{x}} - b^-\left(\frac{{}^k\hat{y} + y}{}^2 - y {}^k\hat{y}\right)_x\right) + \\
 &+ \tilde{a}\left(\tilde{\kappa}\left(\frac{{}^k\hat{y} + y}{}^2\right)_{\bar{x}x} - \tilde{b}^+\left(\frac{{}^k\hat{y} + y}{}^2\right)_{\bar{x}} - \tilde{b}^-\left(\frac{{}^k\hat{y} + y}{}^2\right)_x\right) + \left(\frac{{}^k\hat{y} + y}{}^2 - y {}^k\hat{y}\right)(1 + \lambda),
 \end{aligned}$$

с выбором начального приближения

$$y_i^{n+1} = \begin{cases} \mu_1(t_{n+1}), & i = 0, \\ 2y_i^n - y_i^{n-1}, & i = \overline{1, N-1}, \\ \mu_2(t_{n+1}), & i = N. \end{cases}$$

Так как начальное приближение использует два слоя, то нам нужно предварительно найти решение также и на первом слое. Для этого используем дифференциальное уравнение (19) при  $t = 0$ , т.е.

$$\frac{\partial u}{\partial t}(x, 0) = u_1(x), \quad u_1(x) = u_0'' - h'(u_0)u_0' + u_0(1 - u_0).$$

Аппроксимировав его с порядком  $O(\tau^2)$  относительно точки  $n + 1/2$ , получим

$$y^1 = y^0 + \tau u_1(x).$$

Аналогично рассмотренному выше случаю легко показать, что для погрешности аппроксимации разностной схемы (31) имеет место оценка

$$\|\psi\|_C \leq M_2(h^4 + \tau^2), \quad M_2 = \text{const} > 0.$$

**2.4. Вычислительный эксперимент.** Рассмотрим уравнение Бюргерса–Фишера (28) с начальными и граничными условиями (29).

При расчёте по итерационной схеме потребуем выполнения условия

$$\max_{0 \leq i \leq N} |y_i^{(k+1)} - y_i^{(k)}| \leq \varepsilon, \quad \varepsilon = 10^{-9},$$

где  $k$  – номер итерации,  $k = 0, 1, 2, \dots$

Для определения порядка скорости сходимости по временной и пространственной переменным (4 + 2) в нормах  $L_\infty = C$  и  $L_2$  воспользуемся правилом Рунге (30).

В табл. 3, 4 приведены порядки скорости сходимости по пространственному и временному направлениям в нормах  $L_\infty$  и  $L_2$ .

**Таблица 3.** Скорость сходимости по пространственному направлению в нормах  $L_\infty$  и  $L_2$

$h = 0.1$	$\tau = 0.01$	$\ z\ _{L_\infty}$	$p_\infty^h$	$\ z\ _{L_2}$	$p_{L_2}^h$	$k$
$h$	$\tau$	6.96E-05	–	7.38E-05	–	6
$h/2$	$\tau/2^2$	4.98E-06	3.80658	5.23E-06	3.81834	4
$h/4$	$\tau/4^2$	3.20E-07	3.95658	3.36E-07	3.95920	3
$h/8$	$\tau/8^2$	2.02E-08	3.98687	2.12E-08	3.98745	2
$h/16$	$\tau/16^2$	1.26E-09	3.99809	1.32E-09	3.99840	2

**Таблица 4.** Скорость сходимости по временному направлению в нормах  $L_\infty$  и  $L_2$  ( $h = 0.001$ )

$\tau = 0.1$	$\ z\ _{L_\infty}$	$p_\infty^\tau$	$\ z\ _{L_2}$	$p_{L_2}^\tau$	$k$
$\tau$	6.96E-05	–	7.38E-05	–	6
$\tau/2$	1.90E-05	1.86749	2.01E-05	1.87559	5
$\tau/4$	4.98E-06	1.93923	5.23E-06	1.94311	4
$\tau/8$	1.27E-06	1.97088	1.33E-06	1.97267	3
$\tau/16$	3.20E-07	1.98591	3.36E-07	1.98680	3

**3. Компактные разностные схемы (4 + 1) для двумерного обобщённого уравнения Фишера с нелинейной конвекцией.** Далее построим монотонные и компактные разностные схемы с 4 + 1 порядком аппроксимации на 9-точечном шаблоне для двумерного обобщённого уравнения Фишера с нелинейной конвекцией.

Пусть в области  $\bar{Q}_T = \{(x, t) : x \in \bar{\Omega}, 0 \leq t \leq T\}$ ,  $x = (x_1, x_2)$ ,  $\bar{\Omega} = \{0 \leq x_\alpha \leq l_\alpha, \alpha = 1, 2\}$ ,  $\bar{\Omega} = \Omega \cup \partial\Omega$ , где  $\partial\Omega$  – граница, требуется найти функцию  $u(x, t)$ , удовлетворяющую начально-краевой задаче для двумерного обобщённого уравнения Фишера с нелинейной конвекцией

$$\frac{\partial u}{\partial t} = \sum_{\alpha=1}^2 \left( \frac{\partial^2 u}{\partial x_\alpha^2} - h'_\alpha(u) \frac{\partial u}{\partial x_\alpha} \right) + u(1 - u), \quad x \in \Omega, \quad t \in (0, T], \tag{32}$$

с граничными условиями Дирихле

$$u(x, t) = \mu(x, t), \quad (x, t) \in \partial\Omega \times (0, T] \tag{33}$$

и начальным условием

$$u(x, 0) = u_0(x), \quad x \in \bar{\Omega}. \tag{34}$$

В области  $\bar{Q}_T$  рассмотрим равномерную сетку  $\bar{\omega} = \bar{\omega}_h \times \bar{\omega}_\tau$ , где

$$\bar{\omega}_\tau = \{t_n = n\tau, \quad n = \overline{0, N_0}, \quad \tau = T/N_0, \quad \bar{\omega}_h = \omega_h \cup \gamma_h\},$$

множество внутренних узлов пространственной сетки определяется соотношением

$$\omega_h = \{x = (x_1^{i_1}, x_2^{i_2}), \quad x_\alpha^{i_\alpha} = i_\alpha h_\alpha, \quad i_\alpha = \overline{1, N_\alpha - 1}, \quad h_\alpha N_\alpha = l_\alpha, \quad \alpha = 1, 2\},$$

через  $\gamma_h$  обозначено множество её граничных узлов. На сетке  $\bar{\omega}$  дифференциальную задачу (32)–(34) аппроксимируем разностной схемой

$$\begin{aligned} (1 + \lambda)y_t &= \sum_{\alpha=1}^2 (\Lambda_\alpha^{(b)} \sigma_\alpha \hat{y} + \Lambda_\alpha^{(b)} (1 - \sigma_\alpha)y) + \sum_{\alpha=1}^2 \frac{h_\alpha^2}{12} \Lambda_\alpha^{(b)} \Lambda_{3-\alpha}^{(b)} \hat{y} + \\ &+ \sum_{\alpha=1}^2 \left( \frac{\sigma_0^* h^2}{12} \Lambda_\alpha^{(b)} \hat{y} + \tilde{a}_\alpha \Lambda_\alpha^{(b)} \hat{y} + \frac{h_\alpha^2}{12} \Lambda_\alpha^{(b)} y(1 - \hat{y}) \right) + (1 + \lambda)y(1 - \hat{y}), \end{aligned} \tag{35}$$

$$\begin{aligned} y(x, t) &= \mu(x, t), \quad x \in \gamma_h, \quad t \in \omega_\tau, \quad (x, t) \in \omega_h \times \omega_\tau, \\ y(x, 0) &= u_0(x), \quad x \in \bar{\omega}_h, \end{aligned} \tag{36}$$

где

$$\lambda = \frac{h^2}{12} (\sigma_0^* + \sigma_1^*), \quad |h|^2 = h_1^2 + h_2^2, \quad \sigma_\alpha = 1 - \frac{h_\alpha^2}{12\tau},$$

$$\Lambda_\alpha^{(b)} y = \kappa_\alpha y_{\bar{x}_\alpha x_\alpha} - b_\alpha^+(y) y_{\bar{x}_\alpha} - b_\alpha^-(y) y_{x_\alpha}, \quad \Lambda_\alpha^{(b)} y \hat{y} = \kappa_\alpha (y \hat{y})_{\bar{x}_\alpha x_\alpha} - b_\alpha^+(y) (y \hat{y})_{\bar{x}_\alpha} - b_\alpha^-(y) (y \hat{y})_{x_\alpha},$$

$$\Lambda_\alpha^{(b)} \hat{y} = \kappa_\alpha \hat{y}_{\bar{x}_\alpha x_\alpha} - b_\alpha^+(y) \hat{y}_{\bar{x}_\alpha} - b_\alpha^-(y) \hat{y}_{x_\alpha}, \quad \kappa_\alpha = \frac{1}{1 + R_\alpha + R_\alpha^2 + R_\alpha^3}, \quad R_\alpha = \frac{h_\alpha |b_\alpha|}{2},$$

$$\Lambda_\alpha^{(\tilde{b})} \hat{y} = \tilde{\kappa}_\alpha \hat{y}_{\bar{x}_\alpha x_\alpha} - \tilde{b}_\alpha^+(y) \hat{y}_{\bar{x}_\alpha} - \tilde{b}_\alpha^-(y) \hat{y}_{x_\alpha}, \quad \tilde{\kappa}_\alpha = \frac{1}{1 + \tilde{R}_\alpha + \tilde{R}_\alpha^2 + \tilde{R}_\alpha^3}, \quad \tilde{R}_\alpha = \frac{h_\alpha |\tilde{b}_\alpha|}{2},$$

$$\begin{aligned} \Lambda_\alpha^{(b)} \Lambda_{3-\alpha}^{(b)} &= \kappa_\alpha (\kappa_{3-\alpha} \hat{y}_{\bar{x}_{3-\alpha} x_{3-\alpha}} - b_{3-\alpha}^+(y) \hat{y}_{\bar{x}_{3-\alpha}} - b_{3-\alpha}^-(y) \hat{y}_{x_{3-\alpha}}) - \\ &- b_\alpha^+(y) (\kappa_{3-\alpha} \hat{y}_{\bar{x}_{3-\alpha} x_{3-\alpha}} - b_{3-\alpha}^+(y) \hat{y}_{\bar{x}_{3-\alpha}} - b_{3-\alpha}^-(y) \hat{y}_{x_{3-\alpha}}) - \\ &- b_\alpha^-(y) (\kappa_{3-\alpha} \hat{y}_{\bar{x}_{3-\alpha} x_{3-\alpha}} - b_{3-\alpha}^+(y) \hat{y}_{\bar{x}_{3-\alpha}} - b_{3-\alpha}^-(y) \hat{y}_{x_{3-\alpha}}), \end{aligned}$$

$$\tilde{a}_\alpha = \frac{h^2}{12} \left( \frac{h_\alpha^2}{h^2} r_{1\alpha} + \sigma_1^* \right), \quad r_{1\alpha} = b_\alpha^2 - 2(b_\alpha)_{x_\alpha},$$

$$\tilde{b}_\alpha = \frac{h^2}{12\tilde{a}} (r_{2\alpha} + b\sigma_1^*), \quad r_{2\alpha} = \Lambda_\alpha^{(b)}(b_\alpha), \quad b = h'_\alpha(y_{i_1 i_2}^n).$$

Параметр регуляризации  $\sigma_1^*$  будем выбирать из условия  $\tilde{a}_\alpha > 0$ , т.е.

$$\sigma_1^* = \sigma_1^*(y_{i_1 i_2}^n) = \begin{cases} -\min\{r_{1\alpha}\}, & \text{если } r_{1\alpha} < 0, \\ 0, & \text{если } r_{1\alpha} > 0. \end{cases}$$

В соответствии с работами [23; 24, с. 70] нетрудно показать, что разностная схема (35), (36) аппроксимирует исходную задачу (32)–(34) четвёртым порядком по пространству и первым по времени, т.е. для её невязки

$$\psi = -(1 + \lambda)u_t + \sum_{\alpha=1}^2 (\Lambda_\alpha^{(b)}\sigma_\alpha\hat{u} + \Lambda_\alpha^{(b)}(1 - \sigma_\alpha)u) + \sum_{\alpha=1}^2 \frac{h_\alpha^2}{12}\Lambda_\alpha^{(b)}\Lambda_{3-\alpha}^{(b)}\hat{u} +$$

$$+ \sum_{\alpha=1}^2 \left( \frac{\sigma_0^* h^2}{12}\Lambda_\alpha^{(b)}\hat{u} + \tilde{a}_\alpha\Lambda_\alpha^{(b)}\hat{u} + \frac{h_\alpha^2}{12}\Lambda_\alpha^{(b)}u(1 - \hat{u}) \right) + (1 + \lambda)u(1 - \hat{u})$$

имеет место априорная оценка

$$\|\psi\|_C \leq M_3(|h|^4 + \tau), \quad M_3 = \text{const} > 0.$$

Для применения принципа максимума разностную схему (36) приведём к каноническому виду (2) и проверим достаточные условия на коэффициенты (4), (5). После элементарных преобразований находим

$$A = (1 + \lambda)(1 + \tau y_{ij}^n) + \tau \sum_{\alpha=1}^2 \frac{1}{h_\alpha^2} \left[ \left( \sigma_\alpha - \frac{h_\alpha^2}{12} y_{ij}^n + \frac{h^2}{12} \sigma_0^* \right) \theta_\alpha + \tilde{a}_\alpha \tilde{\theta}_\alpha - \frac{\theta_\alpha \theta_{3-\alpha} h_\alpha^2}{12 h_{3-\alpha}^2} \right],$$

$$B_1 = \frac{\tau}{h_1^2} \left( \left( \sigma_1 - \frac{h_1^2}{12} y_{i-1j}^n \right) \xi_1 + \tilde{a}_1 \tilde{\xi}_1 \right) + \tau \xi_1 \left( \frac{h^2}{12 h_1^2} \sigma_0^* - \frac{1}{12 h_2^2} \theta_2^{(-1_1)} - \frac{1}{12 h_1^2} \theta_2 \right),$$

$$B_2 = \frac{\tau}{h_1^2} \left( \left( \sigma_1 - \frac{h_1^2}{12} y_{i+1j}^n \right) \eta_1 + \tilde{a}_1 \tilde{\eta}_1 \right) + \tau \eta_1 \left( \frac{h^2}{12 h_1^2} \sigma_0^* - \frac{1}{12 h_2^2} \theta_2^{(+1_1)} - \frac{1}{12 h_1^2} \theta_2 \right),$$

$$B_3 = \frac{\tau}{h_2^2} \left( \left( \sigma_2 - \frac{h_2^2}{12} y_{ij-1}^n \right) \xi_2 + \tilde{a}_2 \tilde{\xi}_2 \right) + \tau \xi_2 \left( \frac{h^2}{12 h_2^2} \sigma_0^* - \frac{1}{12 h_2^2} \theta_1 - \frac{1}{12 h_1^2} \theta_1^{(-1_2)} \right),$$

$$B_4 = \frac{\tau}{h_2^2} \left( \left( \sigma_2 - \frac{h_2^2}{12} y_{ij+1}^n \right) \eta_2 + \tilde{a}_2 \tilde{\eta}_2 \right) + \tau \eta_2 \left( \frac{h^2}{12 h_2^2} \sigma_0^* - \frac{1}{12 h_2^2} \theta_1 - \frac{1}{12 h_1^2} \theta_1^{(+1_2)} \right),$$

$$B_5 = \frac{\tau}{12 h_2^2} \xi_1 \xi_2^{-1_1} + \frac{\tau}{12 h_1^2} \xi_1^{-1_2} \xi_2 > 0, \quad B_6 = \frac{\tau}{12 h_2^2} \eta_1 \xi_2^{+1_1} + \frac{\tau}{12 h_1^2} \eta_1^{-1_2} \xi_2 > 0,$$

$$B_7 = \frac{\tau}{12 h_2^2} \xi_1 \eta_2^{-1_1} + \frac{\tau}{12 h_1^2} \xi_1^{+1_2} \eta_2 > 0, \quad B_8 = \frac{\tau}{12 h_2^2} \eta_1 \eta_2^{+1_1} + \frac{\tau}{12 h_1^2} \eta_1^{+1_2} \eta_2 > 0,$$

$$F = (1 + \tau) \left[ \left( 1 + \lambda - \frac{1}{12} (\theta_1 + \theta_2) \right) y_{ij}^n + \xi_1 y_{i-1j}^n + \eta_1 y_{i+1j}^n + \xi_2 y_{ij-1}^n + \eta_2 y_{ij+1}^n \right],$$

$$D = (1 + \lambda)(1 + \tau y_{ij}^n) + \frac{\tau}{12} (\xi_1 y_{i-1j}^n + \eta_1 y_{i+1j}^n + \xi_2 y_{ij-1}^n + \eta_2 y_{ij+1}^n - (\theta_1 + \theta_2) y_{ij}^n),$$

где

$$\begin{aligned}\theta_\alpha &= 2\kappa_\alpha + h_\alpha b_\alpha^+ - h_\alpha b_\alpha^-, & \tilde{\theta}_\alpha &= 2\tilde{\kappa}_\alpha + h_\alpha \tilde{b}_\alpha^+ - h_\alpha \tilde{b}_\alpha^-, \\ \xi_\alpha &= \kappa_\alpha + h_\alpha b_\alpha^+, & \tilde{\xi}_\alpha &= \tilde{\kappa}_\alpha + h_\alpha \tilde{b}_\alpha^+, & \eta_\alpha &= \kappa_\alpha - h_\alpha b_\alpha^-, & \tilde{\eta}_\alpha &= \tilde{\kappa}_\alpha - h_\alpha \tilde{b}_\alpha^-\end{aligned}$$

Параметр регуляризации  $\sigma_0^*$  выбираем из условия неотрицательности коэффициентов  $B_\alpha$ ,  $\alpha = \overline{1, 4}$  (см. работы [10, 11]), т.е.

$$\sigma_0^* \geq \max_\alpha \{2/h_\alpha^2\}, \quad \frac{1}{\sqrt{5}} \leq \frac{h_1}{h_2} \leq \sqrt{5}. \quad (37)$$

Итак, имеет место следующее утверждение.

**Теорема 2.** Пусть выполнены условия (37) и

$$\max_{\alpha=1,2} \{h_\alpha\} < h_0, \quad h_0^2 = \frac{12}{m_3} e^T, \quad \tau > \tau_0, \quad \tau_0 = \max_{\alpha=1,2} \left\{ \frac{h_\alpha^2}{12 - h_\alpha^2 e^T m_3} \right\}.$$

Тогда разностное решение  $y(x, t)$ ,  $(x, t) \in \bar{\omega}$ , неотрицательно и для всех  $i_\alpha = \overline{0, N_\alpha}$ ,  $\alpha = 1, 2$ ,  $k = \overline{0, N_0}$  имеет место двусторонняя оценка

$$0 \leq y_{i_1 i_2}^{k} \leq e^{t k} m_3,$$

где  $m_3 = \max_{(x,t) \in Q_T} \{\max\{\mu_1(t), \mu_2(t), \mu_3(t), \mu_4(t)\}, u_0(x)\}$ .

**Замечание.** При выполнении условий теоремы 2 разностная схема (35) является монотонной.

**Следствие 3.** Для решения разностной схемы (35), (36) имеет место априорная оценка

$$\|y^n\|_{\bar{C}} \leq m_3 e^{t^n}, \quad n = \overline{0, N_0}.$$

## СПИСОК ЛИТЕРАТУРЫ

1. Fisher R.A. The wave of advance of advantageous genes // Ann. Hum. Genetic. 1937. V. 7. № 4. P. 353–369.
2. Kolmogorov A.N., Petrovskii I.G., Piskunov N.S. A study of the equation of diffusion with increase in the quantity of matter, and its application to a biological problem // Moscow University Mathematics Bull. 1937. V. 1. P. 1–26.
3. Шаповалов А.В., Трифонов А.Ю. Метод разложения Адомиана для одномерного нелокального уравнения Фишера–Колмогорова–Петровского–Пискунова // Изв. вузов. Физика. 2019. Т. 62. № 4. С. 135–143.
4. Murray J.D. Mathematical Biology: I. An Introduction. New York, 2002.
5. Толстых А.И. Компактные разностные схемы и их применение в задачах аэрогидродинамики. М., 1990.
6. Матус П.П., Хоанг Тхи Киеу Ань. Компактные разностные схемы на трёхточечном шаблоне для гиперболических уравнений второго порядка // Дифференц. уравнения. 2021. Т. 57. № 7. С. 963–975.
7. Утебаев Б.Д. Компактные разностные схемы для уравнений конвекции–диффузии // Весці НАН Беларусі. Сер. фіз.-мат. навук. 2021. Т. 57. № 3. С. 311–318.
8. Матус П.П., Утебаев Б.Д. Компактные и монотонные разностные схемы для параболических уравнений // Мат. моделирование. 2021. Т. 33. № 4. С. 60–78.
9. Самарский А.А. О монотонных разностных схемах для эллиптических и параболических уравнений в случае несамосопряжённого эллиптического оператора // Журн. вычислит. математики и мат. физики. 1965. Т. 5. № 3. С. 548–551.
10. Полевиков В.К. Схема повышенного порядка точности для задач высокоинтенсивного тепломассообмена // Современные проблемы тепловой гравитационной конвекции. Минск, 1974. С. 84–88.
11. Полевиков В.К. Монотонная разностная схема повышенного порядка точности для двумерных уравнений конвекции–диффузии // Журн. Белорус. гос. ун-та. Математика. Информатика. 2019. № 3. С. 71–83.



12. *Gaspar F.G., Lisbona F.J., Matus P., Tuyen V.T.K.* Monotone finite difference schemes for quasilinear parabolic problems with mixed boundary conditions // *Comput. Methods in Appl. Math.* 2016. V. 16. № 2. P. 231–243.
13. *Matus P., Hieu L.M., Vulkov L.G.* Analysis of second order difference schemes on nonuniform grids for quasilinear parabolic equations // *J. of Comput. and Appl. Math.* 2017. V. 310. P. 186–199.
14. *Matus P., Lemeshevsky S.* Stability and monotonicity of difference schemes for nonlinear scalar conservation laws and multidimensional quasi-linear parabolic equations // *Comput. Methods in Appl. Math.* 2009. V. 9. № 3. P. 253–280.
15. *Самарский А.А.* Теория разностных схем. М., 1983.
16. *Матус П.П., Хиеу Л.М., Волков Л.Г.* Принцип максимума для разностных схем с незначкопостоянными входными данными // *Докл. НАН Беларуси.* 2015. Т. 59. № 5. С. 13–17.
17. *Галанин М.П., Савенков Е.Б.* Методы численного анализа математических моделей. М., 2010.
18. *Самарский А.А., Вабищевич П.Н.* Численные методы решения задач конвекции–диффузии. М., 1999.
19. *Вабищевич П.Н., Самарский А.А.* Монотонные разностные схемы для задач конвекции–диффузии на треугольных сетках // *Журн. вычислит. математики и мат. физики.* 2002. Т. 42. № 9. С. 1368–1382.
20. *Friedrichs K.O., Hyers D.H.* Symmetric hyperbolic linear differential equations // *Comm. on Pure and Appl. Math.* 1954. V. 7. № 2. P. 345–392.
21. *Матус П.П., Поляков Д.Б.* О согласованных двусторонних оценках решений квазилинейных параболических уравнений и их аппроксимаций // *Дифференц. уравнения.* 2017. Т. 53. № 7. С. 991–1000.
22. *Ладыженская О.А., Солонников В.А., Уралъцева Н.Н.* Линейные и квазилинейные уравнения параболического типа. М., 1967.
23. *Самарский А.А.* Схемы повышенного порядка точности для многомерного уравнения теплопроводности // *Журн. вычислит. математики и мат. физики.* 1963. Т. 3. № 5. С. 812–840.
24. *Берковский Б.М., Полевичков В.К.* Вычислительный эксперимент в конвекции. Минск, 1988.
25. *Wang Xinyi, Lu Yuekai.* Exact solutions of the extended Burgers–Fisher equation // *Chinese Phys. Lett.* 1990. V. 7. № 4. P. 145–147.

Институт математики НАН Беларуси,  
г. Минск,  
Католический университет имени Иоанна-Павла II,  
г. Люблин, Польша

Поступила в редакцию 09.03.2022 г.  
После доработки 09.03.2022 г.  
Принята к публикации 25.05.2022 г.

УДК 519.622.2+681.5.015

## ИНТЕРВАЛЬНЫЙ ПОДХОД К РЕШЕНИЮ ЗАДАЧ ПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ ДИНАМИЧЕСКИХ СИСТЕМ

© 2022 г. А. Ю. Морозов, Д. Л. Ревизников

Представлен подход к решению задач параметрической идентификации динамических систем, целью которого является нахождение такой интервальной оценки параметров модели, при которой решение соответствующей задачи моделирования содержало бы исходные (экспериментальные) данные или минимизировало бы отклонение от них. В основе подхода лежит ранее разработанный, апробированный и обоснованный алгоритм адаптивной интерполяции для моделирования динамических систем с интервальными параметрами. Сформулирована задача минимизации расстояния между решением интервальной задачи и экспериментальными значениями фазовых переменных в пространстве границ интервальных оценок параметров модели. Получено выражение градиента минимизируемой целевой функции для дальнейшего применения методов оптимизации первого порядка. Выполнена апробация предложенного подхода на представительном ряде задач.

DOI: 10.31857/S0374064122070081, EDN: SEMWGM

**Введение.** Обратные задачи играют очень важную роль во многих областях. Определение закономерностей по имеющимся экспериментальным данным является ключевым моментом в построении математических моделей. Задача параметрической идентификации возникает на этапе, когда модель того или иного физического процесса уже определена, но неизвестными остаются параметры этой модели. Классический подход к определению параметров заключается в построении и минимизации некоторой целевой функции, которая характеризует отклонение модельного решения от экспериментальных данных.

Важной особенностью обратных задач является то, что они часто бывают некорректными. К признакам некорректности относятся: отсутствие решения, наличие более одного решения, сильная зависимость решения от погрешности входных данных (важность этого фактора связана с тем, что информация о фазовых переменных, как правило, поступает с некоторой погрешностью).

Существует ряд работ, посвящённых рассматриваемому классу задач [1–4]. В статьях [5, 6] предлагаются методы решения обратной коэффициентной задачи для уравнений в частных производных. В [7] решается обратная задача по определению плотности тепловых источников для уравнения субдиффузии. В работе [8] исследуется параметрическая идентификация в обратных задачах теплопроводности в условиях интервальной неопределённости на основе нейронных сетей, а в [9] предлагается структурно-параметрическая идентификация динамических объектов по интервальным исходным данным. В [10] представлен итеративный интервальный метод для прогнозирования границ параметров модели в условиях неопределённости в измеренных данных. Отметим важную роль обратных задач в аэрокосмической отрасли [11], вычислительной астрофизике [12] и электронике [13].

Применение интервального аппарата [14–16] в задачах параметрической идентификации связано с предположением, что параметры модели могут быть интервальными. Преимущество в использовании интервальных моделей заключается в том, что они дают ограничение сверху и снизу на интересующие величины, в отличие от классических моделей, которые их аппроксимируют.

При описании динамики различных объектов и процессов наиболее часто используются математические модели в виде систем обыкновенных дифференциальных уравнений (ОДУ), поэтому далее рассматриваются именно они.

В работе сформулирована задача минимизации расстояния между решением интервальной задачи и исходными значениями фазовых переменных в пространстве границ интервальных оценок параметров модели. Для дальнейшего применения методов оптимизации первого порядка получено выражение градиента соответствующей минимизируемой целевой функции. В процессе определения границ интервальных оценок необходимо многократно решать прямую интервальную задачу, для этого применяется ранее разработанный алгоритм адаптивной интерполяции (см. [17]).

Алгоритм адаптивной интерполяции относится к группе методов, определяющих явную зависимость решения задачи от значений интервальных параметров. В данной группе можно выделить две подгруппы: методы, оперирующие символьными выражениями [18–20], и методы, представляющие решение в виде полинома относительно интервальных параметров [17, 21, 22]. Алгоритм адаптивной интерполяции входит во вторую подгруппу.

Алгоритм имеет теоретическое обоснование [17, 23, 24]. Его идея заключается в построении для каждого момента времени полинома, который интерполирует зависимость решения задачи от значений параметров в заданной области неопределённости. Интерполяционный полином строится по определённому набору узлов, которые образуют сетку. На каждом шаге алгоритма сначала выполняется обновление значений, хранящихся в узлах сетки, а далее в зависимости от погрешности интерполяции происходит адаптация. В тех местах сетки, где погрешность большая, происходит добавление новых узлов, а в тех местах, где погрешность маленькая, – разрежение сетки. В классической версии алгоритма используется интерполяция на полных сетках, и из-за экспоненциального роста числа узлов при увеличении размерности области неопределённости он может применяться только к системам с небольшим числом (6–7) интервальных параметров. В работах [25–27] рассмотрены два подхода, позволяющих расширить область применения на динамические системы с большим количеством интервальных параметров: разреженные сетки [28–30] и тензорные поездки [31, 32]. С использованием данных подходов были промоделированы системы, содержащие до 18 интервальных параметров.

Схематично процесс решения задачи параметрической идентификации можно представить следующим образом. Сначала выполняется инициализация неизвестных параметров некоторыми произвольными интервалами и решается прямая задача. В каждом моменте времени, в котором известна информация о значении фазовых переменных, решается задача минимизации расстояния между параметрическим множеством, полученным в процессе работы алгоритма адаптивной интерполяции, и исходными (экспериментальными) данными. Если точка минимума лежит на одной из границ области неопределённости, то граница соответствующего интервала расширяется. Если точки минимума для всех моментов времени не лежат на определённой границе области неопределённости, то граница соответствующего интервала сужается. Далее снова решается прямая задача и т.д. до тех пор, пока интервальные оценки параметров не сойдутся.

**1. Постановка задачи.** Рассматриваются математические модели, заданные в виде системы ОДУ с интервальными начальными условиями и параметрами:

$$\frac{dy_i(t)}{dt} = f_i(y_1(t), y_2(t), \dots, y_n(t), \theta_1, \theta_2, \dots, \theta_m, t), \quad i = \overline{1, n},$$

$$y_i(t_0) \in [y_i^0, \overline{y_i^0}], \quad i = \overline{1, n},$$

$$\theta_i \in [\underline{\theta}_i, \overline{\theta}_i], \quad i = \overline{1, m}, \quad t \in [t_0, t_N], \quad (1)$$

где  $n$  – количество уравнений в системе,  $m$  – количество параметров,  $y_i^0, \overline{y_i^0}$ ,  $i = \overline{1, n}$ ,  $\underline{\theta}_i, \overline{\theta}_i$ ,  $i = \overline{1, m}$ , – нижние и верхние границы интервальных неопределённостей. Отметим, что интервальные неопределённости могут быть вырожденными, в этом случае границы будут совпадать. Вектор-функция  $\mathbf{f} = (f_1, f_2, \dots, f_n)^T$  удовлетворяет всем условиям, обеспечивающим единственность и существование решения при всех  $y_i(t_0) \in [y_i^0, \overline{y_i^0}]$ ,  $i = \overline{1, n}$ , и  $\theta_i \in [\underline{\theta}_i, \overline{\theta}_i]$ ,  $i = \overline{1, m}$ .

