

Том 59, Номер 7

ISSN 0374-0641  
Июль 2023



# ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ



[www.sciencejournals.ru](http://www.sciencejournals.ru)



# СОДЕРЖАНИЕ

---

---

Том 59, номер 7, 2023

---

---

## ОБЫКНОВЕННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

- О некоторых свойствах отображения сдвига на бесконечномерном торе  
*С. Д. Глызин, А. Ю. Колесов* 867
- Обобщённые решения первой краевой задачи для дифференциально-разностного уравнения  
в дивергентном виде на интервале конечной длины  
*А. Л. Скубачевский, Н. О. Иванов* 881
- 

## УРАВНЕНИЯ С ЧАСТНЫМИ ПРОИЗВОДНЫМИ

- О разрушении решения одной задачи для соболевского уравнения с некоэрцитивным  
источником  
*М. В. Артемьева, М. О. Корпусов* 893
- Существование и единственность классического решения первой краевой задачи  
для параболических систем на плоскости  
*А. Н. Коненков* 904
- 

## ИНТЕГРАЛЬНЫЕ И ИНТЕГРО-ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

- О гиперсингулярных операторах, связанных с перидинамикой  
*Ш. А. Алимов, Ш. Н. Шералиев* 914
- 

## ТЕОРИЯ УПРАВЛЕНИЯ

- Вариационные методы решения задач, связанных с искусственным интеллектом  
*М. В. Бабушкин, В. Ю. Тертыйный-Даури* 919
- Линейная задача группового преследования с дробными производными, простыми  
матрицами и разными возможностями игроков  
*Н. Н. Петров, А. И. Мачтаков* 933
- 

## ЧИСЛЕННЫЕ МЕТОДЫ

- Разностные схемы декомпозиции на основе расщепления решения и оператора задачи  
*П. Н. Вабищевич* 944
- Глобально устойчивые разностные схемы для уравнения Фишера  
*П. П. Матус, Д. Пылак* 960
- К вопросу о численном решении неконсервативных гиперболических систем уравнений  
*Р. Р. Полежаина, М. В. Алексеев, Е. Б. Савенков* 968

Сеточно-характеристический метод повышенного порядка для систем гиперболических уравнений с кусочно-постоянными коэффициентами

*Н. И. Хохлов, И. Б. Петров*

983

---

### КРАТКИЕ СООБЩЕНИЯ

Существование единственной неподвижной точки отображений, порождённых многомерной системой с релейным гистерезисом

*А. М. Камачкин, В. В. Евстафьева, Д. К. Потапов*

996

Об ограниченных траекториях автономной системы с выделенной положительно однородной нелинейностью

*Э. Мухамадиев, А. Н. Наимов, М. В. Быстрецкий*

1001

---

---

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

УДК 517.929

## О НЕКОТОРЫХ СВОЙСТВАХ ОТОБРАЖЕНИЯ СДВИГА НА БЕСКОНЕЧНОМЕРНОМ ТОРЕ

© 2023 г. С. Д. Глызин, А. Ю. Колесов

Исследуется классический в теории динамических систем вопрос о минимальности отображения сдвига на бесконечномерном торе, а точнее, решается проблема отыскания достаточных условий, гарантирующих отсутствие свойства минимальности.

DOI: 10.31857/S0374064123070014, EDN: GTARSE

### 1. Постановка задачи и описание результатов.

**1.1. Основные определения.** История развития современной теории динамических систем подробно описана в обзорах [1, 2] и монографиях [3–12] (это, разумеется, далеко не полный библиографический список). Что же касается бесконечномерных динамических систем, то попытки их изучения предпринимались неоднократно в целом ряде работ (см., например, [13–17]). Упомянутые работы дополняет настоящая статья, являющаяся продолжением серии публикаций [18–22] и посвящённая изучению некоторых свойств стандартного отображения сдвига на бесконечномерном торе. Мотивацией для таких исследований является возможность последующего применения их результатов к динамическим системам с бесконечномерным фазовым пространством, возникающим в различных приложениях. Что же касается бесконечномерного тора, то он представляет собой наиболее естественный модельный пример банахова многообразия.

Поскольку нас интересует отображение сдвига на бесконечномерном торе  $T^\infty$ , то сначала, следуя работам [18–22], дадим определение самого тора  $T^\infty$ . С этой целью фиксируем некоторое бесконечномерное вещественное банахово пространство  $E$  с нормой  $\|\cdot\|$  и в первую очередь сформулируем определение бесконечномерной целочисленной решётки  $\mathbb{Z}^\infty$ .

**Определение 1.1.** *Бесконечномерной целочисленной решёткой* (или просто *целочисленной решёткой*) назовём непустое подмножество  $\mathbb{Z}^\infty \subset E$ , удовлетворяющее следующим аксиомам:

- 1) имеет место свойство линейности: для любых  $l_1, l_2 \in \mathbb{Z}^\infty$ ,  $k_1, k_2 \in \mathbb{Z}$  справедливо включение  $k_1 l_1 + k_2 l_2 \in \mathbb{Z}^\infty$ ;
- 2) выполняется условие дискретности

$$\mu_0 \stackrel{\text{def}}{=} \inf_{\substack{l_1, l_2 \in \mathbb{Z}^\infty \\ l_1 \neq l_2}} \|l_1 - l_2\| > 0; \quad (1.1)$$

3) замыкание линейной оболочки векторов из  $\mathbb{Z}^\infty$  совпадает с исходным пространством  $E$  (это условие естественно назвать *условием максимальнойности*).

При построении конкретных примеров целочисленных решёток оказывается полезным понятие ядра, а именно *ядром*  $\Omega$  решётки  $\mathbb{Z}^\infty$  будем называть непустое подмножество из  $\mathbb{Z}^\infty$  такое, что любой вектор  $l \in \mathbb{Z}^\infty$  представляет собой конечную линейную комбинацию элементов из  $\Omega$  с целочисленными коэффициентами. Очевидно, что всегда имеется так называемое максимальное ядро  $\Omega = \mathbb{Z}^\infty$ . Далее ядро  $\Omega$  назовём *минимальным*, если оно состоит из линейно независимых векторов. Характерная особенность минимального ядра  $\Omega$  заключается в том, что для любого подмножества  $\Omega_0 \subset \Omega$ , также являющегося ядром, выполняется равенство  $\Omega_0 = \Omega$ . В общем случае вопрос о существовании минимального ядра остаётся открытым, однако в некоторых конкретных ситуациях его можно записать явно.

Отдельно остановимся на связи данного нами определения минимального ядра с понятием базиса абелевой группы. Для этого напомним, что *свободная абелева группа* – это абелева группа, имеющая базис, т.е. такое подмножество  $B$  элементов группы, что для любого её элемента

существует единственное его представление в виде линейной комбинации базисных элементов с целыми коэффициентами, из которых только конечное число являются ненулевыми. Очевидно, что решётка  $\mathbb{Z}^\infty$  оказывается абелевой группой по отношению к операции сложения векторов из пространства  $E$ , а минимальное ядро  $\Omega$  (если оно существует) автоматически будет базисом  $B$  в данной группе. Вопрос же о том, каждый ли базис  $B$  автоматически является минимальным ядром, пока открыт. Очевидно лишь, что формально требования, наложенные на минимальное ядро  $\Omega$ , сильнее, чем условия на базис  $B$  (в первом случае мы требуем отсутствия равных нулю нетривиальных конечных линейных комбинаций векторов с вещественными коэффициентами, а во втором – отсутствия таких комбинаций только лишь с целыми коэффициентами).

Примером целочисленной решётки в пространстве  $\ell_p$ ,  $p \geq 1$ , состоящем из векторов

$$\varphi = \text{colon}(\varphi_{(1)}, \varphi_{(2)}, \dots, \varphi_{(k)}, \dots), \quad \varphi_{(k)} \in \mathbb{R}, \quad k \geq 1,$$

$$\|\varphi\| \stackrel{\text{def}}{=} \left( \sum_{k=1}^{\infty} |\varphi_{(k)}|^p \right)^{1/p} < \infty, \tag{1.2}$$

является множество

$$\mathbb{Z}^\infty = \{l = \text{colon}(l_{(1)}, l_{(2)}, \dots, l_{(k)}, \dots) \in \ell_p : l_{(k)} \in \mathbb{Z}, \quad k \geq 1\}. \tag{1.3}$$

В силу сходимости ряда  $\sum_{k=1}^{\infty} |l_{(k)}|^p$  (см. (1.2)) любой вектор  $l \in \mathbb{Z}^\infty$  имеет лишь конечное число ненулевых координат. Что же касается ядра  $\Omega$ , то в данном случае в качестве такового можно взять множество

$$\{e_k, k \in \mathbb{N}\} \tag{1.4}$$

(через  $e_k$  обозначен вектор, у которого  $k$ -я компонента равна единице, а все остальные нулевые). Заметим далее, что поскольку эти векторы линейно независимы, то ядро (1.4) минимально.

В случае пространства  $\ell_\infty$ , элементами которого являются векторы

$$\varphi = \text{colon}(\varphi_{(1)}, \varphi_{(2)}, \dots, \varphi_{(k)}, \dots), \quad \varphi_{(k)} \in \mathbb{R}, \quad k \geq 1,$$

$$\|\varphi\| \stackrel{\text{def}}{=} \sup_{k \geq 1} |\varphi_{(k)}| < \infty, \tag{1.5}$$

аналогичная (1.3) целочисленная решётка имеет вид

$$\mathbb{Z}^\infty = \{l = \text{colon}(l_{(1)}, l_{(2)}, \dots, l_{(k)}, \dots) \in \ell_\infty : l_{(k)} \in \mathbb{Z}, \quad k \geq 1\}. \tag{1.6}$$

Ядром здесь будет множество  $\Omega = \text{Bin}(\ell_\infty)$  так называемых бинарных векторов  $l \in \ell_\infty$ , у которых координаты  $l_{(k)}$ ,  $k \geq 1$ , независимо друг от друга принимают значения 0 или 1.

Действительно, возьмём произвольный вектор  $l \in \mathbb{Z}^\infty$  и заметим, что поскольку последовательность  $l_{(k)}$  из (1.6) ограничена, то найдётся такой конечный набор попарно различных целых чисел  $m_1, m_2, \dots, m_s$ , что  $l_{(k)} \in \{m_1, m_2, \dots, m_s\}$  при всех  $k \geq 1$ . А это означает, что справедливо равенство  $l = m_1 e_{m_1} + m_2 e_{m_2} + \dots + m_s e_{m_s}$ , где бинарные векторы  $e_{m_j}$ ,  $1 \leq j \leq s$ , определяются по правилу

$$e_{m_j} = \text{colon}(e_{m_j}^1, e_{m_j}^2, \dots, e_{m_j}^k, \dots), \quad e_{m_j}^k = \begin{cases} 0 & \text{при } l_{(k)} \neq m_j, \\ 1 & \text{при } l_{(k)} = m_j. \end{cases}$$

Отметим ещё, что множество  $\text{Bin}(\ell_\infty)$  заведомо не является минимальным ядром, так как ядром оказывается, например, и множество  $\text{Bin}(\ell_\infty) \setminus \{l_0\}$ , где  $l_0 = \text{colon}(1, 1, \dots, 1, \dots)$ . Вопрос существования минимального ядра здесь, как и в общем случае, открыт.

Обратимся далее к пространству  $c$ , состоящему из векторов

$$\varphi = \text{colon}(\varphi_{(1)}, \varphi_{(2)}, \dots, \varphi_{(k)}, \dots), \quad \varphi_{(k)} \in \mathbb{R}, \quad \lim_{k \rightarrow +\infty} \varphi_{(k)} = 0, \quad k \geq 1, \quad (1.7)$$

и снабжённому нормой

$$\|\varphi\| \stackrel{\text{def}}{=} \max_{k \geq 1} |\varphi_{(k)}|. \quad (1.8)$$

Как и в двух предыдущих случаях, целочисленная решётка в  $c$  задаётся аналогичным (1.3) и (1.6) равенством

$$\mathbb{Z}^\infty = \{l = \text{colon}(l_{(1)}, l_{(2)}, \dots, l_{(k)}, \dots) \in c : l_{(k)} \in \mathbb{Z}, \quad k \geq 1\}. \quad (1.9)$$

Подчеркнём, что поскольку  $l_{(k)} \rightarrow 0$  при  $k \rightarrow +\infty$ , то любой вектор  $l$  из (1.9) имеет лишь конечное число ненулевых компонент. Следовательно, данная решётка фактически состоит из тех же элементов, что и (1.3). Что же касается минимального ядра  $\Omega$  этой решётки, то таковым здесь будет множество (1.4).

Аналог целочисленной решётки (1.6) можно определить и в лебеговом пространстве  $L_\infty(0, 1)$ , состоящем из классов измеримых функций  $x(t)$ , для которых конечна норма

$$\|x\| = \text{ess sup}_{0 \leq t \leq 1} |x(t)|.$$

Здесь целочисленной решёткой является множество вида

$$\mathbb{Z}^\infty = \{x(t) \in L_\infty(0, 1) : x(t) \in \mathbb{Z} \text{ при почти всех } t \in [0, 1]\}.$$

Ещё один естественный пример целочисленной решётки строится по следующему правилу. Пусть  $E$  – бесконечномерное вещественное гильбертово пространство, а  $\{e_\alpha \in E : \alpha \in \Sigma\}$  – некоторая его полная ортонормированная система (индексное множество  $\Sigma$  заведомо состоит из бесконечного числа элементов). Тогда, как нетрудно проверить, эта система служит минимальным ядром соответствующей целочисленной решётки  $\mathbb{Z}^\infty$ . Элементами данной решётки являются всевозможные конечные линейные комбинации векторов  $e_\alpha$  с целочисленными коэффициентами.

Перейдём теперь к определению тора  $\mathbb{T}^\infty$ . Всюду ниже считаем, что в пространстве  $E$  фиксирована некоторая целочисленная решётка  $\mathbb{Z}^\infty$ . Тогда с её помощью на  $E$  вводится отношение эквивалентности по следующему правилу. Будем говорить, что два вектора  $x, y \in E$  эквивалентны, если существует такой элемент  $l \in \mathbb{Z}^\infty$ , что  $x - y = 2\pi l$ .

**Определение 1.2.** *Бесконечномерным тором*  $\mathbb{T}^\infty$  назовём множество всех классов эквивалентности, порождённых описанным выше отношением.

Иными словами, справедливы равенства  $\mathbb{T}^\infty = E/2\pi\mathbb{Z}^\infty = \text{pr}(E)$ , где отображение  $\text{pr}: E \rightarrow \mathbb{T}^\infty$  – так называемая *естественная проекция*. Эта проекция действует по правилу

$$\text{pr}: \varphi \mapsto \{\varphi\}, \quad (1.10)$$

где  $\varphi$  – произвольный элемент из  $E$ , а  $\{\varphi\}$  – класс эквивалентности из  $\mathbb{T}^\infty$ , содержащий  $\varphi$ . Следует также отметить, что  $\mathbb{T}^\infty$  представляет собой абелеву группу относительно операции сложения, определяющейся по правилу

$$\{\varphi_1\} + \{\varphi_2\} = \{\varphi_1 + \varphi_2\} \quad \text{для любых } \varphi_1, \varphi_2 \in E. \quad (1.11)$$

В дальнейшем для краткости одной и той же буквой  $\varphi$  будем обозначать как вектор из  $E$ , так и соответствующий ему класс  $\{\varphi\} \in \mathbb{T}^\infty$ . Это не вызовет недоразумений, поскольку из контекста всегда будет очевидно о каком именно объекте идёт речь.

Метрику на торе  $\mathbb{T}^\infty$  зададим равенством

$$\rho(\varphi_1, \varphi_2) = \inf_{l \in \mathbb{Z}^\infty} \|\text{pr}^{-1}(\varphi_1) - \text{pr}^{-1}(\varphi_2) + 2\pi l\| \quad \text{для любых } \varphi_1, \varphi_2 \in \mathbb{T}^\infty, \quad (1.12)$$

где, напомним,  $\|\cdot\|$  – норма в  $E$ , а  $\text{pr}^{-1}(\varphi_1), \text{pr}^{-1}(\varphi_2) \in E$  – произвольные прообразы точек  $\varphi_1, \varphi_2 \in \mathbb{T}^\infty$ . Так как упомянутые прообразы определяются с точностью до аддитивных добавок вида  $2\pi l$ ,  $l \in \mathbb{Z}^\infty$ , то метрика (1.12) не зависит от их конкретного выбора. Отметим также, что в силу свойства дискретности (1.1) и формулы (1.12) отображение (1.10) является локальной изометрией, т.е.

$$\rho(\text{pr}(\varphi_1), \text{pr}(\varphi_2)) = \|\varphi_1 - \varphi_2\| \quad \text{для любых } \varphi_1, \varphi_2 \in E, \quad \|\varphi_1 - \varphi_2\| \leq \varepsilon_0, \quad (1.13)$$

где  $\varepsilon_0 = \text{const} \in (0, \pi\mu_0)$ . В свою очередь, из соотношения (1.13) автоматически вытекает полнота пространства  $(\mathbb{T}^\infty, \rho)$ .

В связи с приведённым выше определением бесконечномерного тора следует отметить, что проблема отыскания в произвольном банаховом пространстве  $E$  хотя бы одной целочисленной решётки  $\mathbb{Z}^\infty$  решается с помощью леммы Цорна и леммы Рисса об  $\varepsilon$ -почти перпендикуляре. Для того чтобы убедиться в этом, сделаем ряд построений.

Фиксируем произвольно  $\varepsilon_0 \in (0, 1)$  и рассматриваем всевозможные подмножества  $\mathcal{B} \subset E$ , обладающие свойствами линейности ( $k_1b_1 + k_2b_2 \in \mathcal{B}$  при всех  $k_1, k_2 \in \mathbb{Z}$ ,  $b_1, b_2 \in \mathcal{B}$ ) и дискретности ( $\|b_1 - b_2\| \geq \varepsilon_0$  при всех  $b_1, b_2 \in \mathcal{B}$ ,  $b_1 \neq b_2$ ). Совокупность всех таких подмножеств обозначим через  $\mathcal{L}$  и заметим, что эта совокупность заведомо не пуста. Например, справедливо включение  $\mathcal{B} = \{nb_0 : n \in \mathbb{Z}\} \in \mathcal{L}$ , где  $b_0 \in E$  ( $\|b_0\| \geq \varepsilon_0$ ) – фиксированный вектор.

Далее введём на  $\mathcal{L}$  отношение частичного порядка " $\prec$ ", полагая  $\mathcal{B}_1 \prec \mathcal{B}_2$  для всех  $\mathcal{B}_1, \mathcal{B}_2 \in \mathcal{L}$ ,  $\mathcal{B}_1 \subset \mathcal{B}_2$ , и заметим, что здесь выполняются условия леммы Цорна. Действительно, для любой цепи (т.е. семейства  $\{\mathcal{B}_\alpha \in \mathcal{L} : \alpha \in \Upsilon\}$ , где  $\Upsilon$  – произвольное индексное множество, в котором для любых двух элементов имеем  $\mathcal{B}_{\alpha_1} \prec \mathcal{B}_{\alpha_2}$  или  $\mathcal{B}_{\alpha_2} \prec \mathcal{B}_{\alpha_1}$ ) существует точная верхняя грань  $\mathcal{B}_{\text{sup}} = \bigcup_{\alpha \in \Upsilon} \mathcal{B}_\alpha \in \mathcal{L}$ . Тогда согласно упомянутой лемме в исходном семействе  $\mathcal{L}$  найдётся максимальный элемент  $\mathcal{B}_{\text{max}}$ , который характерен тем, что не существует элемента  $\mathcal{B} \in \mathcal{L}$  такого, что  $\mathcal{B}_{\text{max}} \prec \mathcal{B}$  и  $\mathcal{B}_{\text{max}} \neq \mathcal{B}$ .

Обозначим через  $\mathcal{V}$  замыкание линейной оболочки векторов из  $\mathcal{B}_{\text{max}}$  и покажем, что на самом деле  $\mathcal{V} = E$ , а значит, искомая целочисленная решётка в пространстве  $E$  задаётся равенством  $\mathbb{Z}^\infty = \mathcal{B}_{\text{max}}$ .

В предположении противного в силу леммы об  $\varepsilon$ -почти перпендикуляре найдётся такой элемент  $b_* \in E$ , что

$$\|b_*\| = 1, \quad \rho(b_*, \mathcal{V}) \stackrel{\text{def}}{=} \inf_{v \in \mathcal{V}} \|b_* - v\| \geq \varepsilon_0.$$

Тогда, как нетрудно убедиться, множество  $\mathcal{B}_* = \{kb_* + b : k \in \mathbb{Z}, b \in \mathcal{B}_{\text{max}}\}$  принадлежит семейству  $\mathcal{L}$ . Действительно, свойство линейности для  $\mathcal{B}_*$  выполняется по построению, а дискретность вытекает из неравенств

$$\|kb_* + b\| = \begin{cases} \|b\| \geq \varepsilon_0 & \text{при } k = 0, \quad b \neq 0, \\ |k|\|b_* + b/k\| \geq |k|\rho(b_*, \mathcal{V}) \geq \varepsilon_0 & \text{при } k \neq 0. \end{cases}$$

Кроме того, множество  $\mathcal{B}_*$  таково, что  $\mathcal{B}_{\text{max}} \prec \mathcal{B}_*$ ,  $\mathcal{B}_{\text{max}} \neq \mathcal{B}_*$ . А поскольку в силу определения  $\mathcal{B}_{\text{max}}$  последнее невозможно, то факт существования целочисленной решётки в  $E$  полностью обоснован.

В последующем нам понадобится понятие *фундаментального множества тора*  $\mathbb{T}^\infty$ . Таковым будем называть множество  $\mathcal{U} \subset E$ , для которого  $\text{pr}(\mathcal{U}) = \mathbb{T}^\infty$  и отображение  $\text{pr} : \mathcal{U} \rightarrow \mathbb{T}^\infty$  взаимно однозначно. Подчеркнём, что существование данного множества гарантирует аксиома выбора, применённая к семейству непустых попарно непересекающихся множеств  $\text{pr}^{-1}(\varphi)$ ,  $\varphi \in \mathbb{T}^\infty$  (под  $\text{pr}^{-1}(\varphi)$  здесь понимается полный прообраз элемента  $\varphi \in \mathbb{T}^\infty$ ).

В случае (1.2), (1.3) фундаментальной является область

$$\mathcal{U} = \{\varphi = \text{colon}(\varphi_{(1)}, \varphi_{(2)}, \dots, \varphi_{(k)}, \dots) \in \ell_p : -\pi < \varphi_{(k)} \leq \pi, \quad k \geq 1\}, \quad (1.14)$$

в случае (1.5), (1.6) – область

$$\mathcal{U} = \{\varphi = \text{colon}(\varphi_{(1)}, \varphi_{(2)}, \dots, \varphi_{(k)}, \dots) \in \ell_\infty : -\pi < \varphi_{(k)} \leq \pi, \quad k \geq 1\}, \quad (1.15)$$

а для (1.7)–(1.9) таковой будет область

$$\mathcal{U} = \{\varphi = \text{colon}(\varphi_{(1)}, \varphi_{(2)}, \dots, \varphi_{(k)}, \dots) \in c : -\pi < \varphi_{(k)} \leq \pi, \quad k \geq 1\}. \quad (1.16)$$

Нетрудно увидеть, что множества (1.15) и (1.16) ограничены, а (1.14) этим свойством не обладает.

Необходимо отметить, что, как правило, под понятием ”бесконечномерный тор” подразумевается прямое произведение счётного числа окружностей с тихоновской топологией (см., например, [16, 23–25]). В этом случае после введения на торе  $\mathbb{T}^\infty$  соответствующей метрики он становится компактным метрическим пространством. Но данный вариант нас не устраивает по той причине, что указанный тор не является гладким многообразием. А поскольку теория динамических систем и, в частности, гиперболическая теория, как правило, строятся именно на таких многообразиях, то определение бесконечномерного тора нуждается в доработке.

В нашем случае тор  $\mathbb{T}^\infty$  обладает требуемыми свойствами. Как показано в работах [21, 22], он представляет собой аналитическое банахово многообразие. Это многообразие всегда замкнуто (не имеет края) и некомпактно. В случае же существования у тора  $\mathbb{T}^\infty$  ограниченного фундаментального множества  $\mathcal{U}$  оно является ограниченным (т.е. ограничено соответствующее метрическое пространство  $(\mathbb{T}^\infty, \rho)$ ).

Завершая пояснение необходимых в последующем основных определений, введём понятие так называемого сдвигового отображения на торе  $\mathbb{T}^\infty$ .

**Определение 1.3.** Пусть  $\Delta$  – фиксированный элемент из  $\mathbb{T}^\infty$ . *Отображением сдвига* (или *сдвиговым отображением*) назовём преобразование вида

$$G_\Delta : \varphi \mapsto \varphi + \Delta. \quad (1.17)$$

Здесь  $\varphi$  – произвольная точка тора  $\mathbb{T}^\infty$ , а операция сложения в (1.17) определена по правилу (1.11).

Подчеркнём, что сдвиг  $G_\Delta$  является гомеоморфизмом тора  $\mathbb{T}^\infty$ , так как обратное к нему отображение существует и имеет вид

$$G_\Delta^{-1} : \varphi \mapsto \varphi - \Delta.$$

Здесь ”–” означает операцию, обратную к операции сложения (1.11). Более того, в силу вытекающего из (1.11) соотношения

$$n\{\varphi\} = \{n\varphi\} \quad \text{для любых } \varphi \in E, \quad n \in \mathbb{Z}, \quad (1.18)$$

элемент  $n\Delta \in \mathbb{T}^\infty$  корректно определён при всех  $n \in \mathbb{Z}$ . Тем самым для любой траектории  $\varphi_n = G_\Delta^n(\varphi_0)$ ,  $\varphi_0 \in \mathbb{T}^\infty$ ,  $n \in \mathbb{Z}$ , отображения (1.17) справедливо равенство

$$\varphi_n = \varphi_0 + n\Delta, \quad n \in \mathbb{Z}. \quad (1.19)$$

**1.2. Описание результатов.** В настоящей статье нас будет интересовать наличие или отсутствие у отображения сдвига (1.17) так называемого свойства минимальности. В связи с этим напомним [4, 5], что гомеоморфизм  $f$  полного метрического пространства  $X$  называется минимальным, если у него не существует замкнутых инвариантных подмножеств, отличных от пустого множества и всего пространства  $X$  (эквивалентная формулировка – для любого  $x_0 \in X$  соответствующая траектория  $\{x_n = f^n(x_0), n \in \mathbb{Z}\}$  всюду плотна в  $X$ ).

Приведём сначала классический результат о минимальности отображения сдвига (1.17) в конечномерном случае, т.е. на торе  $\mathbb{T}^m = \mathbb{R}^m / 2\pi\mathbb{Z}^m$ ,  $m \geq 2$ . В канонических координатах

$$(x_1, x_2, \dots, x_m) : \quad -\pi < x_k \leq \pi, \quad k = \overline{1, m},$$

оно записывается в виде

$$x_k \mapsto x_k + \Delta_k \pmod{2\pi}, \quad k = \overline{1, m}. \quad (1.20)$$

Здесь  $\Delta_k \in \mathbb{R}$  – некоторые постоянные, а операция  $x \pmod{2\pi}$  для любого  $x \in \mathbb{R}$  определяется формулой

$$x \pmod{2\pi} = \pi - 2\pi \left\{ -\frac{x}{2\pi} + \frac{1}{2} \right\} \in (-\pi, \pi], \quad (1.21)$$

в данном случае через  $\{\cdot\}$  обозначена дробная часть вещественного числа. Как известно [4, 5], отображение (1.20), (1.21) минимально в том и только в том случае, когда величины  $2\pi$ ,  $\Delta_1$ ,  $\Delta_2$ , ...,  $\Delta_m$  линейно независимы над полем рациональных чисел.

В случае бесконечномерного тора  $\mathbb{T}^\infty$  пока не удалось получить каких-либо достаточных условий, гарантирующих наличие у сдвига (1.17) свойства минимальности. Не очевидно даже, реализуется ли это свойство в принципе. Установлены только некоторые достаточные условия отсутствия минимальности у сдвигового отображения на торе  $\mathbb{T}^\infty = \ell_p/2\pi\mathbb{Z}^\infty$ , где  $p \geq 1$ ,  $\mathbb{Z}^\infty$  – решётка (1.3), и доказано отсутствие минимальности на торе  $\mathbb{T}^\infty = E/2\pi\mathbb{Z}^\infty$  в случае несепарабельного вещественного банахова пространства  $E$ .

Обратимся сначала к тору  $\mathbb{T}^\infty = \ell_p/2\pi\mathbb{Z}^\infty$ , считая, что его фундаментальная область  $\mathcal{U}$  задана равенством (1.14). Далее для произвольного сдвигового отображения (1.17) на торе  $\mathbb{T}^\infty$  введём в рассмотрение вектор сдвига

$$\gamma = \text{colon}(\gamma_{(1)}, \gamma_{(2)}, \dots, \gamma_{(s)}, \dots) \in \mathcal{U}, \quad \gamma = \text{pr}^{-1}(\Delta). \quad (1.22)$$

Первые два наших результата формулируются в терминах компонент  $\gamma_{(s)}$ ,  $s \geq 1$ , вектора (1.22).

**Теорема 1.1.** *Предположим, что при некоторых натуральных числах  $m$ ,  $n_1 < n_2 < \dots < n_m$  и целых  $k_0, k_1, \dots, k_m$  ( $|k_0| + |k_1| + \dots + |k_m| > 0$ ) имеет место равенство*

$$2\pi k_0 + k_1 \gamma_{(n_1)} + k_2 \gamma_{(n_2)} + \dots + k_m \gamma_{(n_m)} = 0. \quad (1.23)$$

Тогда соответствующее отображение сдвига (1.17) не является минимальным на торе  $\mathbb{T}^\infty$ .

Как уже отмечалось выше, аналог сформулированной теоремы справедлив и в конечномерном случае, т.е. для отображения (1.20), а именно при выполнении аналогичного (1.23) равенства

$$2\pi k_0 + k_1 \Delta_1 + k_2 \Delta_2 + \dots + k_m \Delta_m = 0,$$

где  $|k_0| + |k_1| + \dots + |k_m| > 0$ , оно не обладает свойством минимальности. Казалось бы, если, наоборот, потребовать отсутствия между компонентами вектора (1.22) всевозможных нетривиальных соотношений вида (1.23), то по аналогии с конечномерной ситуацией отображение сдвига (1.17) будет минимальным. Но, как показывает следующая теорема, данный факт не верен.

**Теорема 1.2.** *Предположим, что компоненты  $\gamma_{(s)}$  вектора (1.22) удовлетворяют условиям*

$$\gamma_{(s)} \neq 0 \quad \text{для любого } s \geq 1, \quad R_* \stackrel{\text{def}}{=} \sup_{k \geq 1} \frac{\lambda_k}{\lambda_{k+1}} < \infty, \quad (1.24)$$

где  $\lambda_1 > \lambda_2 > \dots > \lambda_k > \dots$  – все занумерованные в порядке убывания попарно различные значения, принимаемые последовательностью  $|\gamma_{(s)}|$ ,  $s \geq 1$ . Тогда  $\omega$ -предельное и  $\alpha$ -предельное множества любой траектории (1.19) отображения (1.17) являются пустыми.

Обратим внимание на следующее обстоятельство: в рамках теоремы 1.2 оператор сдвига (1.17) заведомо не минимален. Связано это с тем, что в данном случае любая его траектория (1.19) представляет собой непустое замкнутое инвариантное множество, отличное от  $\mathbb{T}^\infty$ . Таким образом, даже отсутствие всех резонансных соотношений (1.23) не гарантирует минимальности.

Заключительный результат касается сдвигового отображения (1.17) на торе  $\mathbb{T}^\infty = E/2\pi\mathbb{Z}^\infty$ , где  $E$  – несепарабельное вещественное банахово пространство.

**Теорема 1.3.** *Предположим, что исходное банахово пространство  $E$  несепарабельно. Тогда при любом  $\Delta \in \mathbb{T}^\infty$  отображение сдвига (1.17) не является минимальным на торе  $\mathbb{T}^\infty$ .*

Отметим, что в рамках теоремы 1.3 свойство минимальности отсутствует у любого гомеоморфизма  $f : \mathbb{T}^\infty \rightarrow \mathbb{T}^\infty$ . Причина заключается в том, что (как будет показано ниже) из несепарабельности пространства  $E$  вытекает и несепарабельность метрического пространства  $(\mathbb{T}^\infty, \rho)$ . Тем самым тор  $\mathbb{T}^\infty$  заведомо не может содержать ни одной всюду плотной траектории  $\varphi_n = f^n(\varphi_0)$ ,  $\varphi_0 \in \mathbb{T}^\infty$ ,  $n \in \mathbb{Z}$ .

**2. Обоснование результатов.**

**2.1. Доказательство теоремы 1.1.** Предположим, что найдутся такие натуральные  $m$ ,  $n_1 < n_2 < \dots < n_m$  и целые числа  $k_0, k_1, \dots, k_m$  такие, что  $|k_0| + |k_1| + \dots + |k_m| > 0$ , для которых справедливо равенство (1.23). Тогда введём в рассмотрение непрерывный линейный функционал  $f : \ell_p \rightarrow \mathbb{R}$ , действующий на любой вектор (1.2) по правилу

$$f(\varphi) = k_1\varphi_{(n_1)} + k_2\varphi_{(n_2)} + \dots + k_m\varphi_{(n_m)}. \tag{2.1}$$

Отметим сразу, что этот функционал нетривиален, поскольку соотношение (1.23) не может выполняться в случае  $k_0 \neq 0, k_1 = k_2 = \dots = k_m = 0$ . Далее обозначим через  $\mathbb{T}$  окружность  $\mathbb{R}/2\pi\mathbb{Z}$  и рассмотрим отображение  $F : \mathbb{T}^\infty \rightarrow \mathbb{T}$ , действующее по правилу

$$F : \varphi \mapsto f(\text{pr}^{-1}(\varphi)) \pmod{2\pi}, \tag{2.2}$$

где  $\varphi \in \mathbb{T}^\infty$ ,  $f$  – функционал (2.1),  $\ast \pmod{2\pi}$  – операция (1.21) (поскольку всюду ниже мы отождествляем  $\mathbb{T}$  с полуинтервалом  $(-\pi, \pi]$ , то эта операция задаёт естественную проекцию из  $\mathbb{R}$  в  $\mathbb{T}$ ). Подчеркнём, что в силу очевидного свойства

$$f(2\pi l) = 0 \pmod{2\pi} \quad \text{для любого } l \in \mathbb{Z}^\infty$$

отображение (2.2) определено корректно, т.е. не зависит от конкретного выбора прообраза  $\text{pr}^{-1}(\varphi) \in \ell_p$ . Добавим ещё, что из формул (1.11), (1.18) и равенства (1.23) вытекают свойства

$$F(m_1\varphi_1 + m_2\varphi_2) = m_1F(\varphi_1) + m_2F(\varphi_2) \quad \text{для любых } \varphi_1, \varphi_2 \in \mathbb{T}^\infty, \quad m_1, m_2 \in \mathbb{Z};$$

$$F(\text{pr}(\gamma)) = 0, \tag{2.3}$$

где в зависимости от контекста “+” – операция сложения классов эквивалентности из  $\ell_p/2\pi\mathbb{Z}^\infty$  или из  $\mathbb{R}/2\pi\mathbb{Z}$ , а символ “0” – нейтральный элемент по отношению к операции сложения в  $\mathbb{R}/2\pi\mathbb{Z}$ . Кроме того, отображение (2.2) очевидным образом непрерывно.