Решение системы (1) в каждый момент времени  $t_k$  является параметрическим множеством:

$$\mathbf{Y}^k = \{ \mathbf{y}^k(x_1, x_2, \dots, x_n, z_1, z_2, \dots, z_m) : x_i \in [\underline{y}_i^0, \overline{y}_i^0], \quad i = \overline{1, n}, \quad z_i \in [\underline{\theta}_i, \overline{\theta}_i], \quad i = \overline{1, m} \},$$

где  $\mathbf{y}^k(x_1, x_2, \dots, x_n, z_1, z_2, \dots, z_m) = (y_1(t_k), y_2(t_k), \dots, y_n(t_k))^T$  – решение следующей неинтервальной системы ОДУ:

$$\begin{aligned} \frac{dy_i(t)}{dt} &= f_i(y_1(t), y_2(t), \dots, y_n(t), z_1, z_2, \dots, z_m, t), \quad i = \overline{1, n}, \\ y_i(t_0) &= x_i, \quad i = \overline{1, n}, \quad t \in [t_0, t_k]. \end{aligned} \tag{2}$$

Пусть известны  $N$  экспериментальных точек в фазовом пространстве в различные моменты времени:

$$\hat{\mathbf{y}}(t_k) = (\hat{y}_1(t_k), \hat{y}_2(t_k), \dots, \hat{y}_n(t_k))^T, \quad k = \overline{1, N}. \tag{3}$$

Задача параметрической идентификации заключается в нахождении таких границ интервалов  $\underline{y}_1^0, \overline{y}_1^0, \underline{y}_2^0, \overline{y}_2^0, \dots, \underline{y}_n^0, \overline{y}_n^0, \underline{\theta}_1, \overline{\theta}_1, \underline{\theta}_2, \overline{\theta}_2, \dots, \underline{\theta}_m, \overline{\theta}_m$ , что  $\hat{\mathbf{y}}(t_k) \in \mathbf{Y}^k, k = \overline{1, N}$ . В данной постановке подразумевается, что для каждого  $\hat{\mathbf{y}}(t_k)$  должны существовать такие значения  $x_1, x_2, \dots, x_n$  и  $z_1, z_2, \dots, z_m$ , при которых  $\mathbf{y}^k(x_1, x_2, \dots, x_n, z_1, z_2, \dots, z_m) = \hat{\mathbf{y}}(t_k)$ . Отметим, что не всегда можно добиться того, чтобы любая точка принадлежала множеству решений  $\mathbf{Y}^k$ , но, как правило, в практических задачах это выполнимо.

Для нахождения соответствующих интервальных оценок начальных условий и параметров выполняется переход к задаче минимизации отклонения модельного решения от экспериментальных данных. Минимизируется следующая целевая функция:

$$J(\underline{y}_1^0, \overline{y}_1^0, \dots, \underline{y}_n^0, \overline{y}_n^0, \underline{\theta}_1, \overline{\theta}_1, \dots, \underline{\theta}_m, \overline{\theta}_m) = \sum_{k=1}^N \rho(\mathbf{Y}^k, \hat{\mathbf{y}}(t_k)), \tag{4}$$

где  $\rho(\mathbf{Y}^k, \hat{\mathbf{y}}(t_k)) = \min_{\mathbf{y}^k \in \mathbf{Y}^k} \|\mathbf{y}^k - \hat{\mathbf{y}}(t_k)\|^2$  – квадрат расстояния между множеством  $\mathbf{Y}^k$  и точкой  $\hat{\mathbf{y}}(t_k)$ . Традиционно в качестве нормы вектора  $\|\cdot\|$  используется взвешенная евклидова норма

$$\|\mathbf{a}\|^2 = \sum_{i=1}^n w_i a_i^2,$$

где  $w_i$  – веса, которые позволяют учесть разную природу фазовых переменных (для удобства изложения будем полагать, что  $w_i = 1, i = \overline{1, n}$ ).

Как будет видно в дальнейшем, важную роль в предлагаемом подходе играет явное представление  $\mathbf{y}^k(x_1, x_2, \dots, x_n, z_1, z_2, \dots, z_m), x_i \in [\underline{y}_i^0, \overline{y}_i^0], i = \overline{1, n}, z_i \in [\underline{\theta}_i, \overline{\theta}_i], i = \overline{1, m}$ , которое возможно получить с помощью алгоритма адаптивной интерполяции. В этом случае при вычислении  $\rho(\mathbf{Y}^k, \hat{\mathbf{y}}(t_k))$  минимизацию можно выполнять в пространстве начальных условий и параметров без дополнительного интегрирования системы (2):

$$\rho(\mathbf{Y}^k, \hat{\mathbf{y}}(t_k)) = \min_{\substack{x_i \in [\underline{y}_i^0, \overline{y}_i^0], i = \overline{1, n} \\ z_i \in [\underline{\theta}_i, \overline{\theta}_i], i = \overline{1, m}}} \left[ \sum_{i=1}^n (y_i^k(x_1, \dots, x_n, z_1, \dots, z_m) - \hat{y}_i(t_k))^2 \right]. \tag{5}$$

При минимизации (4) предполагается, что получающиеся интервальные оценки являются правильными:

$$\underline{y}_i^0 \leq \overline{y}_i^0, \quad i = \overline{1, n}, \quad \underline{\theta}_i \leq \overline{\theta}_i, \quad i = \overline{1, m}.$$

**Утверждение.** Если  $\tilde{y}_1^0, \tilde{y}_1^0, \dots, \tilde{y}_n^0, \tilde{y}_n^0, \tilde{\theta}_1, \tilde{\theta}_1, \dots, \tilde{\theta}_m, \tilde{\theta}_m$  – набор значений, на котором достигается минимум целевой функции (4), то на любом другом наборе  $\underline{y}_1^0, \underline{y}_1^0, \dots, \underline{y}_n^0, \underline{y}_n^0, \underline{\theta}_1, \underline{\theta}_1, \dots, \underline{\theta}_m, \underline{\theta}_m$  таком, что

$$\begin{aligned} \underline{y}_1^0 &\leq \tilde{y}_1^0, & \tilde{y}_1^0 &\leq \overline{y}_1^0, & \dots, & \underline{y}_n^0 &\leq \tilde{y}_n^0, & \tilde{y}_n^0 &\leq \overline{y}_n^0, \\ \underline{\theta}_1 &\leq \tilde{\theta}_1, & \tilde{\theta}_1 &\leq \overline{\theta}_1, & \dots, & \underline{\theta}_m &\leq \tilde{\theta}_m, & \tilde{\theta}_m &\leq \overline{\theta}_m, \end{aligned} \quad (6)$$

будет достигаться минимум.

Докажем от противного. Предположим, что существует такой набор  $\underline{y}_1^0, \underline{y}_1^0, \dots, \underline{y}_n^0, \underline{y}_n^0, \underline{\theta}_1, \underline{\theta}_1, \dots, \underline{\theta}_m, \underline{\theta}_m$ , удовлетворяющий условию (6), при котором минимум не достигается, тогда

$$J(\underline{y}_1^0, \underline{y}_1^0, \dots, \underline{y}_n^0, \underline{y}_n^0, \underline{\theta}_1, \underline{\theta}_1, \dots, \underline{\theta}_m, \underline{\theta}_m) > J(\tilde{y}_1^0, \tilde{y}_1^0, \dots, \tilde{y}_n^0, \tilde{y}_n^0, \tilde{\theta}_1, \tilde{\theta}_1, \dots, \tilde{\theta}_m, \tilde{\theta}_m).$$

Целевая функция (4) представляет собой сумму слагаемых, следовательно, существует такое  $k$ -е слагаемое, для которого выполняется неравенство

$$\min_{\substack{x_i \in [\underline{y}_i^0, \underline{y}_i^0], i=1, n \\ z_i \in [\underline{\theta}_i, \underline{\theta}_i], i=1, m}} \left[ \sum_{i=1}^n (y_i^k(\mathbf{x}, \mathbf{z}) - \hat{y}_i(t_k))^2 \right] > \min_{\substack{x_i \in [\tilde{y}_i^0, \tilde{y}_i^0], i=1, n \\ z_i \in [\tilde{\theta}_i, \tilde{\theta}_i], i=1, m}} \left[ \sum_{i=1}^n (y_i^k(\mathbf{x}, \mathbf{z}) - \hat{y}_i(t_k))^2 \right],$$

где  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  и  $\mathbf{z} = (z_1, z_2, \dots, z_m)$ . С учётом того, что  $[\tilde{y}_i^0, \tilde{y}_i^0] \subseteq [\underline{y}_i^0, \underline{y}_i^0]$  и  $[\tilde{\theta}_i, \tilde{\theta}_i] \subseteq [\underline{\theta}_i, \underline{\theta}_i]$ , получаем противоречие. Утверждение доказано.

Сформулированное утверждение можно рассматривать и в обратную сторону – в плане возможности уменьшения найденных интервальных оценок, что является важным при практической реализации. Однако отметим, что целью предлагаемого подхода является не нахождение минимальной интервальной оценки, а только определение таких значений  $\underline{y}_1^0, \underline{y}_1^0, \dots, \dots, \underline{y}_n^0, \underline{y}_n^0, \underline{\theta}_1, \underline{\theta}_1, \dots, \underline{\theta}_m, \underline{\theta}_m$ , при которых целевая функция (4) будет равна нулю или минимальна. Тем не менее если очевидным образом можно уменьшить получаемые интервальные оценки начальных условий и параметров, то эта процедура выполняется.

**2. Методы решения.** Поиск минимума целевой функции (4) в аналитическом виде зачастую затруднителен, поэтому необходимо применять численные методы оптимизации [33, 34]. Наиболее широко использующимися являются градиентные методы, которые, хотя и требуют знания градиента целевой функции, обладают высокой скоростью сходимости, компенсирующей затраты на вычисления градиента.

Одна итерация метода градиентного спуска записывается следующим образом:

$$\mathbf{u}^{(j+1)} = \mathbf{u}^{(j)} - \lambda^{(j)} \nabla F(\mathbf{u}^{(j)}),$$

где  $\mathbf{u}$  – вектор переменных,  $\nabla F$  – градиент минимизируемой функции  $F$ ,  $\lambda^{(j)}$  – скорость градиентного спуска. Данный класс методов является хорошо изученным и исследованным, поэтому подробно рассматриваться в данной работе не будет.

Для того чтобы вычислить градиент целевой функции (4), необходимо продифференцировать (5) по границе области поиска минимума. Обозначим через  $(\tilde{x}_1^k, \dots, \tilde{x}_n^k, \tilde{z}_1^k, \dots, \tilde{z}_m^k)$  точку минимума (5):

$$(\tilde{x}_1^k, \dots, \tilde{x}_n^k, \tilde{z}_1^k, \dots, \tilde{z}_m^k) = \arg \min_{\substack{x_i \in [\underline{y}_i^0, \underline{y}_i^0], i=1, n \\ z_i \in [\underline{\theta}_i, \underline{\theta}_i], i=1, m}} J^k(x_1, \dots, x_n, z_1, \dots, z_m), \quad (7)$$

где

$$J^k(x_1, \dots, x_n, z_1, \dots, z_m) = \sum_{i=1}^n (y_i^k(x_1, \dots, x_n, z_1, \dots, z_m) - \hat{y}_i(t_k))^2.$$

Дополнительно запишем компоненты градиента  $J^k$ :

$$\begin{aligned} & \frac{dJ^k(x_1, \dots, x_n, z_1, \dots, z_m)}{dx_j} = \\ & = 2 \sum_{i=1}^n (y_i^k(x_1, \dots, x_n, z_1, \dots, z_m) - \hat{y}_i(t_k)) \frac{dy_i^k(x_1, \dots, x_n, z_1, \dots, z_m)}{dx_j}, \quad j = \overline{1, n}, \\ & \frac{dJ^k(x_1, \dots, x_n, z_1, \dots, z_m)}{dz_j} = \\ & = 2 \sum_{i=1}^n (y_i^k(x_1, \dots, x_n, z_1, \dots, z_m) - \hat{y}_i(t_k)) \frac{dy_i^k(x_1, \dots, x_n, z_1, \dots, z_m)}{dz_j}, \quad j = \overline{1, m}. \end{aligned} \tag{8}$$

Очевидно, что если точка минимума лежит не на границе рассматриваемой области, то изменение границы никак не будет на неё влиять, т.е. производная по границе будет равна нулю. А если точка лежит на границе, то производная равна частной производной в этой точке. В результате компоненты градиента (4) имеют следующий вид:

$$\begin{aligned} \frac{dJ(\underline{y}_1^0, \overline{y}_1^0, \dots, \underline{y}_n^0, \overline{y}_n^0, \underline{\theta}_1, \overline{\theta}_1, \dots, \underline{\theta}_m, \overline{\theta}_m)}{dy_j^0} &= \sum_{k=1}^N \begin{cases} 0, & \tilde{x}_j^k \neq \underline{y}_j^0, \\ \frac{dJ^k(\tilde{x}_1^k, \dots, \tilde{x}_n^k, \tilde{z}_1^k, \dots, \tilde{z}_m^k)}{dx_j}, & \tilde{x}_j^k = \underline{y}_j^0, \end{cases} \\ \frac{dJ(\underline{y}_1^0, \overline{y}_1^0, \dots, \underline{y}_n^0, \overline{y}_n^0, \underline{\theta}_1, \overline{\theta}_1, \dots, \underline{\theta}_m, \overline{\theta}_m)}{dy_j^0} &= \sum_{k=1}^N \begin{cases} 0, & \tilde{x}_j^k \neq \overline{y}_j^0, \\ \frac{dJ^k(\tilde{x}_1^k, \dots, \tilde{x}_n^k, \tilde{z}_1^k, \dots, \tilde{z}_m^k)}{dx_j}, & \tilde{x}_j^k = \overline{y}_j^0, \end{cases} \\ & j = \overline{1, n}, \\ \frac{dJ(\underline{y}_1^0, \overline{y}_1^0, \dots, \underline{y}_n^0, \overline{y}_n^0, \underline{\theta}_1, \overline{\theta}_1, \dots, \underline{\theta}_m, \overline{\theta}_m)}{d\theta_j} &= \sum_{k=1}^N \begin{cases} 0, & \tilde{z}_j^k \neq \underline{\theta}_j, \\ \frac{dJ^k(\tilde{x}_1^k, \dots, \tilde{x}_n^k, \tilde{z}_1^k, \dots, \tilde{z}_m^k)}{dz_j}, & \tilde{z}_j^k = \underline{\theta}_j, \end{cases} \\ \frac{dJ(\underline{y}_1^0, \overline{y}_1^0, \dots, \underline{y}_n^0, \overline{y}_n^0, \underline{\theta}_1, \overline{\theta}_1, \dots, \underline{\theta}_m, \overline{\theta}_m)}{d\theta_j} &= \sum_{k=1}^N \begin{cases} 0, & \tilde{z}_j^k \neq \overline{\theta}_j, \\ \frac{dJ^k(\tilde{x}_1^k, \dots, \tilde{x}_n^k, \tilde{z}_1^k, \dots, \tilde{z}_m^k)}{dz_j}, & \tilde{z}_j^k = \overline{\theta}_j, \end{cases} \\ & j = \overline{1, m}. \end{aligned} \tag{9}$$

Проанализируем (8). Так как  $(\tilde{x}_1^k, \dots, \tilde{x}_n^k, \tilde{z}_1^k, \dots, \tilde{z}_m^k)$  является точкой минимума, то градиент (8) в этой точке будет равен нулю в случае, когда точка находится внутри рассматриваемой области  $[\underline{y}_1^0, \overline{y}_1^0] \times \dots \times [\underline{y}_n^0, \overline{y}_n^0] \times [\underline{\theta}_1, \overline{\theta}_1] \times \dots \times [\underline{\theta}_m, \overline{\theta}_m]$ , или может быть направлен внутрь области в случае, когда точка находится на границе. Отсюда

$$\frac{dJ^k(\tilde{x}_1^k, \dots, \tilde{x}_n^k, \tilde{z}_1^k, \dots, \tilde{z}_m^k)}{dx_j} \begin{cases} \leq 0, & \tilde{x}_j^k = \overline{y}_j^0, \\ = 0, & \underline{y}_j^0 < \tilde{x}_j^k < \overline{y}_j^0, \\ \geq 0, & \tilde{x}_j^k = \underline{y}_j^0, \end{cases} \quad j = \overline{1, n},$$

$$\frac{dJ^k(\tilde{x}_1^k, \dots, \tilde{x}_n^k, \tilde{z}_1^k, \dots, \tilde{z}_m^k)}{dz_j} \begin{cases} \leq 0, & \tilde{z}_j^k = \overline{\theta}_j, \\ = 0, & \theta_j < \tilde{z}_j^k < \overline{\theta}_j, \quad j = \overline{1, m}, \\ \geq 0, & \tilde{z}_j^k = \underline{\theta}_j, \end{cases}$$

и для (9) получаем

$$\begin{aligned} \frac{dJ(\underline{y}_1^0, \overline{y}_1^0, \dots, \underline{y}_n^0, \overline{y}_n^0, \underline{\theta}_1, \overline{\theta}_1, \dots, \underline{\theta}_m, \overline{\theta}_m)}{dy_j^0} &= \sum_{k=1}^N \max \left[ 0, \frac{dJ^k(\tilde{x}_1^k, \dots, \tilde{x}_n^k, \tilde{z}_1^k, \dots, \tilde{z}_m^k)}{dx_j} \right], \\ \frac{dJ(\underline{y}_1^0, \overline{y}_1^0, \dots, \underline{y}_n^0, \overline{y}_n^0, \underline{\theta}_1, \overline{\theta}_1, \dots, \underline{\theta}_m, \overline{\theta}_m)}{dy_j^0} &= \sum_{k=1}^N \min \left[ 0, \frac{dJ^k(\tilde{x}_1^k, \dots, \tilde{x}_n^k, \tilde{z}_1^k, \dots, \tilde{z}_m^k)}{dx_j} \right], \quad j = \overline{1, n}, \\ \frac{dJ(\underline{y}_1^0, \overline{y}_1^0, \dots, \underline{y}_n^0, \overline{y}_n^0, \underline{\theta}_1, \overline{\theta}_1, \dots, \underline{\theta}_m, \overline{\theta}_m)}{d\theta_j} &= \sum_{k=1}^N \max \left[ 0, \frac{dJ^k(\tilde{x}_1^k, \dots, \tilde{x}_n^k, \tilde{z}_1^k, \dots, \tilde{z}_m^k)}{dz_j} \right], \\ \frac{dJ(\underline{y}_1^0, \overline{y}_1^0, \dots, \underline{y}_n^0, \overline{y}_n^0, \underline{\theta}_1, \overline{\theta}_1, \dots, \underline{\theta}_m, \overline{\theta}_m)}{d\theta_j} &= \sum_{k=1}^N \min \left[ 0, \frac{dJ^k(\tilde{x}_1^k, \dots, \tilde{x}_n^k, \tilde{z}_1^k, \dots, \tilde{z}_m^k)}{dz_j} \right], \quad j = \overline{1, m}. \end{aligned}$$

Следовательно, в процессе минимизации целевой функции (4) нижние границы интервалов не будут увеличиваться, а верхние границы – уменьшаться, т.е. получающаяся область неопределённости  $[\underline{y}_1^0, \overline{y}_1^0] \times \dots \times [\underline{y}_n^0, \overline{y}_n^0] \times [\underline{\theta}_1, \overline{\theta}_1] \times \dots \times [\underline{\theta}_m, \overline{\theta}_m]$  не будет сужаться. Однако если ни одна точка минимума (7) не лежит на определённой границе, то соответствующая производная равна нулю и границу можно передвинуть, тем самым уменьшив рассматриваемую область. Таким образом, получаем правила сужения:

$$\begin{aligned} \underline{y}_i^0 &= \min_k(\tilde{x}_i^k), \quad \text{если } \frac{dJ}{dy_i^0} = 0, \quad \overline{y}_i^0 = \max_k(\tilde{x}_i^k), \quad \text{если } \frac{dJ}{dy_i^0} = 0, \quad i = \overline{1, n}, \\ \underline{\theta}_i &= \min_k(\tilde{z}_i^k), \quad \text{если } \frac{dJ}{d\theta_i} = 0, \quad \overline{\theta}_i = \max_k(\tilde{z}_i^k), \quad \text{если } \frac{dJ}{d\theta_i} = 0, \quad i = \overline{1, m}. \end{aligned}$$

Это соответствует следствию из ранее сформулированного утверждения о том, что получающиеся интервальные оценки можно уменьшить. Отметим, что в контексте рассматриваемых задач данный момент является ключевым, потому что позволяет исключить постоянное расширение интервальных оценок.

**3. Алгоритм адаптивной интерполяции.** Для решения задачи параметрической идентификации требуется знать

$$\mathbf{y}^k(x_1, x_2, \dots, x_n, z_1, z_2, \dots, z_m), \quad k = \overline{1, N}, \quad x_i \in [\underline{y}_i^0, \overline{y}_i^0], \quad i = \overline{1, n}, \quad z_i \in [\underline{\theta}_i, \overline{\theta}_i], \quad i = \overline{1, m}.$$

В свою очередь, для этого необходимо решить прямую задачу (1).

В данном пункте приводится краткое описание алгоритма адаптивной интерполяции для интегрирования систем ОДУ с интервальными начальными условиями и параметрами в соответствии с работами [17, 25, 26].

Цель алгоритма – для каждого момента времени  $t_k$  построить вектор-функцию

$$\mathbf{P}^k(x_1, x_2, \dots, x_n, z_1, z_2, \dots, z_m),$$

интерполирующую  $\mathbf{y}^k(x_1, x_2, \dots, x_n, z_1, z_2, \dots, z_m)$  с контролируемой точностью.

В начальный момент времени  $t_0$  вектор-функция  $\mathbf{P}^0$  определяется тривиальным образом:

$$\mathbf{P}^0(x_1, x_2, \dots, x_n, z_1, z_2, \dots, z_m) = (x_1, x_2, \dots, x_n)^T.$$

Построим с помощью индукции решение для произвольного момента времени. Пусть в момент времени  $t_k$  известна  $\mathbf{P}^k(x_1, x_2, \dots, x_n, z_1, z_2, \dots, z_m)$ . Получение

$$\mathbf{P}^{k+1}(x_1, x_2, \dots, x_n, z_1, z_2, \dots, z_m)$$

сводится к интерполяции неявной функции

$$\mathbf{F}(x_1, x_2, \dots, x_n, z_1, z_2, \dots, z_m) = (y_1(t_{k+1}), y_2(t_{k+1}), \dots, y_n(t_{k+1}))^T,$$

заданной в виде системы ОДУ

$$\frac{dy_i(t)}{dt} = f_i(y_1(t), y_2(t), \dots, y_n(t), z_1, z_2, \dots, z_m, t), \quad i = \overline{1, n},$$

$$y_i(t_k) = P_i^k(x_1, x_2, \dots, x_n, z_1, z_2, \dots, z_m), \quad i = \overline{1, n}, \quad t \in [t_k, t_{k+1}].$$

Традиционно интерполяционный полином  $\mathbf{P}^k(x_1, x_2, \dots, x_n, z_1, z_2, \dots, z_m)$  строится по определённому набору узлов, которые образуют сетку. Поэтому сначала выполняется перенос решений, которые соответствуют узлам, на  $(k + 1)$ -й временной слой, а далее в зависимости от значения погрешности интерполяции происходит адаптация. В тех местах сетки, где погрешность большая, происходит добавление новых узлов, а в тех местах, где погрешность маленькая, – разрежение сетки. Интерполяционный полином  $\mathbf{P}$  может быть любым, необходимо только, чтобы была возможность контролировать погрешность интерполяции.

При решении задачи параметрической идентификации  $\mathbf{y}^k(x_1, x_2, \dots, x_n, z_1, z_2, \dots, z_m)$  заменяется полученным интерполяционным полиномом  $\mathbf{P}^k(x_1, x_2, \dots, x_n, z_1, z_2, \dots, z_m)$ . Отметим, что так как полином представляет собой явную функцию, то требуемые в процессе вычислений частные производные можно вычислить аналитически.

**4. Результаты.** Далее решены несколько задач параметрической идентификации. В качестве критерия останова градиентного спуска применялось ограничение на значение минимизируемой целевой функции  $J < 10^{-12}$ . Во всех примерах вместо экспериментальных точек использовались квазиреальные экспериментальные точки, полученные следующим образом. Параметры и начальные условия задавались интервально и решалась прямая задача. В равномерно распределённые в интервале интегрирования моменты времени  $t_k$  в найденное решение  $\mathbf{P}^k(x_1, x_2, \dots, x_n, z_1, z_2, \dots, z_m)$  подставлялись случайные значения из соответствующих интервалов:

$$\begin{aligned} x_1 = \hat{x}_1^k &\sim U(\underline{y}_1^0, \overline{y}_1^0), & x_2 = \hat{x}_2^k &\sim U(\underline{y}_2^0, \overline{y}_2^0), & \dots, & & x_n = \hat{x}_n^k &\sim U(\underline{y}_n^0, \overline{y}_n^0), \\ z_1 = \hat{z}_1^k &\sim U(\underline{\theta}_1, \overline{\theta}_1), & z_2 = \hat{z}_2^k &\sim U(\underline{\theta}_2, \overline{\theta}_2), & \dots, & & z_m = \hat{z}_m^k &\sim U(\underline{\theta}_m, \overline{\theta}_m), \end{aligned} \tag{10}$$

где  $U(a, b)$  – непрерывное равномерное распределение на отрезке  $[a, b]$ . Таким образом,

$$\hat{\mathbf{y}}(t_k) = \mathbf{P}^k(\hat{x}_1^k, \hat{x}_2^k, \dots, \hat{x}_n^k, \hat{z}_1^k, \hat{z}_2^k, \dots, \hat{z}_m^k).$$

Исходные интервальные оценки начальных условий и параметров:

$$\left[ \min_k(\hat{x}_i^k), \max_k(\hat{x}_i^k) \right], \quad i = \overline{1, n}, \quad \left[ \min_k(\hat{z}_i^k), \max_k(\hat{z}_i^k) \right], \quad i = \overline{1, m}.$$

Вначале рассматривается система ОДУ с двумя неизвестными параметрами, которая соответствует модели Лотки–Вольтерры [35]:

$$\begin{aligned} u' &= \alpha u - 2uv, & v' &= -v + \beta uv, \\ u(0) &= 1, & v(0) &= 3, & t &\in [0, 5.3]. \end{aligned} \tag{11}$$



Из интервала интегрирования были взяты 20 равномерно расположенных точек с шагом по времени 0.265. Квазиреальные экспериментальные точки на фазовой плоскости получены при следующих интервальных оценках параметров:  $\alpha \in [1.95, 2.035]$  и  $\beta \in [0.9652, 1.045]$ . Начальное приближение в методе градиентного спуска:  $\alpha^{(0)} \in [1.5, 1.51]$  и  $\beta^{(0)} \in [0.6, 0.61]$ . На рис. 1 проиллюстрирован процесс параметрической идентификации системы.

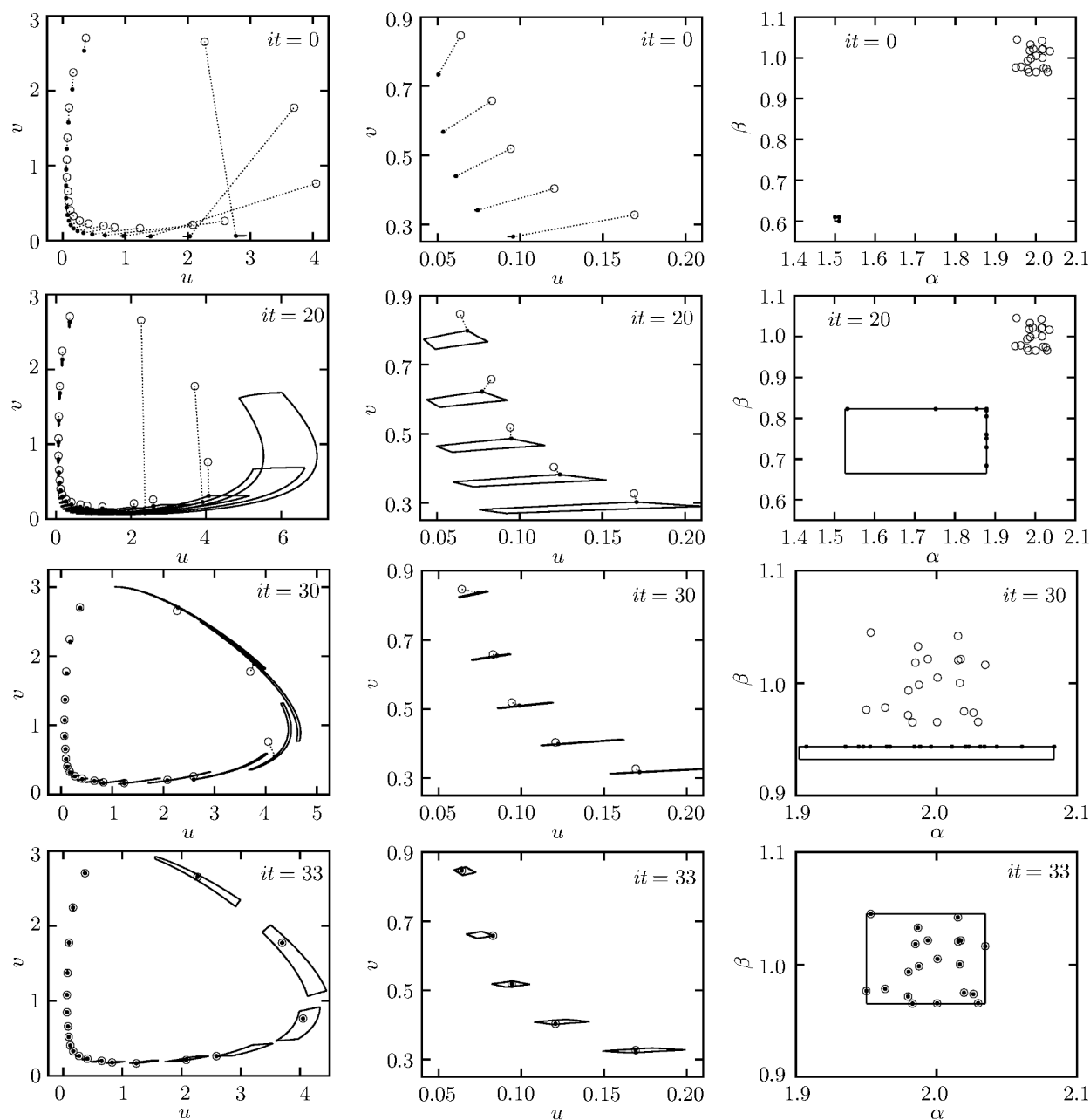


Рис. 1. Результаты решения задачи параметрической идентификации для системы (11).

На рис. 1 и на всех рисунках далее сплошными линиями показаны множества решений прямой задачи в различные моменты времени (первая и вторая колонка на рис. 1), светлые кружки – экспериментальные данные. Жирная точка обозначает ближайшую точку к экспериментальной точке в множестве решений прямой задачи, а пунктирная линия – расстояние между ними. Сумма квадратов расстояний фактически определяет значение минимизируемой целевой функции (4).

Каждая строка рис. 1 соответствует определённой итерации градиентного спуска. Слева показана фазовая плоскость, а по центру – её увеличенная часть ( $u \in [0.04, 0.21]$  и  $v \in [0.25, 0.9]$ ). На рисунках справа продемонстрировано решение задачи в пространстве параметров. Здесь светлыми кружками показаны полученные случайным образом значения параметров (10) при генерации экспериментальных точек; жирными точками – решение соответствующей задачи (7); сплошной линией – интервальная оценка параметров (получающаяся область неопределённости).

В процессе минимизации множества решений прямой задачи стремятся покрыть экспериментальные точки, о чем свидетельствует уменьшение расстояний между ними. При  $it = 30$  наблюдается, что все жирные точки принадлежат верхней границе интервальной оценки параметра  $\beta$ . Это говорит о том, что исходные значения параметров (10) расположились с одной стороны относительно текущей области неопределённости. На последней итерации видно, что полученные множества решений полностью покрыли экспериментальные точки, кроме того, определены соответствующие значения параметров, о чем свидетельствует совпадение жирных точек со светлыми кружками на рисунке справа.

Далее рассматривается система ОДУ, в которой неизвестными являются начальные условия:

$$u' = 4u - \frac{5}{4}uv + \frac{1}{10}u^2, \quad v' = -2v + \frac{1}{2}uv + \frac{1}{10}v^2,$$

$$u(0) = u_0, \quad v(0) = v_0, \quad t \in [0, 4]. \tag{12}$$

Аналогично предыдущему примеру из интервала интегрирования были взяты 20 равномерно расположенных точек с шагом по времени 0.2. Исходные интервальные оценки начальных условий:  $u_0 \in [4.301, 4.639]$  и  $v_0 \in [2.861, 3.180]$ . Начальное приближение:  $u_0^{(0)} \in [3.3, 3.7]$  и  $v_0^{(0)} \in [3.8, 4.2]$ . На рис. 2 показан процесс решения задачи.

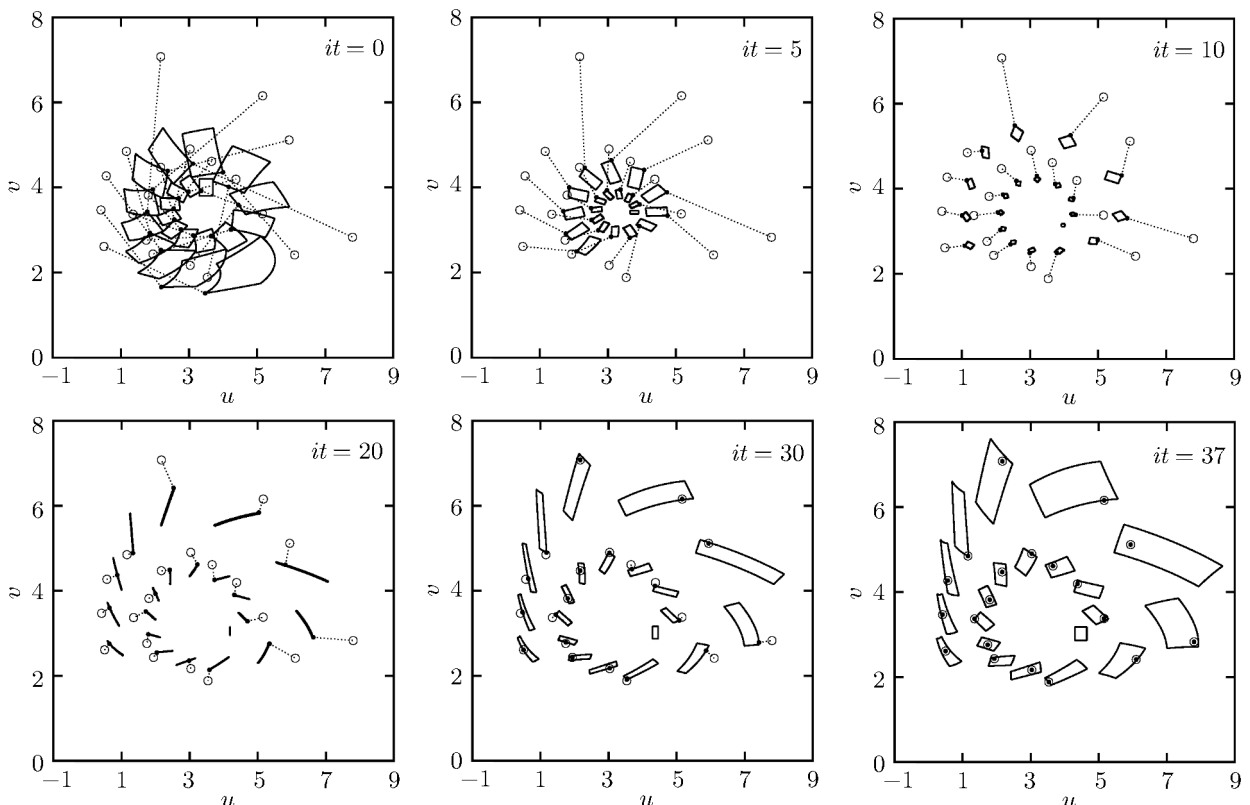


Рис. 2. Иллюстрация решения задачи параметрической идентификации для системы (12).

Траектория решения прямой задачи имеет спиралевидную структуру, что является особенностью рассматриваемой системы ОДУ. С каждой итерацией градиентного спуска получающиеся множества решений приближаются к экспериментальным точкам, уменьшается суммарная длина пунктирных линий. На 37-й итерации все экспериментальные точки содержатся в соответствующих множествах решений прямой задачи. Полученные интервальные оценки начальных условий  $u_0^{(37)} \in [4.301, 4.639]$  и  $v_0^{(37)} \in [2.861, 3.180]$  совпадают с исходными. Однако такая ситуация имеет место не всегда, далее будут рассмотрены примеры, где количество фазовых переменных отличается от количества неизвестных величин.

Отметим, что решение задачи (12) можно получить, выполнив интегрирование в обратную сторону для каждой экспериментальной точки, тем самым найдя соответствующие прообразы в пространстве начальных условий, по которым можно построить интервальную оценку.

Рассмотрим задачу (11) с увеличенным количеством неизвестных величин. Из системы исключим информацию о начальных условиях:

$$\begin{aligned} u' &= \alpha u - 2uv, & v' &= -v + \beta uv, \\ u(0) &= u_0, & v(0) &= v_0, & t &\in [0, 5.3]. \end{aligned} \quad (13)$$

Исходные интервальные оценки начальных условий и параметров, по которым генерировались экспериментальные точки на фазовой плоскости:

$$u_0 \in [0.9505, 1.046], \quad v_0 \in [2.952, 3.050], \quad \alpha \in [1.959, 2.041], \quad \beta \in [0.9732, 1.048].$$

Начальное приближение:

$$u_0^{(0)} \in [0.8, 0.81], \quad v_0^{(0)} \in [2.8, 2.81], \quad \alpha^{(0)} \in [1.8, 1.81], \quad \beta^{(0)} \in [0.6, 0.61].$$

На рис. 3 показан процесс поиска интервальных оценок параметров и начальных условий. Множество решений прямой задачи в каждый момент времени представляет собой проекцию деформирующегося четырёхмерного прямоугольного параллелепипеда.

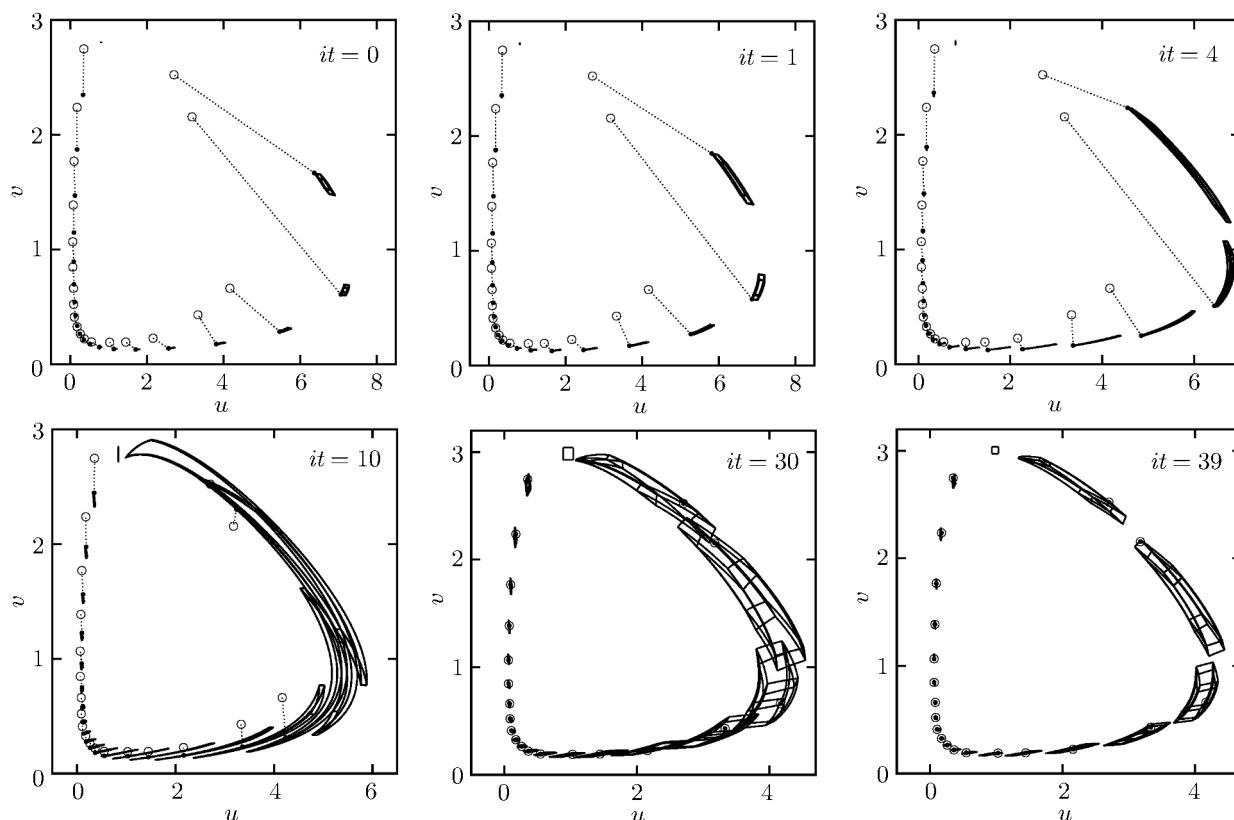


Рис. 3. Процесс решения задачи параметрической идентификации для системы (13).

На последней итерации получены следующие интервальные оценки:  $u_0^{(39)} \in [0.9393, 1.033]$ ,  $v_0^{(39)} \in [2.972, 3.037]$ ,  $\alpha^{(39)} \in [1.970, 2.035]$  и  $\beta^{(39)} \in [0.9749, 1.033]$ . Они отличаются от исходных оценок, но при этом все экспериментальные точки содержатся в соответствующих множествах решений прямой задачи. Отметим, что результат зависит от выбора начального приближения, так как минимизируемая целевая функция является многоэкстремальной.

В заключение рассмотрим модельную задачу: движение тела с неизвестным начальным положением, неизвестными начальными скоростями и неизвестной массой под действием сил гравитации вокруг массивного тела. Система ОДУ в безразмерных переменных имеет следующий вид:

$$\begin{aligned} x_1' &= vx_1, & y_1' &= vy_1, \\ vx_1' &= m_2 \frac{x_2 - x_1}{((x_2 - x_1)^2 + (y_2 - y_1)^2)^{3/2}}, & vy_1' &= m_2 \frac{y_2 - y_1}{((x_2 - x_1)^2 + (y_2 - y_1)^2)^{3/2}}, \\ x_2' &= vx_2, & y_2' &= vy_2, \\ vx_2' &= m_1 \frac{x_1 - x_2}{((x_2 - x_1)^2 + (y_2 - y_1)^2)^{3/2}}, & vy_2' &= m_1 \frac{y_1 - y_2}{((x_2 - x_1)^2 + (y_2 - y_1)^2)^{3/2}}, \\ x_1(0) &= x_{1,0}, & y_1(0) &= y_{1,0}, \\ vx_1(0) &= vx_{1,0}, & vy_1(0) &= vy_{1,0}, \\ x_2(0) &= y_2(0) = 0, & vx_2(0) &= vy_2(0) = 0, \\ m_2 &= 10^2, & t &\in [0, 0.2]. \end{aligned} \quad (14)$$

В отличие от предыдущих примеров здесь с шагом по времени 0.02 было взято меньше точек – 10. Кроме того, экспериментальные данные содержали информацию только о фазовых переменных, соответствующих положению первого тела  $(x_1, y_1)$ . Подобные задачи возникают в астрофизике, когда об исследуемом космическом объекте кроме его положения на небесной сфере ничего не известно. Исходные интервальные оценки начальных условий и массы тела:  $x_{1,0} \in [1.000, 1.018]$ ,  $y_{1,0} \in [3.583 \cdot 10^{-3}, 1.956 \cdot 10^{-2}]$ ,  $vx_{1,0} \in [8.183 \cdot 10^{-3}, 3.794 \cdot 10^{-1}]$ ,  $vy_{1,0} \in [10.04, 10.48]$  и  $m_1 \in [10.07, 10.45]$ . Начальное приближение:  $x_{1,0}^{(0)} \in [0.9, 0.91]$ ,  $y_{1,0}^{(0)} \in [-0.2, -0.1]$ ,  $vx_{1,0}^{(0)} \in [0.0, 0.1]$ ,  $vy_{1,0}^{(0)} \in [9.5, 9.6]$  и  $m_1^{(0)} \in [9.5, 9.6]$ . На рис. 4 продемонстрировано изменение множеств решений прямой задачи в процессе параметрической идентификации.

Здесь показаны две последовательности множеств, которые соответствуют положению двух тел в пространстве в различные моменты времени. Так как второе тело более массивное и изначально находится в состоянии покоя, его перемещение в пространстве незначительно по сравнению с первым телом.

На 40-й итерации градиентного спуска получены следующие интервальные оценки начального положения, скорости и массы первого тела, при которых решение прямой задачи полностью содержит экспериментальные точки:

$$\begin{aligned} x_{1,0}^{(40)} &\in [1.016, 1.027], & y_{1,0}^{(40)} &\in [3.042 \cdot 10^{-2}, 1.011 \cdot 10^{-1}], \\ vx_{1,0}^{(40)} &\in [-0.2047, -0.1264], & vy_{1,0}^{(40)} &\in [9.721, 10.10], & m_1^{(40)} &\in [9.480, 9.483]. \end{aligned}$$

Отметим, что в процессе параметрической идентификации данной системы выполнялась минимизация целевой функции в десятимерном пространстве (пять интервальных неопределённостей, в каждой по две границы).

Таким образом, с помощью предложенного подхода успешно была решена задача интервальной параметрической идентификации для различных систем ОДУ с разным количеством неизвестных значений.

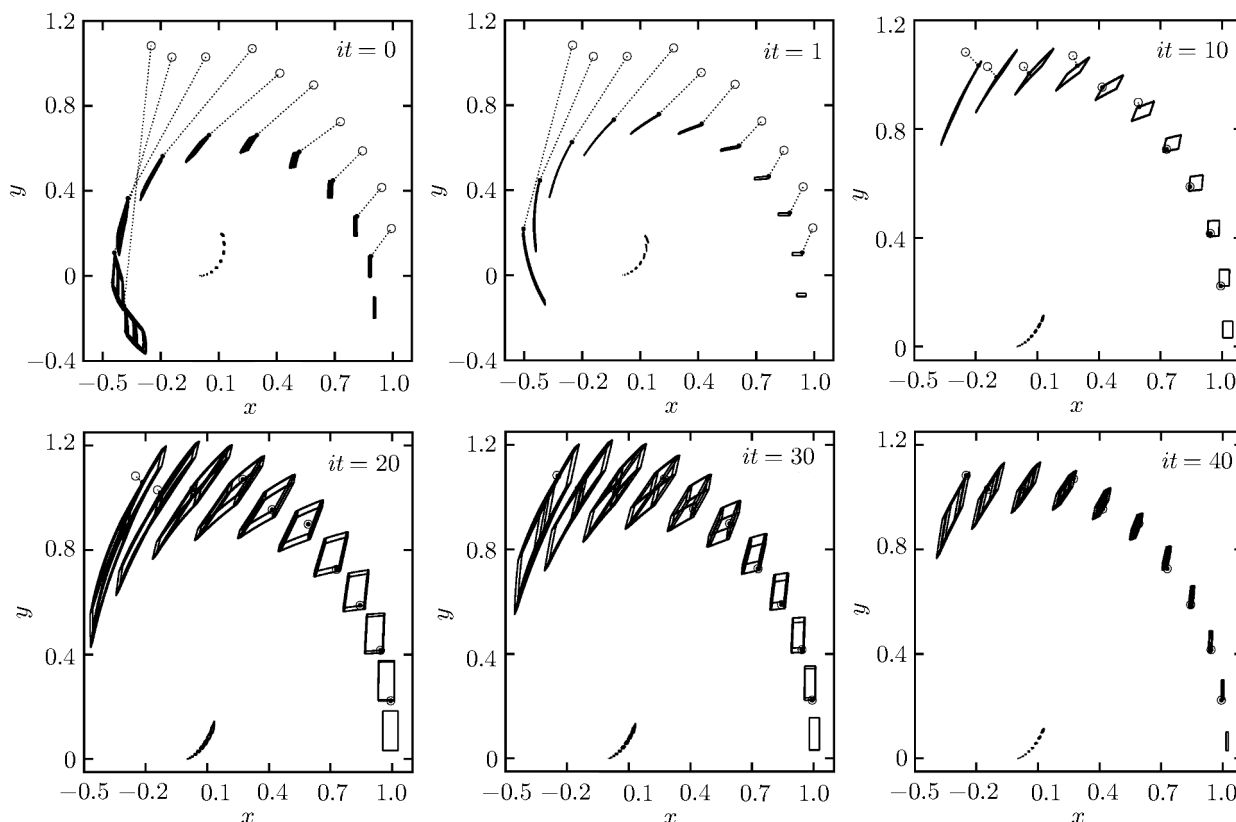


Рис. 4. Параметрическая идентификация системы (14).

**5. Обсуждение.** В общем случае задача (7) может иметь множество решений. В частности, данная ситуация возникает, когда точка  $\hat{y}(t_k)$  принадлежит множеству  $\mathbf{Y}^k$ , тогда (7) превращается в систему

$$\begin{aligned} y_1^k(x_1, \dots, x_n, z_1, \dots, z_m) &= \hat{y}_1(t_k), & y_2^k(x_1, \dots, x_n, z_1, \dots, z_m) &= \hat{y}_2(t_k), & \dots \\ \dots, & & y_n^k(x_1, \dots, x_n, z_1, \dots, z_m) &= \hat{y}_n(t_k), \end{aligned} \quad (15)$$

в которой  $n$  уравнений и  $(n + m)$  неизвестных. На рис. 3 видно, что множество решений системы ОДУ (13) представляет собой проекцию четырёхмерного деформирующегося прямоугольного параллелепипеда. Здесь каждой точке на фазовой плоскости соответствует двумерная поверхность в пространстве параметров и начальных условий, которая определяется решением соответствующей системы (15).

В качестве  $(\tilde{x}_1^k, \dots, \tilde{x}_n^k, \tilde{z}_1^k, \dots, \tilde{z}_m^k)$  имеет смысл брать любую точку, наиболее близкую к центру рассматриваемой области

$$[\underline{y}_1^0, \overline{y}_1^0] \times \dots \times [\underline{y}_n^0, \overline{y}_n^0] \times [\underline{\theta}_1, \overline{\theta}_1] \times \dots \times [\underline{\theta}_m, \overline{\theta}_m],$$

т.е.

$$(\tilde{x}_1^k, \dots, \tilde{x}_n^k, \tilde{z}_1^k, \dots, \tilde{z}_m^k) = \arg \min_{(x_1, \dots, x_n, z_1, \dots, z_m) \in \chi} \left[ \sum_{i=1}^n \left( \frac{y_i^0 + \overline{y}_i^0}{2} - x_i \right)^2 + \sum_{i=1}^m \left( \frac{\theta_i + \overline{\theta}_i}{2} - z_i \right)^2 \right],$$

где  $\chi$  – множество решений (15). Выбор точки таким образом дополнительно способствует сужению получаемых интервальных оценок.

Представленный в работе подход также применим, когда в один и тот же момент  $t_k$  известно несколько экспериментальных точек. В этом случае целевая функция (4) примет следующий вид:

$$J(\underline{y}_1^0, \overline{y}_1^0, \dots, \underline{y}_n^0, \overline{y}_n^0, \underline{\theta}_1, \overline{\theta}_1, \dots, \underline{\theta}_m, \overline{\theta}_m) = \sum_{k=1}^N \sum_{j=1}^{N_k} \rho(\mathbf{Y}^k, \hat{\mathbf{y}}^j(t_k)),$$

где  $N_k$  – количество экспериментальных точек в момент  $t_k$ ,  $\hat{\mathbf{y}}^j(t_k)$  –  $j$ -я экспериментальная точка.

Кроме того, рассматриваемый подход обобщается на случай, когда вместо экспериментальных точек известны интервальные оценки фазовых переменных:

$$\hat{\mathbf{Y}}^k = \left( [\underline{\hat{y}}_0^k, \overline{\hat{y}}_0^k], [\underline{\hat{y}}_1^k, \overline{\hat{y}}_1^k], \dots, [\underline{\hat{y}}_n^k, \overline{\hat{y}}_n^k] \right)^T, \quad k = \overline{1, N}.$$

Целевая функция (4) запишется следующим образом:

$$J(\underline{y}_1^0, \overline{y}_1^0, \dots, \underline{y}_n^0, \overline{y}_n^0, \underline{\theta}_1, \overline{\theta}_1, \dots, \underline{\theta}_m, \overline{\theta}_m) = \sum_{k=1}^N \rho(\mathbf{Y}^k, \hat{\mathbf{Y}}^k), \tag{16}$$

где  $\rho(\mathbf{Y}^k, \hat{\mathbf{Y}}^k)$  характеризует степень непересечения множеств  $\mathbf{Y}^k$  и  $\hat{\mathbf{Y}}^k$ . В частности, величина  $\rho(\mathbf{Y}^k, \hat{\mathbf{Y}}^k)$  может соответствовать расстоянию между множеством  $\mathbf{Y}^k$  и самой удалённой от него точкой множества  $\hat{\mathbf{Y}}^k$ :

$$\rho(\mathbf{Y}^k, \hat{\mathbf{Y}}^k) = \max_{\substack{\hat{y}_i^k \in [\underline{\hat{y}}_i^k, \overline{\hat{y}}_i^k] \\ i = \overline{1, n}}} \min_{\substack{x_i \in [y_i^0, \overline{y}_i^0], i = \overline{1, n} \\ z_i \in [\underline{\theta}_i, \overline{\theta}_i], i = \overline{1, m}}} \left[ \sum_{i=1}^n (y_i^k(x_1, \dots, x_n, z_1, \dots, z_m) - \hat{y}_i^k)^2 \right]$$

или, например, являться интегральной характеристикой:

$$\rho(\mathbf{Y}^k, \hat{\mathbf{Y}}^k) = \int_{\underline{\hat{y}}_1^k}^{\overline{\hat{y}}_1^k} \int_{\underline{\hat{y}}_2^k}^{\overline{\hat{y}}_2^k} \dots \int_{\underline{\hat{y}}_n^k}^{\overline{\hat{y}}_n^k} \min_{\substack{x_i \in [y_i^0, \overline{y}_i^0], i = \overline{1, n} \\ z_i \in [\underline{\theta}_i, \overline{\theta}_i], i = \overline{1, m}}} \left[ \sum_{i=1}^n (y_i^k(x_1, \dots, x_n, z_1, \dots, z_m) - \hat{y}_i^k)^2 \right] d\hat{y}_n^k d\hat{y}_{n-1}^k \dots d\hat{y}_1^k.$$

Решение задачи параметрической идентификации при наличии интервальных оценок фазовых переменных (минимизация (16)) требует применения специального математического аппарата. Данное направление является предметом дальнейших исследований.

**Заключение.** В работе представлен интервальный подход к решению задачи интервальной параметрической идентификации. Цель подхода заключается в нахождении такой интервальной оценки начальных условий и параметров системы ОДУ, при которой решение содержало бы экспериментальные данные или минимизировало бы отклонение от них. Приведена целевая функция, зависящая от границ соответствующих интервальных неопределённостей, минимизация которой эквивалентна решению исходной задачи параметрической идентификации. Выполнен анализ целевой функции и получено выражение градиента для дальнейшего применения методов оптимизации первого порядка. Сформулировано и доказано утверждение о том, что если минимум целевой функции достигается на определённых интервальных оценках, то и на более широких оценках он тоже будет достигаться. При вычислении значения целевой функции и градиента необходимо решать прямую интервальную задачу, для которой применяется ранее разработанный алгоритм адаптивной интерполяции. Выполнена апробация предложенного подхода на нескольких системах ОДУ, содержащих разное количество неизвестных параметров. Продемонстрирована эффективность подхода: для всех задач успешно получены интервальные оценки параметров, при которых решение содержит экспериментальные точки.

## СПИСОК ЛИТЕРАТУРЫ

1. Самарский А.А., Вабищевич П.Н. Численные методы решения обратных задач математической физики. М., 2009.
2. Денисов А.М. Введение в теорию обратных задач. М., 1994.
3. Марчук Г.И. Сопряженные уравнения и анализ сложных систем. М., 1992.
4. Ватутьян А.О. Математические модели и обратные задачи // Соросовский образоват. журн. 1998. № 11. С. 143–148.
5. Денисов А.М. Итерационный метод решения обратной коэффициентной задачи для гиперболического уравнения // Дифференц. уравнения. 2017. Т. 53. № 7. С. 943.
6. Гаврилов С.В., Денисов А.М. Численные методы решения нелинейного операторного уравнения, возникающего в обратной коэффициентной задаче // Дифференц. уравнения. 2021. Т. 57. № 7. С. 900–906.
7. Ашуров Р.Р., Мухиддинова А.Т. Обратная задача по определению плотности тепловых источников для уравнения субдиффузии // Дифференц. уравнения. 2020. Т. 56. № 12. С. 1596–1609.
8. Диллигенская А.Н., Самокиш А.В. Параметрическая идентификация в обратных задачах теплопроводности в условиях интервальной неопределённости на основе нейронных сетей // Вестн. Самарск. гос. техн. ун-та. 2020. Т. 28. № 4 (68). С. 6–18.
9. Петрикевич Я.И. Структурно-параметрическая идентификация динамических объектов по интервальным исходным данным: дис. ... канд. техн. наук. М., 2006.
10. Xiao N., Fedele F., Muhanna R.L. Inverse problems under uncertainties—an interval solution for the beam finite element // Conf. Paper: 11th Intern. Conf. on Structural Safety & Reliability. New York, 2013. (<https://www.researchgate.net/publication/269518192>).
11. Nenarokotov A.V., Alifanov O.M., Krainova I.V., Titov D.M., Morzhukhina A.V. Estimation of environmental influence on spacecraft materials radiative properties by inverse problems technique // Acta Astronautica. 2019. V. 160. P. 323–330.
12. Кабанихин С.И., Куликов И.М., Шишленин М.А. Алгоритм восстановления характеристик начального состояния сверхновой звезды // Журн. вычислит. математики и мат. физики. 2020. Т. 60. № 6. С. 1035–1044.
13. Абгарян К.К., Носков Р.Г., Ревизников Д.Л. Обратная коэффициентная задача теплопереноса в слоистых наноструктурах // Изв. вузов. Материалы электронной техники. 2017. Т. 20. № 3. С. 213–219.
14. Moore R. Interval Analysis. New Jersey, 1966.
15. Moore R.E., Kearfott R.B., Cloud M.J. Introduction to Interval Analysis. Philadelphia, 2009.
16. Шарый С.П. Конечномерный интервальный анализ. Новосибирск, 2019.
17. Морозов А.Ю., Ревизников Д.Л. Алгоритм адаптивной интерполяции на основе kd-дерева для численного интегрирования систем обыкновенных дифференциальных уравнений с интервальными начальными условиями // Дифференц. уравнения. 2018. Т. 54. № 7. С. 963–974.
18. Makino K., Berz M. Models and Their Applications // Numerical Software Verification 2017: conf. July 22–23, 2017. Germany, Heidelberg, 2017. P. 3–13.
19. Nataraj P.S.V., Sundur S. The extrapolated Taylor model // Reliable Computing. July, 2011. P. 251–278.
20. Розалев А.Н. Гарантированные методы решения систем обыкновенных дифференциальных уравнений на основе преобразования символьных формул // Вычислит. технологии. 2003. Т. 8. № 5. С. 102–116.
21. Fu C., Ren X., Yang Y.-F., Lu K., Qin W. Steady-state response analysis of cracked rotors with uncertain but bounded parameters using a polynomial surrogate method // Commun. Nonlin. Sci. Numer. Simul. 2019. V. 68. P. 240–256.
22. Fu C., Xu Y., Yang Y., Lu K., Gu F., Ball A. Response analysis of an accelerating unbalanced rotating system with both random and interval variables // J. Sound Vibrat. 2020. V. 466. P. 115047.
23. Морозов А.Ю., Ревизников Д.Л., Гидаспов В.Ю. Алгоритм адаптивной интерполяции на основе kd-дерева для решения задач химической кинетики с интервальными параметрами // Мат. моделирование. 2018. Т. 30. № 12. С. 129–144.
24. Морозов А.Ю., Журавлев А.А., Ревизников Д.Л. Анализ и оптимизация алгоритма адаптивной интерполяции численного решения систем обыкновенных дифференциальных уравнений с интервальными параметрами // Дифференц. уравнения. 2020. Т. 56. № 7. С. 960–974.
25. Гидаспов В.Ю., Морозов А.Ю., Ревизников Д.Л. Алгоритм адаптивной интерполяции с использованием ТТ-разложения для моделирования динамических систем с интервальными параметрами // Журн. вычислит. математики и мат. физики. 2021. Т. 61. № 9. С. 1416–1430.

26. *Morozov A.Yu., Zhuravlev A.A., Reviznikov D.L.* Sparse grid adaptive interpolation in problems of modeling dynamic systems with interval parameters // *Math.* 2021. V. 9. P. 298.
27. *Морозов А.Ю., Ревизников Д.Л.* Алгоритм адаптивной интерполяции на разреженных сетках для численного интегрирования систем обыкновенных дифференциальных уравнений с интервальными неопределённостями // *Дифференц. уравнения.* 2021. Т. 57. № 7. С. 976–987.
28. *Смоляк С.А.* Квадратурные и интерполяционные формулы на тензорных произведениях некоторых классов функций // *Докл. АН СССР.* 1963. Т. 148. № 5. С. 1042–1045.
29. *Bungartz H.-J., Griebel M.* Sparse grids // *Acta Numerica.* 2004. V. 13. № 1. P. 147–269.
30. *Gerstner T., Griebel M.* Sparse grids // *Encycl. of Quantitat. Finance / Ed. R. Cont.* New York, 2010.
31. *Oseledets I.V.* Tensor-train decomposition // *SIAM J. on Sci. Comput.* 2011. V. 33. № 5. P. 2295–2317.
32. *Oseledets I., Tyrtyshnikov E.* TT-cross approximation for multidimensional arrays // *Linear Algebra and its Appl.* 2010. V. 432. № 1. P. 70–88.
33. *Гилл Ф., Мюррей У., Райт М.* Практическая оптимизация. М., 1985.
34. *Пантелеев А.В., Летова Т.А.* Методы оптимизации в примерах и задачах. М., 2005.
35. *Арнольд В.И.* Обыкновенные дифференциальные уравнения. Ижевск, 2000.