Рассмотрим теперь подмножество  $\Sigma \subset \mathbb{T}^\infty$ , заданное равенством

$$\Sigma = \{\varphi \in \mathbb{T}^\infty : F(\varphi) = 0\}. \tag{2.4}$$

Объединяя соотношения (2.2), (2.3) с очевидными формулами

$$G_\Delta(\varphi) = \varphi + \text{pr}(\gamma), \quad G_\Delta^{-1}(\varphi) = \varphi - \text{pr}(\gamma) \quad \text{для любого } \varphi \in \mathbb{T}^\infty, \tag{2.5}$$

нетрудно убедиться в том, что, во-первых, множество (2.4) не пусто (поскольку  $\Sigma$  всегда содержит нулевой класс эквивалентности из  $\mathbb{T}^\infty$ ), замкнуто (в силу непрерывности  $F$ ) и не совпадает со всем тором  $\mathbb{T}^\infty$  (в силу нетривиальности функционала (2.1)); во-вторых, имеем  $G_\Delta(\Sigma) = \Sigma$ . Тем самым  $\Sigma$  представляет собой собственное замкнутое инвариантное множество оператора  $G_\Delta$ , а значит, последний не является минимальным на торе  $\mathbb{T}^\infty$ . Теорема 1.1 доказана.

**2.2. Доказательство теоремы 1.2.** Прежде чем приступить непосредственно к обоснованию теоремы 1.2, установим в случае тора  $\mathbb{T}^\infty = \ell_p/2\pi\mathbb{Z}^\infty$  некоторое специальное представление для метрики (1.12). В связи с этим для любых двух элементов  $\varphi_1, \varphi_2 \in \mathbb{T}^\infty$  положим

$$\theta_k = \text{pr}^{-1}(\varphi_k) \in \mathcal{U}, \quad \theta_k = \text{colon}(\theta_{(1)}^k, \theta_{(2)}^k, \dots, \theta_{(s)}^k, \dots), \quad k = 1, 2, \tag{2.6}$$

где  $\mathcal{U}$  – фундаментальная область (1.14). Справедливо следующее утверждение.

**Лемма 2.1.** *Имеет место равенство*

$$\rho(\varphi_1, \varphi_2) = \left( \sum_{s=1}^{\infty} \min_{l(s) \in \mathbb{Z}} |\theta_{(s)}^1 - \theta_{(s)}^2 + 2\pi l(s)|^p \right)^{1/p}. \tag{2.7}$$

Кроме того, в случае

$$|\theta_{(s)}^1 - \theta_{(s)}^2| < \pi \tag{2.8}$$

соответствующий минимум из (2.7) достигается только при  $l(s) = 0$ .

**Доказательство.** Прежде всего отметим, что поскольку при каждом  $s \geq 1$  справедливо предельное соотношение

$$|\theta_{(s)}^1 - \theta_{(s)}^2 + 2\pi l(s)| \rightarrow +\infty \quad \text{при} \quad |l(s)| \rightarrow +\infty,$$

то фигурирующие в (2.7) минимумы заведомо достигаются при некоторых

$$l(s) = l_{(s)}^0 \in \mathbb{Z}, \quad s \geq 1. \tag{2.9}$$

Далее в случае (2.8) в силу очевидных неравенств

$$|\theta_{(s)}^1 - \theta_{(s)}^2 + 2\pi l(s)| \geq 2\pi - |\theta_{(s)}^1 - \theta_{(s)}^2| \quad \text{при любых} \quad l(s) \in \mathbb{Z}, \quad l(s) \neq 0, \quad s \geq 1,$$

с необходимостью имеем  $l_{(s)}^0 = 0$ . Что же касается условия (2.8), то согласно включениям  $\theta_k \in \ell_p$ ,  $k = 1, 2$ , и вытекающим из них предельным равенствам

$$\lim_{s \rightarrow +\infty} \theta_{(s)}^k = 0, \quad k = 1, 2,$$

оно заведомо выполняется для всех достаточно больших номеров  $s$ . Тем самым минимизирующий вектор

$$l_0 = \text{colon} (l_{(1)}^0, l_{(2)}^0, \dots, l_{(s)}^0, \dots)$$

с компонентами (2.9) автоматически принадлежит целочисленной решётке (1.3).

Суммируя проделанные построения, приходим к выводу, что

$$\left( \sum_{s=1}^{\infty} \min_{l(s) \in \mathbb{Z}} |\theta_{(s)}^1 - \theta_{(s)}^2 + 2\pi l(s)|^p \right)^{1/p} = \|\theta_1 - \theta_2 + 2\pi l_0\|,$$

где, напомним,  $\|\cdot\|$  – норма в  $\ell_p$ . А отсюда и из (1.12), (2.6), в свою очередь, следует, что

$$\rho(\varphi_1, \varphi_2) \leq \left( \sum_{s=1}^{\infty} \min_{l(s) \in \mathbb{Z}} |\theta_{(s)}^1 - \theta_{(s)}^2 + 2\pi l(s)|^p \right)^{1/p}. \tag{2.10}$$

Противоположное (2.10) нестрогое неравенство получается из очевидной оценки

$$\left( \sum_{s=1}^{\infty} \min_{l(s) \in \mathbb{Z}} |\theta_{(s)}^1 - \theta_{(s)}^2 + 2\pi l(s)|^p \right)^{1/p} \leq \|\theta_1 - \theta_2 + 2\pi l\| \quad \text{для любого} \quad l \in \mathbb{Z}^{\infty}$$

при переходе к инфимуму по  $l \in \mathbb{Z}^{\infty}$ . Лемма 2.1 доказана.

Обратимся теперь к произвольной траектории (1.19) отображения (1.17) и обозначим через  $A$  и  $B$  её  $\omega$ -предельное и  $\alpha$ -предельное множества соответственно. Далее, пусть  $A_0$  и  $B_0$  – аналогичные множества для случая  $\varphi_0 = 0$  (“0” – нулевой класс эквивалентности из  $\mathbb{T}^{\infty}$ ). Тогда, принимая во внимание вытекающее из (1.12) равенство

$$\rho(\varphi_1 + \varphi_0, \varphi_2 + \varphi_0) = \rho(\varphi_1, \varphi_2) \quad \text{при любых} \quad \varphi_0, \varphi_1, \varphi_2 \in \mathbb{T}^{\infty},$$

имеем

$$A = \varphi_0 + A_0 = \{\varphi_0 + a_0 : a_0 \in A_0\}, \quad B = \varphi_0 + B_0 = \{\varphi_0 + b_0 : b_0 \in B_0\}.$$

Тем самым для обоснования теоремы 1.2 достаточно убедиться в пустоте множеств  $A_0$  и  $B_0$ , а так как очевидным образом  $B_0 = -A_0 = \{-a_0 : a_0 \in A_0\}$ , то достаточно показать, что  $A_0 = \emptyset$ .

Согласно равенствам (1.18), (1.19), (2.5) интересующая нас проблема сводится к доказательству отсутствия предельных точек у последовательности

$$\varphi_n = \text{pr}(n\gamma) \in \mathbb{T}^\infty, \quad n \in \mathbb{N}. \quad (2.11)$$

В связи с этим сделаем два допущения, которые не ограничивают общности рассуждений, но упрощают последующий анализ.

Первое допущение заключается в условии положительности

$$\gamma_{(s)} > 0 \quad \text{при каждом } s \geq 1. \quad (2.12)$$

Для того чтобы добиться его справедливости, рассмотрим линейный ограниченный оператор  $\Lambda_1 : \ell_p \rightarrow \ell_p$ , действующий на любой вектор  $\varphi$  из (1.2) по правилу:

$$\Lambda_1 \varphi = \text{colon}((-1)^{r_1} \varphi_{(1)}, (-1)^{r_2} \varphi_{(2)}, \dots, (-1)^{r_s} \varphi_{(s)}, \dots), \quad (2.13)$$

где

$$r_s = \begin{cases} 0 & \text{при } \gamma_{(s)} > 0, \\ 1 & \text{при } \gamma_{(s)} < 0, \end{cases} \quad (2.14)$$

а  $\gamma_{(s)}$  – компоненты вектора (1.22) (в силу условий (1.24) случай  $\gamma_{(s)} = 0$  в (2.14) исключён). Заметим далее, что поскольку справедливо свойство инвариантности  $\Lambda_1 \mathbb{Z}^\infty = \mathbb{Z}^\infty$ , то оператор (2.13), (2.14) порождает отображение  $L_1 : \mathbb{T}^\infty \rightarrow \mathbb{T}^\infty$ , корректно определённое формулой

$$L_1 : \varphi \mapsto \text{pr}[\Lambda_1 \text{pr}^{-1}(\varphi)] \quad (2.15)$$

и являющееся гомеоморфизмом тора  $\mathbb{T}^\infty$ . Это обстоятельство позволяет перейти от последовательности (2.11) к новой последовательности  $\bar{\varphi}_n = L_1 \varphi_n$ ,  $n \geq 1$ . Согласно (2.11), (2.13)–(2.15) элементы указанной последовательности имеют вид

$$\bar{\varphi}_n = \text{pr}(n\bar{\gamma}), \quad n \in \mathbb{N}, \quad (2.16)$$

где

$$\bar{\gamma} = \text{colon}(\bar{\gamma}_{(1)}, \bar{\gamma}_{(2)}, \dots, \bar{\gamma}_{(s)}, \dots) = \Lambda_1 \gamma \in \mathcal{U}. \quad (2.17)$$

Так как в силу (2.13), (2.14) компоненты вектора (2.17) удовлетворяют оценкам  $\bar{\gamma}_{(s)} > 0$ ,  $s \geq 1$ , то, заменяя в (2.16)  $\bar{\gamma}$  на  $\gamma$ ,  $\bar{\varphi}_n$  на  $\varphi_n$ , приходим к требуемому случаю (2.11), (2.12).

Второе допущение, которое также не ограничивает общности, состоит в условии монотонности

$$\gamma_{(s)} \geq \gamma_{(s+1)} \quad \text{при любом } s \geq 1. \quad (2.18)$$

Для того чтобы обеспечить его выполнение, предпримем некоторые дополнительные построения.

Будем считать, что неравенства (2.12) уже имеют место. Далее, поскольку  $\gamma_{(s)} \rightarrow 0$  при  $s \rightarrow +\infty$ , все принимаемые последовательностью  $\gamma_{(s)}$ ,  $s \geq 1$ , попарно различные значения образуют строго монотонную последовательность

$$\lambda_1 > \lambda_2 > \dots > \lambda_k > \dots, \quad \lambda_k \rightarrow 0 \quad \text{при } k \rightarrow +\infty. \quad (2.19)$$

Очевидно, что величины  $\lambda_k$ ,  $k \geq 1$ , из (2.19) те же, что и в (1.24).

Пусть последовательность  $\gamma_{(s)}$ ,  $s \geq 1$ , любое своё значение  $\lambda_k$  из (2.19) принимает только на номерах  $n_1^k < n_2^k < \dots < n_{r_k}^k$ ,  $r_k \geq 1$ , т.е. при каждом  $k \geq 1$  имеем

$$\gamma_{(n_r^k)} = \lambda_k \quad \text{при } r = \overline{1, r_k}, \quad \gamma_{(s)} \neq \lambda_k \quad \text{при } s \notin \{n_1^k, n_2^k, \dots, n_{r_k}^k\}.$$

Учитывая это обстоятельство, введём в рассмотрение аналогичный (2.13) линейный ограниченный оператор  $\Lambda_2 : \ell_p \rightarrow \ell_p$ , действующий на векторы  $\varphi$  из (1.2) по правилу

$$\Lambda_2 \varphi = \text{colon}(\varphi_{(n_1^1)}, \varphi_{(n_2^1)}, \dots, \varphi_{(n_{r_1}^1)}, \varphi_{(n_1^2)}, \varphi_{(n_2^2)}, \dots, \varphi_{(n_{r_2}^2)}, \dots, \varphi_{(n_1^k)}, \varphi_{(n_2^k)}, \dots, \varphi_{(n_{r_k}^k)}, \dots). \quad (2.20)$$

Как и в предыдущем случае, оператор (2.20) обладает свойством инвариантности  $\Lambda_2 \mathbb{Z}^\infty = \mathbb{Z}^\infty$  и порождает на торе  $\mathbb{T}^\infty$  соответствующий гомеоморфизм

$$L_2 : \varphi \mapsto \text{pr} [\Lambda_2 \text{pr}^{-1}(\varphi)].$$

Таким образом, полагая

$$\overline{\varphi}_n = L_2 \varphi_n, \quad n \geq 1, \quad \overline{\gamma} = \text{colon}(\overline{\gamma}_{(1)}, \overline{\gamma}_{(2)}, \dots, \overline{\gamma}_{(s)}, \dots) = \Lambda_2 \gamma \in \mathcal{U},$$

приходим к соотношениям (2.16). А поскольку по построению  $\overline{\gamma}_{(s)} \geq \overline{\gamma}_{(s+1)}$  при всех  $s \geq 1$ , то после соответствующих переобозначений  $\overline{\gamma} \rightarrow \gamma$ ,  $\overline{\varphi}_n \rightarrow \varphi_n$  получаем последовательность (2.11) с вектором  $\gamma$ , удовлетворяющим требованиям (2.12), (2.18). Добавим ещё, что в силу условий (1.24) и неравенств (2.12), (2.18) справедлива формула

$$\sup_{s \geq 1} \frac{\gamma_{(s)}}{\gamma_{(s+1)}} = \sup_{k \geq 1} \frac{\lambda_k}{\lambda_{k+1}} = R_* < \infty. \quad (2.21)$$

Итак, теперь у нас есть всё необходимое для доказательства пустоты  $\omega$ -предельного множества  $A_0$  полутраектории (2.11). Само же доказательство проведём от противного.

Предположим, что существуют такой набор натуральных чисел  $n_1 < n_2 < \dots < n_k < \dots$ ,  $n_k \rightarrow +\infty$  при  $k \rightarrow +\infty$ , и такой элемент  $\alpha \in \mathbb{T}^\infty$ , что

$$\rho(\varphi_{n_k}, \alpha) \rightarrow 0, \quad k \rightarrow +\infty, \quad (2.22)$$

где  $\rho$  – метрика (1.12). Далее рассмотрим векторы

$$\begin{aligned} \text{pr}^{-1}(\varphi_{n_k}) &= \theta_k = \text{colon}(\theta_{(1)}^k, \theta_{(2)}^k, \dots, \theta_{(s)}^k, \dots) \in \mathcal{U}, \\ \text{pr}^{-1}(\alpha) &= \beta = \text{colon}(\beta_{(1)}, \beta_{(2)}, \dots, \beta_{(s)}, \dots) \in \mathcal{U}, \end{aligned} \quad (2.23)$$

где, как и ранее,  $\mathcal{U}$  – множество (1.14). Нетрудно увидеть, что для компонент  $\theta_{(s)}^k$  из (2.23) имеют место равенства

$$\theta_{(s)}^k = n_k \gamma_{(s)} \pmod{2\pi}, \quad \theta_{(s)}^k = n_k \gamma_{(s)} \quad \text{при } n_k \gamma_{(s)} \leq \pi, \quad (2.24)$$

где, напомним ещё раз,  $\ast \pmod{2\pi}$  – операция (1.21).

Зафиксируем две постоянные  $z_1, z_2$ , удовлетворяющие требованиям

$$0 < z_1 < z_2 < \pi/2, \quad z_2/z_1 > R_*, \quad (2.25)$$

где  $R_*$  – величина из (1.24), (2.21), и рассмотрим серию отрезков

$$I_s = [z_1/\gamma_{(s)}, z_2/\gamma_{(s)}], \quad s \geq 1. \quad (2.26)$$

Заметим далее, что в силу соотношений (2.18), (2.21), (2.25) справедливы неравенства

$$\frac{z_1}{\gamma_{(s)}} \leq \frac{z_1}{\gamma_{(s+1)}} < \frac{z_2}{\gamma_{(s)}} \leq \frac{z_2}{\gamma_{(s+1)}}, \quad s \geq 1.$$

В свою очередь, из приведённых неравенств вытекает, что отрезки (2.26) с номерами  $s$  и  $s + 1$  имеют непустое пересечение, а значит,

$$[z_1/\gamma_{(1)}, +\infty) = \bigcup_{s=1}^{\infty} I_s. \tag{2.27}$$

На завершающей стадии доказательства выберем такое достаточно большое  $k_* \in \mathbb{N}$ , чтобы при всех  $k \geq k_*$  выполнялось условие  $n_k \geq z_1/\gamma_{(1)}$ . В этом случае в силу равенства (2.27) при каждом  $k \geq k_*$  существует такой номер  $s = s_k$ ,  $s_k \rightarrow +\infty$  при  $k \rightarrow +\infty$ , что  $n_k \in I_{s_k}$  или, что то же самое,

$$z_1 \leq n_k \gamma_{(s_k)} \leq z_2 \quad \text{при любом } k \geq k_*. \tag{2.28}$$

Заметим ещё, что поскольку  $\beta_{(s_k)} \rightarrow 0$  при  $k \rightarrow +\infty$ , то без ограничения общности можно считать выполненными неравенства

$$|\beta_{(s_k)}| < \pi/2, \quad k \geq k_*. \tag{2.29}$$

Объединяя соотношения (2.24), (2.25), (2.28), (2.29), приходим к выводу, что при всех номерах  $k \geq k_*$

$$\theta_{(s_k)}^k = n_k \gamma_{(s_k)}, \quad |n_k \gamma_{(s_k)} - \beta_{(s_k)}| < \pi,$$

т.е. выполняется условие вида (2.8). Тем самым в силу (2.7), (2.22) имеем

$$|n_k \gamma_{(s_k)} - \beta_{(s_k)}| \leq \rho(\varphi_{n_k}, \alpha) = \left( |n_k \gamma_{(s_k)} - \beta_{(s_k)}|^p + \sum_{\substack{s=1 \\ s \neq s_k}}^{\infty} \min_{l_{(s)} \in \mathbb{Z}} |\theta_{(s)}^k - \beta_{(s)} + 2\pi l_{(s)}|^p \right)^{1/p} \rightarrow 0$$

при  $k \rightarrow +\infty$ . Остаётся добавить, что отсюда с учётом уже упоминавшегося выше свойства  $\beta_{(s_k)} \rightarrow 0$  при  $k \rightarrow +\infty$  автоматически следует, что и  $n_k \gamma_{(s_k)} \rightarrow 0$ ,  $k \rightarrow +\infty$ . Последнее же противоречит оценкам (2.28).

Итак, наше допущение о существовании у последовательности (2.11) частичного предела  $\alpha \in \mathbb{T}^\infty$  не верно, а значит,  $A_0 = \emptyset$ . Теорема 1.2 доказана.

**2.3. Доказательство теоремы 1.3.** Как уже было отмечено в п. 1.2, справедливость теоремы 1.3 вытекает из следующего утверждения.

**Лемма 2.2.** Пусть  $E$  – несепарабельное вещественное банахово пространство. Тогда метрическое пространство  $(\mathbb{T}^\infty, \rho)$ , где  $\mathbb{T}^\infty = E/2\pi\mathbb{Z}^\infty$ ,  $\rho$  – метрика (1.12), также является несепарабельным.

**Доказательство.** Как известно, для того чтобы установить несепарабельность метрического пространства  $X$  достаточно для некоторого  $\varepsilon > 0$  построить в нём несчётную  $\varepsilon$ -цепь. В связи с этим напомним, что  $\varepsilon$ -цепью называется подмножество  $\mathcal{A} \subset X$  такое, что  $\rho(a_1, a_2) \geq \varepsilon$  при всех  $a_1, a_2 \in \mathcal{A}$ ,  $a_1 \neq a_2$ .

Обратимся сначала к исходному пространству  $E$  и покажем существование хотя бы одного такого  $\varepsilon_0 > 0$ , что шар

$$B(0, \varepsilon_0) = \{x \in E : \|x\| \leq \varepsilon_0\} \tag{2.30}$$

представляет собой несепарабельное метрическое пространство (с метрикой, порождённой нормой  $\|\cdot\|$  из  $E$ ). В предположении противного в каждом шаре вида  $B(0, n) \subset E$ ,  $n \in \mathbb{N}$ , найдётся счётное всюду плотное множество  $\Gamma_n$ , а их объединение  $\Gamma = \bigcup_{n \geq 1} \Gamma_n$  будет всюду плотным в  $E$ . Последнее же противоречит предполагаемой несепарабельности пространства  $E$ . Тем самым требуемое  $\varepsilon_0$  действительно существует. Впрочем, в силу подобия несепарабельным будет шар (2.30) любого радиуса.

Итак, зафиксируем произвольно шар (2.30) и заметим, что поскольку он несепарабелен, то в нём найдётся несчётная  $\varepsilon_1$ -цепь  $C$  при некотором  $\varepsilon_1 > 0$ . Что же касается метрического пространства  $(\mathbb{T}^\infty, \rho)$ , то (как будет показано ниже) в нём в качестве требуемой несчётной цепи можно взять множество

$$\mathcal{A} = \{\text{pr}(\varepsilon_2 c) : c \in C\} \subset \mathbb{T}^\infty, \tag{2.31}$$

в котором константа  $\varepsilon_2 > 0$  выбрана из условия  $2\varepsilon_0\varepsilon_2 < \pi\mu_0$ , где  $\mu_0$  – постоянная (1.1).

Действительно, поскольку для любых  $c_1, c_2 \in \mathbb{C}$ ,  $c_1 \neq c_2$ , выполняются неравенства

$$\|\varepsilon_2(c_1 - c_2)\| \leq 2\varepsilon_0\varepsilon_2 < \pi\mu_0,$$

то в силу (1.13) и того факта, что  $\mathbb{C}$  является  $\varepsilon_1$ -цепью, справедлива оценка

$$\rho(\text{pr}(\varepsilon_2c_1), \text{pr}(\varepsilon_2c_2)) = \|\varepsilon_2(c_1 - c_2)\| \geq \varepsilon_1\varepsilon_2. \tag{2.32}$$

Следовательно, множество (2.31) представляет собой  $\varepsilon$ -цепь в пространстве  $(\mathbb{T}^\infty, \rho)$  при  $\varepsilon = \varepsilon_1\varepsilon_2$ . Добавим ещё, что эта  $\varepsilon$ -цепь несчётна, поскольку согласно (2.31), (2.32) между элементами множеств  $\mathcal{A}$  и  $\mathbb{C}$  имеет место взаимно однозначное соответствие. Лемма 2.2 доказана.

**2.4. Заключение.** Подводя итог, отметим, что нами получены три варианта достаточных условий отсутствия у оператора сдвига на торе  $\mathbb{T}^\infty$  свойства минимальности, причём два из них не имеют аналогов в конечномерном случае. Однако остался нерешённым вопрос о принципиальной реализуемости указанного свойства для сдвига на  $\mathbb{T}^\infty$ . Иными словами, проблема заключается либо в получении для отображения (1.17) на торе  $\mathbb{T}^\infty$  достаточных условий минимальности, либо в доказательстве отсутствия минимальности у любого оператора сдвига на  $\mathbb{T}^\infty$ . Какая из этих альтернатив реализуется, пока не очевидно.

Для того чтобы подчеркнуть новизну поставленной проблемы, обратимся к тору  $\tilde{\mathbb{T}}^\infty$ , понимаемому как прямое произведение счётного числа окружностей  $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$  с тихоновской топологией (см. [16, 23–25]). Его элементами являются бесконечномерные векторы вида

$$\tilde{\varphi} = (\varphi_1, \varphi_2, \dots, \varphi_m, \dots), \quad \varphi_m \in \mathbb{T}, \quad m \geq 1,$$

а метрика на  $\tilde{\mathbb{T}}^\infty$  вводится по следующему правилу.

Сначала для любых двух элементов  $\varphi_1, \varphi_2 \in \mathbb{T}$  определяем расстояние между ними по аналогичной (1.12) формуле

$$d(\varphi_1, \varphi_2) = \min_{l \in \mathbb{Z}} |\theta_1 - \theta_2 + 2\pi l|,$$

где  $\theta_1, \theta_2 \in \mathbb{R}$  такие, что  $\varphi_k = \theta_k \pmod{2\pi}$ ,  $k = 1, 2$ . После этого для любых двух элементов

$$\tilde{\varphi}^k = (\varphi_1^k, \varphi_2^k, \dots, \varphi_m^k, \dots), \quad \varphi_m^k \in \mathbb{T}, \quad k = 1, 2, \quad m \geq 1,$$

из  $\tilde{\mathbb{T}}^\infty$  полагаем

$$\rho(\tilde{\varphi}^1, \tilde{\varphi}^2) = \sum_{m=1}^{\infty} \frac{1}{2^m} \frac{d(\varphi_m^1, \varphi_m^2)}{1 + d(\varphi_m^1, \varphi_m^2)}. \tag{2.33}$$

Очевидно, что тор  $\tilde{\mathbb{T}}^\infty$ , снабжённый метрикой (2.33), является компактным метрическим пространством.

Рассмотрим теперь произвольное отображение сдвига на торе  $\tilde{\mathbb{T}}^\infty$ , имеющее в канонических координатах  $x_{(1)}, x_{(2)}, \dots, x_{(s)}, \dots$ ,  $-\pi < x_{(s)} \leq \pi$ ,  $s \geq 1$ , вид

$$x_{(s)} \mapsto x_{(s)} + \gamma_{(s)} \pmod{2\pi}, \tag{2.34}$$

где  $\gamma_{(s)} \in (-\pi, \pi]$ ,  $s \geq 1$ . Справедливо следующее утверждение.

**Лемма 2.3.** *Отображение (2.34) является минимальным на торе  $\tilde{\mathbb{T}}^\infty$  тогда и только тогда, когда между компонентами сдвига  $\gamma_{(s)} \in (-\pi, \pi]$ ,  $s \geq 1$ , отсутствуют всевозможные резонансные соотношения вида (1.23).*

Сформулированный результат вытекает из аналогичного конечномерного результата и из свойства “равномерной малости хвостов” у метрики (2.33). Последнее означает существование для любого  $\varepsilon > 0$  такого (зависящего только от  $\varepsilon$ ) натурального  $m_*$ , что

$$\sum_{m=m_*+1}^{\infty} \frac{1}{2^m} \frac{d(\varphi_m^1, \varphi_m^2)}{1 + d(\varphi_m^1, \varphi_m^2)} < \varepsilon \quad \text{при всех } \tilde{\varphi}^1, \tilde{\varphi}^2 \in \tilde{\mathbb{T}}^\infty.$$

В случае же тора  $\mathbb{T}^\infty = \ell_p/2\pi\mathbb{Z}^\infty$  с метрикой (1.12) соответствующие хвосты (остатки рядов из (2.7)) не являются равномерно малыми по  $\varphi_1, \varphi_2 \in \mathbb{T}^\infty$ . Именно по этой причине при замене  $\widehat{\mathbb{T}}^\infty$  на  $\mathbb{T}^\infty$  утверждение леммы 2.3 не верно (см. теорему 1.2).

Также обсудим теоремы 1.1 и 1.2 на предмет их возможного обобщения. В связи с этим обратим внимание на обстоятельство, что утверждения данных теорем сохраняются при замене пространства  $\ell_p$  на  $c$  (см. (1.7), (1.8)) и соответствующего тора  $\mathbb{T}^\infty = \ell_p/2\pi\mathbb{Z}^\infty$  на тор  $\mathbb{T}^\infty = c/2\pi\mathbb{Z}^\infty$ , где  $\mathbb{Z}^\infty$  – целочисленная решётка (1.9).

Действительно, доказательство теоремы 1.1 в указанном случае сохраняется дословно, а при обосновании для  $c$  теоремы 1.2 используется аналогичное (2.7) представление

$$\rho(\varphi_1, \varphi_2) = \max_{s \geq 1} \left( \min_{l_{(s)} \in \mathbb{Z}} |\theta_{(s)}^1 - \theta_{(s)}^2 + 2\pi l_{(s)}| \right), \quad \varphi_1, \varphi_2 \in \mathbb{T}^\infty,$$

где  $\theta_k = \text{pr}^{-1}(\varphi_k) \in \mathcal{U}$ ,  $\theta_k = \text{colon}(\theta_{(1)}^k, \theta_{(2)}^k, \dots, \theta_{(s)}^k, \dots)$ ,  $k = 1, 2$ ,  $\mathcal{U}$  – множество (1.16).

Интересно также отметить, что аналог теоремы 1.1 можно получить и для абстрактного бесконечномерного банахова пространства  $E$ . Точнее говоря, имеет место следующее утверждение.

**Лемма 2.4.** *Предположим, что существует такой непрерывный и нетривиальный линейный функционал  $f : E \rightarrow \mathbb{R}$ , что*

$$f(2\pi l) = 0 \pmod{2\pi} \quad \text{для любого } l \in \mathbb{Z}^\infty, \quad f(\text{pr}^{-1}(\Delta)) = 0 \pmod{2\pi},$$

где  $\Delta$  – элемент сдвига из (1.17). Тогда оператор (1.17) не является минимальным на торе  $\mathbb{T}^\infty = E/2\pi\mathbb{Z}^\infty$ .

Справедливость сформулированной леммы вытекает из того факта, что и здесь множество вида (2.4), где  $F$  – оператор (2.2), оказывается инвариантным для отображения сдвига  $G_\Delta$ .

Перейдём далее к теореме 1.2. Нетрудно увидеть, что фигурирующие в ней условия (1.24) заведомо справедливы для вектора сдвига (1.22) с компонентами

$$\gamma_{(s)} = 1/s^\delta, \quad \delta = \text{const} > 1/p, \quad s \geq 1 \quad \text{или} \quad \gamma_{(s)} = q^s, \quad q = \text{const} \in (0, 1), \quad s \geq 1,$$

поскольку в обоих случаях  $\lambda_k = \gamma_{(k)}$ ,  $k \geq 1$ , и конечна величина  $R_*$  (см. (2.21)). Если же, например,  $\gamma_{(s)} = \exp(-s^2)$ ,  $s \geq 1$ , то по-прежнему  $\lambda_k = \gamma_{(k)}$ ,  $k \geq 1$ , но  $R_* = +\infty$ , а значит, результаты теоремы 1.2 здесь не применимы. В связи с этим естественным образом возникает вопрос о том, каковым будет множество частичных пределов последовательности (2.11) в случае, когда для компонент  $\gamma_{(s)}$ ,  $s \geq 1$ , вектора  $\gamma$  выполняются требования (2.12), (2.18), а вместо (2.21) имеем

$$\sup_{s \geq 1} \frac{\gamma_{(s)}}{\gamma_{(s+1)}} = +\infty.$$

Обозначенная проблема пока не решена.

Работа выполнена при финансовой поддержке Российского научного фонда (проект 21-71-30011).

## СПИСОК ЛИТЕРАТУРЫ

1. Смейл С. Дифференцируемые динамические системы // Успехи мат. наук. 1970. Т. 25. Вып. 1 (151). С. 113–185.
2. Аносов Д.В., Солодов В.В. Гиперболические множества // Итоги науки и техн. Сер. Современ. проблемы математики. Фунд. направления. М., 1991. Т. 66. С. 12–99.
3. Аносов Д.В. Геодезические потоки на замкнутых римановых многообразиях отрицательной кривизны // Тр. Мат. ин-та имени В.А. Стеклова. 1967. Т. 90. С. 3–210.
4. Каток А.Б., Хасселблат Б. Введение в современную теорию динамических систем. М., 1999.
5. Каток А.Б., Хасселблат Б. Введение в теорию динамических систем с обзором последних достижений. М., 2005.

6. *Пильгогин С.Ю.* Пространства динамических систем. М.; Ижевск, 2008.
7. *Гринес В.З., Починка О.В.* Введение в топологическую классификацию каскадов на многообразиях размерности два или три. М.; Ижевск, 2011.
8. *Grines V., Zhuzhoma E.* Surface Laminations and Chaotic Dynamical Systems. М.; Izhevsk, 2021.
9. *Палис Ж., ду Мелу В.* Геометрическая теория динамических систем. Введение. М., 1986.
10. *Песин Я.Б.* Лекции по теории частичной гиперболичности и устойчивой эргодичности. М., 2006.
11. *Robinson C.* Dynamical Systems. Stability, Symbolic Dynamics, and Chaos. Boca Raton, 1999.
12. *Palis J., Takens F.* Hyperbolicity and Sensitive Chaotic Dynamics at Homoclinic Bifurcations. Cambridge, 1993.
13. *Ruelle D.* Large volume limit of the distribution of characteristic exponents in turbulence // Commun. Math. Phys. 1982. V. 87. P. 287–302.
14. *Māné R.* Ergodic Theory and Differentiable Dynamics. Berlin; Heidelberg, 1987.
15. *Thieullen P.* Entropy and the Hausdorff dimension for infinite-dimensional dynamical systems // J. of Dynamics and Differ. Equat. 1992. V. 4. № 1. P. 127–159.
16. *Hastings H.M.* On expansive homeomorphisms of the infinite torus // The Structure of Attractors in Dynamical Systems. Lecture Notes in Math. Berlin; Heidelberg; New York, 1978. P. 142–149.
17. *Māné R.* Expansive homeomorphisms and topological dimension // Trans. of Amer. Math. Soc. 1979. V. 252. P. 313–319.
18. *Глызин С.Д., Колесов А.Ю., Розов Н.Х.* Растягивающие эндоморфизмы на бесконечномерном торе // Функц. анализ и его приложения. 2020. Т. 54. № 4. С. 17–36.
19. *Глызин С.Д., Колесов А.Ю., Розов Н.Х.* Соленоидальные аттракторы диффеоморфизмов кольцевых множеств // Успехи мат. наук. 2020. Т. 75. Вып. 2 (452). С. 3–60.
20. *Глызин С.Д., Колесов А.Ю., Розов Н.Х.* Об одном классе диффеоморфизмов Аносова на бесконечномерном торе // Изв. РАН. Сер. мат. 2021. Т. 85. № 2. С. 3–59.
21. *Глызин С.Д., Колесов А.Ю.* Критерий гиперболичности одного класса диффеоморфизмов на бесконечномерном торе // Мат. сб. 2022. Т. 213. № 2. С. 50–95.
22. *Глызин С.Д., Колесов А.Ю.* Элементы гиперболической теории на бесконечномерном торе // Успехи мат. наук. 2022. Т. 77. Вып. 3 (465). С. 3–72.
23. *Jessen B.* The theory of integration in a space of an infinite number of dimensions // Acta Math. 1934. V. 63. P. 249–323.
24. *Платонов С.С.* О некоторых задачах теории приближения функций на бесконечномерном торе: аналоги теорем Джексона // Алгебра и анализ. 2014. Т. 26. № 6. С. 99–120.
25. *Kosz D.* On differentiation of integrals in the infinite-dimensional torus // Studia Math. 2021. V. 258. P. 103–119.

Ярославский государственный университет  
имени П.Г. Демидова

Поступила в редакцию 12.03.2023 г.  
После доработки 12.03.2023 г.  
Принята к публикации 19.05.2023 г.