Федеральный исследовательский центр  
“Информатика и управление” РАН, г. Москва,  
Московский авиационный институт  
(национальный исследовательский университет)

Поступила в редакцию 20.02.2022 г.  
После доработки 20.02.2022 г.  
Принята к публикации 25.05.2022 г.



УДК 519.63+517.958:532.5

## ВАЛИДАЦИЯ ВЫЧИСЛИТЕЛЬНОГО АЛГОРИТМА НА ОСНОВЕ РАЗРЫВНОГО МЕТОДА ГАЛЁРКИНА ДЛЯ РЕЛАКСАЦИОННОЙ МОДЕЛИ БАЕРА–НУНИЦАТО

© 2022 г. Р. Р. Полехина, М. В. Алексеев, Е. Б. Савенков

Для моделирования динамики двухфазных сред в рамках полностью неравновесной модели Баера–Нунциато с учётом жёсткой механической релаксации рассматривается разрывный метод Галёркина/Рунге–Кутты. Для монотонизации решения используются лимитер WENO-S, который применяется непосредственно к консервативным переменным модели. Релаксационные процессы моделируются с использованием неявного метода Гира шестого порядка с адаптивным выбором шага интегрирования. Для построения численного потока применяется подход на основе теории DLM, позволяющей обобщить методы гоудновского типа на случай неконсервативных гиперболических систем уравнений. С помощью разработанного метода проводится расчёт одномерных и двумерных задач, даётся анализ результатов расчётов. В двумерном случае данные расчётов сравниваются с результатами лабораторных экспериментов.

DOI: 10.31857/S0374064122070093, EDN: CEQMND

**Введение.** Настоящая работа посвящена численному решению двухфазной полностью неравновесной модели Баера–Нунциато (БН) с релаксационными слагаемыми. Данная модель впервые предложена в работе [1] для анализа процесса перехода дефлаграции в детонацию при моделировании динамики горения гранулированных взрывчатых веществ. В дальнейшем она применялась для решения широкого спектра задач и в настоящее время может рассматриваться как базовая модель для целого ряда обобщений [2–6].

Математическая формулировка модели приведена в п. 1 настоящей работы. Укажем здесь ряд её свойств, которые усложняют задачу построения вычислительных алгоритмов для её решения:

(а) система уравнений БН является гиперболической системой первого порядка и включает в себя члены как в дивергентной, так и в квазилинейной форме. Она не может быть записана полностью в дивергентном виде;

(б) модель включает в себя релаксационные члены, которые описывают процесс релаксации механических и термодинамических параметров фаз к равновесному значению. По крайней мере в ряде приложений характерные времена релаксации могут быть значительно меньше характерных времён протекания других газодинамических процессов. В этом смысле система уравнений БН является “жёсткой”.

В связи с этим численное решение уравнений БН является сложной задачей, решению которой посвящено значительное число работ (см., например, [7–11]). Большая их часть направлена на рассмотрение указанных выше трудностей.

Для построения аппроксимации задачи в настоящей работе используется разрывный метод Галёркина/Рунге–Кутты (RK/DG, Runge–Kutta/Discontinuous Galerkin method) [12]. Его использование предполагает построение слабой постановки задачи. Для неконсервативных систем такая постановка может быть построена с помощью так называемой теории DLM (DalMaso–LeFloch–Murat) [13]. Этот подход позволяет обобщить традиционные методы гоудновского типа на случай неконсервативных систем уравнений. Теория DLM основана на понятии пути – гладкого отображения, которое интерполирует между состояниями решения на разрыве и определяет структуры ударной волны. При этом различным путям соответствуют различные решения задачи Римана. В связи с этим выбор пути однозначно определяет вид решения задачи. Обоснование того или иного способа выбора пути для конкретной постановки задачи является открытым вопросом (см. [14]).

Целью настоящей работы является реализация и исследование возможностей вычислительного алгоритма решения уравнений модели БН, основанного на совместном применении сравнительно простых компонент, которые включают в себя:

- применение разрывного метода Галёркина для построения пространственных аппроксимаций уравнений модели;
- лимитирование консервативных переменных с использованием геометрического лимитера WENO-S;
- применение теории DLM для формулировки слабой постановки задачи;
- использование простейшего расщепления по физическим процессам для учёта жёстких релаксационных слагаемых;
- применение неявного метода Гира для интегрирования правых частей модели.

Эффективность предложенной методики иллюстрируется результатами численных экспериментов. Рассмотренные тесты соответствуют физической постановке задачи, в которой среда эффективно является двухфазной и гетерогенной с прямым разрешением границы раздела фаз. Распределение фаз описывается в рамках так называемой диффузной границы. В этом случае обе фазы формально присутствуют в каждой точке пространства, но в области однородности фазы одна из двух объёмных долей пренебрежимо мала. В рамках такой постановки рассматриваются задачи прохождения ударной волны из лёгкого вещества в тяжёлое (в одномерной постановке) и взаимодействия ударной волны с пузырями более тяжёлого и более лёгкого (по отношению к вмещающей среде) газа.

В первом случае сравнение результатов расчётов проводится с аналитическим решением в рамках модели Капицы [5]. Последняя является равновесной по скоростям и давлениям фаз и может быть получена как предел модели БН при исчезающе малых временах релаксации. Для двумерных тестов результаты расчётов сравниваются с результатами лабораторных экспериментов.

Работа имеет следующую структуру. Математическая модель и её особенности представлены в п. 1. Описанию и особенностям реализации вычислительного алгоритма посвящён п. 2. В п. 3 описаны результаты моделирования и проводится их анализ. В заключении формулируются основные результаты работы.

**1. Модель Баера–Нунциато с релаксацией.** Модель БН описывает двухфазную среду как совокупность двух взаимопроникающих континуумов, каждый из которых характеризуется своей скоростью и совокупностью термодинамических параметров состояния. Скорости и давления фаз в общем случае не равны между собой, и, таким образом, модель является полностью неравновесной. Система уравнений модели БН в варианте, приведённом в [15], имеет вид

$$\frac{\partial \alpha_k}{\partial t} + \mathbf{u}_I \cdot \nabla \alpha_k = \nu(P_k - P_{\bar{k}}), \quad (1a)$$

$$\frac{\partial \alpha_k \rho_k}{\partial t} + \nabla \cdot (\alpha_k \rho_k \mathbf{u}_k) = 0, \quad (1b)$$

$$\frac{\partial \alpha_k \rho_k \mathbf{u}_k}{\partial t} + \nabla \cdot (\alpha_k \rho_k \mathbf{u}_k \otimes \mathbf{u}_k) + \nabla(\alpha_k P_k) - P_I \nabla \alpha_k = \mu(\mathbf{u}_{\bar{k}} - \mathbf{u}_k), \quad (1c)$$

$$\frac{\partial \alpha_k \rho_k E_k}{\partial t} + \nabla \cdot (\alpha_k (\rho_k E_k + P_k) \mathbf{u}_k) - P_I \mathbf{u}_I \nabla \alpha_k = \mu(\mathbf{u}_{\bar{k}} - \mathbf{u}_k) \cdot \mathbf{u}_I + \nu(P_{\bar{k}} - P_k) P_I. \quad (1d)$$

Здесь  $\alpha_k$ ,  $\rho_k$ ,  $\mathbf{u}_k$ ,  $P_k$ ,  $E_k$  – объёмная доля, плотность, поле скоростей, давление и полная энергия фазы  $k = 1, 2$  соответственно;  $\bar{k} = \{1, 2\} \setminus \{k\}$ ; величины  $\nu$ ,  $\mu$  – релаксационные параметры. Для объёмных долей выполнено условие нормировки  $\alpha_1 + \alpha_2 = 1$ . Полная энергия  $k$ -й фазы определена равенством

$$E_k = \mathcal{U}_k + \mathbf{u}_k \cdot \mathbf{u}_k / 2,$$

где  $\mathcal{U}_k$  – внутренняя энергия фазы.

Уравнения модели включают в себя: уравнение динамики для объёмной доли (1a), законы сохранения массы (1b), импульса (1c) и энергии (1d).

Величины  $\mathbf{u}_1$  и  $P_1$  являются так называемыми *интерфейсными* скоростью и давлением и могут быть определены целым рядом способов (см. [1, 16–18]). В настоящей работе используется вариант, предложенный в оригинальной работе [1]:

$$\mathbf{u}_1 = \mathbf{u}_1, \quad P_1 = P_2, \tag{2}$$

где  $\mathbf{u}_1$  – скорость более сжимаемой фазы,  $P_2$  – давление менее сжимаемой фазы.

Термодинамические свойства фаз определяются уравнениями состояния (УрС) и имеют вид  $\mathcal{U}_k = \mathcal{U}_k(P_k, \rho_k)$ . В данной работе рассматриваются УрС двух видов. Первое – это хорошо известное двухчленное УрС

$$\mathcal{U}_k = \frac{P_k + \gamma_k P_{\infty,k}}{(\gamma_k - 1)\rho_k}, \tag{3}$$

где  $P_{\infty,k}$  и  $\gamma_k$  – референсное давление и показатель адиабаты. Вторым является УрС Ми–Грюнайзена [19, раздел 2.7.6, с. 81]

$$\mathcal{U}_k = \frac{P_k - P_k^x(\rho_k)}{(\gamma_k - 1)\rho_k} + \mathcal{U}_k^x(\rho_k), \quad \mathcal{U}_k^x = \int_{\rho_k^0}^{\rho_k} \frac{P_k^x(\rho_k)}{\rho_k^2} d\rho_k, \tag{4}$$

$$P_k^x(\rho_k) = A_{1,k} \left( \frac{\rho_k^0}{\rho_k} \right)^{-E_{1,k}} - A_{2,k} \left( \frac{\rho_k^0}{\rho_k} \right)^{-E_{2,k}},$$

где  $A_k$ ,  $E_k$  – параметры.

Правые части уравнений (1) описывают взаимодействие фаз и часто называются *релаксационными* в силу того, что они обеспечивают релаксацию скоростей и давлений фаз к одинаковому значению. Численные эксперименты, описанные в настоящей работе, соответствуют физической постановке с прямым разрешением межфазных границ. В постановках такого типа характерные времена релаксации много меньше характерных времён протекания других газодинамических процессов. В этом случае параметры  $\mu$  и  $\nu$  велики, и в этом смысле система уравнений (1) является “жесткой”.

Отметим, что система уравнений (1) содержит “неконсервативные произведения” вида  $P_1 \nabla \alpha_k$ ,  $P_1 \mathbf{u}_1 \nabla \alpha_k$  и не может быть записана в консервативной форме. Эта особенность является типичной для целого ряда многофазных многоскоростных моделей (см. [20]). Обобщённое решение и соответствующие соотношения Ранкина–Гюгонио (РГ) для неконсервативных систем уравнений позволяет сформулировать теория DLM [13, 21], которая в дальнейшем также используется для построения численной схемы.

Для системы уравнений (1) может быть построен её предел при  $\mu, \nu \rightarrow +\infty$ . Соответствующая модель является равновесной по скоростям и давлениям и носит название *модели Капилы* [5]. Уравнения модели приведены в приложении. В силу того, что далее рассматривается случай жесткой релаксации, соответствующий большим значениям  $\mu$  и  $\nu$ , представляется естественным сравнивать численные решения уравнений (1) с точным аналитическим решением модели Капилы. Точное решение задачи Римана о распаде разрыва для модели Капилы строится в приложении.

**2. Вычислительный алгоритм.** Использованный алгоритм численного решения системы уравнений (1) основан на применении следующих подходов: расщеплении по физическим процессам для расчёта релаксационных правых частей системы, разрывном методе Галёркина для решения однородной системы (1) и неявном многошаговом методе Гира [22, гл. 3, с. 331] для интегрирования “жестких” правых частей системы (1). Далее в настоящем пункте описаны основные компоненты полного алгоритма решения задачи.

**2.1. Разрывный метод Галёркина.** Рассмотрим трёхмерный вариант неоднородной нелинейной гиперболической системы уравнений, записанной в декартовой системе координат:

$$\frac{\partial \mathbf{Q}}{\partial t} + \sum_{\xi=1}^3 \frac{\partial \mathcal{F}_{\xi}}{\partial x_{\xi}} + \sum_{\xi=1}^3 \mathbf{B}_{\xi}(\mathbf{Q}) \frac{\partial \mathbf{Q}}{\partial x_{\xi}} = \mathbf{S}(\mathbf{Q}), \tag{5}$$

где  $\mathbf{x} = (x_1, x_2, x_3) \in \Omega = [0, L_1] \times [0, L_2] \times [0, L_3] \subset \mathbb{R}^3$ ,  $t \in [0, T] \subset \mathbb{R}^+$ ,  $\mathbf{Q} = \mathbf{Q}(\mathbf{x}, t) = (Q_1, \dots, Q_M) \in \Omega_q \subset \mathbb{R}^M$ ,  $M$  – общее число компонент вектора  $\mathbf{Q}$ ,  $\Omega_q$  – фазовое пространство вектора  $\mathbf{Q}$ ,  $\mathcal{F}_\xi$  – вектор потока,  $\mathcal{B}_\xi$  – матрица, отвечающая квазилинейной неконсервативной части уравнений,  $\mathcal{S}$  – источники, описывающие механическую релаксацию.

Введём разбиение  $\{\omega_{ijk}\}$  (декартова сетка) области  $\Omega$  и обозначим ячейку сетки (конечный элемент)  $\omega_{ijk} = [x_{1,i}, x_{1,i+1}] \times [x_{2,j}, x_{2,j+1}] \times [x_{3,k}, x_{3,k+1}]$ , где  $x_{1,i} = i\Delta x_1$  ( $0 \leq i \leq N_1$ ),  $x_{2,j} = j\Delta x_2$  ( $0 \leq j \leq N_2$ ),  $x_{3,k} = k\Delta x_3$  ( $0 \leq k \leq N_3$ ). Пусть индекс  $\alpha$  каким-либо образом упорядочивает тройки индексов  $(i, j, k)$ . Обозначим через  $\mathcal{V}_h^p(\Omega)$  подпространство элементов из  $\mathcal{L}^1(\Omega)$ , проекции которых на ячейки  $\omega_\alpha$  принадлежат векторному пространству  $\mathcal{P}^p(\omega_\alpha)$  полиномов степени  $p$ :

$$\mathcal{V}_h^p = \{v \in \mathcal{L}^1(\Omega) : v|_{\omega_\alpha} \in \mathcal{P}^p(\omega_\alpha), \quad 1 \leq \alpha \leq N\}, \quad N = N_1 N_2 N_3.$$

Представим решение  $\mathbf{Q}(\mathbf{x}, t)$  в ячейке  $\omega_\alpha$  конечномерной аппроксимацией  $\mathbf{Q}_h \in \mathcal{V}_h^p$ :

$$\mathbf{Q}_h(\mathbf{x}, t)|_{\omega_\alpha} = \sum_{l=0}^p \psi_\alpha^{(l)}(\mathbf{x}) \mathbf{Q}_\alpha^{(l)}(t), \tag{6}$$

где  $\psi_\alpha^{(l)}$  – полином Лежандра степени  $l$ . Для получения полудискретной системы уравнений относительно приближённого решения  $\mathbf{Q}_h(\mathbf{x}, t)$  подставим представление (6) в систему (5), затем умножим (5) на пробную функцию  $v_h \in \mathcal{V}_h^p$  и проинтегрируем получившееся тождество по ячейке  $\omega_\alpha$ :

$$\int_{\omega_\alpha} \frac{\partial \mathbf{Q}_h}{\partial t} v_h(\mathbf{x}) \, d\mathbf{x} + \sum_{\xi=1}^3 \int_{\omega_\alpha} \frac{\partial \mathcal{F}_\xi}{\partial x_\xi} v_h(\mathbf{x}) \, d\mathbf{x} + \sum_{\xi=1}^3 \int_{\omega_\alpha} \left[ \mathcal{B}_\xi(\mathbf{Q}_h) \frac{\partial \mathbf{Q}_h}{\partial x_\xi} \right]_\Psi v_h(\mathbf{x}) \, d\mathbf{x} = \int_{\omega_\alpha} \mathcal{S}(\mathbf{Q}_h) v_h(\mathbf{x}) \, d\mathbf{x}. \tag{7}$$

Для определения обобщённого решения (7) в случае неконсервативных систем ( $\mathcal{B}_\xi \neq 0$ ) в настоящей работе используется распространённый подход, основанный на применении теории DLM [13, 21]. В рамках этого подхода для корректного определения “неконсервативного произведения” вводится липшицево отображение  $\Psi : [0, 1] \times \Omega_q \times \Omega_q \rightarrow \Omega_q$  (путь), которое “соединяет” значения решения по разные стороны от разрыва в точке  $\mathbf{x}_d$  по направлению внешней единичной нормали  $\mathbf{n}$ ,  $\Psi(\mathbf{Q}^-, \mathbf{Q}^+; 0) = \mathbf{Q}^-$ ,  $\Psi(\mathbf{Q}^-, \mathbf{Q}^+; 1) = \mathbf{Q}^+$ ,  $\Psi(\mathbf{Q}, \mathbf{Q}; \eta) = \mathbf{Q}$ ,  $\mathbf{Q}^\pm = \lim_{h \rightarrow 0} \mathbf{Q}(\mathbf{x}_d \pm h\mathbf{n})$ .

Третье слагаемое в левой части системы (5) имеет вид “неконсервативного” произведения  $\mathcal{B}_\xi(\mathbf{Q}) \partial \mathbf{Q} / \partial x_\xi$ . Его значения в слабом смысле могут быть определены следующим образом [13]:

$$\begin{aligned} & \int_{\omega_\alpha} \left[ \mathcal{B}_\xi(\mathbf{Q}_h) \frac{\partial \mathbf{Q}_h}{\partial x_\xi} \right]_\Psi v_h(\mathbf{x}) \, d\mathbf{x} \stackrel{\text{def}}{=} \int_{\omega_\alpha^0} \mathcal{B}_\xi(\mathbf{Q}_h) \frac{\partial \mathbf{Q}_h}{\partial x_\xi} v_h(\mathbf{x}) \, d\mathbf{x} + \\ & + \int_{\partial \omega_\alpha} \left( \int_0^1 \mathcal{B}_\xi(\Psi(\mathbf{Q}^-, \mathbf{Q}^+; s)) \frac{\partial \Psi(\mathbf{Q}^-, \mathbf{Q}^+; \eta)}{\partial \eta} \, d\eta \right) v_h(\mathbf{x}) \, d\sigma, \end{aligned} \tag{8}$$

где  $\omega_\alpha^0 = \omega_\alpha \setminus \partial \omega_\alpha$ ,  $\mathbf{n}$  – внешняя единичная нормаль к  $\partial \omega_\alpha$ . Заметим, что данное определение существенно зависит от выбранного пути  $\Psi$ . В дальнейшем используется линейный путь  $\Psi(\mathbf{Q}^-, \mathbf{Q}^+; \eta) = \mathbf{Q}^- + \eta(\mathbf{Q}^+ - \mathbf{Q}^-)$ . Выбор пути  $\Psi$  является важным компонентом численного алгоритма – он определяет решение задачи Римана системы (1) (см. [13, 21]).

На основе сделанных выше построений численная схема для разрывного метода Галёркина может быть записана в виде

$$\int_{\omega_\alpha} \frac{\partial \mathbf{Q}_h(\mathbf{x}, t)}{\partial t} v_h(\mathbf{x}) \, d\mathbf{x} - \sum_{\xi=1}^3 \int_{\omega_\alpha} \mathcal{F}_\xi \frac{\partial v_h(\mathbf{x})}{\partial x_\xi} \, d\mathbf{x} + \sum_{\xi=1}^3 \int_{\omega_\alpha^0} \mathcal{B}(\mathbf{Q}_h) \frac{\partial \mathbf{Q}_h}{\partial x_\xi} v_h(\mathbf{x}) \, d\mathbf{x} +$$

$$+ \int_{\partial\omega_\alpha} v_h(\mathbf{x}) D_n(\mathbf{x}) d\sigma = \int_{\omega_\alpha} \mathcal{S}(\mathbf{Q}_h) v_h(\mathbf{x}) d\mathbf{x}. \tag{9}$$

В настоящей работе в качестве численного потока используется поток HLLEM [8]:

$$D_n = \frac{s_{out}\mathcal{F}(\mathbf{Q}_{in}) - s_{in}\mathcal{F}(\mathbf{Q}_{out})}{s_{out} - s_{in}} \mathbf{n} + \kappa \frac{s_{out}s_{in}}{s_{out} - s_{in}} (\mathbf{Q}_{out} - \mathbf{Q}_{in}) - \kappa \frac{s_{in}}{s_{in} - s_{out}} [\tilde{B}(\mathbf{Q}_{out}, \mathbf{Q}_*)(\mathbf{Q}_* - \mathbf{Q}_{out}) + \tilde{B}(\mathbf{Q}_*, \mathbf{Q}_{in})(\mathbf{Q}_{in} - \mathbf{Q}_*)],$$

где  $\kappa = n_1 + n_2 + n_3$ ,  $\mathbf{Q}_{in} \in \omega_\alpha$ ,  $\mathbf{Q}_{out} \in \omega_\beta$ , где  $\omega_\beta$  – ячейка, смежная к  $\omega_\alpha$  через элемент её поверхности. Скорость распространения волн в задаче Римана оценивается следующим образом:

$$s_{in/out} = \begin{cases} \min / \max(0, \Lambda(\mathbf{Q}_{in/out}), \Lambda(\hat{\mathbf{Q}})), & \text{если } \kappa > 0, \\ \max / \min(0, \Lambda(\mathbf{Q}_{in/out}), \Lambda(\hat{\mathbf{Q}})) & \text{в противном случае,} \end{cases} \quad \hat{\mathbf{Q}} = 2^{-1}(\mathbf{Q}_{in} + \mathbf{Q}_{out}).$$

Рассматривая произвольное  $v_h \in \mathcal{V}_h^p$ ,  $v_h(x)|_{\omega_\alpha} \in \text{span}\{\psi_\alpha^{(l)}\}$ , получаем следующую полудискретную систему уравнений относительно переменных  $\{\mathbf{Q}_\alpha^{(l)}\}$ , определённых в (6):

$$\mathbf{M} \frac{d\tilde{\mathbf{Q}}_\alpha}{dt} = \mathbf{H}(\tilde{\mathbf{Q}}_\alpha) + \mathbf{I}(\tilde{\mathbf{Q}}_\alpha). \tag{10}$$

Здесь  $\tilde{\mathbf{Q}}_\alpha = (Q_{0,\alpha}^{(0)}, \dots, Q_{0,\alpha}^{(p)}, Q_{1,\alpha}^{(0)}, \dots, Q_{M,\alpha}^{(0)}, \dots, Q_{M,\alpha}^{(p)})$  – вектор неизвестных. Его компоненты  $Q_{m,\alpha}^{(l)}$  в дальнейшем будем называть  $l$ -й гармоникой ( $l = \overline{0, p}$ )  $m$ -й компоненты вектора  $\mathbf{Q}_\alpha$ . Матрица  $\mathbf{M} \in \mathbb{R}^{(p+1) \times (p+1)}$  является матрицей Грама с компонентами  $[\mathbf{M}]_{ml} = \int_{\omega_\alpha} \psi_\alpha^{(l)} \psi_\alpha^{(m)} d\mathbf{x}$ , вектор  $\mathbf{H}(\tilde{\mathbf{Q}}_\alpha)$  – аппроксимация дифференциального оператора в левой части системы (5), вектор  $\mathbf{I}(\tilde{\mathbf{Q}}_\alpha)$  соответствует аппроксимации правой части системы (5):

$$\mathbf{I}_m^{(l)} = \int_{\omega_\alpha} \mathcal{S}_m(\mathbf{Q}_h) \psi_\alpha^{(l)} d\mathbf{x}. \tag{11}$$

Одна из наиболее простых стратегий решения системы (10) – расщепление по процессам первого порядка. Разобьём временной интервал на части точками  $\{t^n\}$ . Далее на каждом временном шаге  $t \in [t^n, t^{n+1}]$  сначала решается задача Коши для системы обыкновенных дифференциальных уравнений (ОДУ)

$$\mathbf{M} \frac{d\tilde{\mathbf{Q}}_\alpha}{dt} = \mathbf{H}(\tilde{\mathbf{Q}}_\alpha), \quad t \in (t^n, t^{n+1}], \tag{12}$$

с начальными данными  $\tilde{\mathbf{Q}}_\alpha^0 = \tilde{\mathbf{Q}}_\alpha(t^n)$  для определения “промежуточного” решения  $\tilde{\mathbf{Q}}_{\alpha, \text{H}} = \tilde{\mathbf{Q}}_\alpha(t^{n+1})$ . Затем решается задача Коши для системы ОДУ

$$\mathbf{M} \frac{d\tilde{\mathbf{Q}}_\alpha}{dt} = \mathbf{I}(\tilde{\mathbf{Q}}_\alpha), \quad t \in (t^n, t^{n+1}], \tag{13}$$

с начальными данными  $\tilde{\mathbf{Q}}_{\alpha, \text{H}}$ . Её решение  $\tilde{\mathbf{Q}}_\alpha$  рассматривается как решение задачи (10).

Для интегрирования по времени однородной системы (12) далее используется вариант метода Рунге–Кутты TVD/RK3 [12] с лимитированием консервативных переменных на каждом шаге метода. В настоящей работе используется многостадийный лимитер на основе лимитера WENO-S и дополнительных лимитирующих процедур, обеспечивающих физическую корректность решения – положительность давления, внутренней энергии и объёмных долей фаз.

Процедура лимитирования будет описана в п. 2.3. Для интегрирования по пространству используется метод квадратур Гаусса–Лежандра.

Для интегрирования системы уравнений (13) используется неявный метод Гира с автоматическим выбором шага интегрирования, описанный в п. 2.2.

**2.2. Алгоритм расчёта релаксационных слагаемых.** Задача Коши для ОДУ (13) решается неявным линейным многошаговым методом Гира [22, гл. 3, с. 331]. на временном интервале  $t \in (t^n, t^{n+1}]$ . Разобьём временной интервал на части точками  $t_0^n, t_1^n, \dots, t_{\chi-1}^n$ , где  $t_{j+1}^n = t_j^n + \tau_j$ ,  $t_0^n = t_n$ ,  $t_{\chi-1}^n = t_{n+1}$ . Таким образом,  $\Delta t = t^{n+1} - t^n$  – шаг интегрирования газодинамической части задачи, а  $\tau_j$  – шаг интегрирования релаксационной части. Разностная схема решения задачи Коши для (13) имеет следующий вид:

$$\tilde{Q}_\alpha^0 = \tilde{Q}_\alpha(t^n); \quad \sum_{k=0}^K \zeta_k \tilde{Q}_\alpha^{j+1-k} = \tau_j \beta_0 \mathbf{R}(\tilde{Q}_\alpha^{j+1}), \quad j = 0, 1, \dots \tag{14}$$

Здесь  $\tau_j = \text{const}$ ,  $\mathbf{R} = \mathbf{M}^{-1}\mathbf{I}$ . Аппроксимации релаксационных членов  $\mathbf{I}$ , определённых в (11), рассчитываются численно методом квадратур Гаусса–Лежандра.

В дальнейшем все численные расчёты проводятся методом Гира 6-го порядка ( $K = 6$ ). Коэффициенты метода представлены в табл. 1. Для начала счёта требуется знать значения  $\tilde{Q}_\alpha^j$ ,  $j = 1, K - 1$ . Для этого рекуррентно используется метод Гира, т.е. сначала для нахождения  $\tilde{Q}_\alpha^1$  используется метод Гира первого порядка, затем второго для  $\tilde{Q}_\alpha^2$  и т.д.

Таблица 1. Коэффициенты метода Гира до 6-го порядка,  $\zeta_0 = 1$

$K$	$\beta$	$\zeta_1$	$\zeta_2$	$\zeta_3$	$\zeta_4$	$\zeta_5$	$\zeta_6$
1	1	-1					
2	2/3	-4/3	1/3				
3	6/11	-18/11	9/11	-2/11			
4	12/25	-48/25	36/25	-16/25	3/25		
5	60/137	-300/137	300/137	-200/137	75/137	-12/137	
6	60/147	-360/147	450/147	-400/147	225/147	-72/147	10/147

Нелинейная система уравнений (14) решается численно методом Ньютона:

$$\left. \frac{\partial \mathbf{F}(\tilde{Q}_\alpha^{j+1})}{\partial \tilde{Q}_\alpha^{j+1}} \right|_{\tilde{Q}_\alpha^{j+1,s}} (\tilde{Q}_\alpha^{j+1,s+1} - \tilde{Q}_\alpha^{j+1,s}) = -\mathbf{F}(\tilde{Q}_\alpha^{j+1,s}), \tag{15}$$

где индекс  $s$  обозначает номер итерации,  $\tilde{Q}_\alpha^{j+1,0} = \tilde{Q}_\alpha^j$  соответствует  $s = 0$ . Функция  $\mathbf{F}$  имеет следующий вид:

$$\mathbf{F}(\tilde{Q}_\alpha^{j+1}) = \zeta_0 \tilde{Q}_\alpha^{j+1} - \tau_j \beta_0 \mathbf{R}(\tilde{Q}_\alpha^{j+1}) - \sum_{k=1}^K \zeta_p \tilde{Q}_\alpha^{j+1-k}.$$

Матрица Якоби  $\partial \mathbf{F}(\tilde{Q}_\alpha^{j+1}) / \partial \tilde{Q}_\alpha^{j+1}$  в (15) вычисляется с применением численного дифференцирования или аналитически (для каждого численного расчёта ниже это указывается отдельно). Итерации метода Ньютона продолжаются до тех пор, пока величина  $r = \max_m |r_m|$  не станет меньше заданного значения. Здесь  $r_m$  определены следующим образом:

$$r_m = \begin{cases} \frac{\tilde{Q}_{\alpha,m}^{j+1,s+1} - \tilde{Q}_{\alpha,m}^{j+1,s}}{\tilde{Q}_{\alpha,m}^{j+1,s+1}}, & \text{если } |\tilde{Q}_{\alpha,m}^{j+1,s+1}| > 1, \\ \tilde{Q}_{\alpha,m}^{j+1,s+1} - \tilde{Q}_{\alpha,m}^{j+1,s} & \text{в противном случае.} \end{cases} \tag{16}$$

Шаг  $\tau_j$  может быть переменным и определяется адаптивно. В начальный момент  $\tau_0 = \Delta t/6$ , где  $\Delta t$  – шаг интегрирования задачи (12). В случае если превышено максимальное

число  $N_{\max}$  ньютоновских итераций (далее  $N_{\max} = 10$ ) или если решение сходится к нефизическим результатам (отрицательное давление или объёмная доля и т.д.), то шаг интегрирования уменьшается. В случае когда число итераций меньше минимального значения  $N_{\min}$  (далее  $N_{\min} = 3$ ), шаг интегрирования увеличивается. При этом если шаг интегрирования меняется, то необходимо вернуться к первому шагу метода Гира и рекуррентно пересчитать значения  $\tilde{Q}_\alpha^j$ ,  $j = \overline{1, K-1}$ , для нового значения временного шага  $\tau$ .