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

УДК 517.929

## ОБОБЩЁННЫЕ РЕШЕНИЯ ПЕРВОЙ КРАЕВОЙ ЗАДАЧИ ДЛЯ ДИФФЕРЕНЦИАЛЬНО-РАЗНОСТНОГО УРАВНЕНИЯ В ДИВЕРГЕНТНОМ ВИДЕ НА ИНТЕРВАЛЕ КОНЕЧНОЙ ДЛИНЫ

© 2023 г. А. Л. Скубачевский, Н. О. Иванов

Рассмотрена задача Дирихле для дифференциально-разностного уравнения второго порядка в дивергентном виде с переменными коэффициентами на конечном интервале  $Q = (0, d)$ . Исследованы условия на правую часть уравнения, обеспечивающие гладкость обобщённого решения на всём интервале. Доказано, что обобщённое решение задачи принадлежит пространству Соболева  $W_2^2(Q)$  в случае ортогональности правой части в пространстве  $L_2(Q)$  конечному числу линейно независимых функций.

DOI: 10.31857/S0374064123070026, EDN: GTLPJF

**Введение.** В работе исследуются разрешимость и гладкость обобщённых решений первой краевой задачи для дифференциально-разностного уравнения с условиями Дирихле на интервале конечной длины  $(0, d)$  со сдвигами независимой переменной в старших производных. Впервые такие задачи для функционально-дифференциальных уравнений нейтрального типа были рассмотрены в статьях [1, 2]. Было показано, что, в отличие от обыкновенных дифференциальных уравнений, решения этих задач обладают рядом принципиально новых свойств. В частности, из-за наличия разностного оператора гладкость обобщённых решений может нарушаться во внутренних точках интервала  $(0, d)$  даже при условии бесконечно дифференцируемой правой части. В работах [3, 4] рассматривалась первая краевая задача для дифференциально-разностного уравнения с постоянными коэффициентами на конечном интервале. Используя теорему об изоморфизме пространства Соболева  $\dot{W}_2^1(Q)$  и подпространства  $W_2^1(Q)$ , состоящего из функций, удовлетворяющих нелокальным краевым условиям, исходная краевая задача для дифференциально-разностного уравнения была сведена к нелокальной краевой задаче для обыкновенного дифференциального уравнения. Такой подход позволил в явном виде получить условия ортогональности правой части уравнения конечному числу линейно независимых функций в  $L_2(Q)$ , обеспечивающие гладкость обобщённого решения на всём интервале. Аналогичные условия были получены для второй краевой задачи для дифференциально-разностного уравнения с переменными коэффициентами (см. [5, 6]). Условия на коэффициенты дифференциально-разностного уравнения, обеспечивающие гладкость обобщённого решения как первой, так и второй краевых задач при любой правой части, были получены в работах [7, 8]. Задачи для сильно эллиптических дифференциально-разностных уравнений со смешанными краевыми условиями рассматривались в статьях [9, 10].

Интерес к функционально-дифференциальным уравнениям связан с их важными приложениями, например, к задаче об успокоении системы управления с последействием [4, 11–15], к теории многослойных пластин и оболочек [4, 16, 17], к теории многомерных диффузионных процессов [4] и др.

На интервале  $Q = (0, d)$ , где  $d = N + \theta$ ,  $N \in \mathbb{N}$ ,  $0 < \theta \leq 1$ , рассматривается задача

$$-(R_Q u)' = f(x), \quad x \in Q, \quad (1)$$

$$u(0) = u(d) = 0, \quad (2)$$

где  $f \in L_2(Q)$  – комплекснозначная функция. Оператор  $R_Q : L_2(Q) \rightarrow L_2(Q)$  задан формулой

$$R_Q = P_Q R I_Q,$$

здесь  $P_Q : L_2(\mathbb{R}) \rightarrow L_2(Q)$ ,  $I_Q : L_2(Q) \rightarrow L_2(\mathbb{R})$ ,  $R : L_2(\mathbb{R}) \rightarrow L_2(\mathbb{R})$ , при этом

$$\begin{aligned} (P_Q u)(x) &= u(x), \quad x \in Q, \\ (I_Q u)(x) &= u(x), \quad x \in Q, \quad (I_Q u)(x) = 0, \quad x \in \mathbb{R} \setminus Q, \\ (Ru)(x) &= \sum_{j=-N}^N b_j(x)u(x+j), \end{aligned}$$

где  $b_j(x) \in C^\infty(\mathbb{R})$  – комплекснозначные функции. В дальнейшем нам понадобятся свойства оператора  $R_Q$  в пространстве  $L_2(Q)$ , которые будут сформулированы в виде лемм. Доказательства этих свойств можно найти в монографии [4].

**Замечание 1.** Заметим, что в случае дифференциально-разностных уравнений с постоянными коэффициентами исследование задачи (1), (2) аналогично исследованию следующей задачи (см. [4]):

$$\begin{aligned} -(R_Q u)'' &= f(x), \quad x \in Q, \\ u(0) &= u(d) = 0. \end{aligned}$$

С другой стороны, в случае задачи (1), (2) для дифференциально-разностного уравнения с переменными коэффициентами, рассматриваемого в настоящей работе, данное утверждение является неверным.

В п. 1 рассматриваются разностные операторы  $R_Q$  на интервале  $Q$ . Вопросу существования обобщённого решения задачи (1), (2) посвящён п. 2. В п. 3 исследуется гладкость обобщённых решений исследуемой задачи на подынтервалах и на всём интервале.

**1. Свойства разностного оператора  $R_Q$  на интервале  $Q$ .** Рассмотрим разбиение интервала  $Q = (0, d)$  на непересекающиеся подынтервалы, образованные выбрасыванием орбит концов интервала  $Q$ , которые порождаются группой целочисленных сдвигов на  $j$ . При  $\theta = 1$  будем рассматривать один класс подынтервалов  $Q_{1k} = (k-1, k)$ ,  $k = \overline{1, N+1}$ , а при  $0 < \theta < 1$  – два класса:  $Q_{1k} = (k-1, k-1+\theta)$ ,  $k = \overline{1, N+1}$ , и  $Q_{2k} = (k-1+\theta, k)$ ,  $k = \overline{1, N}$ .

Через  $s$  обозначим номер класса подынтервалов. Тогда  $s = 1$ , если  $\theta = 1$ , и  $s = 1, 2$ , если  $0 < \theta < 1$ . Обозначим также  $N(s) := N+1$ , если  $s = 1$ , и  $N(s) := N$ , если  $s = 2$ .

Обозначим через  $P_s : L_2(Q) \rightarrow L_2(\bigcup_k Q_{sk})$ ,  $k = \overline{1, N(s)}$ , оператор ортогонального проектирования, где

$$L_2(\bigcup_k Q_{sk}) = \{u \in L_2(Q) : u(x) = 0, \quad x \in Q \setminus \bigcup_k Q_{sk}, \quad k = \overline{1, N(s)}\}.$$

Очевидно, что  $L_2(Q) = \bigoplus_s L_2(\bigcup_k Q_{sk})$ .

**Замечание 2.** При  $\theta = 1$  оператор  $P_1$  является единичным оператором.

**Лемма 1.** Операторы  $R : L_2(\mathbb{R}) \rightarrow L_2(\mathbb{R})$  и  $R_Q : L_2(Q) \rightarrow L_2(Q)$  ограниченные, при этом

$$(R^* u)(x) = \sum_{j=-N}^N \overline{b_j(x-j)} u(x-j), \quad R_Q^* = P_Q R^* I_Q.$$

Доказательство следует из определений операторов  $P_Q$  и  $I_Q$ .

**Лемма 2.** Пространство  $L_2(\bigcup_k Q_{sk})$ ,  $k = \overline{1, N(s)}$ , является инвариантным подпространством оператора  $R_Q$ .

Введём изоморфизм гильбертовых пространств  $U_s : L_2(\bigcup_k Q_{sk}) \rightarrow L_2^{N(s)}(Q_{s1})$  по формуле

$$(U_s u)_k(x) = u(x+k-1), \quad x \in Q_{s1}, \quad k = \overline{1, N(s)}, \tag{3}$$

где

$$L_2^{N(s)}(Q_{s1}) = \prod_{k=1}^{N(s)} L_2(Q_{s1}).$$

Через  $R_s(x)$ ,  $x \in \overline{Q}_{s1}$ , обозначим  $N(s) \times N(s)$ -матрицу с элементами

$$r_{ij}^s(x) = b_{j-i}(x+i-1), \quad x \in \mathbb{R}, \quad i, j = \overline{1, N(s)}. \tag{4}$$

Другими словами,

$$R_1(x) = \begin{pmatrix} b_0(x) & b_1(x) & \dots & b_N(x) \\ b_{-1}(x+1) & b_0(x+1) & \dots & b_{N-1}(x+1) \\ \dots & \dots & \dots & \dots \\ b_{-N}(x+N) & b_{-N+1}(x+N) & \dots & b_0(x+N) \end{pmatrix}, \quad x \in \mathbb{R},$$

$$R_2(x) = \begin{pmatrix} b_0(x) & b_1(x) & \dots & b_{N-1}(x) \\ b_{-1}(x+1) & b_0(x+1) & \dots & b_{N-2}(x+1) \\ \dots & \dots & \dots & \dots \\ b_{-N+1}(x+N-1) & b_{-N+2}(x+N-1) & \dots & b_0(x+N-1) \end{pmatrix}, \quad x \in \mathbb{R}.$$

**Лемма 3.** Оператор  $R_{Q_s} = U_s R_Q U_s^{-1} : L_2^{N(s)}(Q_{s1}) \rightarrow L_2^{N(s)}(Q_{s1})$  является оператором умножения на квадратную матрицу  $R_s(x)$ .

**Доказательство.** Пусть  $V \in L_2^{N(s)}(Q_{s1})$ . Положим  $u = U_s^{-1}V \in L_2(\bigcup_k Q_{sk})$ . Тогда из (3) и (4) следуют равенства

$$(R_{Q_s}V)_i(x) = (U_s R_Q u)_i(x) = \sum_{j=-i+1}^{N(s)-i} b_j(x+i-1)u(x+i-1+j) =$$

$$= \sum_{k=1}^{N(s)} b_{k-i}(x+i-1)u(x+k-1) = \sum_{k=1}^{N(s)} r_{ik}^s(x)V_k(x), \quad x \in Q_{s1}.$$

Лемма доказана.

Из леммы 3 следует

**Лемма 4.** Оператор  $R_Q + R_Q^* : L_2(Q) \rightarrow L_2(Q)$  положительно определён тогда и только тогда, когда матрицы  $R_s(x) + R_s^*(x)$  положительно определены для всех  $x \in \overline{Q}_{s1}$ , где  $R_s^*(x)$  – эрмитово сопряжённые матрицы к  $R_s(x)$ ;  $s = 1, 2$ , если  $0 < \theta < 1$ , и  $s = 1$ , если  $\theta = 1$ .

**Определение 1.** Будем говорить, что уравнение (1) удовлетворяет условию сильной эллиптичности, если матрицы  $R_s(x) + R_s^*(x)$  положительно определены для всех  $x \in \overline{Q}_{s1}$ , где  $s = 1, 2$ , если  $0 < \theta < 1$ , и  $s = 1$ , если  $\theta = 1$ .

**Замечание 3.** Далее будем предполагать, что уравнение (1) удовлетворяет условию сильной эллиптичности. Для доказательства основных результатов нам понадобится эквивалентная формулировка этого условия в виде выполнения неравенства

$$\operatorname{Re} (R_s(x)Y, Y)_{\mathbb{C}^{N(s)}} \geq c_1 \|Y\|_{\mathbb{C}^{N(s)}}^2 \tag{5}$$

для всех  $x \in \overline{Q}_{s1}$  и  $Y \in \mathbb{C}^{N(s)}$ , где  $c_1 > 0$  не зависит от  $x$  и  $Y$ ;  $s = 1, 2$ , если  $0 < \theta < 1$ , и  $s = 1$ , если  $\theta = 1$ .

Через  $W_2^k(Q)$  обозначим пространство Соболева комплекснозначных функций из  $L_2(Q)$ , имеющих все обобщённые производные вплоть до  $k$ -го порядка из  $L_2(Q)$ , со скалярным произведением, определённым по формуле

$$(u, w)_{W_2^k(Q)} = \sum_{i=0}^k \int_Q u^{(i)} \overline{w^{(i)}} dx, \quad u, w \in W_2^k(Q).$$

**Лемма 5.** Пусть  $\det R_s(x) \neq 0$  при  $x \in \overline{Q_{s1}}$ ,  $s = 1, 2$ , если  $0 < \theta < 1$ , и  $s = 1$ , если  $\theta = 1$ . Тогда оператор  $R_Q : L_2(Q) \rightarrow L_2(Q)$  имеет ограниченный обратный. Пусть, кроме того,  $u \in W_2^k(Q_{sj})$ ,  $j = \overline{1, N(s)}$ . Тогда  $R_Q^{-1}u \in W_2^k(Q_{si})$ ,  $i = \overline{1, N(s)}$ , и справедлива оценка

$$\|R_Q^{-1}u\|_{W_2^k(Q_{si})} \leq c_2 \sum_{j=1}^{N(s)} \|u\|_{W_2^k(Q_{sj})},$$

где  $c_2 > 0$  не зависит от функции  $u$ .

Доказательство аналогично доказательству леммы 2.12 из [4].

**2. Обобщённые решения первой краевой задачи для дифференциально-разностного уравнения.** Обозначим через  $\mathring{W}_2^1(Q)$  множество функций из пространства  $W_2^1(Q)$  таких, что  $u(0) = u(d) = 0$ , т.е.

$$\mathring{W}_2^1(Q) = \{u \in W_2^1(Q) : u(0) = u(d) = 0\}.$$

Введём неограниченный оператор  $\mathcal{A}_R : L_2(Q) \supset D(\mathcal{A}_R) \rightarrow L_2(Q)$ , действующий по формуле

$$\mathcal{A}_R u = -(R_Q u')', \quad u \in D(\mathcal{A}_R),$$

где  $D(\mathcal{A}_R) = \{u \in \mathring{W}_2^1(Q) : R_Q u' \in W_2^1(Q)\}$ .

Дадим эквивалентные определения обобщённого решения задачи (1), (2).

**Определение 2.** Функция  $u \in D(\mathcal{A}_R)$  является обобщённым решением задачи (1), (2), если она удовлетворяет уравнению

$$\mathcal{A}_R u = f. \tag{6}$$

**Определение 3.** Функция  $u \in \mathring{W}_2^1(Q)$  является обобщённым решением задачи (1), (2), если для всех  $w \in \mathring{W}_2^1(Q)$  выполнено интегральное тождество

$$(R_Q u', w')_{L_2(Q)} = (f, w)_{L_2(Q)}.$$

**Теорема 1.** Пусть выполнено условие сильной эллиптичности (5). Тогда для любой функции  $f \in L_2(Q)$  существует единственное обобщённое решение  $u \in \mathring{W}_2^1(Q)$  задачи (1), (2), при этом имеет место оценка

$$\|u\|_{W_2^1(Q)} \leq c_3 \|f\|_{L_2(Q)},$$

где  $c_3 > 0$  – постоянная, не зависящая от  $f$ .

Доказательство теоремы см. в [4].

**Следствие 1.** Пусть уравнение (1) удовлетворяет условию сильной эллиптичности. Тогда оператор  $\mathcal{A}_R$  фредгольмов,  $\text{ind } \mathcal{A}_R = 0$  и  $\dim \mathcal{N}(\mathcal{A}_R) = 0$ .

**3. Гладкость обобщённых решений.** Рассмотрим вначале гладкость обобщённых решений задачи (1), (2) на подынтервалах. В этом случае справедлива следующая

**Теорема 2.** Пусть уравнение (1) удовлетворяет условию сильной эллиптичности (5), и пусть  $u \in \mathring{W}_2^1(Q)$  – обобщённое решение задачи (1), (2). Тогда  $u \in W_2^2(Q_{sk})$ ,  $k = \overline{1, N(s)}$ ;  $s = 1, 2$ , если  $0 < \theta < 1$ , и  $s = 1$ , если  $\theta = 1$ , при этом

$$\|u\|_{W_2^2(Q_{sk})} \leq c_4 \|f\|_{L_2(Q)}, \tag{7}$$

где  $c_4 > 0$  не зависит от  $f$ .

Доказательство теоремы аналогично доказательству теоремы 4.1 из работы [5].

**Следствие 2.** Пусть выполняется неравенство (5), а  $u \in \mathring{W}_2^1(Q)$  – обобщённое решение задачи (1), (2), где  $f \in L_2(Q)$ . Тогда уравнение (1) удовлетворяется почти всюду на  $Q$ .

В силу теоремы 2 и следствия 2 область определения оператора  $\mathcal{A}_R$  примет вид

$$D(\mathcal{A}_R) = \{u \in \dot{W}_2^1(Q) : R_Q u' \in W_2^1(Q), u \in W_2^2(Q_{sk}), k = \overline{1, N(s)},$$

$$s = 1, 2, \text{ если } 0 < \theta < 1, \text{ и } s = 1, \text{ если } \theta = 1\}. \tag{8}$$

Рассмотрим ограниченный оператор  $A_R^0 : W_2^2(Q) \supset D(A_R^0) \rightarrow L_2(Q)$  с областью определения  $D(A_R^0) = \{u \in W_2^2(Q) : R_Q u' \in W_2^1(Q), u(0) = u(d) = 0\}$ , действующий по формуле  $A_R^0 u = \mathcal{A}_R u$  при  $u \in D(A_R^0)$ . Очевидно, что

$$D(A_R^0) = D(\mathcal{A}_R) \cap W_2^2(Q). \tag{9}$$

Для формулировки результатов о гладкости обобщённого решения задачи (1), (2) на всём интервале  $Q = (0, d)$  введём некоторые обозначения и докажем вспомогательные результаты.

Будем предполагать, что выполнены условия

$$\sum_{k=1}^N |b_{-k}(k)| \neq 0 \text{ или } \sum_{k=1}^N |b_k(N+1-k)| \neq 0, \text{ если } \theta = 1, \tag{10}$$

$$\sum_{k=1}^N |b_{-k}(k)| \neq 0 \text{ или } \sum_{k=1}^N |b_k(N+\theta-k)| \neq 0, \text{ если } 0 < \theta < 1. \tag{11}$$

Через  $G_j^1 = G_j^1(x)$  ( $G_j^2 = G_j^2(x)$ ),  $j = \overline{1, N+1}$ , обозначим  $j$ -й столбец матрицы размерности  $N \times (N+1)$ , полученной из матрицы  $R_1(x)$  вычёркиванием первой (последней) строки.

**Замечание 4.** Из условий (10) и (11) вытекает, что  $G_1^1(0) \neq 0$  или  $G_{N+1}^2(\theta) \neq 0$  при  $0 < \theta \leq 1$ .

Рассмотрим задачу (1), (2). В силу теоремы 1 операторное уравнение  $\mathcal{A}_R u = f$  имеет единственное решение  $u_f$  при любой правой части  $f \in L_2(Q)$ . Из (9) следует, что  $u_f \in D(A_R^0)$  в том и только том случае, когда

$$u_f \in W_2^2(Q). \tag{12}$$

**Лемма 6.** Пусть  $\theta = 1$  и выполнены условия (5) и (10).

Если столбцы  $G_1^1(0)$  и  $G_{N+1}^2(1)$  линейно независимы, то обобщённое решение  $u$  задачи (1), (2) принадлежит пространству  $W_2^2(Q)$  тогда и только тогда, когда  $u'(0+0) = u'(N+1-0) = 0$ .

Если же существуют такие  $\alpha_1, \alpha_2 \in \mathbb{C}$ , что  $\alpha_1 G_1^1(0) + \alpha_2 G_{N+1}^2(1) = 0$ ,  $\alpha_1^2 + \alpha_2^2 \neq 0$ , т.е. столбцы  $G_1^1(0)$  и  $G_{N+1}^2(1)$  линейно зависимы, то  $u \in W_2^2(Q)$  тогда и только тогда, когда  $\alpha_2 u'(0+0) + \alpha_1 u'(N+1-0) = 0$ .

**Доказательство.** 1. Рассмотрим случай, когда столбцы  $G_1^1(0)$  и  $G_{N+1}^2(1)$  линейно независимы. В силу определения  $D(\mathcal{A}_R)$ , теоремы 2 и вложения  $W_2^2(Q_{1k}) \subset C^1(\overline{Q_{1k}})$ ,  $k = \overline{1, N+1}$ , на концах подынтервалов  $Q_{1k}$  определены значения  $u'(x)$ . Тогда можно записать условие  $u \in W_2^2(Q)$  в виде

$$u'(k+0) = u'(k-0), \quad k = \overline{1, N}. \tag{13}$$

С другой стороны,  $u \in D(\mathcal{A}_R)$ , следовательно,

$$(R_Q u')(k+0) = (R_Q u')(k-0), \quad k = \overline{1, N}. \tag{14}$$

В силу (3), (4) и леммы 3 равенства (14) можно записать в виде

$$\sum_{j=1}^N (-r_{i,j}^1(1)\psi_j + r_{i+1,j+1}^1(0)\varphi_j) = -r_{i+1,1}^1(0)\varphi_0 + r_{i,N+1}^1(1)\psi_{N+1}, \quad i = \overline{1, N}, \tag{15}$$

где  $\varphi_j = (U_1 u')_{j+1}(0 + 0)$ ,  $j = \overline{0, N}$ , и  $\psi_j = (U_1 u')_j(1 - 0)$ ,  $j = \overline{1, N + 1}$ . Так как  $r_{i,j}^1(1) = r_{i+1,j+1}^1(0)$  в силу (4), то при выполнении (13) из (15) следует соотношение

$$\varphi_0 G_1^1(0) - \psi_{N+1} G_{N+1}^2(1) = 0. \tag{16}$$

В силу линейной независимости столбцов  $G_1^1(0)$  и  $G_{N+1}^2(1)$  получаем

$$u'(0 + 0) = \varphi_0 = 0, \tag{17}$$

$$u'(N + 1 - 0) = \psi_{N+1} = 0. \tag{18}$$

Таким образом, из условия  $u \in W_2^2(Q)$  вытекают равенства (17) и (18). Обратное из (15), (17), (18) и условия (5), обеспечивающего невырожденность матрицы  $R_2(1)$ , вытекают равенства (13), т.е. принадлежность решения  $u$  пространству  $W_2^2(Q)$ .

2. Пусть теперь столбцы  $G_1^1(0)$  и  $G_{N+1}^2(1)$  линейно зависимы. В этом случае существуют такие числа  $\alpha_1, \alpha_2 \in \mathbb{C}$ , что  $\alpha_1 G_1^1(0) + \alpha_2 G_{N+1}^2(1) = 0$ ,  $\alpha_1^2 + \alpha_2^2 \neq 0$ . В силу (10), не ограничивая общности, можно предположить, что  $G_{N+1}^2(1) \neq 0$ . Тогда  $\alpha_1 \neq 0$  и

$$G_1^1(0) = -\frac{\alpha_2}{\alpha_1} G_{N+1}^2(1).$$

Следовательно, в силу (16) имеем

$$\left(-\frac{\alpha_2}{\alpha_1} \varphi_0 - \psi_{N+1}\right) G_{N+1}^2(1) = 0,$$

т.е.  $\alpha_2 \varphi_0 + \alpha_1 \psi_{N+1} = \alpha_2 u'(0 + 0) + \alpha_1 u'(N + 1 - 0) = 0$ . Лемма доказана.

**Лемма 7.** Пусть  $0 < \theta < 1$  и выполнены условия (5) и (11).

Если  $G_1^1(0) \neq 0$ ,  $G_{N+1}^2(\theta) \neq 0$ , то обобщённое решение  $u$  задачи (1), (2) принадлежит пространству  $W_2^2(Q)$  тогда и только тогда, когда  $u'(0 + 0) = u'(N + \theta - 0) = 0$ .

Если  $G_1^1(0) = 0$  или  $G_{N+1}^2(\theta) = 0$ , то обобщённое решение  $u$  задачи (1), (2) принадлежит пространству  $W_2^2(Q)$  тогда и только тогда, когда  $u'(0 + 0) = 0$  в случае  $G_1^1(0) \neq 0$  и  $u'(N + \theta - 0) = 0$  в случае  $G_{N+1}^2(\theta) \neq 0$ .

**Доказательство.** Аналогично части 1 доказательства леммы 6 можно записать условие (12) в виде

$$u'(k + 0) = u'(k - 0), \quad k = \overline{1, N}, \tag{19}$$

$$u'(k - 1 + \theta + 0) = u'(k - 1 + \theta - 0), \quad k = \overline{1, N}. \tag{20}$$

С другой стороны, из  $u \in D(\mathcal{A}_R)$  и (8) следует

$$(R_Q u')(k + 0) = (R_Q u')(k - 0), \quad k = \overline{1, N}, \tag{21}$$

$$(R_Q u')(k - 1 + \theta + 0) = (R_Q u')(k - 1 + \theta - 0), \quad k = \overline{1, N}. \tag{22}$$

Рассмотрим, например, соотношения (22). В силу (3), (4) и леммы 3 эти соотношения можно представить в виде

$$\sum_{j=1}^N r_{i,j}^1(\theta)(\varphi^j - \psi^j) = r_{i,N+1}^1(\theta)\psi^{N+1}, \quad i = \overline{1, N}, \tag{23}$$

где  $\varphi^j = (U_2 P_2 u')_j(\theta + 0)$ ,  $j = \overline{1, N}$ , и  $\psi^j = (U_1 P_1 u')_j(\theta - 0)$ ,  $j = \overline{1, N + 1}$ . Поскольку  $\det R_2(\theta) \neq 0$  (см. (5)), то при условии выполнения равенств (20) система (23) имеет единственное решение  $\varphi^j - \psi^j = 0$ ,  $j = \overline{1, N}$ , если  $\psi^{N+1} G_{N+1}^2(\theta) = 0$ .

Пусть, например, выполнено второе из условий (11). Тогда  $G_{N+1}^2(\theta) \neq 0$ . Следовательно,

$$u'(N + \theta - 0) = \psi^{N+1} = 0.$$

Аналогичным образом доказывается, что при выполнении условия (19) вытекающая из равенств (21) система

$$r_{i+1,1}^1(0)\varphi_0 = \sum_{j=1}^N (r_{i,j}^2(1)\psi_j - r_{i+1,j+1}^1(0)\varphi_j), \quad i = \overline{1, N},$$

имеет единственное решение  $\varphi_j - \psi_j = 0, \quad j = \overline{1, N}$ , если  $G_1^1(0) \neq 0$  и

$$u'(0+0) = \varphi_0 = 0,$$

где  $\varphi_j = (U_1 P_1 u')_{j+1}(0+0), \quad j = \overline{0, N}$ , и  $\psi_j = (U_2 P_2 u')_j(1-0), \quad j = \overline{1, N}$ . Лемма доказана.

**Теорема 3.** Пусть  $\theta = 1$  и выполнены условия (5) и (10).

Если столбцы  $G_1^1(0)$  и  $G_{N+1}^2(1)$  линейно независимы, то ограниченный оператор  $A_R^0 : W_2^2(Q) \supset D(A_R^0) \rightarrow L_2(Q)$  фредгольмов и  $\dim \mathcal{N}(A_R^0) = 0, \quad \text{codim } \mathcal{R}(A_R^0) = 2$ .

Если столбцы  $G_1^1(0)$  и  $G_{N+1}^2(1)$  линейно зависимы, то ограниченный оператор  $A_R^0 : W_2^2(Q) \supset D(A_R^0) \rightarrow L_2(Q)$  фредгольмов и  $\dim \mathcal{N}(A_R^0) = 0, \quad \text{codim } \mathcal{R}(A_R^0) = 1$ .

**Доказательство.** 1. В силу теоремы 1 уравнение (6) имеет единственное решение для любой правой части  $f \in L_2(Q)$ , т.е.  $\mathcal{N}(A_R) = \{0\}$ . По определению оператор  $A_R^0$  является сужением оператора  $A_R$  на  $W_2^2(Q)$ , поэтому  $\dim \mathcal{N}(A_R^0) = 0$ .

2. Предположим, что столбцы  $G_1^1(0)$  и  $G_{N+1}^2(1)$  линейно независимы. Докажем, что оператор  $A_R^0$  фредгольмов и  $\text{codim } \mathcal{R}(A_R^0) = 2$ . В силу леммы 6 решение  $u_f$  операторного уравнения (6) принадлежит области определения  $D(A_R^0) = D(A_R) \cap W_2^2(Q)$  тогда и только тогда, когда

$$u'_f(0+0) = 0, \tag{24}$$

$$u'_f(N+1-0) = 0. \tag{25}$$

Из неравенства (7) и вложения  $W_2^2(Q_{sk}) \subset C^1(\overline{Q}_{sk})$  следует, что  $u'_f(0+0)$  и  $u'_f(N+1-0)$  являются линейными ограниченными функционалами на пространстве  $L_2(Q)$ . Тогда в силу теоремы Рисса об общем виде функционала в гильбертовом пространстве существуют единственным образом определённые функции  $g_1, g_2 \in L_2(Q)$  такие, что  $u'_f(0+0) = (f, g_1)_{L_2(Q)}$  и  $u'_f(N+1-0) = (f, g_2)_{L_2(Q)}$ . Таким образом, равенства (24), (25) можно записать в виде

$$(f, g_i)_{L_2(Q)} = 0, \quad i = 1, 2.$$

Остаётся доказать, что функции  $g_1$  и  $g_2$  линейно независимы. Для этого достаточно показать, что существуют функции  $f_1, f_2 \in L_2(Q)$ , обладающие свойством

$$(f_j, g_i)_{L_2(Q)} = \delta_{ij}, \quad i, j = 1, 2,$$

где  $\delta_{ij}$  – символ Кронекера.

Докажем, например, что существует функция  $f_1 \in L_2(Q)$  такая, что

$$(f_1, g_1)_{L_2(Q)} = 1, \quad (f_1, g_2)_{L_2(Q)} = 0.$$

Другими словами, нужно построить функцию  $f_1$ , для которой

$$u'_{f_1}(0+0) := \tilde{\varphi}_0 = 1, \quad u'_{f_1}(N+1-0) := \tilde{\psi}_{N+1} = 0. \tag{26}$$

Положим в (15)  $\varphi_0 = \tilde{\varphi}_0 = 1, \quad \psi_1 = \psi_2 = \dots = \psi_{N+1} = \tilde{\psi}_{N+1} = 0$ . Тогда система уравнений (15) примет вид

$$\sum_{j=1}^N r_{i,j}^1(1)\varphi_j = -r_{i+1,1}^1(0). \tag{27}$$

В силу условия (5) определитель системы уравнений (27) не равен нулю. Поэтому существует единственное решение этой системы, которое обозначим через  $(\tilde{\varphi}_1, \dots, \tilde{\varphi}_N)$ .

Остаётся доказать, что существует функция  $f_1 \in L_2(Q)$  такая, что решение  $u_{f_1}$  уравнения (6) удовлетворяет условиям (26) и  $u'_{f_1}(j+0) = \tilde{\varphi}_j, u'_{f_1}(j-0) = 0, j = \overline{1, N}$ .

Введём функцию

$$w(x) = \begin{cases} \sum_{j=0}^N (x-j)\tilde{\varphi}_j \xi(x-j), & x \in \bigcup_{j=0}^N (j, j+1/2), \\ 0, & x \in \bigcup_{j=1}^{N+1} (j-1/2, j), \end{cases}$$

где  $\xi \in C_0^\infty(\mathbb{R})$  – вещественнозначная функция,  $0 \leq \xi(x) \leq 1; \xi(x) = 1, x \in [-1/8, 1/8], \text{supp } \xi \subset [-1/4, 1/4]; \tilde{\varphi}_0 = 1; \tilde{\varphi}_1, \dots, \tilde{\varphi}_N$  удовлетворяют системе линейных алгебраических уравнений (27).

По построению  $w \in D(\mathcal{A}_R)$ . Обозначим  $f_1 := \mathcal{A}_R w$ . Тогда, положив  $u_{f_1} = w$ , получим равенства (26).

Таким образом, мы доказали, что в случае линейно независимых столбцов  $G_1^1(0)$  и  $G_{N+1}^2(1)$  оператор  $A_R^0$  фредгольмов и  $\text{codim } \mathcal{R}(A_R^0) = 2$ .

3. Предположим теперь, что столбцы  $G_1^1(0)$  и  $G_{N+1}^2(1)$  линейно зависимы, т.е. существуют числа  $\alpha_1, \alpha_2 \in \mathbb{C}, \alpha_1^2 + \alpha_2^2 \neq 0$ , такие, что  $\alpha_1 G_1^1(0) + \alpha_2 G_{N+1}^2(1) = 0$ . Докажем, что оператор  $A_R^0$  фредгольмов и  $\text{codim } \mathcal{R}(A_R^0) = 1$ .

В силу леммы 6 решение  $u_f$  операторного уравнения (6) принадлежит  $D(A_R^0)$  тогда и только тогда, когда

$$\alpha_2 u'_f(0+0) + \alpha_1 u'_f(N+1-0) = 0. \tag{28}$$

Поскольку  $u'_f(0+0)$  и  $u'_f(N+1-0)$  являются линейными ограниченными функционалами на пространстве  $L_2(Q)$ , в силу теоремы Рисса об общем виде функционала в гильбертовом пространстве существует единственная функция  $g \in L_2(Q)$  такая, что

$$\alpha_2 u'_f(0+0) + \alpha_1 u'_f(N+1-0) = (f, g)_{L_2(Q)}.$$

Таким образом, равенство (28) можно записать в виде  $(f, g)_{L_2(Q)} = 0$ .

Остаётся доказать, что существует функция  $f \in L_2(Q)$ , удовлетворяющая равенству

$$(f, g)_{L_2(Q)} = 1. \tag{29}$$

Не ограничивая общности, предположим, что  $\alpha_1 \neq 0$ . В силу (29) достаточно построить функцию  $f_1 \in L_2(Q)$ , для которой выполняются равенства

$$u'_{f_1}(0+0) := \tilde{\varphi}_0 = 0, \quad u'_{f_1}(N+1-0) := \tilde{\psi}_{N+1} = \frac{1}{\alpha_1}. \tag{30}$$

Положим в (15)  $\varphi_0 = \tilde{\varphi}_0 = 0, \varphi_1 = \varphi_2 = \dots = \varphi_N = 0, \psi_{N+1} = \tilde{\psi}_{N+1} = 1/\alpha_1$ . Тогда система уравнений (15) примет вид

$$\sum_{j=1}^N r_{i,j}^1(1)\psi_j = -\frac{r_{i,N+1}^1(1)}{\alpha_1}. \tag{31}$$

Ввиду условия (5) определитель системы уравнений (31) не равен нулю. Поэтому в силу (10) существует единственное нетривиальное решение системы (31), которое обозначим через  $(\tilde{\psi}_1, \dots, \tilde{\psi}_N)$ .

Введём функцию

$$w(x) = \begin{cases} 0, & x \in \bigcup_{j=0}^N (j, j + 1/2), \\ \sum_{j=1}^{N+1} (x - j) \tilde{\psi}_j \xi(x - j), & x \in \bigcup_{j=1}^{N+1} (j - 1/2, j), \end{cases}$$

где  $\tilde{\psi}_{N+1} = 1/\alpha_1$ ,  $\tilde{\psi}_1, \dots, \tilde{\psi}_N$  удовлетворяют системе линейных алгебраических уравнений (31).

По построению  $w \in D(\mathcal{A}_R)$ . Обозначим  $f_1 := \mathcal{A}_R w$ . Тогда, положив  $u_{f_1} = w$ , получим равенства (30).