Коэффициенты релаксации  $\mu$  и  $\nu$  в системе (13) либо задаются постоянными, либо рассчитываются в каждой точке пространства динамически таким образом, чтобы фазовые давления и скорости релаксировали за один газодинамический временной шаг  $\Delta t$ . Соответствующая оценка строится следующим образом. Релаксация давлений и скоростей в случае межфазных соотношений (2) и УрС (3) может быть записана в виде

$$\begin{aligned} \frac{d(\mathbf{u}_1 - \mathbf{u}_2)}{dt} &= -\mu(\mathbf{u}_1 - \mathbf{u}_2) \left( \frac{1}{\alpha_1 \rho_1} + \frac{1}{\alpha_2 \rho_2} \right), \\ \frac{d(P_1 - P_2)}{dt} &= -\nu(P_1 - P_2) \left( \frac{P_1 + C_{1a}P_1 + C_{1b}}{\alpha_1 C_{1a}} + \frac{P_1 + C_{2a}P_1 + C_{2b}}{\alpha_2 C_{2a}} \right), \end{aligned} \quad (17)$$

где  $C_{ka} = 1/(\gamma_k - 1)$ ,  $C_{kb} = \gamma_k P_{\infty,k}/(\gamma_k - 1)$ . Здесь также предполагается, что релаксация скоростей не влияет на релаксацию давлений. Другими словами, в первом уравнении в (17) фиксировано давление. Тогда коэффициенты релаксации можно оценить следующим образом:

$$\mu = \frac{1}{\Delta t} \left( \frac{1}{\alpha_1 \rho_1} + \frac{1}{\alpha_2 \rho_2} \right)^{-1}, \quad \nu = \frac{1}{\Delta t} \left( \frac{P_1 + C_{1a}P_1 + C_{1b}}{\alpha_1 C_{1a}} + \frac{P_1 + C_{2a}P_1 + C_{2b}}{\alpha_2 C_{2a}} \right)^{-1}. \quad (18)$$

Такой выбор значений  $\mu$  и  $\nu$  обеспечивает релаксацию за один временной шаг  $\Delta t$ .

**2.3. Алгоритм монотонизации решения.** В качестве основного ограничителя для монотонизации решения используется ограничитель WENO-S, предложенный в работе [23]. Однако после его использования полученные значения компонент решения могут быть нефизическими – особенно вблизи зоны резкого изменения объёмной доли. Для решения этой проблемы используется дополнительное лимитирование решения. Общая схема лимитирования выглядит следующим образом.

1. Применение ограничителя, обеспечивающего положительность объёмных долей и плотностей фаз. Данный ограничитель обозначим через  $\Lambda_1$  [24].
2. Применение ограничителя WENO-S (обозначим его через  $\Lambda_2$ ).
3. Применение ограничителя, обеспечивающего положительность давлений фаз (обозначим его через  $\Lambda_3$ ).

Рассмотрим подробнее каждый из этапов лимитирования.

**Ограничитель  $\Lambda_1$ .** Пусть имеется вектор  $\mathbf{Q}(\mathbf{x}, t)$  с компонентами  $Q_m(\mathbf{x}, t)$ . В ячейке  $\omega_\alpha$  согласно представлению (6) величины  $Q_m(\mathbf{x}, t)$  имеют вид

$$(Q_m)_h(\mathbf{x}, t)|_{\omega_\alpha} = \sum_{l=1}^k \psi_\alpha^{(l)}(\mathbf{x}) Q_{m,\alpha}^{(l)}(t), \quad m = \overline{1, M}.$$

Введём среднее значение  $Q_m(\mathbf{x}, t)$  равенством

$$\bar{Q}_m = \frac{1}{|\omega_\alpha|} \int_{\omega_\alpha} (Q_m)_h(\mathbf{x}, t)|_{\omega_\alpha} d\mathbf{x}.$$

Процесс лимитирования сводится к пересчёту всех величин по следующей формуле:

$$(\tilde{Q}_m)_h(\mathbf{x}, t) = \bar{Q}_m + \theta((Q_m)_h(\mathbf{x}, t) - \bar{Q}_m), \quad (19)$$

где  $\theta$  – параметр. Если область значений  $Q_m$  лежит на интервале  $(a, b)$ , то для этой величины рассчитывается параметр  $\theta_m$  по формуле

$$\theta_m = \min \left\{ \frac{\bar{Q}_m - (a + \varepsilon)}{\bar{Q}_m - (Q_m)_{\min}}, \frac{\bar{Q}_m - (b - \varepsilon)}{\bar{Q}_m - (Q_m)_{\max}}, 1 \right\}.$$

Если значения  $Q_m$  определены на  $(a, \infty)$ , то

$$\theta_m = \min \left\{ \frac{\bar{Q}_m - (a + \varepsilon)}{\bar{Q}_m - (Q_m)_{\min}}, 1 \right\},$$

где

$$(Q_m)_{\min} = \min_{\omega_\alpha} (Q_m)_h(\mathbf{x}, t)|_{\omega_\alpha} = \min_{\mathbf{x}_{gp}} (Q_m)_h(\mathbf{x}_{gp}, t)|_{\omega_\alpha},$$

$$(Q_m)_{\max} = \max_{\omega_\alpha} (Q_m)_h(\mathbf{x}, t)|_{\omega_\alpha} = \max_{\mathbf{x}_{gp}} (Q_m)_h(\mathbf{x}_{gp}, t)|_{\omega_\alpha},$$

$\mathbf{x}_{gp}$  – квадратурные точки конечного элемента  $\omega_\alpha$ ,  $\varepsilon > 0$  – малая величина (в данной работе  $\varepsilon = 10^{-6}$ ).

Параметр  $\theta_m$  рассчитывается для величин  $\{\alpha_1, \alpha_2, \alpha_1\rho_1, \alpha_2\rho_2\}$ , при этом значения  $\alpha_1$  и  $\alpha_2$  определены на  $(0, 1)$ , значения  $\rho_1$  и  $\rho_2$  определены на  $(0, \infty)$ . Общий коэффициент лимитирования рассчитывается по формуле

$$\theta = \min_m \{\theta_m\}.$$

Затем операция лимитирования  $\Lambda_1$  с общим коэффициентом  $\theta$  применяется ко всем компонентам вектора  $\mathbf{Q}$  по формуле (19).

**Ограничитель  $\Lambda_2$ .** Ограничитель WENO-S применяется в два шага.

*Шаг 1.* Идентификация ячеек, в которых решение подлежит лимитированию. В данной работе используется TVB-индикатор [23].

*Шаг 2.* Применение ограничителя WENO-S для реконструкции решения в отмеченных ячейках.

Рассмотрим лимитер WENO-S. Сначала опишем его одномерный вариант по направлению оси  $Ox_1$ . Рассмотрим ячейки  $\omega_{ijk}$ , где  $i = i_0 - 1, i_0, i_0 + 1$ ,  $j = j_0$ ,  $k = k_0$ , а индекс  $\alpha$  их локально упорядочивает как  $\alpha = -1, 0, 1$  соответственно. Реконструкция решения в центральной ячейке  $\alpha = 0$  происходит в соответствии с выражением

$$Q_{m,0}^{\text{new}}(\mathbf{x}) = \kappa_{-1} Q_{m,-1}^{\text{mod}}(\mathbf{x}) + \kappa_0 Q_{m,0}(\mathbf{x}) + \kappa_1 Q_{m,1}^{\text{mod}}(\mathbf{x}),$$

где

$$\kappa_j = \bar{\kappa}_j \left( \sum_{n=-1,0,1} \bar{\kappa}_n \right)^{-1}, \quad \bar{\kappa}_n = \frac{\gamma_n}{(\varepsilon + \beta_n)^r}, \quad n = -1, 0, 1.$$

Здесь  $\gamma_n$  – линейный вес,  $\varepsilon = 10^{-6}$  и  $r = 2$ . Линейные веса должны удовлетворять следующим требованиям:

$$\gamma_0 \gg \gamma_{\pm 1}, \quad \gamma_{-1} + \gamma_0 + \gamma_{+1} = 1.$$

В настоящей работе  $\gamma_0 = 0.998$ ,  $\gamma_{\pm 1} = 0.001$ . Индикатор гладкости  $\beta_\alpha$  рассчитывается по формуле

$$\beta_\alpha = \sum_{l=1}^p \int_{\omega_\alpha} \Delta x_1^{2l-1} \left( \frac{\partial^l}{\partial x_1^l} Q_{m,\alpha}(\mathbf{x}) \right)^2 dx_1, \quad \alpha = -1, 0, 1.$$

Здесь  $p$  – степень полинома  $Q_{m,\alpha}(\mathbf{x})$ .

Применение ограничителя может быть записано в операторном виде как  $Q_0^{\text{new}} = \Lambda_{2,\xi} Q_0$  для направления вдоль оси  $Ox_{\text{new}}\xi$ . Для многомерного случая и декартовых сеток  $Q_0^{\text{new}} = \Lambda_2 Q_0$ , где  $\Lambda_2 = \Lambda_{2,3} \Lambda_{2,2} \Lambda_{2,1}$ .



**Ограничитель  $\Lambda_3$ .** На данном этапе лимитирования в каждой квадратурной точке ячейки  $\omega_\alpha$  вычисляются давления фаз. Если давление какой-либо фазы отрицательно хотя бы в одной точке, то в этой ячейке оставляются только средние величины консервативных переменных.

**3. Вычислительные эксперименты.** Тесты, рассмотренные ниже, соответствуют физической постановке с прямым разрешением межфазных границ. Напомним, что в постановках такого типа предполагается, что все фазы одновременно присутствуют в каждой точке пространства, но объёмные доли всех фаз, кроме одной, пренебрежимо малы всюду за исключением “диффузных границ” – относительно тонких пространственных областей, соответствующих границам раздела фаз среды.

Тесты 1 и 2, предложенные в [19] (разделы 2.7.5 и 2.7.6 соответственно), являются соответственно одномерной задачей Римана о распаде разрыва и задачей о прохождении ударной волны (УВ) через контактную границу двух веществ при больших соотношениях плотностей (из лёгкого в тяжёлое или из тяжёлого в лёгкое) и разных УрС. Эти тесты допускают аналитическое решение в рамках модели Капилы (см. приложение).

Тесты 3–5 исследуют в рамках пространственно двумерной постановки взаимодействие УВ с пузырьком криптона, азота или гелия в воздухе. Полученные численные расчёты сравниваются с экспериментальными данными [25], а также с численными расчётами, полученными другими авторами [26].

Численные эксперименты проводились с шагом интегрирования, рассчитанным из условия Куранта:

$$\Delta t = \text{CFL} \frac{\min_{\xi,k}(\Delta x_{\xi,k})}{\lambda_{\max}}, \tag{20}$$

где  $\text{CFL} = 0.9$  для  $p = 0$  и  $\text{CFL} = 0.2$  для  $p = 1$  (здесь  $p$  – порядок полинома Лежандра в (6)),  $\lambda_{\max}$  – максимальное собственное значение системы уравнений (1).

**Тест 1.** Тест характеризуется тем, что распад разрыва происходит на контактной границе при больших соотношениях давлений ( $\times 1000$ ). Первая фаза представляет собой воду, а вторая – воздух ( $\rho_1/\rho_2 \approx 20$ ). Начальные данные приведены в табл. 2. Контактный разрыв (КР) в начальный момент времени задан в точке  $x_0$  для всех переменных. Коэффициенты релаксации оцениваются по формуле (18). Численные эксперименты проводились в расчётной области  $[0, 1]$ .

**Таблица 2.** Начальные данные 1D расчётов

Тест	Начальные данные		Тип УрС	Параметры
	$x < x_0$	$x > x_0$		
1	$P_1 = 10^9, P_2 = 10^9,$ $\rho_1 = 1000, \rho_2 = 50,$ $u_1 = 0, u_2 = 0,$ $\alpha_1 = 1 - 10^{-4}$	$P_1 = 10^6, P_2 = 10^6,$ $\rho_1 = 1000, \rho_2 = 50,$ $u_1 = 0, u_2 = 0,$ $\alpha_1 = 10^{-4}$	УрС (3)	$\gamma_1 = 4.4, \pi_1 = 6 \cdot 10^8,$ $\gamma_2 = 1.4, \pi_2 = 0,$ $x_0 = 0.7$
2	$P_1 = P_2 = 10^{-5},$ $\rho_1 = 0.125, \rho_2 = 7.82,$ $u_1 = u_2 = 0$ $\alpha_1 = 1 - 10^{-4}$	$P_1 = P_2 = 10^{-5},$ $\rho_1 = 0.125, \rho_2 = 7.82,$ $u_1 = u_2 = 0,$ $\alpha_1 = 10^{-4}$	фаза 1: УрС (3) фаза 2: УрС (4)	$\gamma_1 = 1.4, \pi_1 = 0,$ $\gamma_2 = 4.54777,$ $A_{1,2} = A_{2,2} = 62.6,$ $E_{1,2} = 4, E_{2,2} = 1,$ $x_0 = 0.4$

Удельная внутренняя энергия  $e$ , представленная на рис. 1 и рис. 2, рассчитывается следующим образом:

$$e = \frac{U_1 \alpha_1 \rho_1 + U_2 \alpha_2 \rho_2}{\alpha_1 \rho_1 + \alpha_2 \rho_2}.$$

На рис. 1 показано сравнение результатов численных расчётов для  $N = 2500$  с аналитическим решением, полученным в приложении, на момент времени  $t = 2.2 \cdot 10^{-4}$  с (см. табл. 2). Видно, что численные результаты достаточно хорошо воспроизводят аналитическое решение. На кривой давления результаты с  $p = 1$  менее диссипативны, чем результаты для  $p = 0$ .

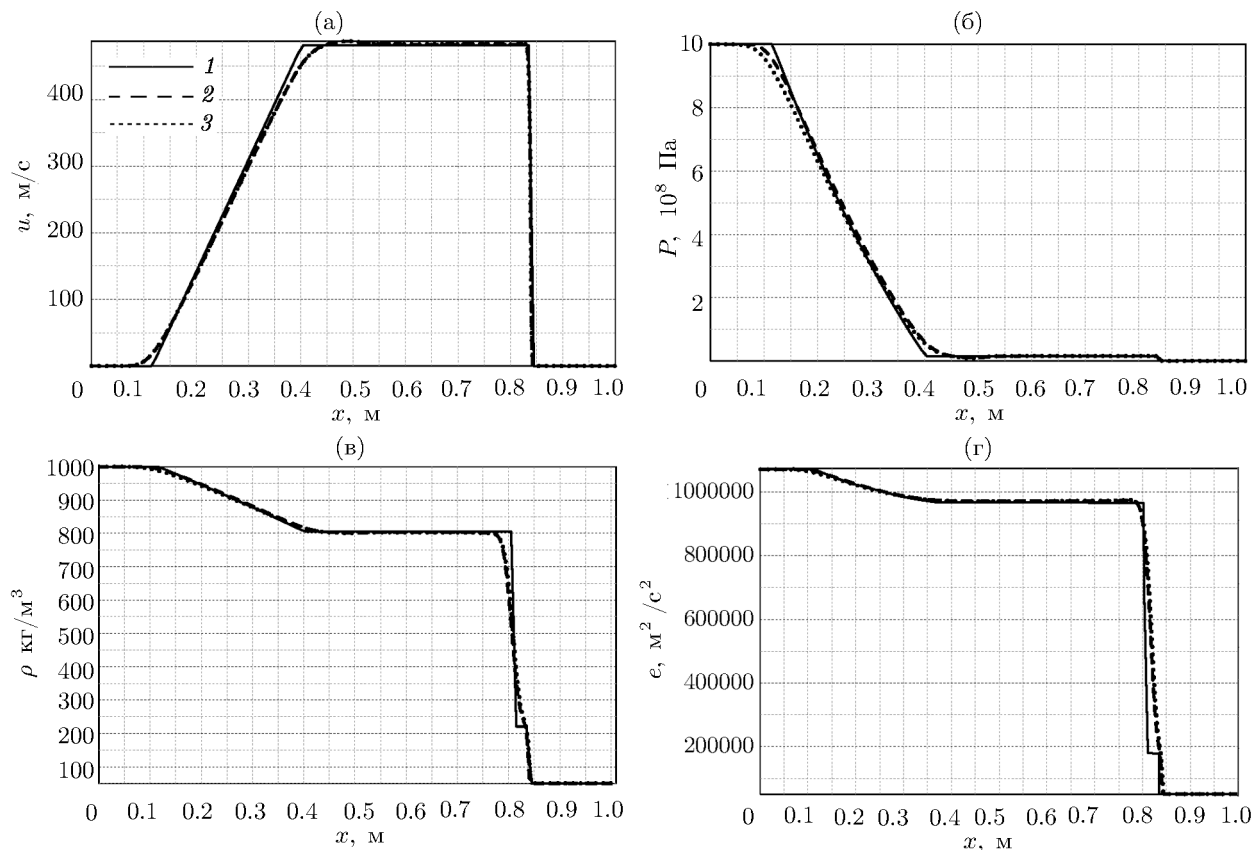


Рис. 1. Распределения для всех переменных в тесте 1 (1 – аналитическое решение, 2 – численный расчёт с  $p = 1$ , 3 – численный расчёт с  $p = 0$ ).

**Тест 2.** Тест характеризуется тем, что УВ формируется в первой фазе, менее плотной, а затем проходит через КР из лёгкого вещества в тяжёлое. Таким образом, в отличие от предыдущего теста положения КР и скачка давления не совпадают в начальный момент времени. В начальный момент времени КР задан в точке  $x_0$  (см. табл. 2). Отношение плотностей сред  $\rho_2/\rho_1 \approx 60$ . Начальные данные представлены в табл. 2. На левой границе задаются следующие краевые условия:

$$\alpha_2 = 10^{-4}, \quad \alpha_1 = 1 - \alpha_2, \quad \alpha_1 \rho_1 = 0.1249875, \quad \alpha_2 \rho_2 = 0.000782,$$

$$\alpha_1 \rho_1 u_1 = 0.288596, \quad \alpha_2 \rho_2 u_2 = 0.0018.$$

На правой границе устанавливаются условия симметрии. Численные эксперименты проводились в расчётной области  $[-1.4, 1]$ , а результаты расчётов представлены для расчётной области  $[0, 1]$ . Коэффициенты релаксации при расчёте на временном интервале  $(t^n, t^{n+1}]$  задаются равными  $\lambda = 5 \cdot 10^9$   $\text{кг}/(\text{м} \cdot \text{с})$ ,  $\nu_k^{(n)} = \alpha_{1,k}^{(n)} \alpha_{2,k}^{(n)} \cdot 10^6$   $\text{Па}^{-1} \cdot \text{с}^{-1}$ , где  $\alpha_{i,k}^p$  – значение объёмной доли фазы  $i = 1, 2$  в ячейке с номером  $k$  на временном слое  $p$ .

Результаты численных экспериментов представлены на рис. 2 на момент времени  $t = 0.606323$  с для  $N = 5000$  ( $p = 0, 1$ ),  $N = 40000$  ( $p = 0$ ) и  $N = 160000$  ( $p = 0$ ). Обратим внимание на пик в окрестности УВ, который возникает в результате взаимодействия УВ с КР. Видно, что численное решение сходится к аналитическому решению, однако очень медленно. Такие результаты связаны с наличием в системе (1) “неконсервативных произведений”.

Проблемы со сходимостью численного решения неконсервативной системы уравнений к обобщённому решению хорошо известны [7, 8] и подробно обсуждаются в работе [14]. Связано это с тем, что в отличие от консервативных систем соотношения РГ для неконсервативных систем зависят от структуры разрыва, т.е. от производных высших порядков [14]. Задавая путь в теории DLM, мы неявно задаём уравнение для определения структуры разрыва. С другой

стороны, первое дифференциальное приближение разностной схемы, описанной в п. 2, зависит не только от заданного пути, но и от численной вязкости. Таким образом, уравнение для определения структуры разрыва, которое используется для получения аналитического решения, не совпадает с первым дифференциальным приближением численной схемы. Системное решение этой проблемы в настоящее время неизвестно.

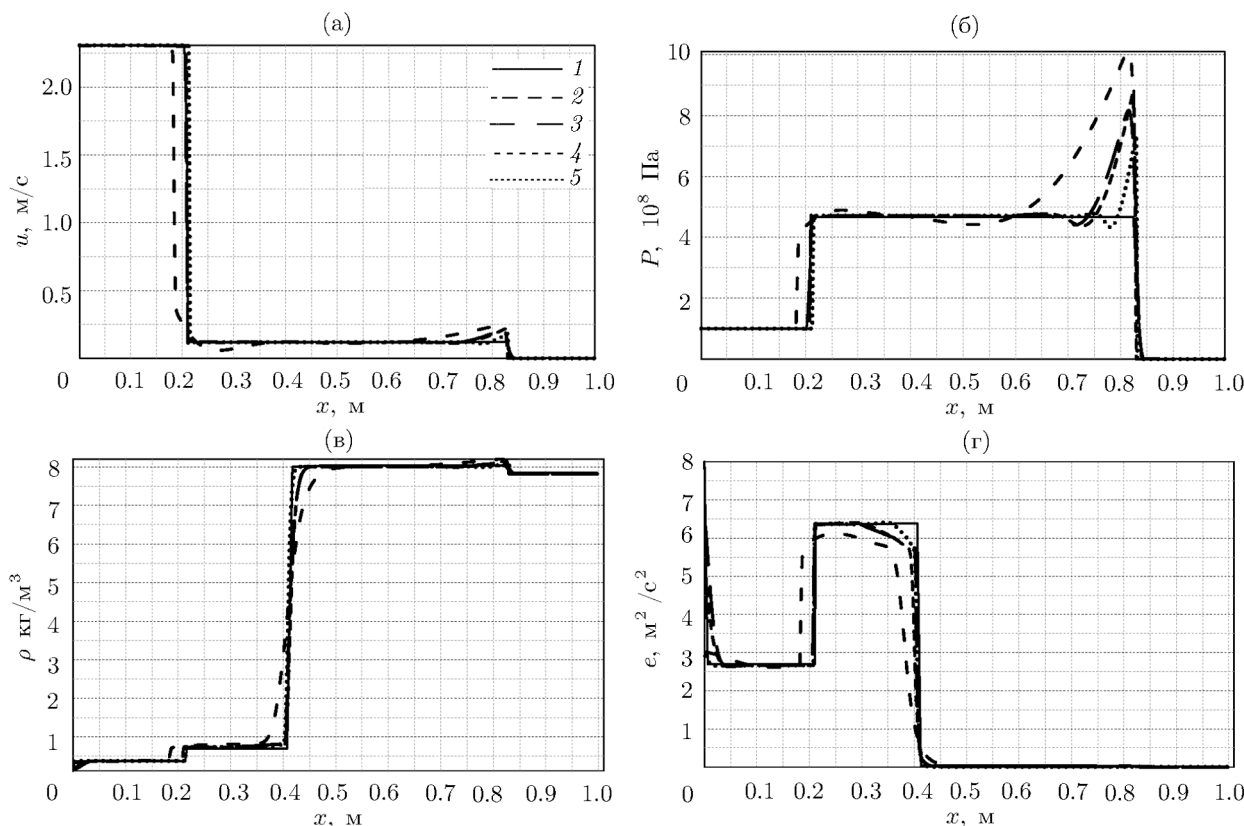


Рис. 2. Распределения для всех переменных в тесте 2: 1 – аналитическое решение; 2 –  $N = 5000, p = 0$ ; 3 –  $N = 5000, p = 1$ ; 4 –  $N = 40000, p = 0$ ; 5 –  $N = 160000, p = 0$ .

**Тест 3–5.** В данной серии экспериментов исследуется прохождение УВ через пузырек гелия, азота или криптона диаметром  $D_0 = 0.4$  м, содержащийся в ударной трубе. Для фаз используется УрС (3) с параметрами, указанными в табл. 3. Термодинамические параметры пузырька и окружающего воздуха приведены в табл. 3 (референсное давление  $P_\infty = 0$ ). Схематическое представление расчётной области с размерами  $L = 0.3$  м,  $H = 0.08$  м,  $X_0 = 0.05$  м представлено на рис. 3. В камере низкого давления (КНД) задаются нормальные начальные условия, а в камере высокого давления (КВД) –  $P_0 = 249\,090$  Па. На твердых стенках устанавливаются условия симметрии. Численные эксперименты проводились на расчётной сетке для  $N_1 = 900, N_2 = 240$  с  $p = 1$ . Коэффициенты релаксации оцениваются по формуле (18).

Таблица 3. Термодинамические свойства воздуха и пузырька

Физические свойства	Воздух	Гелий	Азот	Криптон
Плотность $\rho$ , $\text{кг/м}^3$	1.29	0.167	1.25	3.506
Скорость звука $c$ , $\text{м/с}$	340	1007	367	220
Показатель адиабаты $\gamma$	1.4	1.67	1.67	1.67

Динамика прохождения УВ через пузырёк зависит от числа Атвуда

$$A = \frac{\rho_b - \rho_s}{\rho_b + \rho_s},$$

где  $\rho_b$  – плотность пузырька,  $\rho_s$  – плотность окружающей среды. После того как фронт УВ достигает газового пузырька, он перестаёт быть плоским, так как одна часть волны взаимодействует с пузырьком (в дальнейшем будем называть её проходящей УВ), а другая нет (падающая УВ). Если плотность газового пузырька больше, чем плотность окружающей среды ( $A > 0$ ), то УВ по газовому пузырьку распространяется медленнее. Соответственно, если плотность газового пузырька меньше, чем плотность окружающей среды ( $A < 0$ ), то УВ по газовому пузырьку распространяется быстрее. Результаты численных экспериментов по скорости распространения падающей и проходящей УВ представлены на рис. 4, а (криптон), рис. 4, в (гелий) и рис. 4, г (азот). Видно, что скорость проходящей УВ для криптона меньше, чем скорость падающей УВ ( $A = 0.4621$ ), выше для гелия ( $A = -0.7708$ ) и практически совпадают для азота ( $A = -0.0157$ ). Наблюдается хорошее согласие полученных численных результатов с экспериментальными данными по скорости распространения падающей и проходящей УВ.

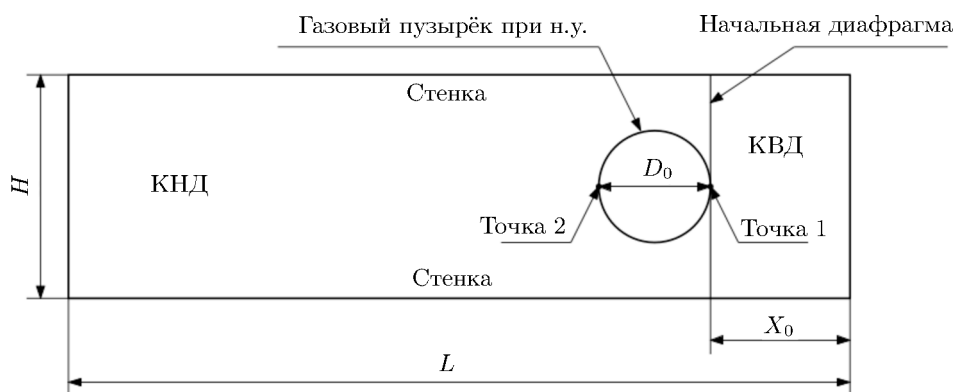


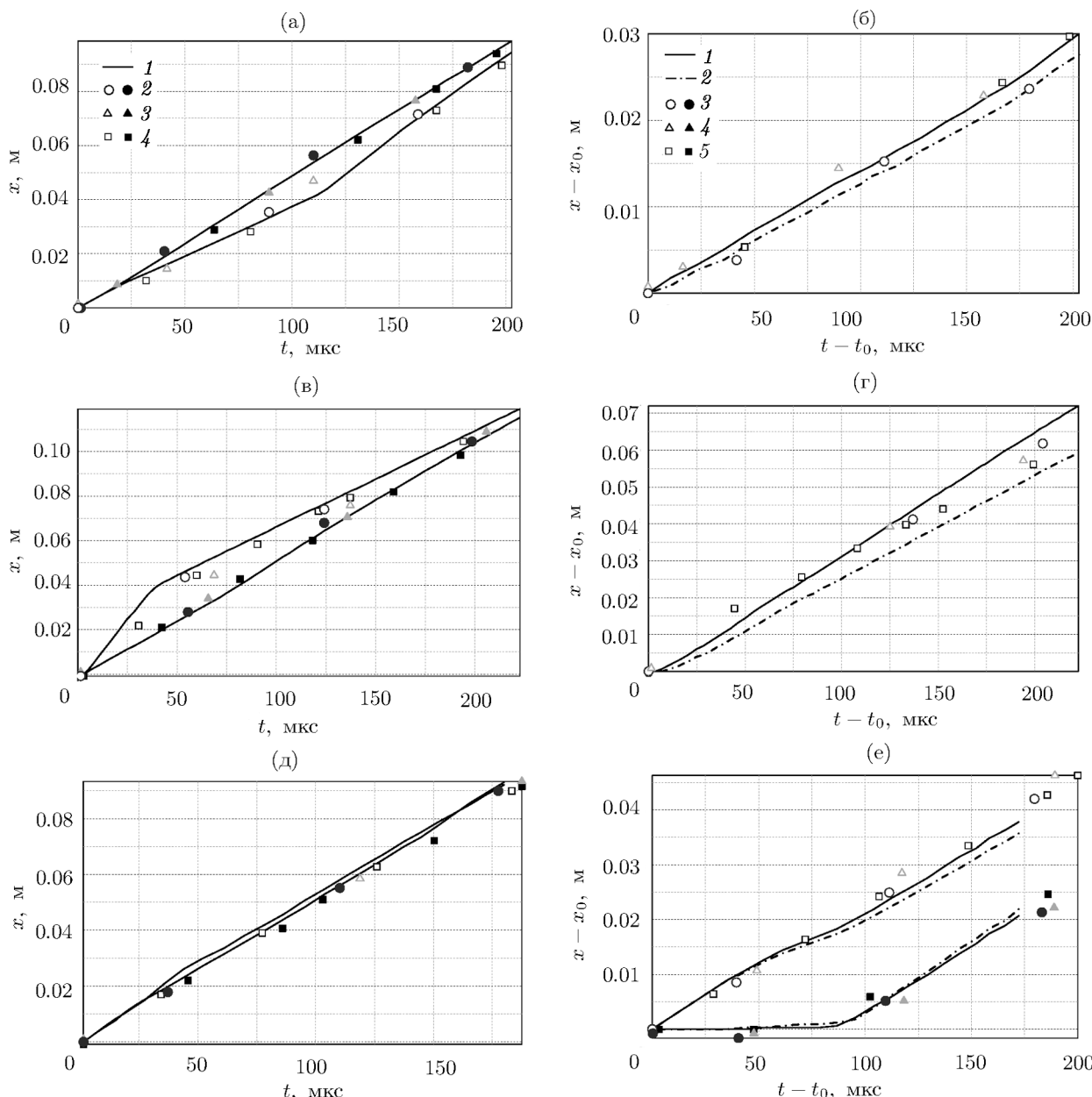
Рис. 3. Схематическое представление расчётной области для тестов 3–5.