Таким образом, доказано, что в случае линейно зависимых столбцов  $G_1^1(0)$  и  $G_{N+1}^2(1)$  оператор  $A_R^0$  фредгольмов и  $\text{codim } \mathcal{R}(A_R^0) = 1$ . Теорема доказана.

**Теорема 4.** Пусть  $0 < \theta < 1$  и выполнены условия (5) и (11).

Если  $G_1^1(0) \neq 0$  и  $G_{N+1}^2(\theta) \neq 0$ , то ограниченный оператор  $A_R^0 : W_2^2(Q) \supset D(A_R^0) \rightarrow L_2(Q)$  фредгольмов и  $\dim \mathcal{N}(A_R^0) = 0$ ,  $\text{codim } \mathcal{R}(A_R^0) = 2$ .

Если  $G_1^1(0) = 0$  или  $G_{N+1}^2(\theta) = 0$ , то ограниченный оператор  $A_R^0 : W_2^2(Q) \supset D(A_R^0) \rightarrow L_2(Q)$  фредгольмов и  $\dim \mathcal{N}(A_R^0) = 0$ ,  $\text{codim } \mathcal{R}(A_R^0) = 1$ .

**Доказательство.** 1. Аналогично доказательству теоремы 3 легко показать, что

$$\dim \mathcal{N}(A_R^0) = 0.$$

2. Предположим, что  $G_1^1(0) \neq 0$ ,  $G_{N+1}^2(\theta) \neq 0$ . В силу леммы 7 решение  $u_f$  операторного уравнения (6) принадлежит пространству  $W_2^2(Q)$  тогда и только тогда, когда

$$u'_f(0+0) = 0, \tag{32}$$

$$u'_f(N + \theta - 0) = 0. \tag{33}$$

Из неравенства (7) и вложения  $W_2^2(Q_{sk}) \subset C^1(\overline{Q}_{sk})$  следует, что  $u'_f(0+0)$  и  $u'_f(N + \theta - 0)$  являются линейными ограниченными функционалами на пространстве  $L_2(Q)$ . Тогда в силу теоремы Рисса об общем виде функционала в гильбертовом пространстве существуют единственным образом определённые функции  $g_1, g_2 \in L_2(Q)$  такие, что  $u'_f(0+0) = (f, g_1)_{L_2(Q)}$  и  $u'_f(N + \theta - 0) = (f, g_2)_{L_2(Q)}$ . Таким образом, равенства (32), (33) можно записать в виде

$$(f, g_i)_{L_2(Q)} = 0, \quad i = 1, 2.$$

Остаётся доказать, что функции  $g_1, g_2$  линейно независимы. Для этого достаточно построить функции  $f_1, f_2 \in L_2(Q)$ , для которых  $(f_j, g_i)_{L_2(Q)} = \delta_{ij}$ ,  $i, j = 1, 2$ .

Докажем, например, существование функции  $f_2 \in L_2(Q)$  такой, что

$$(f_2, g_1)_{L_2(Q)} = 0, \quad (f_2, g_2)_{L_2(Q)} = 1.$$

Другими словами, достаточно построить функцию  $f_2$ , удовлетворяющую равенствам

$$u'_{f_2}(0+0) := \tilde{\varphi}_{2,0} = 0, \quad u'_{f_2}(N + \theta - 0) := \tilde{\psi}_{2,N+1} = 1. \tag{34}$$

В силу (3), (4) и леммы 3 соотношения (21), (22) можно соответственно записать в виде

$$\sum_{j=1}^N r_{i,j}^1(1)(\varphi_{1,j} - \psi_{1,j}) = -r_{i+1,1}^1(0)\varphi_{1,0}, \quad i = \overline{1, N}, \tag{35}$$

$$\sum_{j=1}^N r_{i,j}^1(\theta)(\varphi_{2,j} - \psi_{2,j}) = r_{i,N+1}^1(\theta)\psi_{2,N+1}, \quad i = \overline{1, N}. \tag{36}$$

Здесь

$$\begin{aligned} \varphi_{1,j} &= (U_1 P_1 u'_{f_2})_j(0+0), & j = \overline{0, N}, \\ \psi_{1,j} &= (U_2 P_2 u'_{f_2})_j(1-0), & j = \overline{1, N}, \\ \varphi_{2,j} &= (U_2 P_2 u'_{f_2})_j(\theta+0), & j = \overline{1, N}, \\ \psi_{2,j} &= (U_1 P_1 u'_{f_2})_j(\theta-0), & j = \overline{1, N+1}. \end{aligned}$$

Положим в (36)  $\varphi_{2,1}, \dots, \varphi_{2,N} = 0$ ,  $\psi_{2,N+1} = \tilde{\psi}_{2,N+1} = 1$ . Тогда система уравнений (36) примет вид

$$-\sum_{j=1}^N r_{i,j}^1(\theta) \psi_{2,j} = r_{i,N+1}^1(\theta). \tag{37}$$

В силу условия (5) определитель системы (37) отличен от нуля. Поэтому ввиду (11) существует единственное нетривиальное решение системы (37). Обозначим это решение через  $(\tilde{\psi}_{2,1}, \dots, \tilde{\psi}_{2,N})$ .

Остаётся построить функцию  $f_2 \in L_2(Q)$  такую, что решение  $u_{f_2}$  уравнения (6) удовлетворяет условиям (34), при этом  $u'_{f_2}(\theta + j - 1 + 0) = 0$  и  $u'_{f_2}(\theta + j - 1 - 0) = \tilde{\psi}_{2,j}$ ,  $j = \overline{1, N}$ .

Введём функцию

$$w(x) = \begin{cases} 0, & x \in Q \setminus \bigcup_{j=1}^N (\theta + j - 1 - \sigma, \theta + j - 1), \\ \sum_{j=1}^{N+1} (x - (\theta + j - 1)) \tilde{\psi}_{2,j} \eta(x - (\theta + j - 1)), & x \in \bigcup_{j=1}^N (\theta + j - 1 - \sigma, \theta + j - 1). \end{cases}$$

Здесь  $\eta \in C_0^\infty(\mathbb{R})$  – вещественнозначная функция,  $0 \leq \eta(x) \leq 1$ ;  $\eta(x) = 1$ ,  $x \in [-\sigma/3, \sigma/3]$ ,  $\text{supp } \eta \subset [-\sigma/2, \sigma/2]$ ;  $\tilde{\psi}_{2,N+1} = 1$ ;  $\tilde{\psi}_{2,1}, \dots, \tilde{\psi}_{2,N}$  удовлетворяют системе линейных алгебраических уравнений (37),  $0 < 2\sigma < \min\{\theta, 1 - \theta\}$ .

Очевидно, что в этом случае числа  $\tilde{\varphi}_{1,j} = 0$ ,  $j = \overline{0, N}$ , и  $\tilde{\psi}_{1,j} = 0$ ,  $j = \overline{1, N}$ , удовлетворяют системе уравнений (35).

По построению  $w \in D(\mathcal{A}_R)$ . Положим  $f_2 := \mathcal{A}_R w$ . Тогда, положив  $u_{f_2} = w$ , получим равенства  $(f_2, g_i)_{L_2(Q)} = \delta_{i,2}$ ,  $i = 1, 2$ .

Аналогично можно построить функцию  $f_1 \in L_2(Q)$  такую, что  $(f_1, g_i)_{L_2(Q)} = \delta_{i,1}$ ,  $i = 1, 2$ .

Таким образом, оператор  $A_R^0$  фредгольмов и  $\text{codim } \mathcal{R}(A_R^0) = 2$ .

3. Если  $G_1^1(0) = 0$  или  $G_{N+1}^2(\theta) = 0$ , то доказательство того, что оператор  $A_R^0$  фредгольмов и  $\text{codim } \mathcal{R}(A_R^0) = 1$ , проводится так же, как в части 2. Отличие заключается лишь в том, что в случае  $G_1^1(0) \neq 0$ ,  $G_{N+1}^2(\theta) = 0$  рассматривается система уравнений (35), а в случае  $G_1^1(0) = 0$ ,  $G_{N+1}^2(\theta) \neq 0$  рассматривается система (36), а не две системы (35), (36) одновременно. Теорема доказана.

**Пример 1.** Рассмотрим оператор  $R_Q : L_2(0, 3) \rightarrow L_2(0, 3)$ , где  $Q = (0, 3)$ ,  $(Ru)(x) = b_0 u(x) + (e^{x+1} - 1)u(x+1) + e^{x-1}u(x-1)$ ,  $b_0 \in \mathbb{R}$ . Тогда  $N = 2$ ,  $\theta = 1$ , а матрица  $R_1(x)$  примет вид

$$R_1(x) = \begin{pmatrix} b_0 & e^{x+1} - 1 & 0 \\ e^x & b_0 & e^{x+2} - 1 \\ 0 & e^{x+1} & b_0 \end{pmatrix}, \quad 0 \leq x \leq 1.$$

Предположим, что коэффициент  $b_0 > 0$  удовлетворяет условию

$$b_0^2 > \frac{(e^3 + e^2 - 1)^2 + (e^2 + e - 1)^2}{4}.$$

Тогда матрица  $R_1(x)$ ,  $0 \leq x \leq 1$ , удовлетворяет условию (5). Заметим, что столбцы  $G_1^1(0) = (1, 0)^T$  и  $G_3^2(1) = (0, e^3 - 1)^T$  линейно независимы, а условие (10), очевидно, выполнено.

Таким образом, в силу первой части теоремы 3 оператор  $A_R^0 : W_2^2(0, 3) \supset D(A_R^0) \rightarrow L_2(0, 3)$  фредгольмов,  $\dim \mathcal{N}(A_R^0) = 0$ , при этом  $\text{codim } \mathcal{R}(A_R^0) = 2$ .

**Пример 2.** Рассмотрим оператор  $R_Q : L_2(0, 3) \rightarrow L_2(0, 3)$ , где  $Q = (0, 3)$ ,  $(Ru)(x) = b_0u(x) + (1 - e^{x-2})u(x+1) + u(x-1) + u(x+2)$ ,  $b_0 \in \mathbb{R}$ . Тогда  $N = 2$ ,  $\theta = 1$ , а матрица  $R_1(x)$  имеет вид

$$R_1(x) = \begin{pmatrix} b_0 & 1 - e^{x-2} & 1 \\ 1 & b_0 & 1 - e^{x-1} \\ 0 & 1 & b_0 \end{pmatrix}, \quad 0 \leq x \leq 1.$$

Нетрудно заметить, что условие (10) выполнено, а столбцы  $G_1^1(0) = (1, 0)^T$  и  $G_3^2(1) = (1, 0)^T$  линейно зависимы.

Будем предполагать, что коэффициент  $b_0 > 0$  удовлетворяет неравенству

$$4b_0^3 - b_0(2 + (2 - e^{-1})^2) + (2 - e^{-1}) > 0.$$

Тогда матрица  $R_1(x)$ ,  $0 \leq x \leq 1$ , удовлетворяет условию (5).

Следовательно, в силу второй части теоремы 3 оператор  $A_R^0 : W_2^2(0, 3) \supset D(A_R^0) \rightarrow L_2(0, 3)$  фредгольмов,  $\dim \mathcal{N}(A_R^0) = 0$  и  $\text{codim } \mathcal{R}(A_R^0) = 1$ .

**Пример 3.** Рассмотрим оператор  $R_Q : L_2(0, 11/4) \rightarrow L_2(0, 11/4)$ , где  $Q = (0, 11/4)$ ,  $(Ru)(x) = b_0u(x) + e^{x-1}u(x+1) + e^{x-2}u(x-1)$ ,  $b_0 \in \mathbb{R}$ . Следовательно,  $N = 2$ ,  $\theta = 3/4$ , при этом

$$R_1(x) = \begin{pmatrix} b_0 & e^{x-1} & 0 \\ e^{x-1} & b_0 & e^x \\ 0 & e^x & b_0 \end{pmatrix}, \quad x \in [0, 3/4],$$

$$R_2(x) = \begin{pmatrix} b_0 & e^{x-1} \\ e^{x-1} & b_0 \end{pmatrix}, \quad x \in [3/4, 1].$$

Матрицы  $R_1(x)$  при  $0 \leq x \leq 3/4$  и  $R_2(x)$  при  $3/4 \leq x \leq 1$  удовлетворяют условию (5) при

$$b_0 > 0, \quad b_0^2 > e^{-1/2} + e^{3/2}. \tag{38}$$

Будем предполагать, что условие (38) выполнено.

Очевидно, что столбцы  $G_1^1(0) = (e^{-1}, 0)^T$  и  $G_3^2(\theta) = (0, e^{3/4})^T$  ненулевые, т.е. выполнено условие (11).

Таким образом, в силу теоремы 4 оператор  $A_R^0 : W_2^2(0, 11/4) \supset D(A_R^0) \rightarrow L_2(0, 11/4)$  фредгольмов,  $\dim \mathcal{N}(A_R^0) = 0$  и  $\text{codim } \mathcal{R}(A_R^0) = 2$ .

**Заключение.** Отметим, что используемый в данной работе подход к нахождению числа условий типа ортогональности на правую часть дифференциально-разностного уравнения, обеспечивающих соответствующую гладкость обобщённого решения на всём интервале  $(0, d)$ , является более универсальным, чем разработанный ранее в [3, 4]. В них использовалось сведение дифференциально-разностного уравнения с постоянными коэффициентами и однородными условиями Дирихле к обыкновенному дифференциальному уравнению с многоточечными краевыми условиями. Однако в случае дифференциально-разностного уравнения с переменными коэффициентами такой подход невозможен, так как разностный оператор с переменными коэффициентами и однородными условиями Дирихле не коммутирует с оператором дифференцирования первого порядка. Поэтому здесь был использован подход, основанный на представлении первых производных решения на концах интервала  $(0, d)$  в виде линейных ограниченных функционалов в пространстве  $L_2(0, d)$ , зависящих от правых частей дифференциально-разностного уравнения. Данный подход впервые был разработан для исследования гладкости обобщённых решений задачи Неймана для дифференциально-разностного уравнения (см. [5, 6]). Необходимость использования такого подхода в работах [5, 6] была связана с тем, что дифференциально-разностное уравнение с условиями Неймана (в отличие от условий Дирихле) нельзя свести к обыкновенному дифференциальному уравнению с нелокальными краевыми условиями (доказательство см. в [9, 10]).

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации в рамках государственного задания (проект FSSF-2023-0016).

## СПИСОК ЛИТЕРАТУРЫ

1. Каменский Г.А., Мышкис А.Д. Постановка краевых задач для дифференциальных уравнений с отклоняющимися аргументами в старших членах // Дифференц. уравнения. 1974. Т. 10. № 3. С. 409–418.
2. Каменский А.Г. Краевые задачи для уравнений с формально симметричными дифференциально-разностными операторами // Дифференц. уравнения. 1976. Т. 12. № 5. С. 815–824.
3. Каменский Г.А., Мышкис А.Д., Скубачевский А.Л. О гладких решениях краевой задачи для дифференциально-разностного уравнения нейтрального типа // Укр. мат. журн. 1985. Т. 37. № 5. С. 581–585.
4. Skubachevskii A.L. Elliptic functional differential equations and applications // Operator Theory. Advances and Applications. Basel; Boston; Berlin, 1997. V. 91.
5. Скубачевский А.Л., Иванов Н.О. Об обобщённых решениях второй краевой задачи для дифференциально-разностных уравнений с переменными коэффициентами // Соврем. математика. Фунд. направления. 2021. Т. 67. № 3. С. 576–595.
6. Скубачевский А.Л., Иванов Н.О. Об обобщённых решениях второй краевой задачи для дифференциально-разностных уравнений с переменными коэффициентами на интервале нецелой длины // Мат. заметки. 2022. Т. 111. № 6. С. 873–886.
7. Неверова Д.А., Скубачевский А.Л. О классических и обобщённых решениях краевых задач для дифференциально-разностных уравнений с переменными коэффициентами // Мат. заметки. 2013. Т. 94. № 5. С. 702–719.
8. Neverova D.A. Generalized and classical solutions to the second and third boundary-value problem for differential-difference equations // Funct. Differ. Equat. 2014. V. 21. P. 47–65.
9. Лийко В.В., Скубачевский А.Л. Сильно эллиптические дифференциально-разностные уравнения со смешанными краевыми условиями в цилиндрической области // Соврем. математика. Фунд. направления. 2019. Т. 65. № 4. С. 635–654.
10. Лийко В.В., Скубачевский А.Л. Смешанные задачи для сильно эллиптических дифференциально-разностных уравнений в цилиндре // Мат. заметки. 2020. Т. 107. № 5. С. 693–716.
11. Красовский Н.Н. Теория управления движением. М., 1968.
12. Осипов Ю.С. О стабилизации управляемых систем с запаздыванием // Дифференц. уравнения. 1965. Т. 1. № 5. С. 605–618.
13. Кряжмский А.В., Максимов В.И., Осипов Ю.С. О позиционном моделировании в динамических системах // Прикл. математика и механика. 1983. Т. 47. № 6. С. 883–890.
14. Скубачевский А.Л. К задаче об успокоении системы управления с последствием // Докл. РАН. 1994. Т. 335. № 2. С. 157–160.
15. Адхамова А.Ш., Скубачевский А.Л. Об одной задаче успокоения нестационарной системы управления с последствием // Соврем. математика. Фунд. направления. 2019. Т. 65. № 4. С. 547–556.
16. Onanov G.G., Skubachevskii A.L. Nonlocal problems in the mechanics of three-layer shells // Math. Model. Nat. Phenom. 2017. V. 12. № 6. P. 192–207.
17. Onanov G.G., Tsvetkov E.L. On the minimum of the energy functional with respect to functions with deviating argument in a stationary problem of elasticity theory // Russ. J. Math. Phys. 1995. V. 3. № 4. P. 491–500.

Российский университет дружбы народов  
имени Патриса Лумумбы, г. Москва,  
Московский центр фундаментальной  
и прикладной математики

Поступила в редакцию 31.05.2023 г.  
После доработки 03.06.2023 г.  
Принята к публикации 14.06.2023 г.

УРАВНЕНИЯ С ЧАСТНЫМИ ПРОИЗВОДНЫМИ

УДК 517.957

О РАЗРУШЕНИИ РЕШЕНИЯ ОДНОЙ ЗАДАЧИ  
 ДЛЯ СОБОЛЕВСКОГО УРАВНЕНИЯ  
 С НЕКОЭРЦИТИВНЫМ ИСТОЧНИКОМ

© 2023 г. М. В. Артемьева, М. О. Корпусов

Рассматривается начально-краевая задача для нелинейного дифференциального уравнения второго порядка соболевского типа с некоэрцитивным источником, которая описывает поведение зарядов во внешнем поле полупроводниковой плазмы. Доказана локальная разрешимость этой задачи и получены оценки сверху на время разрушения решения.

DOI: 10.31857/S0374064123070038, EDN: GTQZSM

**Введение.** В работе [1] авторами рассмотрена следующая абстрактная задача:

$$\frac{d^2}{dt^2} \left( A_0 u + \sum_{j=1}^N A_j(u) \right) + \frac{d}{dt} DP(u) + Lu = \frac{d}{dt} F(u), \quad u(0) = u_0, \quad u'(0) = u_1, \quad (1)$$

где  $A_0$ ,  $D$  и  $L$  – линейные, а  $A_j(u)$  и  $F(u)$ ,  $P(u)$  – нелинейные операторы в банаховых пространствах  $V_0$ ,  $V_j$ ,  $W_i$ ,  $j = \overline{1, n}$ ,  $i = \overline{1, 3}$ , с соответствующими нормами  $\|\cdot\|_0$ ,  $\|\cdot\|_j$ ,  $|\cdot|_i$  и с сопряжёнными банаховыми пространствами  $V_0^*$ ,  $V_j^*$ ,  $W_i^*$  относительно соответствующих скобок двойственностей  $\langle \cdot, \cdot \rangle_0$ ,  $\langle \cdot, \cdot \rangle_j$ ,  $(\cdot, \cdot)_i$  и норм  $\|\cdot\|_0^*$ ,  $\|\cdot\|_j^*$ ,  $|\cdot|_i^*$ .

Пусть банаховы пространства  $V_0$ ,  $V_j$ ,  $W_i$ ,  $j = \overline{1, n}$ ,  $i = \overline{1, 3}$ , являются рефлексивными и сепарабельными. Предположим также, что

$$A_0: V_0 \rightarrow V_0^*, \quad A_j: V_j \rightarrow V_j^*, \quad L_1: W_1 \rightarrow W_1^*,$$

$$F: W_2 \rightarrow W_2^*, \quad P: V_0 \rightarrow W_3, \quad D: W_3 \rightarrow V_0^*.$$

Для операторных коэффициентов задачи (1) выполняются следующие условия [1, § 2].

**Условия  $A_0$ :**

(i) оператор  $A_0: V_0 \rightarrow V_0^*$  является линейным, непрерывным и симметричным, причём имеет место неравенство

$$\|A_0 u\|_0^* \leq M_0 \|u\|_0 \quad \text{для всех } u \in V_0;$$

(ii) оператор  $A_0$  является коэрцитивным, причём справедливо неравенство

$$\langle A_0 u, u \rangle_0 \geq m_0 \|u\|_0^2 \quad \text{для всех } u \in V_0.$$

**Условия  $A_j$ :**

(i) оператор  $A_j: V_j \rightarrow V_j^*$  является монотонным и непрерывным;

(ii) оператор  $A_j$  дифференцируем по Фреше, причём его производная Фреше

$$A'_{jf}(u): V_j \rightarrow \mathcal{L}(V_j, V_j^*)$$

является непрерывным, симметричным, монотонным и неотрицательно определённым оператором при любом фиксированном  $u \in V_j$  и  $A'_{jf}(0) = 0$ ;

(iii) оператор  $A_j$  является положительно однородным:

$$A_j(ru) = r^{p_j-1} A_j(u) \quad \text{при } p_j > 2, \quad r \geq 0, \quad u \in V_j;$$

(iv) справедливы неравенства

$$\|A_j(u)\|_j^* \leq M_j \|u\|_j^{p_j-1}, \quad \langle A_j(u), u \rangle_j \geq m_j \|u\|_j^{p_j}, \quad M_j, m_j > 0.$$

**Условия  $F$ :**

(i) оператор  $F: W_2 \rightarrow W_2^*$  является ограниченно Липшиц-непрерывным, т.е. имеет место неравенство

$$|F(u_1) - F(u_2)|_2 \leq \mu(R) |u_1 - u_2|_2 \quad \text{для всех } u_1, u_2 \in W_2,$$

где  $R = \max\{|u_1|_2, |u_2|_2\}$ ,  $\mu = \mu(R)$  есть ограниченная на всяком компакте неубывающая функция своего аргумента;

(ii) оператор  $F$  является положительно однородным, т.е.

$$F(ru) = r^{1+q} F(u) \quad \text{при } q > 0, \quad r \geq 0, \quad u \in W_2;$$

(iii) оператор  $F$  имеет симметричную производную Фреше

$$F'_f(\cdot): W_2 \rightarrow \mathcal{L}(W_2, W_2^*);$$

(iv) оператор  $F$  удовлетворяет неравенству

$$|F(u)|_2^* \leq M |u|_2^{q+1} \quad \text{для всех } u \in W_2.$$

**Условия  $DP$ :**

(i) оператор  $D: W_3 \rightarrow V_0^*$  является линейным и непрерывным, причём

$$\|Du\|_0^* \leq D_3 |u|_3 \quad \text{для всех } u \in W_3;$$

(ii) оператор  $P: V_0 \rightarrow W_3$  является ограниченно Липшиц-непрерывным, т.е.

$$|P(u_1) - P(u_2)|_3 \leq \mu_0(R) \|u_1 - u_2\|_0 \quad \text{для всех } u_1, u_2 \in V_0,$$

где  $\mu_0 = \mu_0(R)$  – ограниченная на всяком компакте неубывающая функция своего аргумента,  $R = \max\{\|u_1\|_0, \|u_2\|_0\}$ ;

(iii) справедливо неравенство

$$|P(u)|_3 \leq D_4 \|u\|_0^{1+q_0}, \quad q_0 \geq 0 \quad \text{для всех } u \in V_0;$$

(iv) для всех  $u \in V_0$  имеет место неравенство

$$\langle DP(u), u \rangle_0 = 0;$$

(v) оператор  $P$  имеет производную Фреше

$$P'_f(\cdot): V_0 \rightarrow \mathcal{L}(V_0, W_3).$$

**Условия  $L$ :**

(i) оператор  $L: W_1 \rightarrow W_1^*$  является линейным, непрерывным и симметричным, причём

$$|Lu|_1^* \leq D_1 |u|_1 \quad \text{для всех } u \in W_1;$$

(ii) оператор  $L$  является коэрцитивным, причём

$$(L_1 u, u)_1 \geq d_1 |u|_1^2 \quad \text{для всех } u \in W_1.$$

Для доказательства существования сильного решения задачи Коши (1) применён метод монотонных операторов Браудера–Минти [2, с. 98], а для доказательства разрушения за конечное

время – метод энергетических оценок, развитый в статье [3]. Отметим, что уравнение (1) содержит некоэрцитивный источник  $dF(u)/dt$ , что сильно усложняет получение достаточных условий разрушения задачи Коши (1) за конечное время.

В настоящей работе рассмотрена представленная в примере 1 из [1] начально-краевая задача для нелинейного дифференциального уравнения второго порядка и доказано существование непродолжаемого во времени её классического решения.

**1. Вывод уравнения.** Пусть  $\Omega \subset \mathbb{R}^3$  – ограниченная область с гладкой границей  $\Gamma$ , причём  $\Gamma$  – односвязная поверхность. Рассмотрим электрическую часть системы уравнений Максвелла в квазистационарном приближении:

$$\operatorname{div} \mathbf{D} = -4\pi en, \quad \mathbf{E} = -\nabla u, \quad \mathbf{D} = \mathbf{E} + 4\pi \mathbf{P}, \tag{2}$$

$$\frac{\partial n}{\partial t} = \operatorname{div} \mathbf{J} + Q_1(u) + \int_0^t Q_2(u)(\tau) d\tau, \quad \mathbf{J} = en_0(u)\mathbf{E}_0, \quad \operatorname{div} \mathbf{P} = \rho(u), \tag{3}$$

где  $\mathbf{E}_0 = E_0 \mathbf{e}_1$  – вектор напряжённости внешнего электрического поля,  $\mathbf{e}_1$  – базисный орт, направленный вдоль оси  $x$ ,  $E_0$  – амплитуда внешнего электрического поля,  $\mathbf{E}$  – вектор напряжённости электрического поля,  $\mathbf{D}$  – вектор электрического смещения,  $\mathbf{P}$  – вектор поляризации,  $\mathbf{J}$  – плотность тока смещения,  $n_0(u)$  – квазистационарная плотность свободных зарядов, зависящая от электрического потенциала  $u$ ,  $Q_1(u)$  – источник свободных зарядов, зависящий от потенциала электрического поля  $u$ ,  $\rho(u)$  – плотность связанных зарядов, зависящая от потенциала электрического поля  $u$ . Кроме того, здесь учтена временная дисперсия, описываемая слагаемым  $\int_0^t Q_2(u)(\tau) d\tau$ . От системы уравнений (2), (3) приходим к уравнению

$$\frac{\partial^2}{\partial t^2}(\Delta u - 4\pi\rho(u)) = 4\pi e^2 E_0 \frac{\partial^2 n_0(u)}{\partial t \partial x_1} + \frac{\partial Q_1(u)}{\partial t} + Q_2(u), \tag{4}$$

где функции  $\rho(u)$ ,  $n_0(u)$ ,  $Q_1(u)$  и  $Q_2(u)$  можно взять следующими:

$$\rho(u) = \rho_0 + \rho_1 u + \rho_2 |u|^{p-2} u, \quad n_0(u) = N_0 u^2, \quad Q_1(u) = q_1 u^3, \quad Q_2(u) = q_2 u. \tag{5}$$

Таким образом, уравнение (4) описывает поведение зарядов во внешнем электрическом поле в полупроводниковой плазме.

**2. Постановка задачи.** Рассмотрим следующую начально-краевую задачу для нелинейного дифференциального уравнения второго порядка:

$$\frac{\partial^2}{\partial t^2} \left( \Delta u - u - \sum_{j=1}^n |u|^{p_j-2} u \right) - u + \frac{\partial^2 |u|^{1+q_0}}{\partial t \partial x_1} - \frac{\partial |u|^q u}{\partial t} = 0, \tag{6}$$

$$u(0) = u_0 \in H_0^1(\Omega), \quad u'(0) = u_1 \in H_0^1(\Omega), \tag{7}$$

$$u(x, t) = 0 \quad \text{при} \quad x \in \partial\Omega, \tag{8}$$

где  $p_j > 2$ ,  $j = \overline{1, n}$ ,  $q_0 > 0$ ,  $q > 0$ . Это задача вида (1), и операторные коэффициенты имеют вид

$$A_0 u \equiv \Delta u - u, \quad A_j(u) \equiv |u|^{p_j-2} u, \quad Lu \equiv u, \quad P(u) \equiv |u|^{1+q_0}, \quad F(u) \equiv |u|^q u. \tag{9}$$

При этом рассматриваются следующие банаховы пространства:

$$V_0 = H_0^1(\Omega), \quad V_j = L^{p_j}(\Omega), \quad W_1 = H = L^2(\Omega), \quad W_2 = L^{q+2}(\Omega),$$

которые мы полагаем рефлексивными и сепарабельными. Пусть выполнены условия  $A_0$ ,  $A_j$ ,  $F$ ,  $DP$  и  $L$  на операторные коэффициенты (9). Дадим определение классического решений этой задачи.

**Определение 1.** Функция  $u(t) \in C^{(2)}([0, T]; H_0^1(\Omega))$  называется *классическим решением задачи (6)–(8)*, если

$$\frac{d^2}{dt^2}|u|^{p_j-2} \in C([0, T]; H_0^{-1}(\Omega)) \quad \text{для всех } j = \overline{1, n},$$

равенство (6) справедливо для каждого  $t \in [0, T]$  и понимается в смысле банахова пространства  $H_0^{-1}(\Omega)$ , причём  $u_0 \in H_0^1(\Omega)$ ,  $u_1 \in H_0^1(\Omega)$ .

Пусть  $\phi(t) \in C[0, T]$  – произвольная функция. Рассмотрим следующую функцию:

$$\psi(t) := \int_t^T \phi(s) ds \in C^{(1)}[0, T], \quad t \in [0, T].$$

Заметим, что  $\psi(T) = 0$  и  $\psi'(t) = -\phi(t)$ . Справедливы следующие формулы интегрирования по частям для интегралов Бохнера в  $H_0^{-1}(\Omega)$ :

$$\begin{aligned} & \int_0^T \frac{d^2}{dt^2} \left( \Delta u(t) - u(t) - \sum_{j=1}^n |u|^{p_j-2} u(t) \right) \psi(t) dt = \\ & = - \left( \Delta u_1 - u_1 - \sum_{j=1}^n (p_j - 1) |u_0|^{p_j-2} u_1 \right) \int_0^T \phi(t) dt + \\ & + \int_0^T \frac{d}{dt} \left( \Delta u(t) - u(t) - \sum_{j=1}^n |u|^{p_j-2} u(t) \right) \phi(t) dt, \end{aligned} \tag{10}$$

$$\int_0^T u(t) \psi(t) dt = \int_0^T \phi(t) \int_0^t u(s) ds dt, \tag{11}$$

$$\int_0^T \frac{d}{dt} |u|^q u(t) \psi(t) dt = -|u_0|^q u_0 \int_0^T \phi(t) dt + \int_0^T |u|^q u(t) \phi(t) dt, \tag{12}$$

$$\int_0^T \frac{\partial^2 |u|^{1+q_0}}{\partial t \partial x_1}(t) \psi(t) dt = -\frac{\partial |u_0|^{1+q_0}}{\partial x_1} \int_0^T \phi(t) dt + \int_0^T \frac{\partial |u|^{1+q_0}}{\partial x_1}(t) \phi(t) dt. \tag{13}$$

Тогда, умножив обе части равенства (6) на функцию  $\psi(t)$ , с учётом (10)–(13) получим равенство

$$\int_0^T \left\{ \frac{d}{dt} \left( \Delta u - u - \sum_{j=1}^n |u|^{p_j-2} u \right) + \frac{\partial |u|^{1+q_0}}{\partial x_1} - \int_0^t u(s) ds - |u|^q u - f \right\} \phi(t) dt = 0 \tag{14}$$

для всех  $\phi(t) \in C[0, T]$ , где

$$f := -|u_0|^q u_0 + \frac{\partial |u|^{1+q_0}}{\partial x_1} + (\Delta u_1 - u_1) + \sum_{j=1}^n (p_j - 1) |u_0|^{p_j-2} u_1.$$

Полученное равенство (14) позволяет сформулировать определение сильного решения начально-краевой задачи (6)–(8).

**Определение 2.** Функция  $u(t) \in \mathbb{C}^{(1)}([0, T]; H_0^1(\Omega))$  называется *сильным решением задачи (6)–(8)*, если для любой функции  $\phi(t) \in \mathbb{C}[0, T]$  выполнено равенство (14), причём  $u(0) = u_0 \in H_0^1(\Omega)$ ,  $u_1 \in H_0^1(\Omega)$ .

Справедлива следующая

**Теорема 1.** В классе сильных решений задачи (6)–(8) равенство (14), выполненное для любых  $\phi(t) \in \mathbb{C}[0, T]$ , эквивалентно равенству

$$\int_0^T \left\langle \frac{d}{dt} \left( \Delta u - u - \sum_{j=1}^n |u|^{p_j-2} u \right) + \frac{\partial |u|^{1+q_0}}{\partial x_1} - \int_0^t u(s) ds - |u|^q u - f, v(t) \right\rangle_0 dt = 0 \quad (15)$$

для всех  $v(t) \in \mathbb{C}([0, T]; H_0^1(\Omega))$ .

С учётом равенства скобок двойственности

$$\langle f, u \rangle_0 = \langle f, u \rangle_j \quad \text{для всех } f \in L^{p_j}(\Omega)^*, \quad u \in H_0^1(\Omega), \quad j = \overline{1, n},$$

$$\langle f, u \rangle_0 = (f, u)_1, \quad f \in L^2(\Omega)^*, \quad \text{и} \quad \langle f, u \rangle_0 = (f, u)_2, \quad f \in L^{q+2}(\Omega)^*, \quad \text{для всех } u \in H_0^1(\Omega)$$

соотношение (15) можно записать в эквивалентном виде

$$\begin{aligned} \int_0^T \left[ \left\langle \frac{d}{dt} (\Delta u - u)(t), v(t) \right\rangle_0 + \sum_{j=1}^n \left\langle \frac{d}{dt} |u|^{p_j-2} u(t), v(t) \right\rangle_j + \left\langle \frac{\partial |u|^{1+q_0}}{\partial x_1}(t), v(t) \right\rangle_0 + \right. \\ \left. + \int_0^t (u(s), v(t))_1 ds - (|u|^q u(t), v(t))_2 - \langle f, v(t) \rangle_0 \right] dt = 0 \end{aligned} \quad (16)$$

для всех  $v(t) \in \mathbb{C}([0, T]; H_0^1(\Omega))$ . Теперь в (16) возьмём  $v(t) = \phi(t)w$ , где  $\phi(t) \in \mathbb{C}[0, T]$ ,  $w \in H_0^1(\Omega)$ , и получим равенство

$$\begin{aligned} \int_0^T \phi(t) \left[ \left\langle \frac{d}{dt} (\Delta u - u)(t), w \right\rangle_0 + \sum_{j=1}^n \left\langle \frac{d}{dt} |u|^{p_j-2} u(t), w \right\rangle_j + \left\langle \frac{\partial |u|^{1+q_0}}{\partial x_1}(t), w \right\rangle_0 + \right. \\ \left. + \int_0^t (u(s), w)_1 ds - (|u|^q u(t), w)_2 - \langle f, w \rangle_0 \right] dt = 0 \end{aligned} \quad (17)$$

для любых  $\phi(t) \in \mathbb{C}[0, T]$  и  $w \in H_0^1(\Omega)$ . В силу условий на операторные коэффициенты (9) имеем

$$\left\langle \frac{d}{dt} (\Delta u - u), w \right\rangle_0 \in \mathbb{C}[0, T], \quad \left\langle \frac{d}{dt} |u|^{p_j-2} u, w \right\rangle_j \in \mathbb{C}[0, T], \quad (18)$$

$$\int_0^t (u(s), w)_1 ds \in \mathbb{C}[0, T], \quad (|u|^q u, w)_2 \in \mathbb{C}[0, T], \quad (19)$$

$$\left\langle \frac{\partial |u|^{1+q_0}}{\partial x_1}, w \right\rangle_0 \in \mathbb{C}[0, T]. \quad (20)$$

Поэтому из (17) и свойств (18)–(20) в силу основной леммы вариационного исчисления вытекает следующее равенство:

$$\left\langle \frac{d}{dt} (\Delta u - u), w \right\rangle_0 + \sum_{j=1}^n \left\langle \frac{d}{dt} |u|^{p_j-2} u, w \right\rangle_j + \left\langle \frac{\partial |u|^{1+q_0}}{\partial x_1}, w \right\rangle_0 +$$

$$+ \int_0^t (u(s), w)_1 ds - (|u|^q u, w)_2 - \langle f, w \rangle_0 = 0 \tag{21}$$

для всех  $w \in H_0^1(\Omega)$  и всех  $t \in [0, T]$ .

Рассмотрим теперь начально-краевую задачу

$$\frac{d}{dt} \left( \Delta u - u - \sum_{j=1}^n |u|^{p_j-2} u \right) + \frac{\partial |u|^{1+q_0}}{\partial x_1} - \int_0^t u(s) ds = |u|^q u + f, \tag{22}$$

$$u(0) = u_0, \quad u(x, t)|_{\partial\Omega} = 0. \tag{23}$$

**Определение 3.** Классическим решением задачи (22), (23) называется функция  $u(t) \in \mathbb{C}^{(1)}([0, T]; H_0^1(\Omega))$ , удовлетворяющая уравнению (22) для каждого  $t \in [0, T]$  в смысле пространства  $H_0^{-1}(\Omega)$ , причём  $u_0 \in H_0^1(\Omega)$  и  $f \in H_0^{-1}(\Omega)$ .

Очевидно, что классическое решение задачи (22), (23) является сильным решением задачи (6)–(8).

**3. Существование непродолжаемого решения.** Справедлива следующая

**Теорема 2.** Пусть выполнены условия  $A_0, A_j, L, DP$  и  $F$  на операторные коэффициенты. Тогда при дополнительном условии, что операторы  $A_j(u)$  дважды непрерывно дифференцируемы по Фреше для всех  $u \in L^{p_j}(\Omega)$ , для любых  $u_0$  и  $u_1$  из  $H_0^1(\Omega)$  найдётся такое число  $T_0 = T_0(u_0, u_1) > 0$ , что существует единственное классическое решение задачи (6)–(8) класса  $u(t) \in \mathbb{C}^{(2)}([0, T_0]; H_0^1(\Omega))$ , причём либо  $T_0 = +\infty$ , либо  $T_0 < +\infty$ , и в последнем случае имеет место предельное свойство

$$\lim_{t \uparrow T_0} \left\| \Delta u - u - \sum_{j=1}^n |u|^{p_j-2} u \right\|_0^* (t) = +\infty.$$

**Доказательство.** Доказательство теоремы аналогично доказательству теоремы 6.4 работы [4].

**4. Разрушение сильного решения задачи (6) при  $q + 2 > \bar{p}$ .** Рассмотрим классическое решение  $u(t) \in \mathbb{C}^{(1)}([0, T_0]; H_0^1(\Omega))$  задачи (22), (23). Прежде всего введём следующие обозначения:

$$\Phi(t) = \frac{1}{2} \langle \Delta u - u, u \rangle_0 + \sum_{j=1}^n \frac{p_j - 1}{p_j} \langle |u|^{p_j-2} u, u \rangle_j, \tag{24}$$

$$J(t) = \langle \Delta u' - u', u' \rangle_0 + \sum_{j=1}^n (p_j - 1) \langle |u|^{p_j-2} u', u' \rangle_j. \tag{25}$$

**Лемма 1.** Имеет место неравенство

$$(\Phi'(t))^2 \leq \bar{p} J(t) \Phi(t) \quad \text{при} \quad \bar{p} = \max_{j=1, n} p_j, \quad t \in [0, T_0]. \tag{26}$$

Доказательство см. в работе [4, лемма 7.1].

Заметим, что определение 2 сильного решения задачи (6)–(8) эквивалентно равенству (21). Положив сначала в (21)  $w = u(t) \in \mathbb{C}^{(1)}([0, T_0]; H_0^1(\Omega))$ , из определения (24) функционала  $\Phi(t) \in \mathbb{C}[0, T_0]$  и свойства (iv) условий  $DP$

$$\left\langle \frac{\partial |u|^{1+q_0}}{\partial x_1}, u \right\rangle_0 = 0 \quad \text{для всех} \quad u \in H_0^1(\Omega), \tag{27}$$

получим первое энергетическое равенство

$$\frac{d\Phi}{dt} + \int_0^t ds (u(s), u(t))_1 = (|u|^q u, u)_2 + \langle f, u \rangle_0. \quad (28)$$

Теперь положим в равенстве (21)  $w = u'(t) \in \mathbb{C}([0, T_0]; H_0^1(\Omega))$  и с учётом определения (25) функционала  $J(t)$  получим второе энергетическое равенство

$$J(t) + \int_0^t ds (u(s), u'(t))_1 + \left\langle \frac{\partial |u|^{1+q_0}}{\partial x_1}, u' \right\rangle_0 = \frac{1}{q+2} \frac{d}{dt} (|u|^q u, u)_2 + \frac{d}{dt} \langle f, u \rangle_0, \quad (29)$$

где мы воспользовались тождеством

$$(|u|^q u, u')_2 = \frac{1}{q+2} \frac{d}{dt} (|u|^q u, u)_2.$$

Выразив из равенства (28) величину  $(|u|^q u, u)_2$  и подставив её в (29), после элементарных преобразований получим следующее выражение для функционала  $J(t)$ :

$$J(t) = \frac{1}{q+2} \frac{d^2 \Phi(t)}{dt^2} + \frac{1}{q+2} (u, u)_1 - \left\langle \frac{\partial |u|^{1+q_0}}{\partial x_1}, u' \right\rangle_0 - \frac{q+1}{q+2} \int_0^t (u(s), u'(t)) ds + \frac{q+1}{q+2} \langle f, u' \rangle_0. \quad (30)$$

Для дальнейшего воспользуемся неравенством Коши–Буняковского

$$ab \leq \varepsilon a^2 + \frac{1}{4\varepsilon} b^2, \quad a, b \geq 0,$$

с произвольным  $\varepsilon > 0$ . Справедливы следующие неравенства:

$$\frac{1}{q+2} (u, u)_1 \leq \frac{l}{q+2} \langle \Delta u - u, u \rangle_0 = \frac{2l}{q+2} \Phi(t), \quad (31)$$

$$\frac{q+1}{q+2} \left| \int_0^t (u(s), u'(t))_1 ds \right| \leq \frac{q+1}{q+2} l \int_0^t \langle (\Delta u - u)(s), u(s) \rangle_0^{1/2} \langle (\Delta u' - u')(t), u'(t) \rangle_0^{1/2} ds \leq$$

$$\leq \varepsilon \langle (\Delta u' - u')(t), u'(t) \rangle_0 + \left( \frac{q+1}{q+2} \right)^2 l^2 \frac{T}{4\varepsilon} \int_0^t \langle (\Delta u - u)(s), u(s) \rangle_0 ds \leq$$

$$\leq \varepsilon J(t) + \left( \frac{q+1}{q+2} \right)^2 l^2 \frac{T}{2\varepsilon} \int_0^t \Phi(s) ds, \quad (32)$$

$$\frac{q+1}{q+2} |\langle f, u' \rangle_0| \leq \frac{q+1}{q+2} \|f\|_0^* \|u'\|_0 \leq \frac{q+1}{q+2} \|f\|_0^* \frac{1}{m_0^{1/2}} \langle \Delta u' - u', u' \rangle_0^{1/2} \leq$$

$$\leq \varepsilon J(t) + \left( \frac{q+1}{q+2} \right)^2 \frac{\|f\|_0^{*2}}{4m_0\varepsilon}, \quad (33)$$

$$\left| \left\langle \frac{\partial |u|^{1+q_0}}{\partial x_1}, u' \right\rangle_0 \right| \leq \varepsilon \langle \Delta u' - u', u' \rangle_0 + \frac{b}{\varepsilon} \langle \Delta u - u, u \rangle_0^{1+q_0} \leq \varepsilon J(t) + \frac{d}{\varepsilon} \Phi^{1+q_0}(t), \quad (34)$$

где  $d = 2^{1+q_0}b$ . Итак, из равенства (30) с учётом (31)–(34) получим оценку

$$(1 - 3\varepsilon)J(t) \leq \frac{1}{q+2} \frac{d^2\Phi(t)}{dt^2} + \frac{2l}{q+2}\Phi(t) + \left(\frac{q+1}{q+2}\right)^2 l^2 \frac{T}{2\varepsilon} \int_0^t \Phi(s) ds + \left(\frac{q+1}{q+2}\right)^2 \frac{\|f\|_0^{*2}}{4m_0\varepsilon} + \frac{d}{\varepsilon}\Phi^{1+q_0}(t). \tag{35}$$

Пусть  $\varepsilon \in (0, 1/3)$ . Тогда из соотношений (26) и (35) получим обыкновенное дифференциальное неравенство второго порядка

$$\Phi\Phi'' - \frac{q+2}{\bar{p}}(1 - 3\varepsilon)(\Phi')^2 + 2l\Phi^2 + \frac{(q+1)^2}{q+2} l^2 \frac{T}{2\varepsilon} \int_0^t \Phi(s) ds \Phi(t) + \frac{(q+1)^2}{q+2} \frac{\|f\|_0^{*2}}{4m_0\varepsilon} \Phi + \frac{d(q+2)}{\varepsilon} \Phi^{2+q_0} \geq 0,$$

которое запишем, сделав замену  $3\varepsilon \mapsto \varepsilon$ , в следующем виде:

$$\Phi\Phi'' - \alpha(\Phi')^2 + \beta\Phi^2 + \gamma_1\Phi + \gamma_2T \int_0^t \Phi(s) ds \Phi(t) + \gamma_3\Phi^{1+\lambda} \geq 0, \tag{36}$$

где

$$\alpha = \frac{q+2}{\bar{p}}(1 - \varepsilon), \quad \beta = 2l, \quad \gamma_1 = \frac{(q+1)^2}{q+2} \frac{3\|f\|_0^{*2}}{4m_0\varepsilon},$$

$$\gamma_2 = l^2 \frac{(q+1)^2}{q+2} \frac{3}{2\varepsilon}, \quad \gamma_3 = \frac{3d(q+2)}{\varepsilon}, \quad \lambda = 1 + q_0. \tag{37}$$

Потребовав выполнения условия  $\alpha > 1$ , отсюда получим неравенства

$$0 < \varepsilon < \frac{q+2 - \bar{p}}{q+2}, \quad q+2 > \bar{p}.$$

Кроме того, имеем

$$2\alpha - 1 = \frac{\alpha_1 - \alpha_2\varepsilon}{\bar{p}}, \quad \alpha_1 = 2(q+2) - \bar{p}, \quad \alpha_2 = 2(q+2).$$

Рассмотрим вспомогательную функцию  $h(x) = x(\alpha_1 - \alpha_2x) \geq 0$  при  $x \in [0, \alpha_1/\alpha_2]$ . Максимум этой функции достигается в точке

$$x_0 = \frac{\alpha_1}{2\alpha_2} = \frac{2(q+2) - \bar{p}}{4(q+2)}.$$

Пусть выполнено неравенство  $q+2 > \bar{p}$ . Теперь рассмотрим следующие два случая:  $q+2 \leq 3\bar{p}/2$  и  $3\bar{p}/2 < q+2$ . В первом случае имеем

$$x_0 = \frac{\alpha_1}{2\alpha_2} = \frac{2(q+2) - \bar{p}}{4(q+2)} \geq \frac{q+2 - \bar{p}}{q+2},$$

а во втором

$$x_0 = \frac{\alpha_1}{2\alpha_2} = \frac{2(q+2) - \bar{p}}{4(q+2)} < \frac{q+2 - \bar{p}}{q+2}.$$

Выберем параметр  $\varepsilon > 0$ , входящий в коэффициенты (37), таким образом, чтобы коэффициент  $2\gamma_1/(2\alpha - 1)$  принял минимальное значение. Это требование необходимо нам для того, чтобы включить в эффект *blow-up* как можно больше элементов  $f \in H_0^{-1}(\Omega)$ . Заметим, что этот коэффициент имеет вид

$$\frac{2\gamma_1}{2\alpha - 1} = \frac{1}{\varepsilon(\alpha_1 - \alpha_2\varepsilon)} \frac{\bar{p}(q + 1)^2}{q + 2} \frac{3\|f\|_0^{*2}}{2m_0}$$

– функции от  $\varepsilon > 0$ , которая принимает минимальное значение в точке

$$\varepsilon_0 = \begin{cases} \frac{q + 2 - \bar{p}}{q + 2}, & \text{если } q + 2 \leq \frac{3}{2}\bar{p}, \\ \frac{2(q + 2) - \bar{p}}{4(q + 2)}, & \text{если } \frac{3}{2}\bar{p} < q + 2. \end{cases}$$

Однако в случае  $q + 2 \leq 3/2\bar{p}$  мы имеем  $\alpha = 1$ , что нам не подходит. Поэтому выберем параметр  $\varepsilon$  следующим образом:

$$\varepsilon = \varepsilon_0 = \begin{cases} \frac{q + 2 - \bar{p}}{q + 2} - \delta_0, & \text{если } q + 2 \leq \frac{3}{2}\bar{p}, \\ \frac{2(q + 2) - \bar{p}}{4(q + 2)}, & \text{если } \frac{3}{2}\bar{p} < q + 2, \end{cases}$$

для любого малого  $\delta_0 > 0$ . Подставим это значение  $\varepsilon = \varepsilon_0$  в коэффициенты (37).

Теперь проверим выполнимость условий следующей теоремы [1, теорема 1].

**Теорема 3.** Пусть функция  $\Phi(t) \in C^{(2)}[0, T_0)$  и удовлетворяет дифференциальному неравенству (36) и  $\Phi(0) > 0$ ,  $\Phi'(0) > 0$ ,  $\alpha > 1$ ,  $1 < \lambda < 2\alpha - 1$ , причём начальные условия  $\Phi(0)$  и  $\Phi'(0)$  таковы, что существует  $T_1$  – наименьший положительный корень уравнения

$$(\Phi'(0))^2 = \frac{1}{T_1^2(\alpha - 1)^2}(\Phi(0))^2 + \frac{\beta + \gamma_2 T_1^2}{\alpha - 1}(\Phi(0))^2 + \frac{2\gamma_1}{2\alpha - 1}\Phi(0) + \frac{2\gamma_3}{(\alpha - 1)\delta}(\Phi(0))^{1+\lambda}. \quad (38)$$

Тогда  $\Phi(t)$  удовлетворяет неравенству

$$\Phi(t) \geq [\Phi^{1-\alpha}(0) - A^{1/2}(T_1)t]^{1/(1-\alpha)} \quad (39)$$

для всех  $t \in [0, \min\{T_1, T_0\})$ , где

$$A(T_1) = (\alpha - 1)^2\Phi^{-2\alpha}(0) \left[ (\Phi'(0))^2 - \frac{\beta + \gamma_2 T_1^2}{\alpha - 1}(\Phi(0))^2 - \frac{2\gamma_1}{2\alpha - 1}\Phi(0) - \frac{2\gamma_3}{(\alpha - 1)\delta}(\Phi(0))^{1+\lambda} \right] > 0.$$

Пусть  $u_0 \in H_0^1(\Omega)$  и  $f \in H_0^{-1}(\Omega)$  – произвольные фиксированные функции, а функции  $u_1 \in H_0^1(\Omega)$  – единственное решение следующего уравнения в  $H_0^{-1}(\Omega)$ :

$$\Delta u_1 - u_1 - \sum_{j=1}^n (p_j - 1)|u_0|^{p_j-2}u_1 = -\frac{\partial|u_0|^{1+q_0}}{\partial x_1} + |u_0|^q u_0 + f \in H_0^{-1}(\Omega). \quad (40)$$

Решение  $u_1 \in H_0^1(\Omega)$  этого уравнения, действительно, существует в силу теоремы Браудера–Минти [2, с. 98]. В нашем случае функционал  $\Phi(t) \in C^{(2)}[0, T_0)$  имеет вид (24). Поэтому при  $t = 0$  имеем

$$\Phi(0) = \frac{1}{2}\langle \Delta u_0 - u_0, u_0 \rangle_0 + \sum_{j=1}^n \frac{p_j - 1}{p_j} \langle |u_0|^{p_j-2}u_0, u_0 \rangle_j,$$

а производная Фреше функционала  $\Phi(t)$  имеет следующий вид:

$$\Phi'(t) = \langle \Delta u' - u', u \rangle_0 + \sum_{j=1}^n \langle (p_j - 1) |u|^{p_j-2} u', u \rangle_j.$$

Тогда получим

$$\Phi'(0) = \langle \Delta u_1 - u_1, u_0 \rangle_0 + \sum_{j=1}^n \langle (p_j - 1) |u|^{p_j-2} u_1, u_0 \rangle_j.$$

Отсюда, с учётом равенства (40) и свойства (27), получим выражение

$$\Phi'(0) = (|u_0|^q u_0, u_0)_2 + \langle f, u_0 \rangle_0.$$

Запишем уравнение (38) в следующем эквивалентном виде:

$$K_1 T_1^4 + K_2 T_1^2 + K_3 = 0,$$

где

$$K_1 = \frac{\gamma_2}{\alpha - 1} (\Phi(0))^2, \quad K_3 = \frac{1}{(\alpha - 1)^2} (\Phi(0))^2, \\ K_2 = \frac{\beta}{\alpha - 1} (\Phi(0))^2 + \frac{2\gamma_1}{2\alpha - 1} \Phi(0) + \frac{2\gamma_3}{(\alpha - 1)\delta} (\Phi(0))^{1+\lambda} - (\Phi'(0))^2. \tag{41}$$

Введём функции

$$I_1(R) = (\Phi'(0)|_{Ru_0})^2 = ((|Ru_0|^q Ru_0, Ru_0)_2 + \langle f, Ru_0 \rangle_0)^2 = (R^{q+2} (|u_0|^q u_0, u_0)_2 + R \langle f, u_0 \rangle_0)^2, \\ I_2(R) = \Phi(0)|_{Ru_0} = R^2 \frac{1}{2} \langle \Delta u_0 - u_0, u_0 \rangle_0 + \sum_{j=1}^n R^{p_j} \frac{p_j - 1}{p_j} \langle |u_0|^{p_j-2} u_0, u_0 \rangle_j. \tag{42}$$

Подставим теперь в правые части равенств (41) вместо  $u_0$  элемент  $Ru_0$  при  $R \geq 0$ . Пусть, кроме того,  $x = T_1^2$ . Тогда биквадратное уравнение примет следующий вид:

$$K_1 x^2 + K_2 x + K_3 = 0. \tag{43}$$

Прежде всего заметим, что в силу условий

$$q + 2 > \bar{p} = \max_{j=1, n} p_j, \quad 2(q + 2) > \bar{p}(1 + \lambda) \quad (\text{т.е. } \lambda < -1 + 2(q + 2)/\bar{p})$$

и формул (42) коэффициент  $K_2$  окажется отрицательным при достаточно большом  $R > 0$  и при условии  $(u_0|^q u_0, u_0)_2 \neq 0$ .

Дискриминант  $\mathcal{D} = K_2^2 - 4K_1 K_3$  является положительным при достаточно большом  $R > 0$ . Итак, при достаточно большом  $R > 0$  существует положительный корень уравнения (43)

$$T_1^2 = x = \frac{-K_2 + \sqrt{K_2^2 - 4K_1 K_3}}{2K_1} > 0.$$

Таким образом, нами доказана следующая

**Теорема 4.** Пусть выполнены неравенства

$$q + 2 > \bar{p}, \quad 0 < q_0 < 2 \frac{q + 2 - \bar{p}}{\bar{p}};$$

$u_0 \in H_0^1(\Omega)$  и  $f \in H_0^{-1}(\Omega)$ ;  $u_1 \in H_0^1(\Omega)$  является решением уравнения (40), причём

$$(|u_0|^q u_0, u_0) \neq 0.$$

Тогда при достаточно большом  $R > 0$  для начальной функции  $Ru_0$  функционал  $\Phi(t)$ , определённый формулой (24), удовлетворяет неравенству (39).

Справедливо следующее утверждение [4, лемма 7.4].

**Лемма 2.** *Имеет место двустороннее неравенство*

$$M_1 \Phi^{1/2}(t) \leq \|A(u)\|_0^* \leq M_2 \Phi^{1/2} + \sum_{j=1}^n B_j \Phi^{(p_j-1)/p_j}(t),$$

где положительные постоянные  $M_1$ ,  $M_2$  и  $B_j$  не зависят от  $u(t)$ ;  $A(u) := \Delta u - u - \sum_{j=1}^n |u|^{p_j-2} u$ .

Из этой леммы вытекает

**Теорема 5.** *Пусть выполнены неравенства*

$$q + 2 > \bar{p}, \quad 0 < q_0 < 2 \frac{q + 2 - \bar{p}}{\bar{p}},$$

в качестве начальной функции  $u_0 \in H_0^1(\Omega)$  взята функция  $Ru_0$ , а  $u_1 \in H_0^1(\Omega)$  – решение уравнения (40) при  $f \in H_0^{-1}(\Omega)$ , в которое вместо  $u_0$  нужно подставить  $Ru_0$ . Тогда при достаточно большом  $R > 0$  время  $T_0 > 0$  существования классического решения задачи (6)–(8) конечно, удовлетворяет предельному свойству

$$\lim_{t \uparrow T_0} \Phi(t) = +\infty,$$

и справедлива оценка сверху  $T_0 \leq T_1$  на время разрушения решения, где число  $T_1$  является положительным решением биквадратного уравнения (38).

## СПИСОК ЛИТЕРАТУРЫ

1. Артемьева М.В., Корпусов М.О. Разрушение решений и локальная разрешимость абстрактной задачи Коши второго порядка с некоэрцитивным источником // Журн. вычислит. математики и мат. физики. 2023. Т. 63. № 4. С. 43–53.
2. Гаевский Х., Греггер К., Захариас К. Нелинейные операторные уравнения и операторные дифференциальные уравнения. М., 1978.
3. Al'shin A.B., Korpusov M.O., Sveshnikov A.G. Blow-up in nonlinear Sobolev type equations // De Gruyter Ser. Nonlin. Anal. Appl. 2011. V. 15. P. 648.
4. Корпусов М.О. Разрушение и глобальная разрешимость в классическом смысле задачи Коши для формально гиперболического уравнения с некоэрцитивным источником // Изв. РАН. Сер. мат. 2020. Т. 84. № 5. С. 119–150.

Московский государственный университет  
имени М.В. Ломоносова,  
Национальный исследовательский ядерный  
университет “МИФИ”, г. Москва

Поступила в редакцию 23.03.2023 г.  
После доработки 31.05.2023 г.  
Принята к публикации 14.06.2023 г.

## УРАВНЕНИЯ С ЧАСТНЫМИ ПРОИЗВОДНЫМИ

УДК 517.956.4

## СУЩЕСТВОВАНИЕ И ЕДИНСТВЕННОСТЬ КЛАССИЧЕСКОГО РЕШЕНИЯ ПЕРВОЙ КРАЕВОЙ ЗАДАЧИ ДЛЯ ПАРАБОЛИЧЕСКИХ СИСТЕМ НА ПЛОСКОСТИ

© 2023 г. А. Н. Коненков

Рассматривается первая краевая задача для равномерно-параболических систем второго порядка с одной пространственной переменной в ограниченных и полуограниченных областях с негладкими боковыми границами. Коэффициенты системы удовлетворяют условию Гёльдера и не зависят от временной переменной. Для непрерывных начальной и граничной функций устанавливается существование и единственность классического решения этой задачи.

DOI: 10.31857/S037406412307004X, EDN: GTUTXN

**Введение.** Для параболических систем однозначная разрешимость краевых задач в областях с гладкими боковыми границами в анизотропных пространствах Гёльдера установлена в работе [1]. В областях на плоскости с негладкими боковыми границами из класса Жевре  $H^{(1+\alpha)/2}$ ,  $0 < \alpha < 1$ , для параболических систем с гёльдеровыми коэффициентами существование решения в анизотропном пространстве Гёльдера  $H^{1+\alpha, (1+\alpha)/2}(\bar{\Omega})$  получено в статье [2]. Для систем с Дини-непрерывными коэффициентами в областях с боковой границей из класса Дини-Гёльдера  $H^{1/2+\omega}$  первая и вторая краевые задачи рассматривались в [3, 4], здесь  $\omega$  – модуль непрерывности, удовлетворяющий условию Дини. При соответствующей гладкости граничной функции доказано существование классических решений этих задач в пространстве  $C_{x,t}^{1,0}(\bar{\Omega})$ .

Единственность решения первой краевой задачи для параболического уравнения второго порядка следует из принципа максимума, а единственность решения второй краевой задачи выводится из теоремы о знаке кривой производной, в доказательстве которой также используется принцип максимума. Для параболических систем принцип максимума, вообще говоря, не имеет места [5]. В пространстве  $C_{x,t}^{1,0}(\bar{\Omega})$  единственность решений первой и второй краевых задач для параболических систем на плоскости в областях с негладкими боковыми границами из классов Жевре и Дини-Гёльдера установлена в работах [6–9].

Вопрос существования классического решения первой краевой задачи с непрерывной граничной функцией для параболического уравнения второго порядка с гёльдеровыми коэффициентами рассмотрен в монографии [10, гл. 3, § 4]. Отметим, что доказательство непрерывности этого решения вплоть до боковой границы области проводится методом барьеров, который основан на принципе максимума. В настоящей работе рассматривается первая краевая задача для параболических систем на плоскости в областях с боковыми границами, удовлетворяющими условию Жевре. Коэффициенты системы не зависят от временной переменной  $t$  и удовлетворяют условию Гёльдера. От граничной функции требуется только непрерывность. Устанавливается существование классического решения этой задачи и его единственность в классе  $C(\bar{\Omega}) \cap C_{x,t}^{2,1}(\Omega)$ , где  $C(\bar{\Omega})$  – пространство непрерывных и ограниченных функций в  $\bar{\Omega}$ .

Статья состоит из трёх пунктов. В п. 1 вводится потенциал, обладающий многими свойствами потенциала двойного слоя, но, в отличие от последнего, не требующий существования производной у старших коэффициентов параболической системы. Далее этот потенциал используется для доказательства существования решения и получения оценок для него. В пп. 2, 3 рассматривается первая краевая задача для ограниченной и полуограниченной областей соответственно.

Результаты настоящей работы анонсированы в [11].

**1. Модифицированный потенциал двойного слоя.** В полосе  $D = \{(x, t) : x \in \mathbb{R}, 0 < t < T < \infty\}$  рассматривается параболическая система

$$Lu = \partial_t u - A(x)\partial_x^2 u - B(x)\partial_x u - C(x)u, \tag{1}$$

где  $u = (u_1, \dots, u_m)^T$ ,  $A(x)$ ,  $B(x)$ ,  $C(x)$  –  $m \times m$ -матрицы с элементами  $a_{ij}(x)$ ,  $b_{ij}(x)$ ,  $c_{ij}(x)$  соответственно. Для оператора  $L$  предполагаются выполненными условия:

а) равномерной параболичности, т.е. собственные значения  $\lambda_k(x)$  матрицы  $A(x)$  удовлетворяют неравенству  $\operatorname{Re} \lambda_k(x) \geq \mu > 0$  для любых  $x \in \mathbb{R}$ ;

б) коэффициенты  $a_{ij}$ ,  $b_{ij}$ ,  $c_{ij}$  принадлежат пространству Гёльдера  $H^\alpha(\mathbb{R})$ ,  $0 < \alpha < 1$ , с нормой

$$\|f\|_{H^\alpha(\mathbb{R})} = \sup_{x \in \mathbb{R}} |f(x)| + \sup_{\substack{x \in \mathbb{R} \\ \Delta x \neq 0}} \frac{|f(x + \Delta x) - f(x)|}{|\Delta x|^\alpha}.$$

Установим предварительно некоторые тождества для фундаментальных матриц решений параболических систем на плоскости с коэффициентами, не зависящими от переменной  $t$ .

Рассмотрим сначала случай, когда младшие члены отсутствуют, положив

$$L_0 = \partial_t - A(x)\partial_x^2.$$

Введём ещё оператор

$$\tilde{L}_0 = \partial_t - A^T(x)\partial_x^2$$

и обозначим фундаментальные матрицы решений операторов  $L_0$  и  $\tilde{L}_0$  через  $\Gamma_0(x, \xi, t - \tau)$  и  $\tilde{\Gamma}_0(x, \xi, t - \tau)$  соответственно.

**Теорема 1.** Пусть для оператора  $L_0$  выполнены условия а), б). Тогда имеет место равенство

$$\Gamma_0(x, \xi, t)A(\xi) = A(x)\tilde{\Gamma}_0^T(\xi, x, t) \tag{2}$$

для всех  $x, \xi \in \mathbb{R}$ ,  $t > 0$ .

**Доказательство.** Зафиксируем  $x, \xi \in \mathbb{R}$ ,  $t > 0$  и рассмотрим интеграл

$$I(\tau) = \int_{-\infty}^{\infty} \tilde{\Gamma}_0^T(y, x, t - \tau)A^{-1}(y)\Gamma_0(y, \xi, \tau) dy.$$

Покажем, что функция  $I(\tau)$  постоянна на интервале  $(0, t)$ . Действительно,

$$\begin{aligned} I'(\tau) &= - \int_{-\infty}^{\infty} \partial_t \tilde{\Gamma}_0^T(y, x, t - \tau)A^{-1}(y)\Gamma_0(y, \xi, \tau) dy + \int_{-\infty}^{\infty} \tilde{\Gamma}_0^T(y, x, t - \tau)A^{-1}(y)\partial_\tau \Gamma_0(y, \xi, \tau) dy = \\ &= - \int_{-\infty}^{\infty} \partial_y^2 \tilde{\Gamma}_0^T(y, x, t - \tau)A(y)A^{-1}(y)\Gamma_0(y, \xi, \tau) dy + \int_{-\infty}^{\infty} \tilde{\Gamma}_0^T(y, x, t - \tau)A^{-1}(y)A(y)\partial_y^2 \Gamma_0(y, \xi, \tau) dy = \\ &= - \int_{-\infty}^{\infty} \partial_y^2 \tilde{\Gamma}_0^T(y, x, t - \tau)\Gamma_0(y, \xi, \tau) dy + \int_{-\infty}^{\infty} \tilde{\Gamma}_0^T(y, x, t - \tau)\partial_y^2 \Gamma_0(y, \xi, \tau) dy = \\ &= \int_{-\infty}^{\infty} \partial_y \tilde{\Gamma}_0^T(y, x, t - \tau)\partial_y \Gamma_0(y, \xi, \tau) dy - \int_{-\infty}^{\infty} \partial_y \tilde{\Gamma}_0^T(y, x, t - \tau)\partial_y \Gamma_0(y, \xi, \tau) dy = 0. \end{aligned}$$

С другой стороны,

$$\lim_{\tau \rightarrow 0+} I(\tau) = \tilde{\Gamma}_0^T(\xi, x, t)A^{-1}(\xi), \quad \lim_{\tau \rightarrow t-0} I(\tau) = A^{-1}(x)\Gamma_0(x, \xi, t).$$

Теорема доказана.

Рассмотрим теперь общий случай, когда оператор (1) имеет младшие коэффициенты. Обозначим через  $l = B(x)\partial_x + C(x)$  его младшую часть, так что  $L = L_0 - l$ , а через  $\Gamma(x, \xi, t - \tau)$  и  $Z(x - \xi, t - \tau; A)$  фундаментальные матрицы решений для  $L$  и оператора  $\partial_t - A\partial_x^2$  с постоянными коэффициентами соответственно.

**Лемма.** Пусть для оператора  $L$  выполнены условия а), б). Тогда

$$\Gamma(x, \xi, t) = \Gamma_0(x, \xi, t) + \int_0^t \int_{-\infty}^{\infty} \Gamma(x, y, t - \tau) l_y \Gamma_0(y, \xi, \tau) dy d\tau \tag{3}$$

для всех  $x, \xi \in \mathbb{R}, t > 0$ .

**Доказательство.** Зафиксируем произвольную точку  $\xi \in \mathbb{R}$  и  $\varepsilon \in (0, T)$ . В полосе  $D \cap \{t > \varepsilon\}$  фундаментальная матрица решений  $\Gamma_0(x, \xi, t)$  как функция переменных  $(x, t)$  является решением следующей задачи Коши:

$$Lu = -l_x \Gamma_0(x, \xi, t), \quad u|_{t=\varepsilon} = \Gamma_0(x, \xi, \varepsilon).$$

Так как решение этой задачи в классе ограниченных функций единственно, то  $\Gamma_0(x, t, \xi, \tau)$  можно представить в виде суммы объёмного потенциала и потенциала типа Пуассона [12, гл. 2, § 1]:

$$\Gamma_0(x, \xi, t) = - \int_{\varepsilon}^t \int_{-\infty}^{\infty} \Gamma(x, y, t - \tau) l_y \Gamma_0(y, \xi, \tau) dy d\tau + \int_{-\infty}^{\infty} \Gamma(x, y, t - \varepsilon) \Gamma_0(y, \xi, \varepsilon) dy.$$

Переходя здесь к пределу при  $\varepsilon \rightarrow 0+$ , получаем равенства

$$\begin{aligned} \Gamma_0(x, \xi, t) &= - \lim_{\varepsilon \rightarrow 0+} \int_{\varepsilon}^t \int_{-\infty}^{\infty} \Gamma(x, y, t - \tau) l_y \Gamma_0(y, \xi, \tau) dy d\tau + \lim_{\varepsilon \rightarrow 0+} \int_{-\infty}^{\infty} \Gamma(x, y, t - \varepsilon) \Gamma_0(y, \xi, \varepsilon) dy = \\ &= - \int_0^t \int_{-\infty}^{\infty} \Gamma(x, y, t - \tau) l_y \Gamma_0(y, \xi, \tau) dy d\tau + \Gamma(x, \xi, t). \end{aligned}$$

Лемма доказана.