Эволюция газового пузырька исследуется путём отслеживания значений решения в точках 1 и 2, связанных с межфазной границей пузырька (см. рис. 3). Экспериментальные данные для точки 2 в [25] представлены только для азота. Для обработки результатов численных расчётов по определению положения межфазной границы устанавливалось пороговое значение  $\alpha_{cr}$ . На рис. 4 представлены результаты с  $\alpha_{cr} = 0.1$  и  $\alpha_{cr} = 0.3$ . Видно, что полученные численные результаты хорошо согласуются с экспериментальными данными по описанию эволюции газового пузырька.

На рис. 5 представлена структура волн, которая формируется при взаимодействии пузырька азота и криптона с ударной волной (УВ). На рис. 5, а видно, что проходящая УВ медленнее падающей УВ (случай  $A = 0.4621$ ), а при прохождении УВ через пузырёк азота её фронт практически не деформируется (случай  $A = -0.0157$ ).

В работе [26] приведены также численные результаты для рассматриваемых постановок. Указанная работа отличается от настоящей тем, что в ней рассматривается равновесная по скорости модель [5]. Напомним, что в настоящей работе использовалась полностью неравновесная модель, а равновесие по скорости и давлению обеспечивается “жесткой” механической релаксацией. На рис. 4 также видно согласие полученных численных расчётов с результатами численных расчётов, полученных в работе [26].

**Заключение.** В работе рассматривается полностью неравновесная модель Баера–Нунциато с релаксационными слагаемыми. Вычислительный алгоритм для решения этой модели включает в себя решение гиперболической части с помощью разрывного метода Галёркина с использованием метода Рунге–Кутты для интегрирования по времени, вычисление релаксационных слагаемых, а также монотонизация решения с использованием ограничителя, обеспечивающего положительность объёмных долей и плотностей, а также ограничителя WENO-S. В работе показано, что полученный вычислительный алгоритм демонстрирует достаточно хорошее совпадение рассчитанных данных с экспериментальными в ряде тестовых расчётов.



**Рис. 4.**  $x - t$  диаграммы в тестах 3 (а, б), 4 (в,г) и 5 (д, е) для падающей и проходящей UV – а, в, д: 1 – численный расчёт; 2 – эксперимент 1 [25]; 3 – эксперимент 2 [25]; 4 – численные расчёты [26] (тёмные значки – падающая UV, светлые – проходящая UV);  $x - t$  диаграммы для точки межфазной границы – б, г, е: 1 – численный расчёт,  $\alpha_{cr} = 0.3$ ; 2 – численный расчёт,  $\alpha_{cr} = 0.1$ ; 3 – эксперимент 1 [25]; 4 – эксперимент 2 [25]; 5 – численные расчёты [26] (тёмные значки – точка 1, светлые – точка 2).

**Приложение. Задача Римана о распаде разрыва.** В данном пункте динамика двухфазной среды в задаче Римана изучается на основе хорошо известной модели Капилы (которая в дальнейшем обозначается как БН5 [5]). Эта модель получена из модели БН7 (1) в предположении мгновенной механической релаксации [5]. Другими словами, скорости и давления фаз в каждой точке пространства для модели БН5 равны. Решение модели БН5 может быть использовано для валидации численных расчётов, полученных на основе модели БН7, для физических постановок, которые соответствуют прямому разрешению межфазных границ.

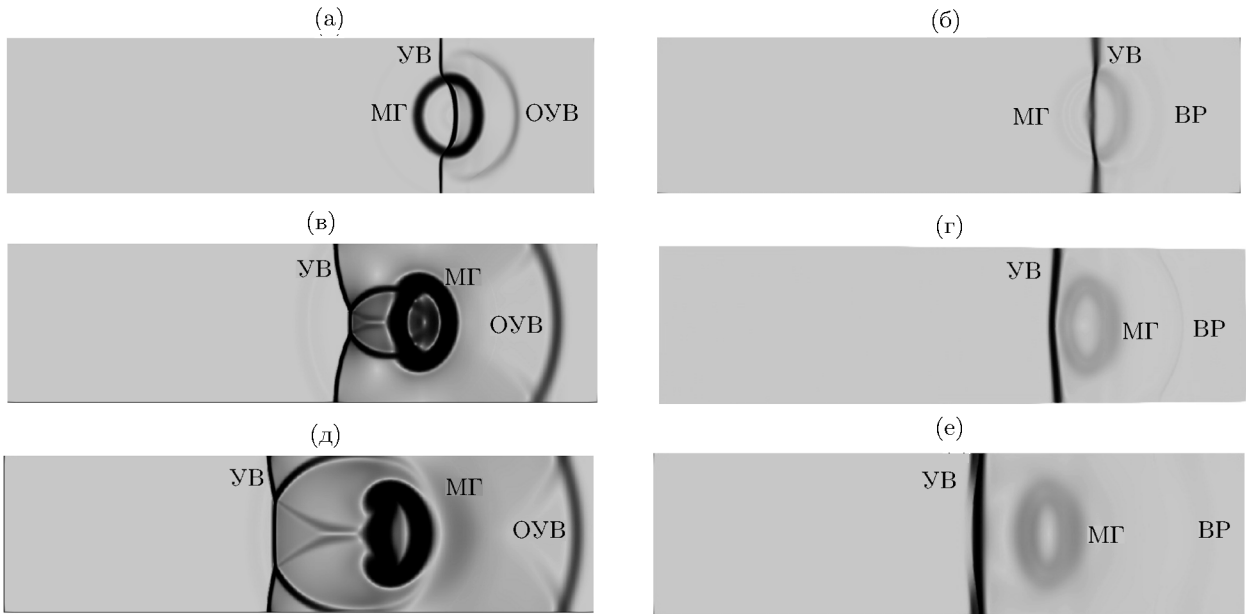
В модели БН5 законы сохранения массы записываются для каждой фазы отдельно, а законы сохранения импульса и энергии – для смеси в целом:

$$\frac{\partial \alpha_1}{\partial t} + \mathbf{u} \cdot \nabla \alpha_1 = K \nabla \cdot \mathbf{u}, \quad \frac{\partial \alpha_k \rho_k}{\partial t} + \nabla \cdot (\alpha_k \rho_k \mathbf{u}) = 0,$$

$$\frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u} + P) = 0, \quad \frac{\partial \rho E}{\partial t} + \nabla \cdot ((\rho E + P)\mathbf{u}) = 0, \quad (21)$$

где  $\mathbf{u}$  и  $P$  – скорость и давление смеси,  $\rho = \alpha_1 \rho_1 + \alpha_2 \rho_2$  и  $E = e + \mathbf{u}^2/2$  – плотность и полная энергия смеси, коэффициент  $K$  определяет сжимаемость фаз [5]:

$$K = (\rho_2 c_2^2 - \rho_1 c_1^2) / \left( \frac{\rho_1 c_1^2}{\alpha_1} + \frac{\rho_2 c_2^2}{\alpha_2} \right).$$



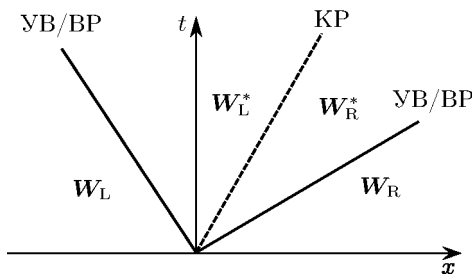
**Рис. 5.** Структура волн в тестах 3 (а,б), 4 (в, г), 5 (д,е) для пузырька криптона (а, в, д) и азота (б, г, е), которая рассчитывалась как  $f(\rho_2) = 1 - \exp(-\|\nabla \rho_2\|_2/25)$  (МГ – межфазная граница, ОУВ – отражённая УВ, ВР – волна разрежения): а –  $t = 60$  мкс, б –  $t = 50$ , в –  $t = 171$ , г –  $t = 102$ , д –  $t = 240$ , е –  $t = 179$ .

Система уравнений (21) дополняется нормировочным соотношением  $\alpha_1 + \alpha_2 = 1$ , а также уравнением состояния смеси, которое имеет вид

$$P = \left( \rho e - \left( \frac{\alpha_1 \gamma_1 \pi_1}{\gamma_1 - 1} + \frac{\alpha_2 \gamma_2 \pi_2}{\gamma_2 - 1} \right) \right) / \left( \frac{\alpha_1}{\gamma_1 - 1} + \frac{\alpha_2}{\gamma_2 - 1} \right) \quad (22)$$

в случае, если для описания фаз используется двучленное УрС вида (3). Начальные условия имеют следующий вид:

$$\mathbf{W}|_{t=0} = \begin{cases} \mathbf{W}_L, & x < 0, \\ \mathbf{W}_R, & x > 0, \end{cases}$$



**Рис. 6.** Расположение линий разрыва для задачи Римана.

где  $\mathbf{W} = (u, P, \rho_1, \rho_2, \alpha_1)$  – вектор примитивных переменных задачи.

С течением времени начальный разрыв распадается на несколько. На каждом из них выполняются соотношения РГ. Схематически автомодельная картина решения в одномерном случае на плоскости  $x-t$  представлена на рис. 6. Предполагается, что всего возможно пять конфигураций, а именно: левая и правая линии разрыва могут быть либо УВ, либо ВР, а центральная линия разрыва, если она есть, является КР.

Запишем соотношения РГ для модели БН5 вдоль некоторого пути  $\Psi$  ( $k = 1, 2$ ):

$$\xi[\alpha_1] = \int_0^1 (u)_\Psi \frac{\partial(\alpha_1)_\Psi}{\partial s} ds - \int_0^1 (K)_\Psi \frac{\partial(u)_\Psi}{\partial s} ds, \tag{23a}$$

$$\xi[\alpha_k \rho_k] = [\alpha_k \rho_k u], \tag{23b}$$

$$\xi[\rho u] = [\rho u^2 + P], \tag{23c}$$

$$\xi[\rho E] = [\rho E u + P u], \tag{23d}$$

где скобки обозначают скачок на линии разрыва:  $[f] = f^+ - f^-$ , верхний индекс “+” (“-”) указывает, что состояние вычисляется перед (за) разрывом,  $\xi$  – скорость разрыва.

В дальнейшем используем следующие обозначения. Невозмущённые области слева от разрыва будем обозначать индексом “L”, невозмущённые области справа от него – индексом “R”. Индекс 0 может обозначать как правое невозмущённое состояние, так и левое. Состояния слева и справа от контактного разрыва обозначим  $\mathbf{W}_L^*$  и  $\mathbf{W}_R^*$  соответственно (см. рис. 6).

**Утверждение.** *Решение задачи Римана  $(P^*, v_{1,L}^*, v_{1,R}^*)$  для системы уравнений (21), (22) вдоль линейного пути*

$$\Psi(s; \mathbf{W}_L, \mathbf{W}_R) = (1 - s)\mathbf{W}_L + s\mathbf{W}_R \tag{24}$$

удовлетворяют системе уравнений

$$F_1(P^*, v_{1,L}^*, v_{1,R}^*; \mathbf{W}_L, \mathbf{W}_R) = u_R - u_L + \Phi(P^*, v_{1,L}^*; \mathbf{W}_L) + \Phi(P^*, v_{1,R}^*; \mathbf{W}_R) = 0,$$

$$F_2(P^*, v_{1,L}^*; \mathbf{W}_L) = 0, \quad F_2(P^*, v_{1,R}^*; \mathbf{W}_R) = 0,$$

где

$$F_2(P, v_1; \mathbf{W}_0) = \begin{cases} (\alpha_1^0 - \alpha_1) \left( \frac{v^0 + v}{2} \right) - \left( \frac{K^0 + K}{2} \right) (v^0 - v) = 0, & \text{если } P > P^0, \\ 0, & \text{если } P \leq P^0, \end{cases}$$

$$\Phi(P, v_1; \mathbf{W}_0) = \begin{cases} (P - P^0)/m(P, v_1), & \text{если } P > P^0, \\ \int_{P^0}^P \frac{dP}{I(P)}, & \text{если } P \leq P^0. \end{cases} \tag{25}$$

Здесь функции  $\alpha_1(P, v_1)$ ,  $v(P, v_1)$ ,  $K(P, v_1)$  и  $m(P, v_1)$  определяются соотношениями

$$\alpha_1(v_1, v) = \frac{\alpha_1^0 \rho_1^0 v_1}{\rho^0 v}, \quad v(v_1, v_2) = Y_1^0 v_1 + Y_2^0 v_2,$$

$$v_2(v_1, P) = \left\{ Y_1^0 v_1^0 \left( \frac{P^0 + \gamma_1 \pi_1}{\gamma_1 - 1} + \frac{P^0 + P}{2} \right) + Y_2^0 v_2^0 \left( \frac{P^0 + \gamma_2 \pi_2}{\gamma_2 - 1} + \frac{P^0 + P}{2} \right) + Y_1^0 v_1 \left( \frac{P + \gamma_1 \pi_1}{\gamma_1 - 1} + \frac{P^0 + P}{2} \right) \right\} / \left\{ Y_2^0 \left( \frac{P + \gamma_2 \pi_2}{\gamma_2 - 1} + \frac{P^0 + P}{2} \right) \right\},$$

$$K(P, \alpha_1) = \frac{P(\gamma_2 - \gamma_1) + \gamma_2 \pi_2 - \gamma_1 \pi_1}{P(\gamma_1/\alpha_1 + \gamma_2/\alpha_2) + \gamma_1 \pi_1/\alpha_1 + \gamma_2 \pi_2/\alpha_2}, \quad m(P, v) = \sqrt{\frac{P - P^0}{v^0 - v}}. \tag{26}$$

**Доказательство.** Преобразуем систему уравнений (23) с учётом пути (24) к следующему виду:

$$[\alpha_1]\bar{v} = \bar{K}[v], \tag{27a}$$

$$Y_k^+ = Y_k^-, \tag{27b}$$

$$\rho^+(u^+ - \xi) = \rho^-(u^- - \xi) = m, \tag{27c}$$

$$P^+ - P^- + m(u^+ - u^-) = 0, \tag{27d}$$

$$P^+ - P^- + m^2(v^+ - v^-) = 0, \tag{27e}$$

$$e^+ - e^- + \frac{p^+ + p^-}{2}(v^+ - v^-) = 0, \tag{27f}$$

где  $k = 1, 2$ ,  $m$  – массовая скорость,  $Y_k = \alpha_k \rho_k / \rho$  – массовые доли,  $v_k = 1 / \rho_k$  – удельный объём  $k$ -й фазы,  $v = Y_1 v_1 + Y_2 v_2$  – удельный объём смеси,  $\bar{f} = (f^+ + f^-) / 2$ . Отметим, что уравнения (27b) и (27c) получены в результате сложения уравнений (23b) для  $k = 1, 2$ , а уравнения (27d) и (27e) эквивалентны друг другу (в дальнейшем используются оба соотношения).

Если  $m \neq 0$ , то разрыв является УВ, если  $m = 0$ , то КР. В последнем случае соответствующие соотношения (27) упрощаются до следующих:

$$u^+ = u^- = \xi, \quad P^+ = P^-. \tag{28}$$

Соотношения в ВР для модели БН5 получены в работе [27]:

$$\rho_k = \rho_k^0 \left( \frac{P + \pi_k}{P^0 + \pi_k} \right)^{1/\gamma_k}, \quad dP = \pm I(P) du, \quad Y_k = Y_k^0, \tag{29}$$

где волна разрежения распространяется со скоростью  $\xi = u \pm c_w$ ,  $k = 1, 2$ ; знаки “+” и “–” относятся к правой и левой волнам разрежения соответственно; индекс 0 обозначает невозмущённое состояние, а  $I(P)$  имеет следующий вид:

$$I(P) = \rho c_w = \left( \sum_{k=1,2} Y_k^0 / (\rho_k^0 c_k^0)^2 \left( \frac{P + \pi_k}{P^0 + \pi_k} \right)^{(\gamma_k+1)/\gamma_k} \right)^{-1/2}.$$

Выражения (27)–(29) позволяют получить решение задачи Римана общего вида. В самом деле, из соотношений (28) следует, что на КР давление и скорость непрерывны при переходе через разрыв:

$$P^* = P_L^* = P_R^*, \quad u^* = u_L^* = u_R^*. \tag{30}$$

Вид решения на разрыве (УВ или ВР) зависит только от  $P^*$ . Если  $P^* > P^0$ , то разрыв является УВ, если  $P^* \leq P^0$ , то ВР. Для скорости имеем  $u^* = u^0 \pm \Phi(P, v_1; \mathbf{W}_0)$ , где  $\Phi(P, v_1; \mathbf{W}_0)$  определяется равенством (25). Поэтому вне зависимости от конфигурации на линии разрыва для скорости справедливы соотношения

$$u_L^* = u_L - \Phi(P^*, v_{1,L}^*; \mathbf{W}_L), \quad u_R^* = u_R + \Phi(P^*, v_{1,R}^*; \mathbf{W}_R). \tag{31}$$

Далее из равенств (30) и (31) вытекает, что

$$F_1(P^*, v_{1,L}^*, v_{1,R}^*; \mathbf{W}_L, \mathbf{W}_R) = u_R^* - u_L^* = u_R - u_L + \Phi(P^*, v_{1,L}^*; \mathbf{W}_L) + \Phi(P^*, v_{1,R}^*; \mathbf{W}_R) = 0.$$

Для определения  $v_1$  используем уравнение (27a), которое запишем в виде

$$F_2(v_1, P; \mathbf{W}_0) = \begin{cases} (\alpha_1^0 - \alpha_1) \left( \frac{v^0 + v}{2} \right) - \left( \frac{K^0 + K}{2} \right) (v^0 - v) = 0, & \text{если } P > P^0, \\ 0, & \text{если } P \leq P^0. \end{cases}$$



Выражая переменные  $\alpha_1$ ,  $v$ ,  $K$  через переменные  $P$  и  $v_1$  и используя систему (27), получаем

$$\alpha_1(v_1, v) = \frac{\alpha_1^0 \rho_1^0 v_1}{\rho^0 v}, \quad v(v_1, v_2) = Y_1^0 v_1 + Y_2^0 v_2,$$

$$v_2(v_1, P) = \left\{ Y_1^0 v_1^0 \left( \frac{P^0 + \gamma_1 \pi_1}{\gamma_1 - 1} + \frac{P^0 + P}{2} \right) + Y_2^0 v_2^0 \left( \frac{P^0 + \gamma_2 \pi_2}{\gamma_2 - 1} + \frac{P^0 + P}{2} \right) + \right. \\ \left. + Y_1^0 v_1 \left( \frac{P + \gamma_1 \pi_1}{\gamma_1 - 1} + \frac{P^0 + P}{2} \right) \right\} / \left\{ Y_2^0 \left( \frac{P + \gamma_2 \pi_2}{\gamma_2 - 1} + \frac{P^0 + P}{2} \right) \right\},$$

$$K(P, \alpha_1) = \frac{P(\gamma_2 - \gamma_1) + \gamma_2 \pi_2 - \gamma_1 \pi_1}{P(\gamma_1/\alpha_1 + \gamma_2/\alpha_2) + \gamma_1 \pi_1/\alpha_1 + \gamma_2 \pi_2/\alpha_2}.$$

В итоге приходим относительно трёх неизвестных  $P^*$ ,  $v_{1,L}^*$ ,  $v_{1,R}^*$  к системе трёх нелинейных уравнений

$$F_1(P^*, v_{1,L}^*, v_{1,R}^*; \mathbf{W}_L, \mathbf{W}_R) = 0, \quad F_2(P^*, v_{1,L}^*; \mathbf{W}_L) = 0, \quad F_2(P^*, v_{1,R}^*; \mathbf{W}_R) = 0, \quad (32)$$

что доказывает утверждение.

Система уравнений (32) относительно переменных  $(P^*, v_{1,L}^*, v_{1,R}^*)$  может быть решена, например, методом Ньютона. Далее по любой из формул (31) определяется скорость, а оставшиеся переменные находятся по формулам (26).

Исследование Полехиной (Тухватуллиной) Р.Р. (п.п. 2.2, 3, приложение) выполнено при финансовой поддержке Российского научного фонда (проект 19-71-30004). Исследование Савенкова Е.Б. и Алексева М.В. (введение, п.п. 2.1, 2.3, заключение) выполнено при поддержке Московского центра фундаментальной и прикладной математики (соглашение с Министерством науки и высшего образования Российской Федерации № 075-15-2019-1623).

### СПИСОК ЛИТЕРАТУРЫ

1. *Baer M.R., Nunziato J.W.* A two-phase mixture theory for the deflagration-to-detonation transition (ddt) in reactive granular materials // *Int. J. of Multiphase Flow*. 1986. V. 12. № 6. P. 861–889.
2. *Drew D., Passman S.* Theory of Multicomponent Fluids. New York, 2014.
3. *Favrie N., Gavriluk S., Saurel R.* Solid-fluid diffuse interface model in cases of extreme deformations // *J. Comput. Phys.* 2009. V. 228. № 16. P. 6037–6077.
4. *Kapila A., Son S., Bdzil J., Menikoff R.* Two-phase modeling of DDT: structure of the velocity-relaxation zone // *Phys. Fluids*. 1997. V. 9. № 12. P. 3885–3897.
5. *Kapila A., Melnikoff R., Bdzil J., Son S., Stewart S.* Two-phase modeling of deflagration-to-detonation transition in granular materials: reduced equations // *Phys. Fluids*. 2001. V. 13. № 10. P. 3002–3024.
6. *Murrone A., Guillard H.* A five-equation reduced model for compressible two phase flow problems // *J. Comput. Phys.* 2005. V. 202. № 2. P. 664–698.
7. *De Lorenzo M., Pelanti M., Lafon P.* HLLC-type and path-conservative schemes for a single-velocity six-equation two-phase flow model: a comparative study // *Appl. Math. and Comput.* 2018. V. 333. P. 95–117.
8. *Dumbser M., Balsara D.S.* A new efficient formulation of the hllm Riemann solver for general conservative and non-conservative hyperbolic systems // *J. of Comput. Phys.* 2016. V. 304. P. 275–319.
9. *Kemm F., Gaburro E., Thein F., Dumbser M.* A simple diffuse interface approach for compressible flows around moving solids of arbitrary shape based on a reduced Baer–Nunziato model // *Computers & Fluids*. 2020. V. 204. P. 104536.
10. *Chiocchetti S., Müller C.* A solver for stiff finite-rate relaxation in Baer–Nunziato two-phase flow models // *Droplet Interactions and Spray Processes* / Eds. G. Lamanna, S. Tonini, G.E. Cossali, B. Weigand. Cham, 2020. P. 31–44.
11. *Serezhkin A., Menshov I.* On solving the Riemann problem for non-conservative hyperbolic systems of partial differential equations // *Computers & Fluids*. 2020. V. 210. P. 104675.
12. *Cockburn B., Shu C.-W.* The Runge–Kutta local projection–discontinuous–Galerkin finite element method for scalar conservation laws // *ESAIM Math. Model. Numer. Anal.* 1991. V. 25. № 3. P. 337–361.

13. *Dal Maso G., Lefloch P.G., Murat F.* Definition and weak stability of nonconservative products // *J. de Mathématiques Pures et Appliquées*. 1995. V. 74. № 6. P. 483–548.
14. *Castro M.J., LeFloc C.h P.G., Luz Muñoz-Ruiz M., Parés C.* Why many theories of shock waves are necessary: convergence error in formally path-consistent schemes // *J. of Comput. Phys.* 2008. V. 227. № 17. P. 8107–8129.
15. *Bdzil J.B., Menikoff R., Son S.F., Kapila A.K., Scott Stewart D.* Two-phase modeling of deflagration-to-detonation transition in granular materials: A critical examination of modeling issues // *Phys. of Fluids*. 1999. V. 11. № 2. P. 378–402.
16. *Andrianov N., Warnecke G.* The Riemann problem for the Baer–Nunziato two-phase flow model // *J. of Comput. Phys.* 2004. V. 195. № 2. P. 434–464.
17. *Daude F., Berry R.A., Pascal Galon.* A finite-volume method for compressible non-equilibrium two-phase flows in networks of elastic pipelines using the Baer–Nunziato model // *Comput. Meth. in Appl. Mech. and Eng.* 2019. V. 354. P. 820–849.
18. *Saurel R., Abgrall R.* A simple method for compressible multifluid flows // *SIAM J. Sci. Comput.* 1999. V. 21. № 3. P. 1115–1145.
19. *Янилкин Ю.В., Бондаренко Ю.А., Гончаров Е.А., Гужова А.Р., Колобянин В.Ю., Софронов В.Н., Стаценко В.П.* Тесты для гидрокодов, моделирующих ударноволновые течения в многокомпонентных средах. Саров, 2017.
20. *Nigmatulin R.* Dynamics of Multiphase Media. New York, 1990.
21. *Le Floch P., Liu T.-P.* Existence Theory for Nonlinear Hyperbolic Systems in Nonconservative Form. V. 5. Berlin; New York, 1993. P. 261–280.
22. *Ваннер Г., Хайрер Э., Нёрсетт С.* Решение обыкновенных дифференциальных уравнений. Нежесткие задачи. М., 1990.
23. *Zhong X., Shu C.-W.* A simple weighted essentially nonoscillatory limiter for Runge–Kutta discontinuous Galerkin methods // *J. Comput. Phys.* 2013. V. 232. № 1. P. 397–415.
24. *Zhong X., Shu C.-W.* On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes // *J. Comput. Phys.* 2010. V. 229. № 23. P. 8918–8934.
25. *Layes G., Le Métayer O.* Quantitative numerical and experimental studies of the shock accelerated heterogeneous bubbles motion // *Physics of Fluids*. 2007. V. 19. № 4. P. 042105.
26. *Nowakowski A.F., Ballil A., Nicolleau F.C.G.A.* Passage of a shock wave through inhomogeneous media and its impact on gas-bubble deformation // *Phys. Rev. E*. 2015. V. 92. № 2. P. 023028.
27. *Petitpas F., Franquet E., Saurel R., Le Metayer O.* A relaxation–projection method for compressible flows. Part ii: Artificial heat exchanges for multiphase shocks // *J. of Comput. Phys.* 2007. V. 225. № 2. P. 2214–2248.

Институт прикладной математики  
имени М.В. Келдыша РАН, г. Москва

Поступила в редакцию 04.03.2022 г.  
После доработки 04.03.2022 г.  
Принята к публикации 25.05.2022 г.

УДК 519.6+517.9

## ОЦЕНКИ ТОЧНОСТИ ПРОЕКЦИОННОГО МЕТОДА СО СТАБИЛИЗАТОРОМ ДРОБНОЙ ГЛАДКОСТИ В ЗАДАЧЕ ВОССТАНОВЛЕНИЯ ВОЛНОВОГО ФРОНТА ПО ЕГО НАКЛОНАМ

© 2022 г. А. В. Разгулин, Н. Г. Ирошников, А. В. Ларичев,  
С. А. Турганбаев, Т. Е. Романенко

Рассматривается проекционный метод со стабилизатором дробной гладкости для аппроксимации задачи восстановления волнового фронта по его наклонам. Показана устойчивость метода в шкале периодических функций двух аргументов. Найдены условия согласования параметров стабилизатора с шагом сетки, позволяющие провести вывод согласованных с гладкостью наклонов оценок скорости сходимости проекционного метода. С помощью теории интерполяции получены оценки точности метода при естественных требованиях на гладкость наклонов из шкалы анизотропных пространств Соболева дробной гладкости.

DOI: 10.31857/S037406412207010X, EDN: CERLCH

**Введение.** Одной из важных задач адаптивной оптики является восстановление волнового фронта по данным об измерениях его локальных наклонов. Такие измерения проводятся с помощью датчика Шака–Гартмана (см., например, [1]), который также с успехом применяется для оценки параметров атмосферной турбулентности [2], при конструировании оптоэлектронных инструментов [3], в прикладных задачах медицинской физики [4] и в др. областях. Поскольку на практике измерения наклонов сопровождаются искажениями различной природы, то данные о наклонах обычно являются негладкими функциями координат. Кроме того, сами волновые фронты зачастую являются негладкими и даже разрывными. Разработка устойчивых методов восстановления волнового фронта в этих условиях является весьма актуальной задачей (см. работы [5–10]).

Ранее в статьях [9, 10] было предложено семейство новых методов реконструкции волнового фронта, в основе которых лежит вариационный подход в сочетании с проекционным методом на основе билинейных сплайнов. Такая комбинация методов наилучшим образом подходит для случая восстановления волнового фронта, не обладающего высокой гладкостью и не обязательно непрерывного. Отличительной особенностью схем работы [10] является использование стабилизатора с разностными производными второго порядка в дробной степени, нацеленного на улучшение пространственно-частотной характеристики (ПЧХ) метода, причём параметры стабилизатора выбираются также на основе вариационного подхода. В сравнении с широко используемым методом Симпсона [5] семейство методов в статье [10] позволяет достичь лучшего профиля ПЧХ в области высоких частот, а по сравнению с семейством кумулятивных методов в работе [6] в практическом применении не приводит к накоплению погрешности и устойчиво работает в случае потери данных о наклонах.

Представляемые в данной статье результаты посвящены исследованию точности проекционного метода, описанного в [10], используемого для конечномерной аппроксимации вариационной задачи восстановления волнового фронта. Сложностью рассматриваемой задачи является возможная негладкость волновых фронтов и, соответственно, необходимость развивать такие способы исследования погрешности, которые позволяли бы получать оценки скорости сходимости аппроксимаций в естественных функциональных классах исходных данных (наклонов) без априорных предположений об их дополнительной гладкости.

**1. Вариационный метод восстановления волнового фронта по его наклонам.** Опишем и исследуем вариационный подход к восстановлению волнового фронта  $u(x, y)$  по измерениям его наклонов  $g_1(x, y)$ ,  $g_2(x, y)$  вдоль соответствующих направлений  $x$  и  $y$ . Отметим,

что непосредственное восстановление волнового фронта из условия равенства градиентов волнового фронта (или их конечномерных аппроксимаций) измеряемым наклонам приводит к операторному уравнению первого рода, которое представляет собой плохообусловленную задачу (см. [7]), устойчивое решение которой строится на основе регуляризованного варианта метода сингулярных разложений. Вместо операторного уравнения сформулируем задачу восстановления волнового фронта в пределах квадратной апертуры  $\mathbb{T}^2 = [-\pi, \pi] \times [-\pi, \pi]$  в виде задачи минимизации функционала невязки  $J(u)$ :

$$J(u) = \iint_{\mathbb{T}^2} ((u_x(x, y) - g_1(x, y))^2 + (u_y(x, y) - g_2(x, y))^2) dx dy. \tag{1}$$

Функционал  $J(u)$  задаёт среднеквадратичное отклонение между наблюдаемыми наклонами волнового фронта  $g_1(x, y)$ ,  $g_2(x, y)$  и соответствующими производными  $u_x = \partial u / \partial x$ ,  $u_y = \partial u / \partial y$  искомого волнового фронта.