Как известно, для параболических систем с одной пространственной переменной, если старшие коэффициенты  $A_{ij}^{(2)}$  достаточно гладкие, первую краевую задачу можно решать с помощью потенциала двойного слоя, ядром которого является функция  $\partial_{\xi} \Gamma(x, \xi, t)$ . Однако если у старших коэффициентов нет производной по  $x$ , то фундаментальная матрица решений не имеет, вообще говоря, производной по переменной  $\xi$ .

В качестве ядра потенциала, который определён, в отличие от потенциала двойного слоя, даже когда старшие коэффициенты  $a_{ij}$  параболического оператора удовлетворяют только условию Гёльдера, возьмём функцию  $K(x, \xi, t) = \partial_{\xi}(\Gamma(x, \xi, t)A(\xi))$ .

**Теорема 2.** Если для оператора  $L$  выполнены условия а), б), то функция  $\Gamma(x, \xi, t)A(\xi)$  непрерывно дифференцируема по переменной  $\xi$  в полупространстве  $t > 0$ . Функция  $K$  удовлетворяет уравнению  $L_{x,t}K(x, \xi, t) = 0$  и справедливы оценки

$$|\partial_x^m K(x, \xi, t)| < C_m t^{-(m+2)/2} e^{-c_m(x-\xi)^2/t}, \quad m \leq 2, \tag{4}$$

при  $t > 0, x, \xi \in \mathbb{R}$ , а также представление

$$K(x, \xi, t) = -\partial_x Z(x - \xi, t; A(\xi))A(\xi) + \tilde{\Phi}(x, \xi, t), \tag{5}$$

где

$$|\tilde{\Phi}(x, \xi, t)| < Ct^{(\alpha-2)/2} e^{-c(x-\xi)^2/t}. \tag{6}$$

**Доказательство.** Нам понадобятся оценки [12, гл. 1, § 3] фундаментальной матрицы решений

$$|\partial_x^m \Gamma(x, \xi, t)| < C_m t^{-(m+1)/2} e^{-c_m(x-\xi)^2/t}, \quad m \leq 2. \tag{7}$$

Установим (4) сначала для функции

$$K_0(x, \xi, t) = \partial_\xi(\Gamma_0(x, \xi, t)A(\xi)) = A(x)\partial_\xi \tilde{\Gamma}_0(\xi, x, t),$$

которая удовлетворяет тем же оценкам, что и первая производная по  $x$  фундаментальной матрицы решений, а именно,

$$|K_0(x, \xi, t)| < Ct^{-1}e^{-c(x-\xi)^2/t}.$$

Используя свойство свёртки [12, гл. 1, § 3]

$$\Gamma_0(x, \xi, t) = \int_{-\infty}^{\infty} \Gamma_0(x, y, t-\lambda)\Gamma_0(y, \xi, \lambda) dy, \quad 0 < \lambda < t,$$

получаем  $L_0 K_0(x, \xi, t) = 0$  и при  $\lambda = t/2$

$$\partial_x^m K_0(x, \xi, t) = \int_{-\infty}^{\infty} \partial_x^m \Gamma_0(x, y, t/2)K_0(y, \xi, t/2) dy,$$

откуда

$$|\partial_x^m K_0(x, \xi, t)| \leq Ct^{-(3+m)/2} \int_{-\infty}^{\infty} \exp\left(-c\frac{(x-y)^2}{t} - c\frac{(y-\xi)^2}{t}\right) dy = C_1 t^{-1-m/2} e^{-c(x-\xi)^2/t}, \quad m \leq 2.$$

Умножая равенство (5) справа на  $A(\xi)$ , имеем

$$K(x, \xi, t) = K_0(x, \xi, t) + \int_0^t \int_{-\infty}^{\infty} \Gamma(x, y, t-\tau)l_y K_0(y, \xi, \tau) d\tau \equiv K_0(x, \xi, t) + \Psi(x, \xi, t).$$

Из теоремы 1 и оценки (7) следуют неравенства

$$|\partial_\xi^m \partial_x^k (\Gamma_0(x, \xi, t)A(\xi))| \leq Ct^{-(1+k+m)/2} C e^{-c(x-\xi)^2/t}, \quad k, m = 0, 1,$$

$$|\Delta_\xi \partial_x^k (\Gamma_0(x, \xi, t)A(\xi))| \leq C|\Delta\xi|^\alpha Ct^{-(\alpha+1+k)/2} C e^{-c(x-\xi^*)^2/t}, \quad k = 0, 1,$$

где  $\xi^*$  – ближайшая из промежутка  $(\xi, \xi + \Delta\xi)$  точка к  $x$ . Применив последнее неравенство для оценки интеграла в представлении (3), получим

$$|\Gamma(x, y, t) - \Gamma(x, \xi, t)| \leq C(t-\tau)^{-(\alpha+1)/2} |y-\xi|^\alpha (e^{-c(x-y)^2/(t-\tau)} + e^{-c(x-\xi)^2/(t-\tau)}).$$

В силу гёльдеровости коэффициентов системы такая же оценка справедлива и для разности  $\Gamma(x, y, t)B(y) - \Gamma(x, \xi, t)B(\xi)$ .

Имеем соотношения

$$|\Psi(x, \xi, t)| = \left| \int_0^t \int_{-\infty}^{\infty} [(\Gamma(x, y, t-\tau)B(y) - \Gamma(x, \xi, t-\tau)B(\xi)) \times \right.$$

$$\begin{aligned} & \times \partial_y K_0(y, \xi, \tau) + \Gamma(x, y, t - \tau)C(y)K_0(y, \xi, \tau)] d\tau \Big| \leq \\ & \leq C \int_0^t \int_{-\infty}^{\infty} [(t - \tau)^{-(\alpha+1)/2} |y - \xi|^\alpha \tau^{-3/2} + (t - \tau)^{-1} \tau^{-1/2}] \times \\ & \times (e^{-c(x-y)^2/(t-\tau)} + e^{-c(x-\xi)^2/(t-\tau)}) e^{-c(y-\xi)^2/(t-\tau)} dy d\tau \leq Ct^{-1/2} e^{-c(x-\xi)^2/t}. \end{aligned}$$

Кроме того, матрицу  $\tilde{\Gamma}_0$  можно представить в виде [12, гл. 1, § 3]

$$\tilde{\Gamma}_0(x, \xi, t) = Z(x - \xi, t; A^T(\xi)) + \Phi_0(x, \xi, t),$$

где  $\Phi_0$  имеет более слабую особенность по сравнению с первым слагаемым:

$$|\partial_x^m \Phi_0(x, \xi, t)| < Ct^{(\alpha-1-m)/2} e^{-c(x-\xi)^2/t}, \quad m \leq 2. \tag{8}$$

Используя равенство (2) для фундаментальной матрицы решений с постоянными коэффициентами, получаем

$$\begin{aligned} K(x, \xi, t) &= K_0(x, \xi, t) + \Psi(x, \xi, t) = A^T(x) \partial_\xi \tilde{\Gamma}_0^T(\xi, x, t) + \Psi(x, \xi, t) = \\ &= A^T(x) \partial_\xi Z^T(\xi - x, t; A^T(x)) + A^T(x) \partial_\xi \Phi_0^T(\xi, x, t) + \Psi(x, \xi, t) = \\ &= -\partial_x Z(x - \xi, t; A(x))A(x) + A^T(x) \partial_\xi \Phi_0^T(\xi, x, t) + \Psi(x, \xi, t) = \\ &= -\partial_x Z(x - \xi, t; A(\xi))A(\xi) - [\partial_x Z(x - \xi, t; A(x))A(x) - \partial_x Z(x - \xi, t; A(\xi))A(\xi)] + \\ &+ A^T(x) \partial_\xi \Phi_0^T(\xi, x, t) + \Psi(x, \xi, t) = -\partial_x Z(x - \xi, t; A(\xi))A(\xi) + \tilde{\Phi}(x, \xi, t), \end{aligned}$$

причём из (7) и (8) следует, что

$$|\tilde{\Phi}(x, \xi, t)| < Ct^{(\alpha-2)/2} e^{-c(x-\xi)^2/t}.$$

Таким образом, получено представление (5), (6). Теорема доказана.

В области

$$\Omega = \{(x, t) \in D : x > g(t), \quad 0 < t < T\}$$

с боковой границей  $\Sigma = \{(x, t) \in D : x = g(t), \quad 0 < t < T\}$ , где функция  $g$  удовлетворяет условию  $g \in C^{(1+\alpha)/2}([0, T])$ ,  $0 < \alpha < 1$ , рассмотрим потенциал

$$W\varphi(x, t) = \int_0^t K(x, g(\tau), t - \tau)\varphi(\tau) d\tau.$$

Для множества  $E \subset \bar{D}$  обозначим через  $C(E)$  пространство непрерывных и ограниченных вектор-функций  $f : E \rightarrow \mathbb{R}^m$  с нормой

$$|f|_{0,E} = \sup_{(x,t) \in E} |f(x, t)|$$

и положим  $\underset{\circ}{C}(E) = \{f \in C(E) : f|_{t=0} = 0\}$ .

**Теорема 3.** Функция  $W : \varphi \rightarrow W\varphi$  отображает пространство  $\underset{\circ}{C}(\Sigma)$  в  $\underset{\circ}{C}(\bar{\Omega})$ , причём

$$|W\varphi|_{0,\Omega} \leq C|\varphi|_{0,\Sigma}. \tag{9}$$

Для плотности  $\varphi \in C(\Sigma)$  справедлива формула скачка

$$W^\pm \varphi(g(t), t) = \pm \frac{\varphi(t)}{2} + W^0 \varphi(g(t), t), \quad 0 < t \leq T, \tag{10}$$

где  $W^\pm \varphi$  – предельные значения при приближении к точке  $(g(t), t)$  справа и слева, а  $W^0 \varphi$  – прямое значение потенциала  $W\varphi$  на кривой  $x = g(t)$ .

**Доказательство.** По теореме 2 потенциал может быть представлен в виде

$$W\varphi(x, t) = \int_0^t \partial_x Z(x - g(\tau), t - \tau; A(\xi)) A(\xi) \varphi(\tau) d\tau + \int_0^t \tilde{\Phi}(x, \xi, t - \tau) \varphi(\tau) d\tau.$$

Формула скачка для первого слагаемого, также как и оценка (9), получена в [13]. Функция  $\tilde{\Phi}$  имеет слабую особенность, а потому второе слагаемое непрерывно в замыкании области и

$$\left| \int_0^t \tilde{\Phi}(x, \xi, t - \tau) \varphi(\tau) d\tau \right| \leq C |\varphi|_{0, \Sigma} \int_0^t (t - \tau)^{\alpha/2 - 1} d\tau \leq C |\varphi|_{0, \Sigma}.$$

Теорема доказана.

**2. Первая краевая задача в ограниченной области.** В полосе  $D$  рассматриваем область

$$\Omega = \{(x, t) \in D : g_1(t) < x < g_2(t), \quad 0 < t < T\}$$

с основанием  $\Omega_0 = \bar{\Omega} \cap \{t = 0\}$  и боковыми границами  $\Sigma_i = \{(x, t) \in D : x = g_i(t), \quad 0 < t < T\}$ ,  $i = 1, 2$ , где функции  $g_i$  удовлетворяют условиям

$$\begin{aligned} |g_i(t + \Delta t) - g_i(t)| &\leq C |\Delta t|^{(1+\alpha)/2}, \quad t, t + \Delta t \in [0, T], \\ g_1(t) &< g_2(t), \quad t \in [0, T], \quad \alpha \in (0, 1). \end{aligned} \tag{11}$$

В области  $\Omega$  рассмотрим первую краевую задачу

$$Lu = 0 \text{ в } \Omega, \quad u|_{\Sigma_1} = \psi_1, \quad u|_{\Sigma_2} = \psi_2, \quad u|_{t=0} = h. \tag{12}$$

**Теорема 4.** Пусть для оператора  $L$  выполнены условия а), б), а для боковых границ области  $\Omega$  – условия (11). Если  $\psi_1, \psi_2 \in C(\Sigma_i)$ ,  $h \in C(\Omega_0)$  и справедливы условия согласования  $\psi_1(0) = h(g_1(0))$ ,  $\psi_2(0) = h(g_2(0))$ , то существует и единственно решение  $u \in C_{x,t}^{2,1}(\Omega) \cap C(\bar{\Omega})$  первой краевой задачи (12), причём имеет место оценка

$$|u|_{0, \Omega} \leq C (|\psi_1|_{0, \Sigma_1} + |\psi_2|_{0, \Sigma_2} + |h|_{0, \Omega_0}).$$

При  $h \equiv 0$  существуют функции  $\varphi_i \in C(\Sigma_i)$ ,  $i = 1, 2$ , такие, что решение представляется в виде суммы потенциалов:  $u = W_1 \varphi_1 + W_2 \varphi_2$  в  $\Omega$ , где

$$W_i \varphi_i(x, t) = \int_0^t K(x, g_i(\tau), t - \tau) \varphi_i(\tau) d\tau,$$

причём  $|\varphi_i|_{0, [0, T]} \leq C (|\psi_1|_{0, \Sigma_1} + |\psi_2|_{0, \Sigma_2})$ ,  $i = 1, 2$ .

**Доказательство.** Продолжим функцию  $h$  до функции  $\tilde{h} \in C(\mathbb{R})$  с сохранением нормы:  $|\tilde{h}|_{0, \mathbb{R}} = |h|_{0, \Omega_0}$ . Вычитая из решения задачи (12) потенциал типа Пуассона

$$\Pi \tilde{h}(x, t) = \int_{-\infty}^{\infty} \Gamma(x, y, t) \tilde{h}(y) dy,$$

получаем задачу с нулевой начальной функцией:

$$Lv = 0 \text{ в } \Omega, \quad v|_{\Sigma_1} = \tilde{\psi}_1, \quad v|_{\Sigma_2} = \tilde{\psi}_2, \quad v|_{t=0} = 0, \tag{13}$$

где  $\tilde{\psi}_i = \psi_i - \tilde{h}$ , причём  $\tilde{\psi}_1(0) = \tilde{\psi}_2(0) = 0$  и в силу (7)

$$|\Pi\tilde{h}|_{0,D} \leq C|\tilde{h}|_{0,\mathbb{R}} = C|h|_{0,\Omega_0}.$$

Решение задачи (13) будем искать в виде суммы двух потенциалов:  $u = W_1\varphi_1 + W_2\varphi_2$ . Воспользовавшись формулой скачка (10), получим систему уравнений типа Вольтерры

$$\left(\frac{\varphi_1}{2} + W_1^0\varphi_1 + W_2^0\varphi_2\right)\Big|_{\Sigma_1} = \tilde{\psi}_1, \quad \left(-\frac{\varphi_2}{2} + W_1^0\varphi_1 + W_2^0\varphi_2\right)\Big|_{\Sigma_2} = \tilde{\psi}_2 \tag{14}$$

со слабо сингулярными ядрами.

Решив эту систему методом последовательных приближений, получим её единственное решение  $\varphi_1, \varphi_2 \in C([0, T])$ , причём  $\varphi_1(0) = \varphi_2(0) = 0$  и

$$|\varphi_i|_{0,[0,T]} \leq C(|\tilde{\psi}_1|_{0,[0,T]} + |\tilde{\psi}_2|_{0,[0,T]}) \leq C(|\psi_1|_{0,[0,T]} + |\psi_2|_{0,[0,T]} + |h|_{0,\Omega_0}). \tag{15}$$

Отметим характер зависимости константы в этом неравенстве от области. Пусть квадрат параболического расстояния между кривыми отделён от нуля некоторой константой:

$$\min_{0 \leq t, \tau \leq T} [(g_2(t) - g_1(\tau))^2 + |t - \tau|] \geq \delta > 0.$$

Из оценок для слабо сингулярных ядер  $K_{ij}^0$  системы (14)

$$\begin{aligned} |K_{ii}^0(t, \tau)| &= |K_{ii}^0(g_i(t), g_j(\tau), t - \tau)| \leq \\ &\leq C(t - \tau)^{\alpha/2-1} \exp\left(-c\frac{(g_i(t) - g_j(\tau))^2}{t - \tau}\right) \leq C(t - \tau)^{\alpha/2-1}, \quad i = 1, 2, \\ |K_{ij}^0(t, \tau)| &= |K(g_i(t), g_j(\tau), t - \tau)| \leq C(t - \tau)^{-1} \exp\left(-c\frac{(g_i(t) - g_j(\tau))^2}{t - \tau}\right) \leq \\ &\leq C_1(t - \tau)^{-1} \exp\left(-c\frac{(g_i(t) - g_j(\tau))^2 + |t - \tau|}{t - \tau}\right) \leq C_1(t - \tau)^{-1} \exp\left(-c\frac{\delta}{t - \tau}\right), \quad i \neq j, \end{aligned}$$

следует, что постоянная в (15), помимо  $T$ , константы параболичности  $\mu$ , норм коэффициентов системы в пространстве  $H^\alpha(\mathbb{R})$  и константы  $C$  из (11), зависит от  $\delta$ .

Установим единственность классического решения. Пусть  $u \in C_{x,t}^{2,1}(\Omega) \cap C(\bar{\Omega})$  – решение задачи (12) с нулевыми начальной и граничными функциями. Для малого числа  $\varepsilon > 0$  рассмотрим область

$$\Omega^\varepsilon = \{(x, t) \in D : g_1(t) + \varepsilon < x < g_2(t) - \varepsilon, \quad \varepsilon < t < T - \varepsilon\}$$

с боковыми границами  $\Sigma_i^\varepsilon = \{(x, t) \in D : x = g_i(t) - (-1)^i\varepsilon, \quad \varepsilon < t < T - \varepsilon\}$ ,  $i = 1, 2$ , и основанием  $Q^\varepsilon = \bar{\Omega}^\varepsilon \cap \{t = \varepsilon\}$ .

Так как  $u \in C_{x,t}^{2,1}(\Omega) \cap C(\bar{\Omega})$ , то  $u \in C_{x,t}^{2,1}(\bar{\Omega}^\varepsilon)$ . Решение задачи

$$Lv = 0 \text{ в } \Omega^\varepsilon, \quad v|_{\Sigma_1} = u|_{\Sigma_1}, \quad v|_{\Sigma_2} = u|_{\Sigma_2}, \quad v|_{t=\varepsilon} = u|_{t=\varepsilon},$$

в этом классе единственно [7], откуда следует, что  $v \equiv u$  в  $\Omega^\varepsilon$ . По доказанному это решение удовлетворяет оценке

$$|u|_{0,\Omega^\varepsilon} \leq C(|u|_{0,\Sigma_1^\varepsilon} + |u|_{0,\Sigma_2^\varepsilon} + |u|_{0,Q^\varepsilon}),$$

причём константа  $C$  может быть взята одной и той же для всех достаточно малых  $\varepsilon > 0$ . Но по условию функция  $u$  непрерывна в  $\bar{\Omega}$  и равна нулю на параболической границе области, поэтому при  $\varepsilon \rightarrow 0+$  правая часть неравенства стремится к нулю и, следовательно,  $u \equiv 0$  в  $\bar{\Omega}$ . Теорема доказана.

**3. Первая краевая задача в полуограниченной области.** Рассматриваем область

$$\Omega = \{(x, t) \in D : x > g(t), \quad 0 < t < T\}$$

с боковой границей

$$\Sigma = \{(x, t) \in D : x = g(t), \quad 0 < t < T\}$$

и основанием  $\Omega_0 = \bar{\Omega} \cap \{t = 0\}$ , где функция  $g$  удовлетворяет условию Жевре

$$|g(t + \Delta t) - g(t)| \leq C|\Delta t|^{(1+\alpha)/2}, \quad t, t + \Delta t \in [0, T], \quad 0 < \alpha < 1. \quad (16)$$

В области  $\Omega$  рассмотрим первую краевую задачу

$$Lu = 0 \text{ в } \Omega, \quad u|_{\Sigma} = \psi, \quad u|_{t=0} = h. \quad (17)$$

**Теорема 5.** Пусть для оператора  $L$  выполнены условия а), б), а для боковой границы области  $\Omega$  – условие (16). Если  $\psi \in C(\Sigma)$ ,  $h \in C(\Omega_0)$  и выполнено условие согласования  $\psi(0) = h(g(0))$ , то существует и единственно классическое решение  $u \in C_{x,t}^{2,1}(\Omega) \cap C(\bar{\Omega})$  задачи (17), которое удовлетворяет оценке

$$|u|_{0,\Omega} \leq C(|\psi|_{0,\Sigma} + |h|_{0,\Omega_0}).$$

При  $h \equiv 0$  существует функция  $\varphi \in C(\Sigma)$  такая, что решение представляется в виде потенциала

$$u(x, t) = \int_0^t K(x, g(\tau), t - \tau) \varphi(\tau) d\tau, \quad (18)$$

причём  $|\varphi|_{0,[0,T]} \leq C|\psi|_{0,\Sigma}$ .

**Доказательство.** Существование классического решения устанавливается так же, как и в случае ограниченной области. А именно, с помощью потенциала типа Пуассона задача (17) сводится к задаче с нулевой начальной функцией и граничной функцией  $\tilde{\psi}$ . Затем решение ищется в виде (18). Плотность  $\varphi$  находится из интегрального уравнения

$$\left( \frac{\varphi}{2} + W^0 \varphi \right) \Big|_{\Sigma} = \tilde{\psi},$$

откуда получается оценка на  $\varphi$ , а затем, с помощью теоремы 3, на решение  $u$ .

Докажем единственность. Без ограничения общности считаем, что боковая граница  $\Sigma$  лежит в левой полуплоскости, т.е.  $\max_{t \in [0,T]} g(t) \leq 0$ . Пусть  $u \in C_{x,t}^{2,1}(\Omega) \cap C(\bar{\Omega})$  – ограниченное

решение задачи (17) с нулевыми начальной и граничной функциями. Для  $R \geq 1$  в области  $\Omega^R = \Omega \cap \{x < R\}$  с боковыми границами  $\Sigma$  и  $\Sigma^R = \{x = R, \quad 0 < t < T\}$  рассмотрим первую краевую задачу

$$Lv = 0 \text{ в } \Omega^R, \quad v|_{\Sigma} = 0, \quad v|_{\Sigma^R} = u|_{\Sigma^R}, \quad v|_{t=0} = 0.$$

По теореме 4 существует решение этой задачи  $v \in C_{x,t}^{2,1}(\Omega^R) \cap C(\bar{\Omega}^R)$ , представимое в виде суммы двух потенциалов:

$$v(x, t) = \int_0^t K(x, g(\tau), t - \tau) \varphi_1(\tau) d\tau + \int_0^t K(x, R, t - \tau) \varphi_2(\tau) d\tau = W_1 \varphi_1(x, t) + W_2 \varphi_2(x, t).$$

В силу единственности решения этой задачи  $v \equiv u$  в  $\bar{\Omega}^R$ . Пусть

$$M = \sup_{\Omega} |u|.$$

Тогда по теореме 4  $|\varphi_2(t)| \leq CM$ , где  $C$  не зависит от  $R$  и

$$\begin{aligned} |W_2\varphi_2(x, t)| &\leq CM \int_0^t (t - \tau)^{-1} \exp\left(-\frac{c(x - R)^2}{t - \tau}\right) d\tau \leq \\ &\leq CM \exp\left(-\frac{c(x - R)^2}{2(t - \tau)}\right) \int_0^t (t - \tau)^{-1} \exp\left(-\frac{c(x - R)^2}{2(t - \tau)}\right) d\tau = \\ &= CM \exp\left(-\frac{c(x - R)^2}{2(t - \tau)}\right) \int_0^{t/(x-R)^2} \tau^{-1} \exp\left(-\frac{c}{2\tau}\right) d\tau \leq \\ &\leq CM \exp\left(-\frac{c(x - R)^2}{2(t - \tau)}\right) \int_0^{T/(x-R)^2} \tau^{-1} \exp\left(-\frac{c}{2\tau}\right) d\tau \leq \\ &\leq CM \exp\left(-\frac{c(x - R)^2}{2T}\right) \left(\left|\ln \frac{T}{(x - R)^2}\right| + 1\right). \end{aligned} \quad (19)$$

Краевое условие на левой границе даёт интегральное уравнение

$$\left(-\frac{\varphi_1}{2} + W_1^0\varphi_1\right)\Big|_{\Sigma_1} = -W_2\varphi_2\Big|_{\Sigma_1} = p,$$

причём, в силу предыдущего неравенства и теоремы 3,

$$|p|_{0,[0,T]} \leq C_1 M e^{-cR^2/(2T)} \left(\left|\ln \frac{T}{R^2}\right| + 1\right) \leq C_2 M e^{-cR^2/(4T)}.$$

Отсюда  $|\varphi_1|_{0,[0,T]} \leq C_1 |p|_{0,[0,T]} \leq C_2 M e^{-cR^2/(4T)}$  и  $|W_1\varphi_1(x, t)| \leq C_3 M e^{-cR^2/(4T)}$  в  $\Omega^R$ , где константа  $C_3$  не зависит от  $R$  и точки  $(x, t) \in \Omega^R$ .

Пусть  $(x, t)$  – произвольная точка области  $\Omega$ . Из последнего неравенства и (19) получаем

$$u(x, t) = \lim_{R \rightarrow +\infty} (W_1\varphi_1(x, t) + W_2\varphi_2(x, t)) = 0.$$

Теорема доказана.

#### СПИСОК ЛИТЕРАТУРЫ

1. Солонников В.А. О краевых задачах для линейных параболических систем дифференциальных уравнений общего вида // Тр. Мат. ин-та имени В.А. Стеклова. 1965. Т. 83. С. 3–163.
2. Бадерко Е.А., Черепова М.Ф. Потенциал простого слоя и первая краевая задача для параболической системы на плоскости // Дифференц. уравнения. 2016. Т. 52. № 2. С. 198–208.
3. Бадерко Е.А., Черепова М.Ф. Задача Дирихле для параболических систем с Дини-непрерывными коэффициентами на плоскости // Докл. РАН. 2017. Т. 476. № 1. С. 7–10.
4. Baderko E.A., Cherepova M.F. Dirichlet problem for parabolic systems with Dini continuous coefficients // Appl. Anal. 2021. V. 100. № 13. P. 2900–2910.

5. Мазья В.Г., Кресин Г.И. О принципе максимума для сильно эллиптических и параболических систем второго порядка с постоянными коэффициентами // Мат. сб. 1984. Т. 125 (167). № 4. С. 458–480.
6. Бадерко Е.А., Черепова М.Ф. Единственность решений начально-краевых задач для параболических систем в плоских ограниченных областях с негладкими боковыми границами // Докл. РАН. Математика, информатика, процессы управления. 2020. Т. 494. № 5. С. 5–8.
7. Бадерко Е.А., Черепова М.Ф. О единственности решений первой и второй начально-краевых задач для параболических систем в ограниченных областях на плоскости // Дифференц. уравнения. 2021. Т. 57. № 8. С. 1039–1048.
8. Бадерко Е.А., Сахаров С.И. Единственность решения первой начально-краевой задачи для параболической системы с дифференцируемыми коэффициентами в полуполосе с негладкой боковой границей // Дифференц. уравнения. 2021. Т. 57. № 5. С. 625–634.
9. Бадерко Е.А., Сахаров С.И. О единственности решений начально-краевых задач для параболических систем с Дини-непрерывными коэффициентами в полуограниченной области на плоскости // Журн. вычислит. математики и мат. физики. 2023. Т. 63. № 4. С. 54–65.
10. Фридман А. Уравнения с частными производными параболического типа. М., 1968.
11. Коненков А.Н. Классические решения первой краевой задачи для параболических систем на плоскости // Докл. РАН. 2022. Т. 503. С. 67–69.
12. Эйдельман С.Д. Параболические системы. М., 1964.
13. Тверитинов В.А. Гладкость потенциала простого слоя для параболической системы второго порядка // Деп. в ВИНТИ АН СССР 02.09.88. № 6850-В88.

Рязанский государственный университет  
имени С.А. Есенина

Поступила в редакцию 27.03.2023 г.  
После доработки 27.03.2023 г.  
Принята к публикации 19.05.2023 г.

## ИНТЕГРАЛЬНЫЕ И ИНТЕГРО-ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

УДК 517.923+517.925.54+517.968

### О ГИПЕРСИНГУЛЯРНЫХ ОПЕРАТОРАХ, СВЯЗАННЫХ С ПЕРИДИНАМИКОЙ

© 2023 г. Ш. А. Алимов, Ш. Н. Шералиев

Для гиперсингулярного интегрального оператора типа Кальдерона–Зигмунда, связанного с задачами перидинамики, найдено гильбертово пространство, которое переводится данным оператором в пространство квадратично суммируемых периодических функций.

DOI: 10.31857/S0374064123070051, EDN: GUNJTY

Основное уравнение перидинамики, предложенное в работе [1], в линеаризованном варианте имеет следующий вид:

$$\frac{\partial^2 u(x, t)}{\partial t^2} - A_s u(x, t) = f(x, t), \quad x \in D \subset \mathbb{R}^n, \quad t > 0,$$

где сингулярный интегральный оператор  $A_s$  определяется равенством

$$A_s u(x) = \int_D K_s(x, y)[u(y) - u(x)] dy,$$

$D$  – ограниченная  $n$ -мерная ( $n \geq 3$ ) область с кусочно-гладкой границей;  $u : D \times [0, T] \rightarrow \mathbb{R}^n$  – неизвестная функция;  $n \times n$ -матрица-функция  $K_s$ , определённая в  $D \times D$ , является заданным ядром интегрального оператора; функция  $f : D \times [0, T] \rightarrow \mathbb{R}^n$  представляет собой внешнюю силу.

Ядро  $K_s(x, y)$ , описывающее распределённое взаимодействие между частицами твёрдого тела  $D$ , имеет носитель в  $\rho$ -окрестности диагонали  $\{x = y\}$ , число  $\rho$  при этом называется *горизонтом взаимодействия*. На диагонали ядро  $K_s(x, y)$  может иметь неинтегрируемую особенность, компенсируемую разностью  $u(y) - u(x)$ . В этом случае интегральный оператор  $A_s$  является сингулярным и может оказаться неограниченным.

В настоящей работе рассматривается интегральный оператор  $A_s$  в виде свёртки в пространстве периодических функций, а именно:

$$A_s u(x) = \int_{\mathbb{T}^n} K_s(x - y)[u(y) - u(x)] dy, \quad x \in \mathbb{T}^n = [-\pi, \pi]^n. \quad (1)$$

Важный класс операторов вида (1) составляют операторы с сингулярным ядром вида

$$K_s(x) = \frac{\Omega(x)}{|x|^n}, \quad x \in \mathbb{T}^n.$$

Будем предполагать, что матрица-функция  $\Omega(x)$  является гладкой в области  $\mathbb{R}^n \setminus \{0\}$  и для любого  $\lambda > 0$  удовлетворяет условию

$$\Omega(\lambda x) = \Omega(x), \quad x \in \mathbb{R}^n \setminus \{0\}.$$

Рассмотрим сингулярный интегральный оператор

$$Sf(x) = \int_{\mathbb{T}^n} \frac{\Omega(y)}{|y|^n} f(x - y) dy. \quad (2)$$

Определим матрицу  $\Omega^*$ , представляющую собой среднее значение ядра  $\Omega$  по единичной сфере:

$$\Omega^* = \frac{1}{\omega_n} \int_{\theta} \Omega(\theta) d\theta. \tag{3}$$

Когда  $\Omega^* = 0$ , т.е. все элементы  $\Omega_{ij}^*$ ,  $i, j = \overline{1, n}$ , этой матрицы равны нулю, интегральный оператор (2) является оператором типа Кальдерона–Зигмунда (см. [2]). Известно, что в этом случае оператор (2) естественным образом определяется в классе гладких функций и продолжается до оператора, непрерывного из  $L_p(\mathbb{T}^n)$  в  $L_p(\mathbb{T}^n)$ . Отметим, что условие  $\Omega^* = 0$  является необходимым для справедливости данного утверждения.

В настоящей работе рассматривается случай  $\Omega^* \neq 0$ . Тогда оператор (2) становится гиперсингулярным, в связи с чем необходимо откорректировать его определение. Такую корректировку можно произвести, опираясь на основное уравнение перидинамики (1).

Зафиксируем  $\rho$  из интервала  $0 < \rho < \pi$  и обозначим символом  $\chi(r)$  неотрицательную функцию, принадлежащую пространству  $C^\infty(\mathbb{R})$ , равную единице при  $r \leq \rho/2$  и нулю при  $r \geq \rho$ .

Основной целью работы является изучение интегрального оператора

$$Af(x) = \int_{\mathbb{T}^n} \frac{\Omega(y)}{|y|^n} \chi(|y|) [f(x - y) - f(x)] dy. \tag{4}$$

Очевидно, что в случае  $\Omega^* = 0$  оператор (4) совпадает с сингулярным оператором типа Кальдерона–Зигмунда.

Для того чтобы сформулировать основной результат, введём функциональное пространство периодических функций с логарифмической гладкостью.

Рассмотрим самосопряжённый оператор

$$\Lambda f(x) = \sum_{k \in \mathbb{Z}^n} f_k e^{ikx} \sqrt{1 + |k|^2},$$

где

$$f_k = (2\pi)^{-n} \int_{\mathbb{T}^n} f(x) e^{-ikx} dx.$$

Иными словами,  $\Lambda = \sqrt{1 - \Delta}$ , где  $\Delta$  – самосопряжённое расширение в  $L_2(\mathbb{T}^n)$  оператора Лапласа, отвечающее периодическим граничным условиям. Отметим, что область определения оператора  $\Delta$  является пространство Соболева  $W_2^2(\mathbb{T}^n)$ , следовательно,  $D(\Lambda) = W_2^1(\mathbb{T}^n)$ . При этом функция  $f : \mathbb{T}^n \rightarrow \mathbb{R}^n$  принадлежит пространству вектор-функций  $W_2^1(\mathbb{T}^n)$ , если каждая компонента  $f_j(x)$  принадлежит обычному пространству  $W_2^1(\mathbb{T}^n)$ .

Для любого натурального  $m$  рассмотрим положительный самосопряжённый в  $L_2(\mathbb{T}^n)$  оператор  $\log^m(1 + \Lambda)$ . Область определения этого оператора обозначим  $H_{\log}^m(\mathbb{T}^n)$ :

$$H_{\log}^m(\mathbb{T}^n) = D(\log^m(1 + \Lambda)).$$

Каждое пространство  $H_{\log}^m$  является гильбертовым пространством со скалярным произведением

$$(f, g)_m = \sum_{k \in \mathbb{Z}^n} f_k \overline{g_k} \log^{2m}(1 + \sqrt{1 + |k|^2}).$$

Ассоциированную с этим скалярным произведением норму элемента  $f \in H_{\log}^m(\mathbb{T}^n)$  обозначим через  $\|f\|_m$ .

Положим  $H_{\log}^0(\mathbb{T}^n) = L_2(\mathbb{T}^n)$ . Очевидно, что для любого натурального  $m$  выполняется равенство

$$\log(1 + \Lambda) H_{\log}^m(\mathbb{T}^n) = H_{\log}^{m-1}(\mathbb{T}^n), \quad m \in \mathbb{N}. \tag{5}$$

Под нормой матрицы  $\Omega^*$  будем понимать величину

$$\|\Omega^*\| = \left( \sum_{i,j=1}^n |\Omega_{ij}^*|^2 \right)^{1/2}.$$

Справедливы следующие утверждения.

**Теорема 1.** Для любого  $m \in \mathbb{N}$  пространство  $C^\infty(\mathbb{T}^n)$  плотно в  $H_{\log}^m(\mathbb{T}^n)$ .

**Теорема 2.** Для любого натурального  $m$  оператор  $A$ , определённый равенством (4), действует из  $H_{\log}^m(\mathbb{T}^n)$  в  $H_{\log}^{m-1}(\mathbb{T}^n)$  и удовлетворяет оценке

$$\|Af\|_{m-1} \leq C\|\Omega^*\| \|f\|_m + C\|f\|_{m-1}. \tag{6}$$

В наиболее важном случае  $m = 1$  данное утверждение может быть сформулировано следующим образом.

**Следствие.** Оператор  $A$  действует непрерывно из  $D(\log(1 + \Delta))$  в  $L_2(\mathbb{T}^n)$ , при этом выполняется неравенство

$$\|Af\|_{L_2(\mathbb{T}^n)} \leq C\|\Omega^*\| \|f\|_1 + C\|f\|_{L_2(\mathbb{T}^n)}.$$

Приведём краткую схему доказательства теорем 1 и 2. Рассмотрим вспомогательный гиперсингулярный оператор

$$A^*f(x) = \int_{\mathbb{T}^n} \frac{\chi(|y|)}{|y|^n} \Omega^*[f(x - y) - f(x)] dy. \tag{7}$$

Справедливы следующие утверждения.

**Лемма 1.** Пусть  $2\pi$ -периодическая функция  $L(x)$  имеет вид

$$L(x) = -\frac{1}{n-2} \frac{\ln|x|}{|x|^{n-2}} \chi(|x|), \quad x \in \mathbb{T}^n.$$

Тогда справедливо равенство

$$\Delta L(x) = \frac{\chi(|x|)}{|x|^n} + Q(x),$$

где функция  $Q$  принадлежит  $C^\infty(\mathbb{T}^n)$  и её носитель расположен в слое  $\rho/2 \leq |x| \leq \rho$ .

**Доказательство** проводится непосредственным дифференцированием с применением свойств функции  $\chi(r)$ , определённой перед формулой (4).

**Лемма 2.** Для любой функции  $f \in C^\infty(\mathbb{T}^n)$  выполняется равенство

$$A^*f(x) = \int_{\mathbb{T}^n} L(y)\Omega^*\Delta f(x - y) dy + Bf(x) + Mf(x), \tag{8}$$

где  $M$  – постоянная матрица, а  $B$  – оператор свёртки с ядром из пространства  $C^\infty(\mathbb{T}^n)$ .

**Доказательство.** Согласно определению (7) и лемме 1 можно записать

$$A^*f(x) = \int_{\mathbb{T}^n} [\Delta L(y)]\Omega^*[f(x - y) - f(x)] dy - \int_{\mathbb{T}^n} Q(y)\Omega^*[f(x - y) - f(x)] dy. \tag{9}$$

Из условия леммы следует, что равномерно по  $x \in \mathbb{T}^n$  выполняется оценка

$$f(x - y) - f(x) = O(|y|).$$

Стало быть, первый интеграл в равенстве (9) сходится абсолютно и равномерно, и мы можем применить к нему вторую формулу Грина (см. [3, гл. IV, § 2, п. 1, формула (5)]). В результате интеграл преобразуется следующим образом:

$$\int_{\mathbb{T}^n} [\Delta L(y)] \Omega^* [f(x - y) - f(x)] dy = \int_{\mathbb{T}^n} L(y) \Omega^* \Delta_y f(x - y) dy. \tag{10}$$

Далее воспользуемся леммой 1. Положим

$$Bf(x) = - \int_{\mathbb{T}^n} Q(y) \Omega^* f(x - y) dy, \tag{11}$$

где  $Q \in C^\infty(\mathbb{T}^n)$ .

Обозначим

$$M = \int_{\mathbb{T}^n} Q(y) dy \Omega^*.$$

В таком случае из (9)–(11) следует требуемое соотношение (8). Лемма доказана.

**Лемма 3.** Коэффициенты Фурье  $L_k$  функции  $L(x)$  для любого натурального  $N$  удовлетворяют равенству

$$L_k = \frac{1}{|k|^2} [\alpha_n \ln |k| + \beta_n] + O(|k|^N), \quad k \in \mathbb{Z}^n, \quad |k| \geq 1.$$

Постоянные  $\alpha_n$  и  $\beta_n$  зависят только от размерности  $n$  области.

**Доказательство** леммы проводится методом формулы среднего значения, впервые применённым В.А. Ильиным для исследования ядра интегрального оператора  $\Lambda^{-\alpha}$  (см. [4]). Основанная на этом методе схема доказательства леммы приведена в работе [5].

**Лемма 4.** Для любой функции  $f \in C^\infty(\mathbb{T}^n)$  коэффициенты Фурье функции  $A^*f(x)$  удовлетворяют оценке

$$|(A^*f)_k| \leq C |\Omega^*| |f_k| \ln(1 + \sqrt{1 + |k|^2}) + C |f_k|. \tag{12}$$

**Доказательство.** Воспользуемся равенством (8). Первый интеграл в этом соотношении представляет собой свёртку, поэтому

$$(A^*f)_k = (2\pi)^n L_k \Omega^* (\Delta f)_k + (Bf)_k + Mf_k. \tag{13}$$

Аналогично из (11) получаем

$$(Bf)_k = (2\pi)^n Q_k f_k.$$

Так как  $(\Delta f)_k = -|k|^2 f_k$ , то равенство (13) можно записать следующим образом:

$$(A^*f)_k = -(2\pi)^n |k|^2 L_k \Omega^* f_k + Q_k f_k + Mf_k.$$

Для получения требуемой оценки (12) остаётся воспользоваться леммой 3 и неравенством

$$\ln |k| \leq \ln(1 + \sqrt{1 + |k|^2}), \quad |k| \geq 1.$$

Лемма доказана.

**Лемма 5.** Для любой функции  $f \in C^\infty(\mathbb{T}^n)$  выполняется оценка (6).

**Доказательство.** Введём следующую вспомогательную функцию:

$$\Omega_1(x) = \Omega(x) - \Omega^*,$$

где постоянная матрица  $\Omega^*$  определена равенством (3). Очевидно, что выполняется равенство

$$\int_{\theta} \Omega_1(\theta) d\theta = 0.$$

В таком случае оператор  $A_1 = A - A^*$  является оператором Кальдерона–Зигмунда и поэтому для него имеет место оценка

$$\|A_1 f\| \leq C \|f\|, \quad f \in H_{\log^0}(\mathbb{T}^n) = L_2(\mathbb{T}^n).$$

Следовательно, для любой функции  $f \in C^\infty(\mathbb{T}^n)$  выполняется неравенство

$$\|A f\| \leq \|A^* f\| + C \|f\|, \quad f \in L_2(\mathbb{T}^n).$$

Основное неравенство (6) теперь может быть получено непосредственно из оценки (12) и соотношения (5). Лемма доказана.

**Доказательство теоремы 1** основано на том, что частичные суммы ряда Фурье любой функции  $f \in H_{\log^m}^m(\mathbb{T}^n)$  принадлежат пространству  $C^\infty(\mathbb{T}^n)$  и сходятся к  $f$  в метрике  $H_{\log^m}^m(\mathbb{T}^n)$ .

**Доказательство теоремы 2** следует из леммы 5. По условию этой леммы оценка (6) справедлива лишь для функций из  $C^\infty(\mathbb{T}^n)$ . Тем не менее, согласно теореме 1, мы можем замкнуть оператор  $A$  в гильбертовом пространстве  $H_{\log^1}^1(\mathbb{T}^n)$  и получить неравенство (6) в условиях теоремы 2.

Подчеркнём, что в основе доказательства приведённых лемм лежит метод формулы среднего значения, разработанный В.А. Ильиным (см. [6]).

## СПИСОК ЛИТЕРАТУРЫ

1. *Silling S.A.* Reformulation of elasticity theory for discontinuities and long-range forces // *J. Mech. Phys. Solids.* 2000. V. 48. № 1. P. 175–209.
2. *Calderon A.P., Zygmund A.* On the existence of certain singular integrals // *Acta Math.* 1952. V. 88. P. 85–139.
3. *Тихонов А.Н., Самарский А.А.* Уравнения математической физики. М., 1999.
4. *Ильин В.А.* Ядра дробного порядка // *Мат. сб.* 1957. Т. 41 (83). № 4. С. 459–480.
5. *Alimov Sh., Sheraliev Sh.* On the solvability of the singular equation of peridynamics // *Complex Variables and Elliptic Equat.* 2019. V. 64. № 5. P. 873–887.
6. *Ильин В.А.* Спектральная теория дифференциальных операторов. Самосопряжённые дифференциальные операторы. М., 1991.

Национальный университет Узбекистана  
имени Мирзо Улугбека, г. Ташкент,  
Филиал Московского государственного университета  
имени М.В. Ломоносова, г. Ташкент, Узбекистан

Поступила в редакцию 04.04.2023 г.  
После доработки 04.04.2023 г.  
Принята к публикации 19.05.2023 г.

УДК 517.97

## ВАРИАЦИОННЫЕ МЕТОДЫ РЕШЕНИЯ ЗАДАЧ, СВЯЗАННЫХ С ИСКУССТВЕННЫМ ИНТЕЛЛЕКТОМ

© 2023 г. М. В. Бабушкин, В. Ю. Тертычный-Даури

Введены понятия алгоритма обучения, целевой функции, распознающей системы, класса изображений, тренировочного множества, алгоритма с поощрением, конечно-сходящегося алгоритма, адаптивной управляющей системы, цели управления, тактики управления, времени адаптации и т.д., относящиеся к проблеме искусственного интеллекта в процессах обучения и адаптации. Поставлена общая задача о самообучении (обучении “без учителя”) – о разделении множеств – в терминах классического вариационного исчисления. Общность задачи обусловлена введением в анализ дополнительной переменной времени. Задача решена с помощью определения экстремальных условий, при которых будет достигнута минимизация функционала общего среднего риска. Рассмотрены задачи, соответствующие нефиксированному и фиксированному промежуткам времени. Для двух этих случаев найдены выражения для расчёта вариаций функционалов качества. Указаны необходимые условия для определения экстремальных значений процесса самообучения (разделения классов множества изображений) во времени.

DOI: 10.31857/S0374064123070063, EDN: GUMYBE

**Введение.** Начнём с известного тезиса о том, что имеются разночтения по поводу смыслового содержания понятия “искусственный интеллект” (ИИ). Тем не менее доминирует достаточно общая и универсальная точка зрения на ИИ как на задачу обучения. При этом, конечно же, речь не идёт об обучении (дрессировке) биологических существ. Имеются в виду лишь искусственно созданные человеком вычислительные комплексы (роботы, компьютеры, цифровые имитаторы и т.д.), наделённые памятью, сенсорными устройствами, регуляторами в виде обратных связей, механическими и электрическими приводами и т.д.

Часто под ИИ понимают задачу машинного, т.е. компьютерного обучения различного типа нейросетей умению производить классификацию текстов, детекции, сегментации изображений. При всём при этом задачи управления динамическими объектами здесь не рассматриваются – дело сводится только к формированию сходящихся алгоритмов и процедур самообучения (обучения “без учителя”) с учётом экстремизации некоторого функционала качества. “Распознавание образов представляет собой часть общей проблемы искусственного интеллекта, возникшей под влиянием работ Н. Винера, Дж. фон Неймана, А.М. Тьюринга, У. Росс Эшби и других основоположников кибернетики” [1, с. 218].

Подробнее о задачах распознавания образов, в которых изучаются вопросы функционирования опознающих систем, можно узнать из обширной литературы, например работ [1–3]. Итак, проблема ИИ может рассматриваться, как было упомянуто, по-разному, например, как задача машинной классификации изображений, а может изучаться как задача адаптивного обучения через задачу динамического оценивания неизвестных параметров управляемого объекта или параметров воздействующей на него среды (различных возмущений). Так, хорошо известная задача обучения езде робота на велосипеде [1, 4, 5] попадает под эту трактовку проблемы ИИ.

**1. Искусственный интеллект как процесс самообучения с использованием адаптивных обратных связей.** Задача формирования и создания обучаемых, адаптивных, опознающих систем, как теперь становится отчётливо понятно, возникла при создании систем, сходных по своим возможностям и эффективности использования с системами искусственного интеллекта.

Речь идёт о задаче создания систем, работающих в условиях существенной априорной неопределённости о свойствах внешней среды и некоторых внутренних параметров. Эта задача

приводит к необходимости построения таких систем, которые могли бы при своём функционировании на основе поступающей и обрабатываемой информации менять свои параметры или свою структуру так, чтобы обеспечить с течением времени целевые задания. Такие системы стали называться *обучаемыми (самообучаемыми) адаптивными системами*.

Многие известные адаптивные системы используют в режиме своего функционирования итеративные *алгоритмы обучения*, которые также называются *алгоритмами настройки* или *адаптивными алгоритмами*, которые определяют правило (закон) целенаправленного изменения параметров (структуры) системы, зависящей от поступающей на вход информации. Большинство таких алгоритмов представляют собой градиентные либо поисковые процедуры, связанные с поставленной *целевой функцией*.

Эффективность применения данных алгоритмов напрямую зависит от характера выполнения их сходимости. Отметим, что в основе доказательства сходимости многих алгоритмов обучения лежат идеи, использующие метод функций Ляпунова в теории дифференциальных уравнений, при котором положительная функция Ляпунова убывает на траекториях изучаемого процесса.

В настоящее время имеется большое количество работ, посвящённых решению задач обучения, самообучения опознающих систем, исследованию алгоритмов адаптивного управления. Следует также отметить, что перцептронные системы Розенблатта при исследовании проблемы автоматической классификации изображений предназначались для изучения механизмов работы мозга и их возможностей. Этот и другие подходы в виде использования теории групп и групповых инвариантов, экстраполяционных и лингвистических схем часто не позволяют моделировать способность мозга к обучению на очень коротких обучающих последовательностях.

**2. Процесс обучения.** Изначально обучаемая опознающая система трактовалась как перцептронная система, где задача обучения ставилась в виде задачи построения по обучающей последовательности разделяющих поверхностей в пространстве различных признаков. Такой подход к задаче распознавания можно назвать геометрическим или, иначе, экстраполяционным. При этом опознающие системы различаются как алгоритмами обучения, так и их предназначением.

Модель перцептрона (системы нейронных сетей) Розенблатта послужила прототипом многих моделей обучаемых опознающих систем. Будем под *системой* иметь в виду некоторое устройство, предназначенное для однозначного отображения множества входных сигналов в множество выходных сигналов. Входные сигналы называют *изображениями*, а выходные – *ответами системы*.

Далее *опознающая*, или, иначе, *распознающая, система*, которая часто называется *классификатором*, представляет собой систему, классифицирующую множество изображений исходя из некоторой доступной классификации. Здесь под классификацией подразумевается разбиение множества изображений или его части на непересекающиеся подмножества, именуемые *классами изображений*.

Система, способная стать опознающей с помощью введения в неё некоторой информации о части изображений и классах изображений, называется *обучаемой опознающей системой*. Это означает, что после введения информации лишь о части изображений такая система способна стать классификатором всех изображений. Тем самым обнаруживаем *свойство экстраполяции*, присущее обучаемым системам.

Отметим, что множество изображений (с информацией о них, вводимой в систему) значительно беднее всего множества изображений. Это “вводимое” множество (вернее, вводимая информация о нём) состоит из конечного числа изображений. Система в результате такого ввода может по-разному производить классификацию, т.е. может менять свою структуру. Процесс ввода информации об изображениях, как результат, приводящий к изменению структуры, называется *обучением* системы, где множество изображений, которые система использует в обучении, называется *тренировочным* или *обучающим множеством*.

Важно обратить внимание, что в обучаемую опознающую систему заранее не вводится полная информация о классах изображений; требуется, однако, чтобы способность к правильной классификации изображений возникла, как итог, в процессе обучения.

Ещё одно важное замечание. Конкретная обучаемая опознающая система – это система, зависящая от некоторого конкретного параметра, определяющего её структуру. Обучение тогда сводится к нахождению или оценке требуемого значения этого параметра.

Входящие сигналы воспринимаются системой на основе конечного набора признаков – функций, определённых на множестве изображений, причём набор значений всех признаков, соответствующих данному изображению, называется *описанием изображения*. Точки в пространстве признаков называются *изображениями*, а множества в пространстве признаков, отвечающие классам изображений, называются *классами в пространстве признаков* или *образами*.

Обозначим множество изображений через  $\Omega$  – абстрактное множество с элементами  $\omega$ . Пусть  $\Omega_1, \dots, \Omega_l$  – непересекающиеся подмножества множества  $\Omega$ ; здесь  $\Omega_k, k = \overline{1, l}$ , – классы изображений. Вещественные функции  $s_k(\omega), k = \overline{1, l}$ , такие, что  $s_k(\omega) > \max_{i \neq k} s_i(\omega)$

для любых  $k$  и  $\omega \in \Omega_k$ , называются *дискриминантными* или *разделяющими* по отношению к подмножествам  $\Omega_1, \dots, \Omega_l$ .

В случае двух классов изображений ( $l = 2$ ) функция  $s(\omega) = s_1(\omega) - s_2(\omega)$  принимает значения разных знаков на разных классах: множества  $\Omega_1$  и  $\Omega_2$  разделяются по знаку функции  $s(\omega)$ .

Для опознающей системы с переменной структурой можно рассмотреть задачу, называемую *задачей обучения*, которая состоит в отыскании такого параметра  $\tau_* \in T$ , при котором  $\Omega_k(\tau_*) = \Omega_k, k = \overline{1, l}$ .

Найти требуемый параметр  $\tau_*$  можно с помощью перебора элементов множества  $T$ . Такая процедура перебора называется *алгоритмом обучения*, а опознающая система с переменной структурой и с соответствующим алгоритмом обучения – *обучаемой опознающей системой*.

Задача обучения опознающей системы, основанной на *методе конечно-сходящихся алгоритмов* (КСА) решения целевых неравенств, изучается как задача решения счётной системы *целевых неравенств*. Приходим тем самым к задаче оценивания неизвестного параметра, определяющего структуру опознающей системы; эта задача решается при помощи специальных рекуррентных процедур, названных *алгоритмами с поощрением*, которые сходятся при определённых условиях за конечное число итераций.

Пусть алгоритм обучения имеет вид

$$\tau_{n+1} = \Phi(x_n, \tau_n, \sigma_n), \quad n \in \mathbb{N}, \quad (1)$$

формирование оценок векторного параметра  $\tau$  (вектора весов) происходит последовательно с учётом текущей информации. Здесь  $\tau_1$  – произвольный начальный вектор весов;  $x_1, x_2, \dots$  – элементы тренировочной последовательности;  $\sigma_n$  – некоторые параметры, характеризующие класс векторных функций  $\Phi$ .

Для функционирования алгоритма (1) необходимо знать изображение  $x_n$ , поступающее на вход системы в данный момент времени. Рекуррентный алгоритм (1) называется *алгоритмом с поощрением*, если его можно представить в виде

$$\tau_{n+1} = \tau_n + \theta_n \Phi(x_n, \tau_n, \sigma_n), \quad n \in \mathbb{N}, \quad (2)$$

где

$$\theta_n = \begin{cases} 1, & \text{если } \varphi(x_n, \tau_n, \sigma_n) \leq 0, \\ 0, & \text{если } \varphi(x_n, \tau_n, \sigma_n) > 0, \end{cases} \quad (3)$$

а  $\varphi(x, \tau, \sigma)$  – некоторая *целевая функция*, характеризующая качество работы алгоритма.

Видим, что в алгоритме (2), (3) вектор весов на  $n$ -м шаге не изменяется, если система работает правильно и  $\theta_n = 0$ ; в противном случае он изменяется по некоторому правилу. Эта информация о характере работы системы в процессе обучения (которая является информацией об изображениях тренировочного множества) сообщается опознающей системе и может учитываться при формировании алгоритмов обучения.

Рекуррентные алгоритмы без поощрения называются также *алгоритмами с принудительным обучением*, отличительной чертой которых является наличие целевого условия, при котором их работа в процессе обучения состоит в накоплении некоторых статистических данных о внешних возмущениях с целью нахождения вектора весов.

Введём понятие о КСА обучения опознающей системы [1, 4, 5]. Алгоритм (2), (3) называется *конечно-сходящимся*, если для произвольного начального вектора  $\tau_1$  и произвольной бесконечной тренировочной последовательности  $x_1, x_2, \dots$  выполнено условие

$$\lim_{n \rightarrow \infty} \varphi(x_n, \tau_n, \sigma_n) > 0,$$

означающее существование такого номера  $r$ , что при  $n \geq r$  будет выполнено неравенство  $\varphi(x_n, \tau_n, \sigma_n) > 0$ . Для алгоритма (2), (3) это означает, что  $\tau_r = \tau_{r+1} = \dots$ , т.е. последовательность  $\tau_n$  сходится к пределу за *конечное* число шагов алгоритма.

Если на тренировочном множестве  $\{x_n\}_{n \in \mathbb{N}}$  определена счётная последовательность целых неравенств относительно искомого вектора  $\tau$ :

$$\varphi(x_n, \tau, \sigma_n) > 0, \quad n \in \mathbb{N}, \quad (4)$$

то тогда задача об обучении опознающей системы представляет собой задачу нахождения алгоритма решения  $\tau$  счётной системы неравенств (4).

Пусть задана вещественная функция  $\varphi(x, \tau)$ ,  $x \in X \subset \mathbb{R}^p$ ,  $\tau \in \mathbb{R}^p$ , такая, что выполнены условия:

1)  $\varphi(x, \tau)$  ограничена равномерно по  $x \in X$ ,  $|\varphi(x, \tau)| < C_1$ , дифференцируема по  $\tau$  для любого  $x$ ; евклидова норма её градиента  $\|\nabla_{\tau} \varphi(x, \tau)\| < C_2$ , где  $C_1, C_2$  – постоянные, не зависящие от выбора  $x \in X$ ;

2) существуют  $\tau_* \in \mathbb{R}^p$  и число  $\varepsilon_* > 0$ , для которых справедливы неравенства  $\varphi(x, \tau_*) \geq \varepsilon_* > 0$  при всех  $x \in X$ ;

3) для любых пар точек  $(x, \tau)$ ,  $x \in X$ ,  $\tau \in \mathbb{R}^p$ , удовлетворяющих условию  $\varphi(x, \tau) \leq 0$ , справедливо неравенство

$$(\nabla_{\tau} \varphi(x, \tau), \tau_* - \tau) \geq \varphi(x, \tau_*) - \varphi(x, \tau).$$

Пусть, кроме того,  $\tau_1 \in \mathbb{R}^p$  – произвольный вектор и  $\{x_n\} \in X$  – произвольное счётное множество. Зададим алгоритм с поощрением вида

$$\tau_{n+1} = \tau_n + \alpha_n \theta_n \nabla_{\tau} \varphi(x_n, \tau_n), \quad (5)$$

где

$$\theta_n = \begin{cases} 1, & \text{если } \varphi(x_n, \tau_n) \leq 0, \\ 0, & \text{если } \varphi(x_n, \tau_n) > 0, \end{cases}$$

и  $\alpha_n$  – некоторая числовая последовательность

Из утверждения теоремы 2.1.1 [1, с. 34] следует, что в случае выполнения условий 1)–3) алгоритм (5) является конечно-сходящимся, если  $\alpha_n = A_n + B_n$ , либо  $\alpha_n = \max\{A_n, B_n\}$ , где

$$A_n = \frac{1}{\varkappa_n}, \quad B_n = -\frac{\beta_n \varphi(x_n, \tau_n)}{\|\nabla_{\tau} \varphi(x_n, \tau_n)\|^2},$$

$$\varkappa_n = 1 + \sum_{k=1}^n \theta_k, \quad \varkappa_{n+1} = \varkappa_n + \theta_{n+1}, \quad \varkappa_0 = 1, \quad \beta_n \in [0, 2], \quad n \in \mathbb{N} \cup \{0\}.$$

Алгоритм с поощрением (5) имеет характерную форму: вектор  $\tau$  не меняется, если очередное неравенство выполнено, и изменяется в направлении наибольшего увеличения функции  $\varphi(x, \tau)$  в противном случае.

Конкретные КСА предназначены для поиска решения линейных систем неравенств. Их можно использовать в задаче обучения опознающей системы перцептронного типа относительно разделения множеств. Вопросы, связанные с таким разделением (алгоритмы “Ява”, “Полоска”, “Отражение”, “Надёжное разделение”, “Оптимальное разделение”, процедуры построения комитета неравенств и др.) подробно рассмотрены в работе [1, § 2.2]. Там же можно

найти результаты, относящиеся к теории обучения: стохастическим аналогам КСА, алгоритмам случайного поиска, теории принятия статистических решений, минимизации эмпирического риска, методу стохастической аппроксимации и процедуре Роббинса–Монро, методу потенциальных функций и самообучению (обучению “без учителя”).

**3. Процесс адаптации.** Воспользуемся некоторыми рассуждениями и доводами о применении метода КСА для решения задач адаптивного управления, изложенными в работах [1, 4–6]. Важно сразу же отметить, что задачи адаптивной стабилизации и слежения могут быть переформулированы как задачи нахождения решения системы неравенств, задаваемых целевым условием и способом формирования управления. Поэтому для нахождения требуемого решения можно использовать рекуррентные алгоритмы с поощрением.

Пусть имеется непрерывный динамический объект управления (ОУ), снабжённый системой управления (СУ), причём синтез СУ должен осуществляться в условиях априорной неопределённости о параметрах самого ОУ либо внешних условий в виде неконтролируемых возмущений. Тогда естественно обратиться к *адаптивным управляющим системам* для восполнения недостающей и необходимой информации в процессе формирования СУ и функционирования ОУ.

Зададим ОУ с помощью уравнения

$$F[y(t), \dots, y^{(m_1)}(t), u(t), \dots, u^{(m_2)}(t), \xi, t] = 0, \quad (6)$$

где  $F(\cdot)$  – известная вещественная функция своих аргументов,  $y(t)$  – вектор-функция выходов ОУ,  $u(t)$  – вектор-функция управлений,  $y^{(k)}(t)$ ,  $u^{(k)}(t)$  –  $k$ -е производные по  $t$  соответствующих векторных функций,  $m_1$  и  $m_2$  – натуральные числа,  $\xi$  – неизвестный параметр из множества  $\Xi$ ,  $t \in [t_0, \infty)$  – время. Множество  $\Xi$  характеризует степень априорной неопределённости об ОУ и условиях его работы. Считаем функцию  $F$  в уравнении (6) такой, что при задании достаточно гладкой вектор-функции  $u(t)$  и фиксации параметра  $\xi \in \Xi$  уравнение (6) однозначно определяет выходной процесс  $y(\xi, t)$ . Всё это в совокупности позволяет говорить о *классе ОУ*, определяемом множеством  $\Xi$  с параметрами  $\xi$  в виде его элементов.

При выборе управления нужно ориентироваться на *цель управления* (ЦУ) – требование обеспечения некоторых ограничений на выходной процесс. Понятно, что ЦУ определяются особенностями каждой конкретной задачи управления. К примеру, зададим ЦУ в виде обеспечения предельного неравенства

$$\liminf_{t \rightarrow \infty} \varphi[\sigma(s), s] > 0, \quad (7)$$

где  $\sigma(t)$  – набор известных к моменту времени  $t$  величин, связанных с ОУ и СУ ( $\sigma(t)$ , в частности, может содержать выход  $y(t)$  и вход  $u(t)$  и их производные до определённого порядка),  $\varphi[\sigma(t), t]$  – известная вещественная функция своих аргументов.

Предположим, что для каждого ОУ из класса  $\Xi$  управление  $u(t)$ , обеспечивающее выполнение ЦУ, существует и его можно найти из уравнения

$$U[\sigma(t), u(t), \dots, u^{(m_3)}(t), \omega(\xi), t] = 0, \quad (8)$$

где  $\omega(\xi)$  – отображение  $\Xi$  в конечномерное пространство  $\Omega$ ,  $U$  – известная функция своих аргументов. Тем самым считаем, что если модель ОУ определена, параметр  $\xi$  известен, то можно синтезировать управление, обеспечивающее ЦУ (7) в форме (8).

Отметим, что в ряде практических задач синтез требуемых управлений можно произвести в виде некоторого правила, зависящего не от всех характеристик ОУ, а лишь от некоторого числа их комбинаций, что и служит объяснением появления величины  $\omega(\xi)$  в соотношении (8). Если  $\omega(\xi)$  – это тождественное преобразование, то  $\Xi \subseteq \Omega$ .

Управлением (8) нельзя, однако, воспользоваться, так как оно зависит от неизвестного вектора  $\omega(\xi)$ . Поэтому следует искать требуемое управление из уравнения

$$U[\sigma(t), u(t), \dots, u^{(m_3)}(t), \tau(t), t] = 0, \quad (9)$$

где  $\tau(t) \in \Omega$  – вектор-функция, которую также называют *тактикой управления* (ТУ). Соотношение (9) дополняют алгоритмом изменения ТУ в виде

$$\tau(t+0) = Q\{\tau(t-0) + \theta(t)\psi[\sigma(t-0), \tau(t-0), t-0]\}, \quad (10)$$

где обозначено

$$f(t \pm 0) = \lim_{\varepsilon \rightarrow 0} f(t \pm \varepsilon), \quad \varepsilon > 0, \quad \theta(t) = \begin{cases} 1, & \text{если } \varphi[\sigma(t), t] \leq 0, \\ 0, & \text{если } \varphi[\sigma(t), t] > 0. \end{cases} \quad (11)$$

Здесь  $\psi(\sigma, \tau, t)$  – функция со значениями в  $\Omega$ ;  $Q(\tau)$  – отображение множества  $\Omega$  в его подмножество  $T$ ,  $\tau(t_0) \in T$  – произвольный элемент.

Предполагается, что для любого  $\tau \in T$  из уравнения (9) можно определить вектор-функцию  $u(t)$ , не обязательно обеспечивающую выполнение ЦУ (7), и  $\omega(\Xi) \subseteq T$ . В нашем случае достаточно ограничиться алгоритмом (11), реализующим кусочно-постоянную ТУ со значениями в  $T$ . Полагаем, что функции  $Q(\tau)$  и  $\psi(\sigma, \tau, t)$  определяются, чтобы обеспечить ЦУ (7); понятно, что эти функции не должны зависеть от неизвестного параметра  $\xi$ .

Соотношения (9)–(11) задают полностью СУ, предназначенную для управления любым ОУ из класса  $\Xi$ . СУ (9)–(11) называется *адаптивной в классе  $\Xi$  ОУ по отношению к ЦУ (7)*, если неравенство (7) выполнено при всех  $\xi \in \Xi$ .

Ещё отметим, что наименьшее время  $t(\xi)$ , начиная с которого во все дальнейшие моменты времени выполнено неравенство  $\inf_{s \geq t} \varphi[\sigma(s), s] > 0$ , называется *временем адаптации* или, иначе, *временем настройки, временем обучения* адаптивной системы управления (АдСУ). Процедура (10), (11) называется *алгоритмом адаптации*.

В случае когда АдСУ предназначена для поддержания некоторых характеристик ОУ на каком-то конкретном уровне в заданных пределах, говорят об *адаптивном регуляторе* (АР). АР включает в себя сам *регулятор* и *адаптор* – устройство, назначение которого состоит в настройке регулятора. *Системой адаптивного управления* называют систему, объединяющую ОУ и АдСУ.

К АдСУ предъявляется требование практической приемлемости, а именно время  $t(\xi)$  адаптации не должно быть слишком большим. После завершения процесса адаптации АдСУ значения параметра  $\omega(\xi)$  не найдёт, а только получит некоторую его оценку  $\tau[t(\xi)]$ , при которой будет выполнено ЦУ. Построение АдСУ в виде соотношений (9)–(11) для конструирования  $\psi(\sigma, \tau, t)$  и  $Q(\tau)$  называют *синтезом АдСУ*. На практике, очевидно, нужно уметь так формулировать закон управления (9) и ЦУ (7), чтобы синтез АдСУ был возможен.

Обратим теперь внимание на важную задачу об адаптивной стабилизации линейного ОУ. Предположим, что ОУ задаётся уравнением

$$\sum_{k=0}^{m_1} B_k y^{(k)}(t) = \sum_{k=0}^{m_2} C_k u^{(k)}(t) + v(t), \quad (12)$$

где  $y(t)$  –  $m$ -я компонента вектор-функции выходов ОУ,  $u(t)$  –  $l$ -я компонента вектор-функции управлений,  $v(t)$  –  $m$ -я компонента вектор-функции внешних возмущений.

Будем считать, что  $v(t)$  равномерно ограничена по  $t$ :

$$\sup_t \|v(t)\| < \infty.$$

Считаем также, что матрицы  $B_k \in \mathbb{R}^m \times \mathbb{R}^m$ ,  $C_k \in \mathbb{R}^m \times \mathbb{R}^l$  не зависят от  $t$ . Видим, что ОУ (12) является частным случаем ОУ (6). Пусть матрица  $B_{m_1}$  известна и невырождена. Тогда можно положить, что  $B_{m_1} = I_m$  – единичная матрица порядка  $m$ . Кроме того, пусть матрица  $C_{m_2}$  известна, причём матрица  $C_{m_2}^* C_{m_2}^*$  невырождена. Остальные коэффициенты ОУ считаются неизвестными и их объединение даёт вектор неизвестных параметров  $\xi$ . Пусть также элементы вектор-функции  $v(t)$  не измеряются.

Задача о стабилизации ОУ (12) ставится так: требуется найти такие управления  $u(t)$ , при которых выход ОУ удовлетворял бы условию  $\|y(t)\| < \varepsilon_0$ , где  $\varepsilon_0$  – заданное положительное число.

Приходим при синтезе АР к задаче о нахождении матричного решения  $\tau$  неравенств  $\varphi(\tau, t) > 0$ , где для целевой функции  $\varphi(\tau, t)$  существует решение  $\tau_*$  такое, что для всех  $t$  имеет место “усиленное” неравенство  $\varphi(\tau_*, t) \geq \varepsilon_*$ ,  $\varepsilon_* > 0$ . Значения  $\varphi(\tau, t)$  известны для любых  $t$ ; функция  $\varphi(\tau, t)$  выпукла по  $\tau$ . Поэтому для решения неравенств  $\varphi(\tau, t) > 0$  можно воспользоваться матричным аналогом КСА с произвольной матрицей  $\tau(t_0)$  соответствующей размерности.

Возьмём рекуррентную процедуру для нахождения матричной функции  $\tau(t)$  вида

$$\tau(t+) = \tau(t-0) + \gamma(t-0)\theta(t)\nabla_{\tau}\varphi(\tau-0, t-0), \tag{13}$$

где

$$\theta(t) = \begin{cases} 1, & \text{если } \varphi[\tau(t), t] \leq 0, \\ 0, & \text{если } \varphi[\tau(t), t] > 0, \end{cases}$$

$\gamma(t)$  – некоторые положительные функции,  $\nabla_{\tau}\varphi(\tau, t)$  – градиент функции  $\varphi$  по матричному аргументу  $\tau$ . Выражения для  $\varphi(\cdot)$ ,  $\nabla_{\tau}\varphi(\cdot)$  можно найти в работе [1, с. 34].