Для дальнейшего изложения потребуются некоторые обозначения и вспомогательные утверждения. Сокращения  $\langle u, \phi \rangle = \iint_{\mathbb{T}^2} u(x, y)\phi(x, y) dx dy$  и  $\|u\| = (\langle u, u \rangle)^{1/2}$  обозначают стандартные скалярное произведение и норму в  $L_2(\mathbb{T}^2)$ . Для вектора наклонов  $g = (g_1, g_2) \in L_2(\mathbb{T}^2) \times L_2(\mathbb{T}^2)$  используется евклидова норма  $\|g\| = (\|g_1\|^2 + \|g_2\|^2)^{1/2}$ . Для полной ортонормированной в  $L_2(\mathbb{T}^2)$  системы комплексных экспонент  $e_{k,m}(x, y) = (2\pi)^{-1} \exp(i(kx + my))$ ,  $k, m \in \mathbb{Z}$ , и коэффициентов  $\hat{u}_{k,m} = \langle u, e_{k,m}^* \rangle$ ,  $e_{k,m}^* = e_{-k,-m}$  разложения функции  $u \in L_2(\mathbb{T}^2)$  по этой системе в силу равенства Парсеваля имеем  $\|u\|^2 = \sum_{(k,m) \in \mathbb{Z}^2} |\hat{u}_{k,m}|^2$ .

Будем использовать анизотропные пространства функций

$$H^{s,r}(\mathbb{T}^2) = \{u \in L_2(\mathbb{T}^2) : \|u\|_{H^{s,r}(\mathbb{T}^2)}^2 = \sum_{(k,m) \in \mathbb{Z}^2} (1 + k^{2s} + m^{2r}) |\hat{u}_{k,m}|^2 < +\infty\}, \quad s, r > 0.$$

Нетрудно видеть, что по теореме о следах [11, гл. 1, § 4] используемые далее пространства  $H^{s,0}(\mathbb{T}^2)$  ( $H^{0,r}(\mathbb{T}^2)$ ) при  $s > 0.5$  ( $r > 0.5$  соответственно) состоят из  $2\pi$ -периодических функций по переменной  $x$  ( $y$  соответственно), которые Гёльдер-непрерывны по  $x$  ( $y$  соответственно) со значениями в  $L_2(0, 2\pi)$  с показателем  $s - 0.5$  ( $r - 0.5$  соответственно).

Будем использовать изотропное пространство функций  $H^s(\mathbb{T}^2) = H^{s,s}(\mathbb{T}^2)$ . В частности, пространство  $H^1(\mathbb{T}^2)$  является замкнутым подпространством пространства Соболева  $H^1$  на множестве  $\mathbb{T}^2$  (см. [12, гл. 3]) с эквивалентными нормами

$$\|u\|_{H^1} = (\|u\|^2 + \|u_x\|^2 + \|u_y\|^2)^{1/2}, \quad \|u\|_{H^1(\mathbb{T}^2)}^2 = \sum_{(k,m) \in \mathbb{Z}^2} (1 + k^2 + m^2) |\hat{u}_{k,m}|^2.$$

Для пространства Соболева второго порядка  $H^2$  используется норма

$$\|u\|_{H^2} = (\|u\|^2 + \|u_x\|^2 + \|u_y\|^2 + \|u_{xx}\|^2 + \|u_{xy}\|^2 + \|u_{yy}\|^2)^{1/2}.$$

Рассмотрим подпространство  $\tilde{H}^\theta(\mathbb{T}^2)$  функций с нулевым средним (или, что то же самое,  $\hat{u}_{0,0} = 0$ ), т.е.  $\tilde{H}^\theta(\mathbb{T}^2) = \{u \in H^\theta(\mathbb{T}^2) : \iint_{\mathbb{T}^2} u(x, y) dx dy = 0\}$ . Для функций из  $\tilde{H}^1(\mathbb{T}^2)$  имеет место неравенство Пуанкаре

$$\iint_{\mathbb{T}^2} (|u_x|^2 + |u_y|^2) dx dy \geq \iint_{\mathbb{T}^2} |u|^2 dx dy, \tag{2}$$

проверка которого в рассматриваемом случае легко проводится с помощью разложения по комплексным экспонентам  $e_{k,m}(x, y)$ . Отметим, что обзору современных результатов по точным константам в неравенствах типа Пуанкаре посвящена работа [13].

В качестве допустимого множества функционала (1) будем использовать пространство  $\tilde{H}^1(\mathbb{T}^2)$ . Такой выбор допустимого множества функционала невязки представляется достаточно естественным для рассматриваемой задачи восстановления волнового фронта, поскольку,

во-первых, позволяет естественным образом вычислять его наклоны и, во-вторых, охватывает широкий круг негладких поверхностей, в том числе разрывных (пример см. в [11, п. 9.4]).

Необходимое условие минимума функционала невязки (1) на  $\tilde{H}^1(\mathbb{T}^2)$  получается из равенства нулю его первой вариации и приводит к интегральному тождеству

$$\langle u_x, \phi_x \rangle + \langle u_y, \phi_y \rangle = \langle g_1, \phi_x \rangle + \langle g_2, \phi_y \rangle \quad \text{для любых функций } \phi \in \tilde{H}^1(\mathbb{T}^2), \quad (3)$$

которому должен удовлетворять искомый волновой фронт  $u(x, y) \in \tilde{H}^1(\mathbb{T}^2)$ .

**Теорема 1.** Пусть  $g_1, g_2 \in L_2(\mathbb{T}^2)$ . Тогда вариационная задача (3) имеет единственное решение  $u \in \tilde{H}^1(\mathbb{T}^2)$ . Решение непрерывно зависит от вектора наклонов и справедлива оценка

$$\|u\|_{H^1(\mathbb{T}^2)} \leq \sqrt{2}\|g\|. \quad (4)$$

**Доказательство.** Задача (3) может быть записана в стандартном для вариационных задач виде

$$a(u, \phi) = \ell(\phi) \quad \text{для любой } \phi \in \tilde{H}^1(\mathbb{T}^2), \quad (5)$$

где введены обозначения для билинейной формы  $a(u, \phi) = \langle u_x, \phi_x \rangle + \langle u_y, \phi_y \rangle$  и линейного функционала  $\ell(\phi) = \langle g_1, \phi_x \rangle + \langle g_2, \phi_y \rangle$ . Непрерывность билинейной формы и функционала вытекает из оценок

$$|a(u, \phi)| \leq \|u_x\| \|\phi_x\| + \|u_y\| \|\phi_y\| \leq \|u\|_{H^1(\mathbb{T}^2)} \|\phi\|_{H^1(\mathbb{T}^2)}, \quad |\ell(\phi)| \leq \|g\| \|\phi\|_{H^1(\mathbb{T}^2)}. \quad (6)$$

Свойство коэрцитивности билинейной формы следует из неравенства Пуанкаре (2) и записывается в виде

$$a(u, u) \geq 0.5\|u\|_{H^1(\mathbb{T}^2)}^2 \quad \text{для всех функций } u \in \tilde{H}^1(\mathbb{T}^2). \quad (7)$$

Тогда существование и единственность решения вариационной задачи (5) следует из непрерывности функционала, непрерывности и коэрцитивности билинейной формы и теоремы Лакса–Мильграма (см., например, [14, с. 19; 15, с. 43]). Оценка (4) является следствием неравенства (2) и вытекающей из (5) и неравенства Коши–Буняковского оценки  $\|u_x\|^2 + \|u_y\|^2 \leq \|g\|^2$ . Теорема доказана.

Исследуем гладкость решения вариационной задачи в зависимости от гладкости входных данных  $g_1, g_2$ .

**Теорема 2.** Пусть  $g_1 \in H^{s,0}(\mathbb{T}^2)$ ,  $g_2 \in H^{0,s}(\mathbb{T}^2)$ ,  $s \in (0, 1]$ . Тогда решение вариационной задачи (3) принадлежит  $\tilde{H}^{1+s}(\mathbb{T}^2)$  и имеет место оценка

$$\|u\|_{H^{1+s}(\mathbb{T}^2)} \leq C(\|g_1\|_{H^{s,0}(\mathbb{T}^2)} + \|g_2\|_{H^{0,s}(\mathbb{T}^2)}). \quad (8)$$

**Доказательство.** Рассмотрим разложение  $u = \sum_{(0,0) \neq (n,m) \in \mathbb{Z}^2} \hat{u}_{n,m} e_{n,m}(x, y)$  полученного в теореме 1 решения задачи (3) (или, что то же самое, (5)). Возьмём в (5) пробную функцию  $\phi = e_{k,l}^*$ . Несложные вычисления показывают, что  $a(u, e_{k,l}^*) = (k^2 + l^2)u_{k,l}$  и  $\ell(e_{k,l}^*) = -i(k\hat{g}_{1,k,l} + l\hat{g}_{2,k,l})$ , где  $\hat{g}_{j,k,l} = \langle g_j, e_{k,l}^* \rangle$  – коэффициенты разложения наклонов,  $j = 1, 2$ . Получаем

$$\hat{u}_{k,l} = -i \frac{k\hat{g}_{1,k,l} + l\hat{g}_{2,k,l}}{k^2 + l^2}, \quad (0,0) \neq (k,l) \in \mathbb{Z}^2, \quad \hat{u}_{0,0} = 0. \quad (9)$$

Воспользуемся (9) и оценим норму

$$\begin{aligned} \|u\|_{H^{1+s}(\mathbb{T}^2)}^2 &= \sum_{(0,0) \neq (k,l) \in \mathbb{Z}^2} (1 + k^{2(1+s)} + l^{2(1+s)}) |\hat{u}_{k,l}|^2 \leq \\ &\leq 2 \sum_{(0,0) \neq (k,l) \in \mathbb{Z}^2} \frac{(1 + k^{2(1+s)} + l^{2(1+s)})(k^2 |\hat{g}_{1,k,l}|^2 + l^2 |\hat{g}_{2,k,l}|^2)}{(k^2 + l^2)^2}. \end{aligned}$$

Заметим, что для всех индексов, по которым проводится суммирование, справедливо

$$\frac{(1 + k^{2(1+s)} + l^{2(1+s)})k^2}{(k^2 + l^2)^2} = \frac{(1 + k^{2(1+s)} + l^{2(1+s)})k^{2(1-s)}}{(k^2 + l^2)^2} k^{2s} \leq 3k^{2s}.$$

Аналогично получаем

$$\frac{(1 + k^{2(1+s)} + l^{2(1+s)})l^2}{(k^2 + l^2)^2} \leq 3l^{2s}.$$

С учётом этих оценок имеем

$$\|u\|_{H^{1+s}(\mathbb{T}^2)}^2 \leq C_1 \sum_{(0,0) \neq (k,l) \in \mathbb{Z}^2} (k^{2s} |\hat{g}_{1,k,l}|^2 + l^{2s} |\hat{g}_{2,k,l}|^2) = C_1 (\|g_1\|_{H^{s,0}(\mathbb{T}^2)}^2 + \|g_2\|_{H^{0,s}(\mathbb{T}^2)}^2).$$

Отсюда вытекают включение  $u \in \tilde{H}^{1+s}(\mathbb{T}^2)$  и искомая оценка (9). Теорема доказана.

**2. Проекционная схема аппроксимации вариационной задачи.** Для построения конечномерной аппроксимации вариационной задачи (3) на основе проекционного метода возьмём натуральное чётное  $N$ , введём множество индексов

$$\mathbb{I} = \{-N/2, -N/2 + 1, \dots, N/2 - 1\}$$

и рассмотрим подпространство

$$S^1(\mathbb{T}^2) = \text{Lin} \{ \phi_{k,l}(x, y) \} \subset H^1(\mathbb{T}^2)$$

размерности  $N \times N$  кусочно-линейных по каждому направлению сплайнов, являющихся линейными комбинациями функций  $\phi_{k,l}(x, y) = \phi_k(x)\phi_l(y)$ ,  $(k, l) \in \mathbb{I}^2$ , где  $\phi_k(t) = \max(0, 1 - |t - kh|h^{-1})$ ,  $k \in \mathbb{I} \setminus \{-N/2\}$ ,  $\phi_{-N/2}(t) = \phi_{1-N/2}(t+h) + \phi_{1-N/2}(t+h-2\pi)$ ,  $h = 2\pi N^{-1}$ .

Обозначим  $\tilde{S}^1(\mathbb{T}^2) = S^1(\mathbb{T}^2) \cap \tilde{H}^1(\mathbb{T}^2)$ . Представив далее аппроксимацию искомого волнового фронта в виде конечной суммы  $u^h(x, y) = \sum_{(k,l) \in \mathbb{I}^2} u_{k,l}^h \phi_{k,l}(x, y) \in \tilde{S}^1(\mathbb{T}^2)$  и ограничившись в тождестве (3) пробными функциями из подпространства  $\tilde{S}^1(\mathbb{T}^2)$ , приходим к конечномерной вариационной задаче

$$\langle u_x^h, \phi_x \rangle + \langle u_y^h, \phi_y \rangle = \langle g_1, \phi_x \rangle + \langle g_2, \phi_y \rangle \quad \text{при всех } \phi \in \tilde{S}^1(\mathbb{T}^2). \tag{10}$$

Условие принадлежности решения  $u^h$  пространству  $\tilde{S}^1(\mathbb{T}^2)$  накладывает условие равенства нулю среднего значения сеточной функции

$$\sum_{(k,l) \in \mathbb{I}^2} u_{k,l}^h = 0. \tag{11}$$

Отметим, что поскольку в записи равенств (10) пробные функции  $\phi \in \tilde{S}^1(\mathbb{T}^2)$  участвуют только под знаком производных, то фактически можно расширить класс пробных функций до всего пространства  $S^1(\mathbb{T}^2)$ . Тогда задача (10) допускает эквивалентную запись в виде операторного уравнения относительно матрицы искомых коэффициентов  $u^h = \{u_{kl}^h\}$  (см. [10])

$$B_2 \Lambda_1 u^h + B_1 \Lambda_2 u^h = f, \quad f = f^1 + f^2 \tag{12}$$

с дополнительным условием (11), где

$$f_{k,l}^1 = h^{-3} \int_{y_{l-1}}^{y_{l+1}} \phi_l(y) \int_{x_{k-1}}^{x_k} (g_1(x, y) - g_1(x+h, y)) dx dy,$$

$$f_{k,l}^2 = h^{-3} \int_{x_{k-1}}^{x_{k+1}} \phi_k(x) \int_{y_{l-1}}^{y_l} (g_2(x, y) - g_2(x, y+h)) dx dy.$$

Операторы  $B_j, \Lambda_j$  задаются с помощью матриц Грама (см. монографию [16, гл. 2, § 2]) в  $L_2(-\pi, \pi)$  рассматриваемых одномерных сплайнов и их градиентов,  $B_j = E_j - h^2 \Lambda_j / 6$ ,  $E_j$  – единичный оператор по соответствующему направлению,  $j = 1, 2$ . Отметим особенность записи слагаемых в уравнении (12): операторы с индексом  $j = 1$  действуют по переменной  $x$  (по строкам), операторы с индексом  $j = 2$  действуют по переменной  $y$  (по столбцам). Таким образом, например, запись  $B_2 \Lambda_1$  означает суперпозицию двух операторов, действующих по разным переменным.

Далее каждой функции  $u^h(x, y) = \sum_{(k,l) \in \mathbb{I}^2} u_{k,l}^h \phi_{k,l}(x, y)$  из пространства сплайнов  $S^1(\mathbb{T}^2)$  сопоставим заданную на  $\mathbb{I}^2$  сеточную функцию  $u^h = \{u_{k,l}^h\}$ , составленную из коэффициентов разложения по базисным сплайнам. Для сеточных функций будем использовать дискретное скалярное произведение и соответствующую дискретную евклидову норму

$$\langle u^h, v^h \rangle_h = \sum_{(k,l) \in \mathbb{I}^2} u_{k,l}^h v_{k,l}^h h^2, \quad \|u^h\|_h = \sqrt{\langle u^h, u^h \rangle_h}.$$

Сеточные комплексные экспоненты

$$e_k(x_j) = \frac{1}{\sqrt{2\pi}} \exp\left(i \frac{2\pi k j}{N}\right), \quad e_m(y_l) = \frac{1}{\sqrt{2\pi}} \exp\left(i \frac{2\pi m l}{N}\right)$$

являются собственными функциями как операторов  $\Lambda_1, \Lambda_2$  с собственными значениями

$$\lambda_k = \frac{4}{h^2} \sin^2\left(\frac{kh}{2}\right), \quad \lambda_m = \frac{4}{h^2} \sin^2\left(\frac{mh}{2}\right), \quad k, m \in \mathbb{I}, \tag{13}$$

так и операторов  $B_1, B_2$  с собственными значениями

$$\mu_k = 1 - \frac{h^2}{6} \lambda_k, \quad \mu_m = 1 - \frac{h^2}{6} \lambda_m, \quad k, m \in \mathbb{I}.$$

Собственные значения допускают известные двусторонние оценки

$$0 \leq \lambda_k \leq \frac{4}{h^2}, \quad k \in \mathbb{I}, \tag{14}$$

$$\frac{1}{3} \leq \mu_k \leq 1, \quad k \in \mathbb{I}. \tag{15}$$

Сеточные комплексные экспоненты  $e_{k,m}(x_j, y_l) = e_k(x_j) e_m(y_l)$  представляют собой ортонормированный базис в пространстве сеточных функций на  $\mathbb{I}^2$ :  $\langle e_{k,m}, e_{p,q}^* \rangle_h = \delta_{k,m} \delta_{p,q}$ , где  $\delta_{k,m}$  – символ Кронекера. Любая заданная на  $\mathbb{I}^2$  сеточная функция  $\{z_{j,l}\}$  однозначно раскладывается в сумму

$$z_{j,l} = \sum_{(k,m) \in \mathbb{I}^2} \hat{z}_{k,m} e_{k,m}(x_j, y_l), \quad \hat{z}_{k,m} = \langle z, e_{k,m}^* \rangle_h.$$

Для функции  $u^h(x, y)$  из пространства сплайнов  $\tilde{S}^1(\mathbb{T}^2)$  и её сеточного варианта  $u^h$  справедливо соотношение  $\|u^h\|^2 = \langle u^h, u^h \rangle = \langle B_1 B_2 u^h, u^h \rangle_h$ , где  $B_1 B_2$  означает суперпозицию двух операторов по соответствующим переменным. Отсюда в силу (15) вытекает эквивалентность норм для функций из  $\tilde{S}^1(\mathbb{T}^2)$ :

$$\frac{1}{3} \|u^h\|_h \leq \|u^h\| \leq \|u^h\|_h. \tag{16}$$

Задачу (10) можно записать в аналогичном задаче (5) виде:

$$a(u^h, v^h) = \ell(v^h), \quad u^h \in \tilde{S}^1(\mathbb{T}^2) \quad \text{для любой функции} \quad v^h \in \tilde{S}^1(\mathbb{T}^2). \tag{17}$$

**Теорема 3.** Пусть  $g_1, g_2 \in L_2(\mathbb{T}^2)$ . Тогда задача (17) имеет единственное решение  $u^h \in \tilde{S}^1(\mathbb{T}^2)$  и справедлива оценка

$$\|u^h\|_{H^1(\mathbb{T}^2)} \leq \sqrt{2}\|g\|. \quad (18)$$

**Доказательство** аналогично доказательству теоремы 1.

Следующие теоремы дают оценки скорости сходимости  $u^h$  к функции  $u$  в различных нормах и при различных требованиях к гладкости измеряемых наклонов  $g_1, g_2$ .

**Теорема 4.** Пусть  $g_1 \in H^{1,0}(\mathbb{T}^2)$ ,  $g_2 \in H^{0,1}(\mathbb{T}^2)$ . Тогда имеют место оценки скорости сходимости решения  $u^h$  задачи (17) к решению  $u$  задачи (3):

$$\|u^h - u\|_{H^1(\mathbb{T}^2)} \leq C_1 h (\|g_1\|_{H^{1,0}(\mathbb{T}^2)} + \|g_2\|_{H^{0,1}(\mathbb{T}^2)}), \quad (19)$$

$$\|u^h - u\| \leq C_2 h^2 (\|g_1\|_{H^{1,0}(\mathbb{T}^2)} + \|g_2\|_{H^{0,1}(\mathbb{T}^2)}). \quad (20)$$

**Доказательство.** В силу симметричности билинейной формы, оценок (6), (7) и замечания 2.3.1 к лемме Сеа [14, с. 109] имеем

$$\|u^h - u\|_{H^1(\mathbb{T}^2)} \leq \sqrt{2} \inf_{v^h \in \tilde{S}^1(\mathbb{T}^2)} \|u - v^h\|_{H^1(\mathbb{T}^2)}.$$

Используя аппроксимационные свойства тензорного произведения кусочно-линейных сплайнов [17, с. 37], приходим к оценке с независимой от решения  $u$  и шага сетки константой  $C_1 > 0$ :

$$\|u^h - u\|_{H^1(\mathbb{T}^2)} \leq C_1 h \|u\|_{H^2(\mathbb{T}^2)}. \quad (21)$$

Воспользовавшись далее оценкой (8) из теоремы 2, окончательно получаем (19). При тех же требованиях к гладкости наклонов для получения оценки в норме  $L_2(\mathbb{T}^2)$  воспользуемся леммой Обэна–Нитше (см., например, [14, с. 139]). Имеем

$$\|u^h - u\| \leq M \|u^h - u\|_{H^2(\mathbb{T}^2)} \times \sup_{q \in L_2(\mathbb{T}^2)} \left( \frac{1}{\|q\|} \inf_{\phi^h \in \tilde{S}^1(\mathbb{T}^2)} \|\phi_q - \phi^h\|_{H^1(\mathbb{T}^2)} \right), \quad (22)$$

где  $\phi_q$  – единственное решение вариационной задачи

$$a(\phi_q, v) = \langle q, v \rangle \quad \text{для всех } v \in \tilde{H}^1(\mathbb{T}^2). \quad (23)$$

Инфимум в (22) оценивается аналогично (21) и подобно доказательству оценки (8) теоремы 2 имеем

$$\inf_{\phi^h \in \tilde{S}^1(\mathbb{T}^2)} \|\phi_q - \phi^h\|_{H^1(\mathbb{T}^2)} \leq C_1 h \|\phi_q\|_{H^2(\mathbb{T}^2)} \leq C_2 h \|q\|. \quad (24)$$

Подставив полученную оценку в (22) с учётом неравенства (19), окончательно получаем (20). Теорема доказана.

Исследуем сходимость метода в норме  $L_2(\mathbb{T}^2)$  при негладких наклонах  $g_1, g_2 \in L_2(\mathbb{T}^2)$ .

**Теорема 5.** Пусть  $g_1, g_2 \in L_2(\mathbb{T}^2)$ . Тогда справедлива оценка скорости сходимости

$$\|u^h - u\| \leq C_3 h \|g\|. \quad (25)$$

**Доказательство.** Обратимся снова к неравенству (22). Поскольку задача (23) не зависит от наклонов  $g_1, g_2$ , то для оценки инфимума воспользуемся (24). В силу неравенства треугольника и (4), (18) имеем

$$\|u^h - u\|_{H^1(\mathbb{T}^2)} \leq \|u^h\|_{H^1(\mathbb{T}^2)} + \|u\|_{H^1(\mathbb{T}^2)} \leq 2\sqrt{2}\|g\|.$$

Подставив полученные оценки в (22), приходим к (25). Теорема доказана.



**3. Проекционная схема со стабилизатором дробной гладкости.** Рассмотрим следующую модификацию схемы (5):

$$B_2 \Lambda_1 w^h + B_1 \Lambda_2 w^h + \gamma (\Lambda_1 \Lambda_2)^s w^h = f. \quad (26)$$

По сравнению с (5) схема (26) содержит зависящий от двух параметров  $\gamma > 0$ ,  $s > 0$  дополнительный стабилизатор  $\gamma (\Lambda_1 \Lambda_2)^s w^h$ . Это слагаемое отвечает за выравнивание частотной характеристики метода в области средних частот и приносит в уравнение смешанные производные дробного порядка. Различные способы выбора параметров стабилизатора обсуждаются в работе [10].

Действие стабилизатора реализуется в пространстве фурье-коэффициентов разложения по комплексным экспонентам  $e_{m,n}$  в виде мультипликатора  $\gamma (\lambda_m \lambda_n)^s$ , а для самих коэффициентов  $\hat{w}_{m,n}^h$  разложения искомого решения задачи (26) имеет место формула

$$\hat{w}_{m,n}^h = \frac{\hat{f}_{m,n}}{\mu_n \lambda_m + \mu_m \lambda_n + \gamma (\lambda_m \lambda_n)^s}, \quad 0 \neq (m, n) \in \mathbb{I}^2,$$

где  $\hat{f}_{m,n}$  – коэффициенты Фурье правой части  $f$ . Из (14), (15) вытекает корректность вычислений по формуле (26).

Далее потребуется вариационный аналог операторного уравнения (26):

$$a_s(w^h, v^h) = \ell(v^h), \quad w^h \in \tilde{S}^1(\mathbb{T}^2) \quad \text{при всех} \quad v^h \in \tilde{S}^1(\mathbb{T}^2), \quad (27)$$

где отвечающая задаче (26) билинейная форма имеет вид

$$a_s(w^h, v^h) = \langle u_x^h, v_x^h \rangle + \langle w_y^h, v_y^h \rangle + \gamma \langle B_1^{-1} B_2^{-1} (\Lambda_1 \Lambda_2)^s u^h, v^h \rangle. \quad (28)$$

Отметим, что в формуле (28) в скалярном произведении в четвертом слагаемом использовано обозначение для восполнения сеточной функции

$$B_1^{-1} B_2^{-1} (\Lambda_1 \Lambda_2)^s w^h = \sum_{(k,l) \in \mathbb{I}^2} (B_1^{-1} B_2^{-1} (\Lambda_1 \Lambda_2)^s w^h)_{k,l} \phi_{k,l}(x, y) \in \tilde{S}^1(\mathbb{T}^2).$$

**Теорема 6.** Пусть  $g_1, g_2 \in L_2(\mathbb{T}^2)$ . Тогда задача (26) (задача (27)) имеет единственное решение  $w^h \in \tilde{S}^1(\mathbb{T}^2)$  и справедлива оценка

$$\|w^h\|_{H^1(\mathbb{T}^2)} \leq C_1 \|g\|. \quad (29)$$

Пусть  $g_1 \in H^{1,0}(\mathbb{T}^2)$ ,  $g_2 \in H^{0,1}(\mathbb{T}^2)$ . Тогда справедлива оценка

$$\|\Lambda_j w^h\|_h \leq C_2 (\|g_1\|_{H^{1,0}(\mathbb{T}^2)} + \|g_2\|_{H^{0,1}(\mathbb{T}^2)}), \quad j = 1, 2. \quad (30)$$

**Доказательство** проводится аналогично доказательству теоремы 1. Существование вытекает из (26), где знаменатель не обращается в нуль. Для вывода оценки (29) возьмём в (27)  $v^h = w^h$  в качестве пробной функции. Имеем

$$a(w^h, w^h) + \gamma \langle B_1^{-1} B_2^{-1} (\Lambda_1 \Lambda_2)^s w^h, w^h \rangle = \ell(w^h).$$

Перейдём к дискретным функциям во втором слагаемом левой части и получим

$$a(w^h, w^h) + \gamma \langle (\Lambda_1 \Lambda_2)^s w^h, w^h \rangle_h = \ell(w^h).$$

Воспользовавшись далее коэрцитивностью билинейной формы  $a(w^h, w^h)$  (оценка (7)), непрерывностью функционала  $\ell(w^h)$  (оценка (6)), а также вытекающей из оценки (14) неотрицательностью операторов  $\Lambda_1, \Lambda_2$ , приходим к неравенству (29).

Для доказательства оценки (30) воспользуемся (26), равенством Парсеваля и неравенствами (15):

$$\begin{aligned} \|\Lambda_1 w^h\|_h^2 &= \sum_{(0,0) \neq (m,n) \in \mathbb{I}^2} \lambda_m^2 |\hat{w}_{m,n}^h|^2 = \sum_{(0,0) \neq (m,n) \in \mathbb{I}^2} \frac{\lambda_m^2 |\hat{f}_{m,n}|^2}{(\mu_n \lambda_m + \mu_m \lambda_n + \gamma(\lambda_m \lambda_n)^s)^2} \leq \\ &\leq C_1 \sum_{(0,0) \neq (m,n) \in \mathbb{I}^2} |\hat{f}_{m,n}|^2 = C_1 \|f\|_h^2 \leq C_2 (\|g_1\|_{H^{1,0}(\mathbb{T}^2)}^2 + \|g_2\|_{H^{0,1}(\mathbb{T}^2)}^2). \end{aligned}$$

Теорема доказана.

Исследуем близость решения  $w^h$  к точному решению  $u$  задачи (3).

**Теорема 7.** Пусть  $g_1, g_2 \in L_2(\mathbb{T}^2)$ . Тогда справедлива оценка скорости сходимости

$$\|w^h - u\| \leq (C_3 h + C_4 \gamma \lambda_{N/2}^{\sigma(s)}) \|g\|, \quad \sigma(s) = \max\{2s - 3/2, 0\}. \tag{31}$$

Пусть  $g_1 \in H^{1,0}(\mathbb{T}^2)$ ,  $g_2 \in H^{0,1}(\mathbb{T}^2)$ . Тогда при  $0.25 \leq s$  справедлива оценка скорости сходимости

$$\|w^h - u\|_{H^1(\mathbb{T}^2)} \leq (C_5 h + C_6 \gamma \lambda_{N/2}^{\sigma(s)}) (\|g_1\|_{H^{1,0}(\mathbb{T}^2)} + \|g_2\|_{H^{0,1}(\mathbb{T}^2)}). \tag{32}$$

При  $0 < s < 0.25$  оценка имеет вид

$$\|w^h - u\|_{H^1(\mathbb{T}^2)} \leq C_7 h (\|g_1\|_{H^{1,0}(\mathbb{T}^2)} + \|g_2\|_{H^{0,1}(\mathbb{T}^2)}) + C_8 \gamma \|g\|. \tag{33}$$

Также справедлива оценка

$$\|w^h - u\| \leq (C_9 h^2 + C_{10} \gamma \lambda_{N/2}^{\eta(s)}) (\|g_1\|_{H^{1,0}(\mathbb{T}^2)} + \|g_2\|_{H^{0,1}(\mathbb{T}^2)}), \quad \eta(s) = \max\{2s - 2, 0\}. \tag{34}$$

**Доказательство.** В силу неравенства треугольника имеем

$$\|w^h - u\| \leq \|w^h - u^h\| + \|u^h - u\|, \quad \|w^h - u\|_{H^1(\mathbb{T}^2)} \leq \|w^h - u^h\|_{H^1(\mathbb{T}^2)} + \|u^h - u\|_{H^1(\mathbb{T}^2)}.$$

В правых частях этих неравенств нормы разности  $u^h - u$  оцениваются с помощью (19), (20) и (25). Осталось оценить нормы разности  $z^h = w^h - u^h$  решения  $w^h$  задачи (26) и решения  $u^h$  задачи (12). Вычитая почленно равенства (26) и (12), приходим к уравнению

$$B_2 \Lambda_1 z^h + B_1 \Lambda_2 z^h + \gamma (\Lambda_1 \Lambda_2)^s w^h = 0 \tag{35}$$

или, в эквивалентной вариационной форме,

$$a(z^h, v^h) + \gamma \langle B_1^{-1} B_2^{-1} (\Lambda_1 \Lambda_2)^s w^h, v^h \rangle = 0 \quad \text{при каждой } v^h \in \tilde{S}^1(\mathbb{T}^2).$$

Полагая  $v^h = z^h$  и используя коэрцитивность билинейной формы, в случае  $0.25 \leq s$  имеем

$$\begin{aligned} 0.5 \|z^h\|_{H^1(\mathbb{T}^2)}^2 &\leq a(z^h, z^h) = -\gamma \langle B_1^{-1} B_2^{-1} (\Lambda_1 \Lambda_2)^s w^h, z^h \rangle = \\ &= -\gamma \langle (\Lambda_1 \Lambda_2)^s w^h, z^h \rangle_h = -\gamma \langle (\Lambda_1 \Lambda_2)^{s-1/4} w^h, (\Lambda_1 \Lambda_2)^{1/4} z^h \rangle_h \leq \\ &\leq \gamma \|(\Lambda_1 \Lambda_2)^{s-1/4} w^h\|_h \|(\Lambda_1 \Lambda_2)^{1/4} z^h\|_h \leq C_1 \gamma \|(\Lambda_1 \Lambda_2)^{s-1/4} w^h\|_h \|z^h\|_{H^1(\mathbb{T}^2)}. \end{aligned}$$

Отсюда следуют неравенства

$$\|z^h\|_{H^1(\mathbb{T}^2)} \leq \gamma C_2 \|(\Lambda_1 \Lambda_2)^{s-1/4} w^h\|_h \leq \gamma \lambda_{N/2}^{\sigma(s)} C_3 \|(\Lambda_1 \Lambda_2)^{1/2} w^h\|_h.$$

С помощью оценки (30) имеем

$$\|(\Lambda_1 \Lambda_2)^{1/2} w^h\|_h \leq C_3 (\|\Lambda_1 w^h\|_h + \|\Lambda_2 w^h\|_h) \leq C_4 (\|g_1\|_{H^{1,0}(\mathbb{T}^2)} + \|g_2\|_{H^{0,1}(\mathbb{T}^2)}).$$

Следовательно,

$$\|z^h\|_{H^1(\mathbb{T}^2)} \leq \gamma \lambda_{N/2}^{\sigma(s)} C_5 (\|g_1\|_{H^{1,0}(\mathbb{T}^2)} + \|g_2\|_{H^{0,1}(\mathbb{T}^2)}).$$

В более простом случае  $0 < s < 0.25$  воспользуемся оценкой (29):

$$0.5 \|z^h\|_{H^1(\mathbb{T}^2)}^2 \leq \gamma C_1 \|w^h\|_h (\Lambda_1 \Lambda_2)^{1/4} z^h \|z^h\|_h \leq \gamma C_2 \|w^h\|_{H^1(\mathbb{T}^2)} \|z^h\|_{H^1(\mathbb{T}^2)} \leq \gamma C_3 \|g\| \|z^h\|_{H^1(\mathbb{T}^2)},$$

в результате получим

$$\|z^h\|_{H^1(\mathbb{T}^2)} \leq \gamma C_3 \|g\|.$$

Применив далее оценку (19), приходим к (32) и (33).

Получим для  $z^h$  оценки дискретной нормы  $L_2$ . Из уравнения (35) вытекает выражение для соответствующих коэффициентов Фурье

$$\hat{z}_{m,n}^h = -\frac{\gamma (\lambda_m \lambda_n)^s \hat{w}_{m,n}^h}{\mu_n \lambda_m + \mu_m \lambda_n}.$$

Отсюда в случае  $s > 0.75$  с использованием (14), (15) и неравенства  $(\lambda_m \lambda_n)^{1/2} \leq (\lambda_m + \lambda_n)/\sqrt{2}$  имеем

$$\begin{aligned} \|z^h\|_h^2 &= \sum_{(0,0) \neq (m,n) \in \mathbb{I}^2} |\hat{z}_{m,n}^h|^2 = \gamma^2 \sum_{(0,0) \neq (m,n) \in \mathbb{I}^2} \frac{(\lambda_m \lambda_n)^{2s} |\hat{w}_{m,n}^h|^2}{(\mu_n \lambda_m + \mu_m \lambda_n)^2} \leq \\ &\leq C_5 \gamma^2 \sum_{(0,0) \neq (m,n) \in \mathbb{I}^2} \frac{(\lambda_m \lambda_n)^{2s-1/2} (\mu_n \lambda_m + \mu_m \lambda_n) |\hat{w}_{m,n}^h|^2}{(\mu_n \lambda_m + \mu_m \lambda_n)^2} \leq C_5 \gamma^2 \lambda_{N/2}^{4s-3} \|w^h\|_{H^1(\mathbb{T}^2)}^2 \leq C_6 \gamma^2 \lambda_{N/2}^{4s-3} \|g\|^2. \end{aligned}$$

Тогда с учётом (16) получаем

$$\|z^h\| \leq C_7 \gamma \lambda_{N/2}^{2s-3/2} \|g\|.$$

При  $0 < s \leq 0.75$  имеем  $(\lambda_m \lambda_n)^{2s} \leq C_9 (\mu_n \lambda_m + \mu_m \lambda_n)^3$  и оценка нормы принимает вид

$$\|z^h\| \leq C_9 \gamma \|g\|.$$

Объединив полученные оценки с неравенством (25), заключаем, что при  $g \in L_2(\mathbb{T}^2)$  справедливо неравенство

$$\|w^h - u\| \leq (C_3 h + C_{10} \gamma \lambda_{N/2}^{\sigma(s)}) \|g\|.$$

Оценка (31) доказана.

При дополнительной гладкости наклонов  $g_1 \in H^{1,0}(\mathbb{T}^2)$ ,  $g_2 \in H^{0,1}(\mathbb{T}^2)$  воспользуемся неравенством

$$(\lambda_m \lambda_n)^{2s} \leq C_{12} \lambda_{N/2}^{2\eta(s)} (\mu_n \lambda_m + \mu_m \lambda_n)^2 (\lambda_m + \lambda_n)^2, \quad \eta(s) = \max\{2s - 2, 0\}.$$

Тогда с учётом неравенств (30) имеем

$$\begin{aligned} \|z^h\|_h^2 &= \gamma^2 \sum_{(0,0) \neq (m,n) \in \mathbb{I}^2} \frac{(\lambda_m \lambda_n)^{2s} |\hat{w}_{m,n}^h|^2}{(\mu_n \lambda_m + \mu_m \lambda_n)^2} \leq \\ &\leq C_{11} \gamma^2 \lambda_{N/2}^{2\eta(s)} \sum_{(0,0) \neq (m,n) \in \mathbb{I}^2} (\lambda_m + \lambda_n)^2 |\hat{w}_{m,n}^h|^2 \leq C_{12} \gamma^2 \lambda_{N/2}^{2\eta(s)} (\|\Lambda_1 w^h\|_h^2 + \|\Lambda_2 w^h\|_h^2) \leq \\ &\leq C_{13} \gamma^2 \lambda_{N/2}^{2\eta(s)} (\|g_1\|_{H^{1,0}(\mathbb{T}^2)}^2 + \|g_2\|_{H^{0,1}(\mathbb{T}^2)}^2). \end{aligned}$$

С помощью (20) получаем неравенство (34). Теорема доказана.

Обсудим согласование параметров в оценках теоремы 7. Полученные в теореме 7 оценки показывают, что для сходимости проекционного метода выбор параметров  $\gamma$  и  $s$  необходимо согласовывать с шагом сетки  $h$  в зависимости от гладкости наклонов и выбора нормы. Учтём, что в силу (13)  $\lambda_{N/2} = 4h^{-2}$ . Тогда для случая  $g_1, g_2 \in L_2(\mathbb{T}^2)$  при согласовании  $0 < \gamma < C_{14}h^{1+2\sigma(s)}$  в силу (31) имеет место оценка

$$\|w^h - u\| \leq C_{15}h\|g\|. \tag{36}$$

Если же  $g_1 \in H^{1,0}(\mathbb{T}^2)$ ,  $g_2 \in H^{0,1}(\mathbb{T}^2)$ , то при  $0 < \gamma < C_{16}h^{2+2\eta(s)}$  в силу (34) справедлива оценка скорости сходимости

$$\|w^h - u\| \leq C_{17}h^2(\|g_1\|_{H^{1,0}(\mathbb{T}^2)} + \|g_2\|_{H^{0,1}(\mathbb{T}^2)}), \tag{37}$$

а при  $0 < \gamma < C_{18}h^{1+2\sigma(s)}$  в силу (32) имеем

$$\|w^h - u\|_{H^1(\mathbb{T}^2)} \leq C_{19}h(\|g_1\|_{H^{1,0}(\mathbb{T}^2)} + \|g_2\|_{H^{0,1}(\mathbb{T}^2)}). \tag{38}$$

Таким образом, при выполнении вычислительных экспериментов для наилучшего использования свойства выравнивания частотной характеристики схемы при одновременном сохранении точности вычислений параметр  $\gamma$  следует брать достаточно малым. Результаты вычислительных экспериментов подтверждают справедливость сделанных теоретических оценок точности с согласованным выбором параметров.

**4. Оценки скорости сходимости в пространствах дробной гладкости.** В данном пункте с помощью интерполяции оператора погрешности выведем оценки скорости сходимости в анизотропных пространствах дробной гладкости. Введём необходимые обозначения. Пусть  $X$  и  $Y$  – гильбертовы пространства со скалярными произведениями  $\langle \cdot, \cdot \rangle_X$ ,  $\langle \cdot, \cdot \rangle_Y$  и соответствующими евклидовыми нормами  $\|\cdot\|_X$ ,  $\|\cdot\|_Y$ , причём  $X$  вложено в  $Y$ , всюду плотно и непрерывно. Из [11, п. 2.1] известно, что существует  $\Lambda = \Lambda^* > 0$  – некоторый, вообще говоря, неограниченный оператор в  $Y$  с  $D(\Lambda) = X$ , причём норма в  $X$  эквивалентна норме графика  $\Lambda$ . Следуя [11], для  $\theta \in [0, 1]$  рассмотрим промежуточные пространства Лионса  $[X, Y]_\theta = D(\Lambda^{1-\theta})$  с нормой графика  $\|u\|_{[X, Y]_\theta} = (\|u\|_Y^2 + \|\Lambda^{1-\theta}u\|_Y^2)^{1/2}$ . Очевидно, что  $[X, Y]_0 = X$ ,  $[X, Y]_1 = Y$ , т.е. пространства  $[X, Y]_\theta$  непрерывно соединяют крайние пространства  $X$  и  $Y$ . В частности, аналогично [11, гл. 1, ш. 9 и 13] с использованием теоремы об интерполяции пересечений пространств проверяется, что для  $\theta \in [0, 1]$  справедливы равенства

$$[H^1(\mathbb{T}^2), L_2(\mathbb{T}^2)]_\theta = H^{1-\theta}(\mathbb{T}^2), \tag{39}$$

$$[H^{1,0}(\mathbb{T}^2) \times H^{0,1}(\mathbb{T}^2), L_2(\mathbb{T}^2) \times L_2(\mathbb{T}^2)]_\theta = H^{1-\theta,0}(\mathbb{T}^2) \times H^{0,1-\theta}(\mathbb{T}^2). \tag{40}$$

Также потребуются интерполяционные пространства, построенные с помощью К-метода. Рассмотрим функционалы

$$K_p(t, a; X, Y) = \inf_{a_0+a_1=a, a_0 \in X, a_1 \in Y} (\|a_0\|_X^p + t^p\|a_1\|_Y^p)^{1/p}, \quad p \in \{1, 2\}.$$

Согласно [18, п. 1.3.2] при  $\theta \in (0, 1)$  определим интерполяционные пространства

$$(X, Y)_{\theta, 2} = \left\{ a \in Y : \|a\|_{(X, Y)_{\theta, 2}} = \left( \int_0^{+\infty} t^{-1-2\theta} K_1^2(t, a; X, Y) dt \right)^{1/2} < +\infty \right\}.$$

Поскольку  $K_2 \leq K_1 \leq \sqrt{2}K_2$ , то использование функционала  $K_2$  в предыдущем определении даёт эквивалентную норму. Отсюда в силу [11, теорема 15.1] при каждом  $\theta \in (0, 1)$  вытекает

эквивалентность норм пространств  $[X, Y]_\theta$  и  $(X, Y)_{\theta,2}$ . Однако равномерной по параметру  $\theta$  эквивалентности нет в силу полученного в упомянутой теореме равенства

$$\int_0^{+\infty} t^{-1-2\theta} K_2^2(t, a; X, Y) dt = C(\theta) \|\Lambda^{1-\theta} a\|_Y^2 \tag{41}$$

с неограниченным при  $\theta \rightarrow 0 + 0$ ,  $\theta \rightarrow 1 - 0$  множителем

$$C(\theta) = \int_0^{+\infty} \frac{s^{1-2\theta}}{1+s^2} ds = \frac{\pi}{2 \sin(\pi\theta)}$$

(см. также [19, с. 116]).

Если  $\mathcal{X}$  и  $\mathcal{Y}$  – другая пара пространств со свойствами, аналогичными свойствам пары  $X$  и  $Y$ , то имеет место следующее утверждение (см., например, [18, теорема 1.3.3]).

**Теорема 8.** Пусть для двух пар пространств  $X, Y$  и  $\mathcal{X}, \mathcal{Y}$  задан линейный оператор  $T \in L(X \rightarrow \mathcal{X}) \cap L(Y \rightarrow \mathcal{Y})$ . Тогда  $T \in L((X, Y)_{\theta,2} \rightarrow (\mathcal{X}, \mathcal{Y})_{\theta,2})$  и справедлива оценка

$$\|T\|_{L((X, Y)_{\theta,2} \rightarrow (\mathcal{X}, \mathcal{Y})_{\theta,2})} \leq M_0^{1-\theta} M_1^\theta, \quad M_0 = \|T\|_{L(X \rightarrow \mathcal{X})}, \quad M_1 = \|T\|_{L(Y \rightarrow \mathcal{Y})}. \tag{42}$$

Нам потребуется следствие из этой теоремы для промежуточных пространств Лионса. Для любого  $a \in [X, Y]_\theta$  в силу (42) имеем неравенство

$$\|Ta\|_{(\mathcal{X}, \mathcal{Y})_{\theta,2}} \leq M_0^{1-\theta} M_1^\theta \|a\|_{(X, Y)_{\theta,2}}. \tag{43}$$

Поскольку при переходе от нормы пространства К-метода к норме пространства Лионса фигурирующий в (41) множитель  $C(\theta)$  входит в обе части неравенства (43), то после несложных перенормировок получаем неравенство в пространствах Лионса с независимой от  $\theta \in [0, 1]$  константой  $C_{20} > 0$ :

$$\|Ta\|_{[\mathcal{X}, \mathcal{Y}]_\theta} \leq C_{20} M_0^{1-\theta} M_1^\theta \|a\|_{[X, Y]_\theta}. \tag{44}$$

**Теорема 9.** Пусть выполнено условие согласования параметров схемы и шага сетки  $0 < \gamma < C_{21} h^{1+2\sigma(s)}$ ,  $0 < h < 1$ . Тогда для любых  $\alpha \in [0, 1]$  справедливы следующие оценки:

$$\|w^h - u\|_{H^\alpha(\mathbb{T}^2)} \leq C_{22} h (\|g_1\|_{H^{\alpha,0}(\mathbb{T}^2)} + \|g_2\|_{H^{0,\alpha}(\mathbb{T}^2)}), \tag{45}$$

$$\|w^h - u\|_{L_2(\mathbb{T}^2)} \leq C_{23} h^{1+\alpha} (\|g_1\|_{H^{\alpha,0}(\mathbb{T}^2)} + \|g_2\|_{H^{0,\alpha}(\mathbb{T}^2)}), \tag{46}$$

$$\|w^h - u\|_{H^\alpha(\mathbb{T}^2)} \leq C_{24} h^{2-\alpha} (\|g_1\|_{H^{1,0}(\mathbb{T}^2)} + \|g_2\|_{H^{0,1}(\mathbb{T}^2)}). \tag{47}$$

**Доказательство.** Рассмотрим оператор погрешности проекционной схемы

$$R : g = (g_1, g_2) \mapsto w^h - u$$

на разных сочетаниях пар пространств, участвующих в оценках (36)–(38).

Начнём с оценок (36), (38). Имеем  $R \in L(L_2(\mathbb{T}^2) \times L_2(\mathbb{T}^2) \rightarrow L_2(\mathbb{T}^2))$  с операторной нормой  $M_1 \leq C_{15} h$  и  $R \in L(H^{1,0}(\mathbb{T}^2) \times H^{0,1}(\mathbb{T}^2) \rightarrow H^1(\mathbb{T}^2))$  с операторной нормой  $M_0 \leq C_{19} h$ . Применим неравенство (44), в котором в роли  $T$  выступает оператор погрешности  $R$  на парах пространств  $Y = L_2(\mathbb{T}^2) \times L_2(\mathbb{T}^2)$ ,  $\mathcal{Y} = L_2(\mathbb{T}^2)$  и  $X = H^{1,0}(\mathbb{T}^2) \times H^{0,1}(\mathbb{T}^2)$ ,  $\mathcal{X} = H^1(\mathbb{T}^2)$ . Тогда с учётом соотношений (39), (40) неравенство (44) принимает вид

$$\|w^h - u\|_{H^{1-\theta}(\mathbb{T}^2)} \leq C_{22} h (\|g_1\|_{H^{1-\theta,0}(\mathbb{T}^2)} + \|g_2\|_{H^{0,1-\theta}(\mathbb{T}^2)}).$$

Обозначив  $\alpha = 1 - \theta$ , получим (45).

Из рассмотрения оценок (36), (37) вытекает возможность применить для оператора погрешности  $R$  неравенство (44) на парах пространств  $Y = L_2(\mathbb{T}^2) \times L_2(\mathbb{T}^2)$ ,  $\mathcal{Y} = L_2(\mathbb{T}^2)$  с операторной нормой  $M_1 \leq C_{15}h$  и  $X = H^{1,0}(\mathbb{T}^2) \times H^{0,1}(\mathbb{T}^2)$ ,  $\mathcal{X} = L_2(\mathbb{T}^2)$  с операторной нормой  $M_0 \leq C_{17}h^2$ . Положив  $\theta = 1 - \alpha$ , приходим к оценке (46).

Наконец, оценки (37), (38) позволяют рассмотреть оператор погрешности на парах пространств  $Y = H^{1,0}(\mathbb{T}^2) \times H^{0,1}(\mathbb{T}^2)$ ,  $\mathcal{Y} = L_2(\mathbb{T}^2)$  с операторной нормой  $M_1 \leq C_{17}h^2$  и  $X = H^{1,0}(\mathbb{T}^2) \times H^{0,1}(\mathbb{T}^2)$ ,  $\mathcal{X} = H^1(\mathbb{T}^2)$  с операторной нормой  $M_0 \leq C_{19}h$ . Положив  $\theta = 1 - \alpha$  и применив (44), получим (47). Теорема доказана.

**5. Пример восстановления негладкого волнового фронта по неточно заданным наклонам.** Качество восстановления негладкого волнового фронта продемонстрируем на примере поверхности, задаваемой уравнением

$$u(x, y, n, \varkappa) = u_0 + \exp\{-(r/a)^{2M}\}U(x, y, n, \varkappa),$$

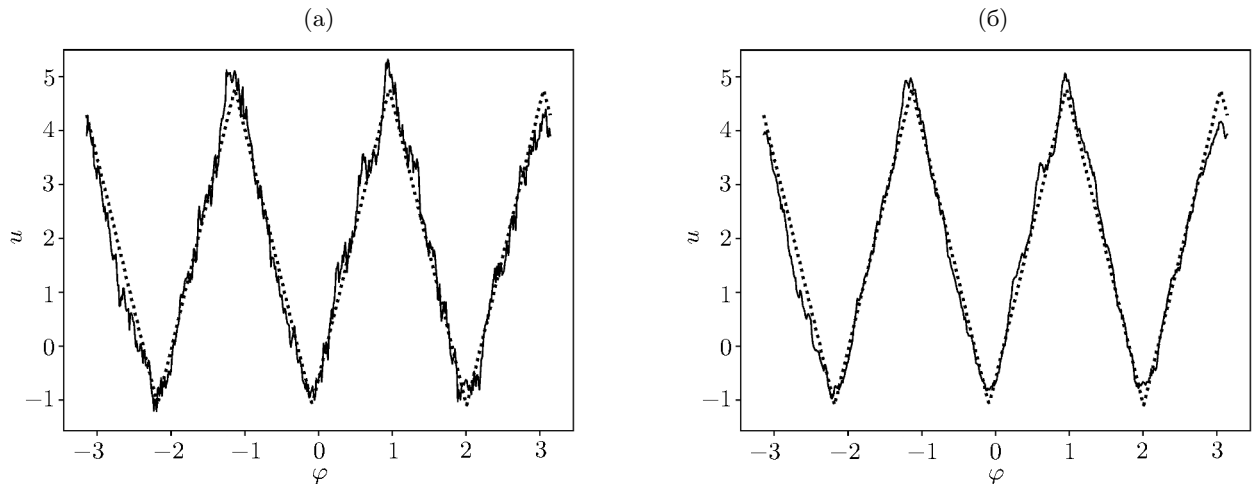
$$U(x, y, n, \varkappa) = \begin{cases} (\varkappa r + 2n(\varphi - \pi/2) + \pi/6) \bmod 2\pi, & \text{если } \left[ \frac{\varkappa r + 2n(\varphi - \pi/2) + \pi/6}{2\pi} \right] \text{ нечётное;} \\ 2\pi - (\varkappa r + 2n(\varphi - \pi/2) + \pi/6) \bmod 2\pi, & \text{если } \left[ \frac{\varkappa r + 2n(\varphi - \pi/2) + \pi/6}{2\pi} \right] \text{ чётное,} \end{cases} \quad (48)$$

где  $\varphi = \operatorname{argctg}(y/x)$ ,  $r = \sqrt{x^2 + y^2}$ ,  $n \in \mathbb{N}$ ,  $\varkappa \in \mathbb{R}$ ,  $[\cdot]$  – целая часть числа. Функция из (48) описывает многолепестковую (при  $\varkappa = 0$ ) и спиральную (при  $\varkappa \neq 0$ ) структуры с  $n$  рукавами, при этом параметр  $\varkappa$  отвечает за степень закрученности спирали. Такая функция служит адекватной моделью для кусочно-гладких фазовых фронтов, встречающихся при работе датчиков Шака–Гартмана. Константа  $u_0$  выбирается из условия равенства нулю среднего значения функции,  $2\pi$ -периодичность рассматриваемой функции по переменным  $x$ ,  $y$  обеспечивалась приближённым равенством нулю самой функции и её производных при достаточно большом  $M$  (в расчётах брались значения  $a = 0.7\pi$ ,  $M = 4$ ).

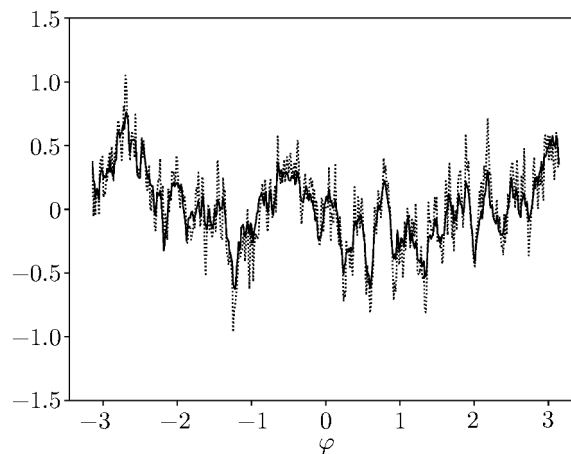
В приведённом примере для численного моделирования в качестве исходных данных брались дискретные значения точных наклонов волнового фронта (48), к которым далее добавлялся 5%-й гауссов шум для имитации искажений измерений наклонов на сетке  $512 \times 512$ . Далее проводилось восполнение дискретных наклонов из пространства сплайнов  $S^1(\mathbb{T}^2)$  и находилась правая часть в (12), которая в этом случае выражается через разностные производные от дискретных наклонов (подробнее см. в [9]). Результаты вычислений на примере волнового фронта (48) с параметрами  $\varkappa = 2$ ,  $n = 3$  иллюстрируются на рис. 1, а для схемы без стабилизатора ( $\gamma = 0$ ) и на рис. 1, б для схемы со стабилизатором ( $\gamma = 0.03$ ,  $s = 0.9$ ) в виде сечений волнового фронта окружностью  $r = \pi/2$ ,  $0 \leq \varphi \leq 2\pi$ . На рис. 2 приведены графики сечений отклонений от искомого волнового фронта (48) полученного решения со стабилизатором и без него. Сравнение показывает, что схема уверенно справляется с восстановлением кусочно-гладких фаз. При этом линии разрывов производной точно локализуются и в областях вне разрывов имеет место хорошее поточечное приближение. Как видно из графиков, использование стабилизатора приводит к эффективному локальному сглаживанию паразитных искажений волнового фронта, вызванных неточностью измерений наклонов.

**Заключение.** В работе рассмотрена проблема восстановления волнового фронта по его наклонам, которая возникает в адаптивной оптике и связана с определением фазы световой волны на основе датчика Шака–Гартмана. Вариационная постановка задачи приводит к уравнению Пуассона на подпространстве периодических функций с нулевым средним, при этом свойство коэрцитивности задачи вытекает из неравенства Пуанкаре для таких функций. При условии согласования шага сетки с параметрами стабилизатором дробного порядка получены оценки точности проекционного метода на основе билинейных сплайнов при естественных требованиях к гладкости волнового фронта из анизотропных пространств Соболева с целыми показателями гладкости. На основе выбора специальных анизотропных пространств с дробными показателями гладкости, составляющих непрерывную шкалу, соединяющую пространства

с целыми показателями гладкости, и использования связанных с ними методов интерполяции оператора погрешности проекционного метода получены шкалы оценок точности метода, согласованные с дробной гладкостью наклонов. Теоретические оценки иллюстрируются результатами численного эксперимента по эффективному восстановлению кусочно-негладких спиральных волновых фронтов в условиях неточно заданных дискретных наклонов.



**Рис. 1.** Графики сечений при  $r = \pi/2$  искомого фронта  $u(x, y, 3, 2)$  (пунктирная линия) и  $w^h$  (сплошная линия) без стабилизатора (а) и со стабилизатором (б).



**Рис. 2.** Графики сечений при  $r = \pi/2$  отклонений  $w^h$  от искомого фронта  $u(x, y, 3, 2)$  без стабилизатора (пунктирная линия) и со стабилизатором (сплошная линия).

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект 18-29-02103) (А.В. Разгулин, Н.Г. Ирошников, А.В. Ларичев) и Министерства образования и науки Российской Федерации в рамках реализации программы Московского центра фундаментальной и прикладной математики (соглашение 075-15-2022-284) (А.В. Разгулин, Т.Е. Романенко).

#### СПИСОК ЛИТЕРАТУРЫ

1. Platt B.C., Shack R. History and principles of Shack–Hartmann wavefront sensing // J. Refract. Surg. 2001. V. 17. № 5. P. S573–S577.
2. Andreeva M.S., Iroshnikov N.G., Koryabin A.B., Larichev A.V., Shmalgauzen V.I. Usage of wavefront sensor for estimation of atmospheric turbulence parameters // Optoelectronics, Instrumentation and Data Processing. 2012. V. 48. P. 197–204.

3. *Baryshnikov N.V., Denisov D.G., Dzhumamuratova A.A., Larichev A.V.* Development and research of an optoelectronic device based on a wavefront sensor to control the form parameters of intraocular lenses // *Measurement Techniques*. 2019. V. 62. P. 31–35.
4. *Goncharov A.S., Iroshnikov N.G., Larichev A.V., Nikolaev I.P.* The impact of speckle on the measurement of eye aberrations // *J. of Modern Optics*. 2015. V. 62. № 21. P. 1775–1780.
5. *Bahk S.W.* Highly accurate wavefront reconstruction algorithms over broad spatial-frequency bandwidth // *Optics Express*. 2011. V. 19. № 20. P. 18997–19014.
6. *Zhariy M., Neubauer A., Rosensteiner M., Ramlau R.* Cumulative wavefront reconstructor for the Shack–Hartmann sensor // *Inverse Problems and Imaging*. 2011. V. 5. № 4. P. 893–913.
7. *Neubauer A.* On the ill-posedness and convergence of the Shack–Hartmann based wavefront reconstruction // *J. of Inverse and Ill-posed Problems*. 2010. V. 18. № 5. P. 551–576.
8. *Kindermann S., Neubauer A., Ramlau R.* A singular value decomposition for the Shack–Hartmann based wavefront reconstruction // *J. of Comput. and Appl. Math.* 2012. V. 236. № 8. P. 2186–2199.
9. *Razgulin A.V., Kuzhamaliyev Y.Z., Goncharov A.S., Larichev A.V.* A variational method for wavefront reconstruction from Shack–Hartman sensor measurements // *Atmospheric and Oceanic Optics*. 2017. V. 30. № 4. P. 399–403.
10. *Razgulin A.V., Kuzhamaliyev Y.Z., Iroshnikov N.G., Larichev A.V.* A variational method of wavefront reconstruction from local slope measurements using a fractional order of smoothness stabilizer // *Comput. Math. and Model.* 2019. V. 30. № 2. P. 164–176.
11. *Лионс Ж.-Л., Маджисенс Э.* Неоднородные граничные задачи и их приложения. М., 1971.
12. *Adams R., Fournier J.* Sobolev Spaces. Pure and Applied Mathematics. V. 140. Amsterdam, 2003.
13. *Kuznetsov N., Nazarov A.* Sharp constants in the Poincare, Steklov and related inequalities (a survey) // *Mathematika*. 2015. V. 61. № 2. P. 328–344.
14. *Сьярле Ф.* Метод конечных элементов для эллиптических задач. М., 1980.
15. *Обэн Ж.-П.* Приближенное решение эллиптических краевых задач. М., 1977.
16. *Марчук Г.И., Агошков В.И.* Введение в проекционно-сеточные методы. М., 1981.
17. *Варга Р.* Функциональный анализ и теория аппроксимации в численном анализе. М., 1974.
18. *Трибель Х.* Теория интерполяции, функциональные пространства, дифференциальные операторы. М., 1980.
19. *Tartar L.* An introduction to Sobolev spaces and interpolation spaces. Berlin; Heidelberg, 2007.

Московский государственный университет  
имени М.В. Ломоносова,  
ООО “ВК Цифровые технологии”,  
г. Москва

Поступила в редакцию 18.02.2022 г.  
После доработки 18.02.2022 г.  
Принята к публикации 25.05.2022 г.