Функции  $\gamma(t)$  вычисляются по формуле

$$\gamma(t) = \frac{1}{\varkappa(t)} - \frac{\beta(t)\varphi[\tau(t), t]}{\|\nabla_{\tau}\varphi[\tau(t), t]\|^2}, \tag{14}$$

где функция  $\varkappa(t)$  определяется согласно рекуррентной процедуре

$$\varkappa(t+0) = \varkappa(t-0) + \theta(t), \quad \varkappa(t_0) = 1, \quad \beta(t) \in [0, 2].$$

Здесь  $\beta(t)$  – произвольная функция со значениями из промежутка  $[0, 2]$ ,

$$\|\nabla_{\tau}\varphi(\tau, t)\|^2 = \text{Sp} \{[\nabla_{\tau}\varphi(\tau, t)]^*\nabla_{\tau}\varphi(\tau, t)\}.$$

Можно показать, что при всех  $t, s$  таких, что  $t > s$ , справедливо неравенство

$$\|\tau_* - \tau(s)\| \geq \|\tau_* - \tau(t)\|,$$

где  $\|\tau\|^2 = \text{Sp}(\tau^*\tau)$ , откуда следует, что с течением времени  $\tau(t) \rightarrow \tau_*$ ;  $\tau_*$  – искомый матричный параметр. Если

$$\sup_t \|\nabla_{\tau}\varphi[\tau(t), t]\| < \infty,$$

то легко показать, что алгоритм (13), (14) – это КСА, так как

$$\lim_{t \rightarrow \infty} \theta(t) = 0, \tag{15}$$

и в процессе обучения будет лишь конечное число коррекций коэффициентов регулятора. Из двузначности  $\theta(t)$ , а также из условия (15) вытекает, что  $\theta(t) \equiv 0$  при достаточно больших  $t$ .

**4. Общая задача о самообучении.** Ставится задача о самообучении, т.е. обучении “без учителя”, при условии, что в анализ добавляется ещё одна переменная, принимающая смысл времени. Это условие в постановке задачи существенным образом её усложняет. Происходит тем самым не только разбиение пространства признаков на конечное число подмножеств, но и учёт происходящих при этом изменений во временной области. Такая задача может считаться новой.

Перейдём теперь к пояснениям, необходимым для формулировки основной задачи о самообучении. Требуется всё множество изображений разбить на  $l$  классов. Введём в рассмотрение случайную функцию (случайный процесс), характеризующую расстояние от точки  $x(\omega, t)$  до

некоторой неизвестной пока особой точки (центра) класса. В этом случае классификация “во времени” может быть основана на требовании, чтобы каждый элемент  $x$  каждого класса был удалён от особой точки класса на расстояние меньшее, чем от особых точек других классов.

Важно отметить, что данное условие можно связать с минимизацией *функционала среднего риска* на рассматриваемом промежутке времени  $[t_0, t_1]$ . Более подробную информацию об этом функционале можно найти в работах [1, 2]. Задачу самообучения будем считать решённой, если по тренировочному множеству удастся найти особые точки (центры)  $\tau^{(1)}, \dots, \tau^{(l)}$  и границы соответствующих множеств, для которых риск будет минимален.

Итак, пусть пространство признаков  $X(t) \subset \mathbb{R}^q$  разбито на  $l$  непересекающихся подмножеств  $X^{(1)}(t), \dots, X^{(l)}(t)$ ; оно подвержено деформации во времени, на что указывает зависимость от времени  $t$ . Обозначим через  $\tau^{(k)}(t), k = \overline{1, l}$ , особые точки (центры) выделенных множеств.

Зададим далее  $R_k[x], k = \overline{1, l}$ , – произвольные функции, характеризующие расстояние от изображения  $x(t) = x(\omega, t)$  до особой точки (центра)  $\tau^{(k)}(t), k = \overline{1, l}$ , соответствующего класса, используя выражение  $\rho_k[x(t), \tau^{(k)}(t)] = R_k[x(t) - \tau^{(k)}(t)]$ . Здесь функция  $\rho_k[x(t), \tau^{(k)}(t)]$  играет роль *функции потерь* или *функции штрафов* для изображений  $k$ -го класса,  $k = \overline{1, l}$ . С учётом того, что  $x(t) = x(\omega, t)$  – случайный вектор процессов, средний риск (средний штраф) можно представить в виде

$$W = \int_{t_0}^{t_1} \sum_{k=1}^l p_k(t) \int_{X^{(k)}(t)} R_k[x - \tau^{(k)}(t)] p_k[x, t] dx dt, \tag{16}$$

где с помощью вероятностей введены обозначения для плотностей распределения непрерывных случайных величин (изображений)  $x(t)$  при данном  $t$ :

$$p_k(t) = P\{x(\omega, t) \in X^{(k)}(t)\}, \quad k = \overline{1, l}, \quad t \in [t_0, t_1],$$

$$p_k[x, t] = \lim_{\varepsilon \rightarrow 0} \frac{P\{[x(\omega, t) \in D_\varepsilon(x)] \cap X^{(k)}(t)\}}{P[x(\omega, t) \in D_\varepsilon(x)] V_\varepsilon(x)},$$

здесь  $D_\varepsilon(x)$  – шар радиуса  $\varepsilon$  с центром в точке  $x$ ,  $V_\varepsilon(x)$  – объём этого шара, в предположении, что написанный предел существует.

Пользуясь структурой функций  $p_k(t)$  и  $p_k[x, t]$ , можно ввести плотность распределения изображений  $p[x, t] = \sum_{k=1}^l p_k(t) p_k[x, t]$ . Тогда средний риск (16) запишется как

$$W = \int_{t_0}^{t_1} \sum_{k=1}^l \int_{X^{(k)}(t)} R_k[x - \tau^{(k)}(t)] p[x, t] dx dt. \tag{17}$$

Переход от (16) к (17) возможен в силу сделанного предположения о непересекаемости множеств  $X^{(k)}(t)$ . При формировании общей функции потерь в виде

$$Q[x, \tau, t] = \sum_{k=1}^l R_k[x - \tau^{(k)}(t)] I^{(k)}[x, t],$$

где  $\tau(t) = (\tau^{(1)}(t), \dots, \tau^{(l)}(t))$  – набор векторов  $\tau^{(k)}(t), I^{(k)}[x, t]$  – индикатор (характеристическая функция) множества  $X^{(k)}(t), k = \overline{1, l}, t \in [t_0, t_1]$ , *общий средний риск* можно записать следующим образом:

$$W(\tau) = \int_{t_0}^{t_1} \int_X \sum_{k=1}^l R_k[x - \tau^{(k)}(t)] I^{(k)}[x, t] p[x, t] dx dt = \int_{t_0}^{t_1} M_x Q[x, \tau, t] dt, \tag{18}$$

где  $M_x Q[x, \tau, t]$  – среднее по  $x$  функции  $Q[x, \tau, t]$ . Задача, таким образом, ставится так: требуется с учётом введённых функций потерь и общей функции потерь произвести разбиение множества изображений на классы путём минимизации соответствующего функционала общего среднего риска (18), добываясь тем самым построения разделяющих функций, с помощью которых решается поставленная задача разбиения пространства признаков на классы. В этом, собственно, и заключается основная решаемая задача в рамках концепции минимизации функционала (18) в задаче о самообучении [1, 2].

В работе [1. с. 174] *особыми точками (центрами)* множеств названы параметры с экстремальными значениями. Множество разделяющих поверхностей параметризуют, что приводит к вариационной задаче для нахождения экстремальных значений параметров разделяющих поверхностей и отвечающих им центров (особых точек) классов. При варьировании обозначенных параметров получим уравнения, решив которые, найдём экстремальные параметры. В случае когда плотность  $p[x, t]$  распределения изображений не известна, следует оценивать параметры с использованием тренировочного множества в виде некоторой обучающей выборки [1].

Теорема ниже помогает определить условия, которым должны удовлетворять экстремальные параметры задачи. Пользуясь стандартными приёмами вариационного исчисления [7, 8], эти условия можно найти из равенства нулю вариации функционала (18). Особенностью здесь является то, что варьированию подлежит не только подынтегральное выражение, но и меняющиеся пределы интегрирования.

**5. Основные результаты.** Будем считать, что промежутки времени  $[t_0, t_1]$  не фиксирован и точка  $t_1$ , где  $t_1 > t_0$ , может принимать произвольные переменные значения, т.е. величина  $t_1$  варьируема. Через  $o(z)$  обозначаем величину, для которой  $\lim_{|z| \rightarrow 0} (o(z)/|z|) = 0$ .

**Теорема 1.** Пусть  $R: \mathbb{R}_u^q \times \mathbb{R}_x^q \times \mathbb{R}_t^1 \rightarrow \mathbb{R}$  – непрерывная по всем переменным функция, гладкая по переменной  $u$ , финитная по переменным  $x$  и  $t$ ;  $g: \mathbb{R}_x^q \times \mathbb{R}_t^1 \rightarrow \mathbb{R}^q$  – непрерывная вектор-функция;  $f$  – гладкая вещественная функция, задающая множество  $S_+(t) = \{x \in \mathbb{R}^q: f(x, t) > 0\}$ . Тогда вариация функционала

$$W = W(t_1, g) = \int_{t_0}^{t_1} \int_{S_+(t)} R(g(x, t), x, t) dx dt$$

имеет вид

$$\delta W = \int_{S_+(t_1)} R(g(x, t_1), x, t_1) dx \delta t_1 + \int_{t_0}^{t_1} \int_{S_+(t)} \nabla_u R(g(x, t), x, t) \cdot \delta g(x, t) dx dt,$$

где  $\nabla_u R(u, x, t)$  – градиент функции  $R$  по переменной  $u$ , а знак “ $\cdot$ ” обозначает скалярное произведение.

**Доказательство.** Запишем приращение функционала

$$\begin{aligned} \Delta W &= W(t_1 + \delta t_1, g + \delta g) - W(t_1, g) = \\ &= \int_{t_0}^{t_1 + \delta t_1} \int_{S_+(t)} R(g(x, t) + \delta g(x, t), x, t) dx dt - \int_{t_0}^{t_1} \int_{S_+(t)} R(g(x, t), x, t) dx dt = \\ &= \int_{t_1}^{t_1 + \delta t_1} \int_{S_+(t)} R(g(x, t) + \delta g(x, t), x, t) dx dt + \int_{t_0}^{t_1} \int_{S_+(t)} (R(g(x, t) + \delta g(x, t), x, t) - R(g(x, t), x, t)) dx dt. \end{aligned} \quad (19)$$

Первый интеграл в правой части формулы (19) представим в силу теоремы о среднем в виде

$$\int_{t_1}^{t_1+\delta t_1} \int_{S_+(t)} R(g(x,t) + \delta g(x,t), x, t) dx dt = \int_{S_+(t)} R(g(x,t) + \delta g(x,t), x, t) dx \Big|_{t=t_1+\theta\delta t_1} \delta t_1, \quad (20)$$

где  $\theta \in (0, 1)$ . Так как функция  $R$  по условию финитна по  $x$  и  $t$ , то найдётся такой компакт  $K$ , что  $R(u, x, t) = 0$  при  $x \notin K$ . Тогда

$$\int_{S_+(t)} R(g(x,t) + \delta g(x,t), x, t) dx \Big|_{t=t_1+\theta\delta t_1} = \int_{S_+(t) \cap K} R(g(x,t) + \delta g(x,t), x, t) dx \Big|_{t=t_1+\theta\delta t_1}.$$

Учитывая гладкость функции  $f$  и равномерную непрерывность функций  $R$  и  $g$ , можно записать равенство

$$\int_{S_+(t) \cap K} R(g(x,t) + \delta g(x,t), x, t) dx \Big|_{t=t_1+\theta\delta t_1} = \int_{S_+(t_1) \cap K} R(g(x, t_1), x, t_1) dx + \sigma,$$

где  $\sigma \rightarrow 0$  при  $\delta t_1, \delta g \rightarrow 0$ . Поэтому, принимая во внимание формулу (20), имеем

$$\int_{t_1}^{t_1+\delta t_1} \int_{S_+(t)} R(g(x,t) + \delta g(x,t), x, t) dx dt = \int_{S_+(t_1)} R(g(x, t_1), x, t_1) dx \cdot \delta t_1 + \sigma \cdot \delta t_1. \quad (21)$$

Рассмотрим подынтегральную функцию второго слагаемого правой части формулы (19). Пользуясь формулой Тейлора, получаем

$$\int_{S_+(t)} (R(g(x,t) + \delta g(x,t), x, t) - R(g(x,t), x, t)) dx = \int_{S_+(t)} (\nabla_u R(g(x,t), x, t), \delta g(x,t)) dx + o(\delta g).$$

Следовательно,

$$\begin{aligned} & \int_{t_0}^{t_1} \int_{S_+(t)} (R(g(x,t) + \delta g(x,t), x, t) - R(g(x,t), x, t)) dx dt = \\ & = \int_{t_0}^{t_1} \int_{S_+(t)} (\nabla_u R(g(x,t), x, t), \delta g(x,t)) dx dt + o(\delta g). \end{aligned} \quad (22)$$

Сопоставив формулы (19), (21) и (22), имеем

$$\Delta W = \int_{S_+(t_1)} R(g(x, t_1), x, t_1) dx \cdot \delta t_1 + \sigma \cdot \delta t_1 + \int_{t_0}^{t_1} \int_{S_+(t)} (\nabla_u R(g(x,t), x, t), \delta g(x,t)) dx dt + o(\delta g).$$

Главная часть этого выражения, линейная по  $\delta t_1$  и  $\delta g$ , даёт требуемую формулу для вариации  $\delta W$ . Теорема доказана.

Сформулируем следствие из этой теоремы для фиксированного промежутка времени  $[t_0, t_1]$  и специального вида подынтегральной функции.

**Следствие.** Пусть  $f$  – гладкая вещественная функция, задающая множество  $S_+(t) = \{x \in \mathbb{R}^q: f(x, t) > 0\}$ ;  $R: \mathbb{R}^q \rightarrow \mathbb{R}$  – гладкая функция;  $\tau: [t_0, t_1] \rightarrow \mathbb{R}^q$  – непрерывная вектор-функция;  $p: \mathbb{R}_x^q \times \mathbb{R}_t^1 \rightarrow \mathbb{R}$  – непрерывная финитная функция. Тогда вариация функционала

$$W = W(\tau) = \int_{t_0}^{t_1} \int_{S_+(t)} R(x - \tau(t))p(x, t) dx dt$$

имеет вид

$$\delta W = - \int_{t_0}^{t_1} \int_{S_+(t)} (\nabla R(x - \tau(t)), \delta \tau(t))p(x, t) dx dt.$$

**Доказательство.** Для доказательства достаточно в теореме 1 положить  $\delta t_1 = 0$ ,  $g(x, t) = x - \tau(t)$ , а в качестве  $R(u, x, t)$  взять функцию  $R(u)p(x, t)$ .

В следующей теореме полагаем, что промежуток времени  $[t_0, t_1]$  фиксирован и полностью определён.

**Теорема 2.** Пусть  $R: \mathbb{R}_x^q \times \mathbb{R}_t^1 \rightarrow \mathbb{R}$  – гладкая финитная функция;  $f$  – гладкая вещественная функция, задающая поверхность  $S(t) = \{x \in \mathbb{R}^q: f(x, t) = 0\}$ ;  $\nabla_x f(x, t) \neq 0$  при  $x \in S(t)$ . Положим

$$S_+(t) = \{x \in \mathbb{R}^q: f(x, t) > 0\}.$$

Тогда вариация функционала

$$W = W(f) = \int_{t_0}^{t_1} \int_{S_+(t)} R(x, t) dx dt$$

имеет вид

$$\delta W = \int_{t_0}^{t_1} \int_{S(t)} \frac{\delta f(x, t)}{\|\nabla_x f(x, t)\|} R(x, t) ds dt.$$

**Доказательство.** (См. также работы [6, 9, 10].) Положим

$$S'_+(t) = \{x: f(x, t) + \delta f(x, t) > 0\}.$$

Запишем приращение функционала  $W$ :

$$\begin{aligned} \Delta W &= W(f + \delta f) - W(f) = \int_{t_0}^{t_1} \int_{S'_+(t)} R(x, t) dx dt - \int_{t_0}^{t_1} \int_{S_+(t)} R(x, t) dx dt = \\ &= \int_{t_0}^{t_1} \left( \int_{S'_+(t)} R(x, t) dx - \int_{S_+(t)} R(x, t) dx \right) dt. \end{aligned}$$

Обозначив через  $V(t)$  объём, заключённый между поверхностями  $S(t)$  и

$$S'(t) = \{x \in \mathbb{R}^q: f(x, t) + \delta f(x, t) = 0\},$$

получим

$$\Delta W = \int_{t_0}^{t_1} \int_{V(t)} R(x, t) dx dt. \tag{23}$$

Взяв произвольную точку  $x \in S(t)$ , обозначим через  $x'$  такую точку на  $S'(t)$ , что

$$x' = x + \alpha \frac{\nabla_x f(x, t)}{\|\nabla_x f(x, t)\|} = x + \delta x, \quad (24)$$

где  $\alpha$  – некоторая малая скалярная величина. Для интеграла по объёму в соотношении (23) имеем

$$\int_{V(t)} R(x, t) dx = \int_{S(t)} \int_x^{x'} R(x, t) dz \vec{ds},$$

где  $\vec{ds}$  – ориентированный по антиградиенту  $-\nabla_x f(x, t)$  элемент площади на поверхности  $S(t)$ .

В силу того, что функция  $R(x, t)$  равномерно непрерывна, а точка  $x'$  близка к точке  $x$ , т.е. вариация  $\delta x$  мала, получим равенства

$$\int_{t_0}^{t_1} \int_{V(t)} R(x, t) dx dt = \int_{t_0}^{t_1} \int_{S(t)} R(x, t) \int_x^{x'} dz \vec{ds} dt + o(\|\delta x\|) = \int_{t_0}^{t_1} \int_{S(t)} R(x, t) (\delta x, \vec{ds}) dt + o(\|\delta x\|). \quad (25)$$

Формула (24) позволяет вычислить  $\delta x$ . В силу дифференцируемости функций  $f$  и  $\delta f$  справедливы соотношения

$$f(x + \delta x, t) = f(x, t) + (\nabla_x f(x, t), \delta x) + o(\|\delta x\|),$$

$$\delta f(x + \delta x, t) = \delta f(x, t) + (\nabla_x \delta f(x, t), \delta x) + o(\|\delta x\|).$$

Поэтому, принимая во внимание определение множества  $S'(t)$ , находим

$$0 = f(x', t) + \delta f(x', t) = f(x + \delta x, t) + \delta f(x + \delta x, t) = f(x, t) + (\nabla_x f(x, t), \delta x) + \delta f(x, t) + o(\|\delta x\|).$$

Так как  $f(x, t) = 0$  при  $x \in S(t)$ , то

$$0 = (\nabla_x f(x, t), \delta x) + \delta f(x, t) + o(\|\delta x\|).$$

Отсюда с помощью формулы (24) найдём

$$\delta x = -\frac{\delta f(x, t) \cdot \nabla_x f(x, t)}{\|\nabla_x f(x, t)\|^2} + \frac{\nabla_x f(x, t)}{\|\nabla_x f(x, t)\|^2} o(\|\delta x\|). \quad (26)$$

Так как функция  $f$  гладкая и  $\nabla_x f(x, t) \neq 0$  при  $x \in S(t)$ , то величина  $\|\nabla_x f(x, t)\|^{-1}$  ограничена некоторым числом  $C$  на множестве  $\{(x, t) : t \in [t_0, t_1], x \in S(t) \cap K\}$ , где  $K$  – компакт, такой что  $R(x, t) = 0$  при  $x \notin K$ . Следовательно,

$$\|\delta x\| \leq C |\delta f(x, t)| + o(\|\delta x\|) \leq C \max_{\substack{x \in K \\ t \in [t_0, t_1]}} |\delta f(x, t)| + o(\|\delta x\|).$$

Учитывая, что  $\delta x$  стремится к нулю вместе с величиной  $\delta f$ , при  $\delta f \rightarrow 0$  получаем

$$o(\|\delta x\|) = o\left(\max_{\substack{x \in K \\ t \in [t_0, t_1]}} |\delta f(x, t)|\right).$$

Таким образом, из формулы (26) следует

$$\delta x = -\frac{\delta f(x, t) \cdot \nabla_x f(x, t)}{\|\nabla_x f(x, t)\|^2} + o\left(\max_{\substack{x \in K \\ t \in [t_0, t_1]}} |\delta f(x, t)|\right).$$

Подставив последние два выражения в соотношение (25), получим

$$\int_{t_0}^{t_1} \int_{V(t)} R(x, t) dx dt = \int_{t_0}^{t_1} \int_{S(t)} \frac{\delta f(x, t)}{\|\nabla_x f(x, t)\|} R(x, t) ds dt + o\left(\max_{\substack{x \in K \\ t \in [t_0, t_1]}} |\delta f(x, t)|\right), \quad (27)$$

где  $ds = -(\vec{ds}, \nabla_x f(x, t)) / \|\nabla_x f(x, t)\|$ .

Сопоставляя соотношения (23) и (27), оставляя главную линейную по  $\delta f$  часть приращения  $\Delta W$ , получаем требуемую формулу для вариации функционала  $W$ . Теорема доказана.

Возвращаясь к функционалу общего среднего риска (18), представим вариацию  $\delta W$  в виде суммы двух вариаций:  $\delta W = \delta_1 W + \delta_2 W$ , где вариация  $\delta_1 W$  связана с изменением особой точки (центров, параметров)  $\tau^{(1)}(t), \dots, \tau^{(l)}(t)$ , а вариация  $\delta_2 W$  – с изменением множеств  $X^{(1)}(t), \dots, X^{(l)}(t), t \in [t_0, t_1]$ .

Просчитаем эти вариации. Пользуясь выражением (17) и следствием, находим

$$\delta_1 W = - \int_{t_0}^{t_1} \sum_{k=1}^l \int_{X^{(k)}(t)} (\nabla R_k[x - \tau^{(k)}(t)], \delta \tau^{(k)}(t)) p[x, t] dx dt. \quad (28)$$

Чтобы вычислить вариацию  $\delta_2 W$ , обратимся к теореме 2. Для записи функционала  $W$  в требуемой форме введём в рассмотрение дискриминантные (разделяющие) функции  $f^{(k)}$ :

$$f^{(k)}(x, t) = \begin{cases} \inf_{y \notin X^{(k)}(t)} \|x - y\|, & \text{если } x \in X^{(k)}(t), \\ - \inf_{y \in X^{(k)}(t)} \|x - y\|, & \text{если } x \notin X^{(k)}(t), \end{cases} \quad (29)$$

где  $t \in [t_0, t_1]$ ;  $f^{(k)}(x, t)$  – разность расстояний от изображения  $x$  до множества  $X \setminus X^{(k)}(t)$  и от изображения  $x$  до множества  $X^{(k)}(t)$ .

Из соотношений (29) имеем:  $f^{(k)}(x, t) \geq 0$ , если  $x \in X^{(k)}(t)$ , и  $f^{(k)}(x, t) < 0$ , если  $x \in X \setminus X^{(k)}(t)$ , а знак равенства достигается лишь на границе  $S_k(t)$  множества  $X^{(k)}(t)$ . Тогда выражение для  $W$  можно уточнить (ср. с выражением (17)):

$$W = \int_{t_0}^{t_1} \sum_{k=1}^l \int_{S_+^{(k)}(t)} R_k[x - \tau^{(k)}(t)] p[x, t] dx dt,$$

где  $S_+^{(k)}(t) = \{x : f^{(k)}(x, t) > 0\}$ ; считается также, что граница  $S_k(t) = \{x : f^{(k)}(x, t) = 0\}$  множества  $X^{(k)}(t)$  имеет нулевую меру, т.е. интегралы по множествам  $S_+^{(k)}(t)$  и  $S_{\oplus}^{(k)}(t) = \{x : f^{(k)}(x, t) \geq 0\}$  совпадают.

По теореме 2 требуется, чтобы  $\nabla_x f^{(k)}(x, t) \neq 0, x \in S_k(t)$ . С учётом формул (29) получим, что  $\nabla_x f^{(k)}(x, t)|_{x \in S_k(t)} = -n_k$ , где  $n_k$  – единичный вектор внешней нормали к границе  $S_k(t)$  множества  $X^{(k)}(t)$  в предположении, что этот вектор существует для всех точек  $S_k(t)$ . Про-варьируем разделяющие функции  $f^{(k)}(x, t)$ ; их гладкие вариации обозначим через  $\delta f^{(k)}(x, t)$ . Тогда по теореме 2 найдём

$$\delta_2 W = \int_{t_0}^{t_1} \sum_{k=1}^l \int_{S_k(t)} \delta f^{(k)}(x, t) R_k[x - \tau^{(k)}(t)] p[x, t] ds dt, \quad (30)$$

ввиду того, что  $\|\nabla_x f^{(k)}(x, t)\|_{x \in S_k(t)} = 1, k = \overline{1, l}$ . Формула (30) даёт соответствующую запись для вариации  $\delta_2 W$ , связанной с изменением множеств  $X^{(k)}(t)$ .

Если точка  $x$  принадлежит границе двух множеств  $X^{(q)}(t)$  и  $X^{(r)}(t)$ , т.е.  $x \in S_q(t) \cap S_r(t)$ , то имеет место равенство

$$\delta f^{(q)}(x, t) = -\delta f^{(r)}(x, t). \quad (31)$$

Для решения экстремальной задачи необходимо, чтобы  $\delta_1 W = 0$ ,  $\delta_2 W = 0$ . Из соотношения (28) следует необходимое условие

$$\int_{S_+^{(k)}(t)} \nabla R_k[x - \tau^{(k)}(t)]p[x, t] dx = 0. \quad (32)$$

Из равенства  $\delta_2 W = 0$  с учётом соотношений (30) и (31) получим, что для любого  $x$ , принадлежащего границе множеств  $X^{(q)}(t)$  и  $X^{(r)}(t)$ , когда  $p[x, t] \neq 0$ , должно быть выполнено также необходимое условие

$$R_q[x - \tau^{(q)}(t)] = R_r[x - \tau^{(r)}(t)]. \quad (33)$$

Таким образом, экстремальные множества  $X^{(k)}(t)$  и их особые точки (центры, параметры)  $\tau^{(k)}(t)$  должны удовлетворять условиям (32), (33). При том, что  $x(t) = x(\omega, t)$  – случайная функция (случайный процесс), требование (32) можно записать в виде

$$M\{\nabla R_k[x(\omega, t) - \tau^{(k)}(t)]I^{(k)}[x(\omega, t), t]\} = 0, \quad k = \overline{1, l},$$

где  $M\{\cdot\}$  – математическое ожидание (операция усреднения). Эти условия дают уравнения для нахождения особых точек (центров, параметров)  $\tau^{(k)}(t)$  классов множества изображений, зная которые можно затем однозначно из (33) определить множества  $X^{(k)}(t)$ ,  $t \in [t_0, t_1]$ .

**Заключение.** К задаче самообучения, или к задаче обучения “без учителя”, имеются различные подходы, объединённые общей идеей экстремизации некоторого функционала среднего риска. Эту концепцию авторы данной работы дополнили важной переменной времени, призванной придать изучаемой задаче ИИ динамический характер. В терминах классического вариационного исчисления определены необходимые условия, при которых рассматриваемый функционал общего среднего риска достигает экстремального значения в процессе самообучения на всём промежутке времени.

## СПИСОК ЛИТЕРАТУРЫ

1. Фомин В.Н. Математическая теория обучаемых опознающих систем. Л., 1976.
2. Цыпкин Я.З., Кельманс Г.К. О рекуррентных алгоритмах самообучения // Изв. АН СССР. Техн. кибернетика. 1967. № 5. С. 78–87.
3. Якубович В.А. Об одной задаче обучения целесообразному поведению // Автоматика и телемеханика. 1969. № 8. С. 119–139.
4. Якубович В.А. Метод рекуррентных целевых неравенств в теории адаптивных систем // Вопросы кибернетики. Адаптивные системы. Науч. совет по комплексной проблеме “Кибернетика” АН СССР. М., 1976. С. 32–64.
5. Фомин В.Н., Фрадков А.Л., Якубович В.А. Адаптивное управление динамическими объектами. М., 1981.
6. Tertychny-Dauri V.Yu. Adaptive Mechanic. Dordrecht; Boston; London, 2002.
7. Гельфанд И.М., Фомин С.В. Вариационное исчисление. М., 1961.
8. Эльсгольц Л.Э. Дифференциальные уравнения и вариационное исчисление. М., 1969.
9. Тertychny-Dauri В.Ю. Решение вариационных динамических задач в условиях параметрической неопределённости // Проблемы передачи информации. 2005. Т. 41. Вып. 1. С. 53–67.
10. Тertychny-Dauri В.Ю. Оптимальная стабилизация в задачах адаптивной ядерной кинетики // Дифференц. уравнения. 2006. Т. 42. № 3. С. 374–384.

Национальный исследовательский университет ИТМО,  
г. Санкт-Петербург

Поступила в редакцию 27.12.2022 г.  
После доработки 16.05.2023 г.  
Принята к публикации 19.05.2023 г.

УДК 517.977.8

# ЛИНЕЙНАЯ ЗАДАЧА ГРУППОВОГО ПРЕСЛЕДОВАНИЯ С ДРОБНЫМИ ПРОИЗВОДНЫМИ, ПРОСТЫМИ МАТРИЦАМИ И РАЗНЫМИ ВОЗМОЖНОСТЯМИ ИГРОКОВ

© 2023 г. Н. Н. Петров, А. И. Мачтакова

В конечномерном евклидовом пространстве рассмотрена задача преследования группой преследователей одного убегающего, описываемая системой уравнений с производной по Капуто порядка  $\alpha$ , где множествами допустимых управлений являются выпуклые компакты. Получены достаточные условия разрешимости задач преследования и уклонения, при исследовании которых использовался метод разрешающих функций.

DOI: 10.31857/S0374064123070075, EDN: GUVVLF

**Введение.** Работы Р. Айзекса (см., например, [1]) заложили основы теории дифференциальных игр преследования–уклонения двух лиц, которая к настоящему времени представляет собой фундаментальную содержательную теорию с различными подходами к анализу конфликтных ситуаций [2–10]. Обобщением этой теории преследования–уклонения являются задачи конфликтного взаимодействия преследователей и одного или нескольких убегающих [11–14]. Следует отметить, что для анализа игровых задач группового преследования до настоящего времени нет конструктивных подходов, позволяющих использовать теорию дифференциальных игр двух лиц. Одна из причин этого состоит в том, что объединение множеств достижимости всех преследователей и объединение всех целевых множеств представляют собой множества, не являющиеся выпуклыми и, более того, не являющиеся связными. С другой стороны, есть некоторые приложения этих игр к задачам движения транспортных средств, избежания столкновений кораблей и других. Одним из направлений современного развития теории группового преследования является поиск новых задач, к которым применимы ранее разработанные методы. В частности, в работах [15–18] рассматривались задачи преследования двух лиц, описываемые уравнениями с дробными производными, где были получены достаточные условия поимки. В статье [19] доказано существование цены игры в дифференциальной игре, описываемой уравнением с дробными производными. Задачи группового преследования и уклонения с дробными производными при условии, что все участники обладают равными возможностями, рассматривались в [20–24].

В данной работе рассматривается задача о поимке убегающего группой преследователей в дифференциальной игре с дробными производными без предположения о равенстве всех возможностей участников конфликта. Получены достаточные условия поимки и уклонения.

## 1. Постановка задачи.

**Определение 1** [25]. Пусть  $\alpha \in (0, 1)$ ,  $f: [0, \infty) \rightarrow \mathbb{R}^k$  такая функция, что  $f'$  абсолютно непрерывна на промежутке  $[0, \infty)$ . Производной по Капуто порядка  $\alpha$  функции  $f$  называется функция  $D^{(\alpha)}f$  вида

$$(D^{(\alpha)}f)(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{f'(s)}{(t-s)^\alpha} ds, \quad \text{где} \quad \Gamma(\beta) = \int_0^\infty e^{-s} s^{\beta-1} ds.$$

В пространстве  $\mathbb{R}^k$ ,  $k \geq 2$ , рассматривается дифференциальная игра  $G(n+1)$   $(n+1)$ -го лица:  $n$  преследователей  $P_1, \dots, P_n$  и убегающего  $E$ . Закон движения каждого из преследователей  $P_i$  имеет вид

$$D^{(\alpha)}x_i = a_i x_i + u_i, \quad x_i(0) = x_i^0, \quad u_i \in U_i, \quad (1)$$

а закон движения убегающего  $E$  – вид

$$D^{(\alpha)}y = ay + v, \quad y(0) = y^0, \quad v \in V. \tag{2}$$

Здесь  $i \in I = \{1, \dots, n\}$ ,  $x_i, y, u_i, v \in \mathbb{R}^k$ ,  $U_i, V$  – выпуклые компакты в  $\mathbb{R}^k$ ,  $a_i, a$  – вещественные числа. Считаем, что  $x_i^0 - y^0 \notin M_i$  для всех  $i \in I$ , где  $M_i, i \in I$ , – заданные выпуклые компакты. Если  $x_i^0 - y^0 \in M_i$  при некотором  $i$ , то в игре  $G(n + 1)$  происходит поимка в начальный момент времени.

**2. Достаточные условия поимки.** Назовём *предысторией*  $v_t(\cdot)$  управления  $v(\cdot)$  убегающего  $E$  в момент  $t, t > 0$ , сужение функции  $v$  на  $[0, t]$ .

**Определение 2.** Будем говорить, что задана *квазистратегия*  $U_i$  преследователя  $P_i$ , если определено отображение  $U_i^0(t, z^0, v_t(\cdot))$ , ставящее в соответствие начальным позициям  $z^0 = (x_1^0, \dots, x_n^0, y^0)$ , моменту  $t$  и произвольной предыстории управления  $v_t(\cdot)$  убегающего  $E$  измеримую по Лебегу функцию  $u_i(t) = U_i^0(t, z^0, v_t(\cdot))$ ,  $t \geq 0$ , со значениями в  $U_i$ .

**Определение 3.** В игре  $G(n + 1)$  происходит *поимка*, если существуют момент  $T > 0$  и квазистратегии  $U_1, \dots, U_n$  преследователей  $P_1, \dots, P_n$  такие, что для любой измеримой по Лебегу функции  $v(\cdot), v(t) \in V, t \in [0, T]$ , существуют момент  $t_0 \in [0, T]$  и номер  $l \in I$ , для которых  $x_l(t_0) - y(t_0) \in M_l$ , где  $x_i(t), i \in I$ , – решения задачи Коши (1), порождённые квазистратегиями  $U_1, \dots, U_n$ , а  $y(t)$  – решение задачи Коши (2), порождённое функцией  $v(\cdot)$ .

Введём следующие обозначения:

$$E_\rho(z, \mu) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(k\rho^{-1} + \mu)}, \quad \rho > 0, \quad z, \mu \in \mathbb{R}^1,$$

– обобщённая функция Миттаг-Лёффлера,

$$\begin{aligned} f_i(t) &= E_{1/\alpha}(a_i t^\alpha, 1), \quad f(t) = E_{1/\alpha}(a t^\alpha, 1), \\ g_i(t, \tau) &= (t - \tau)^{\alpha-1} E_{1/\alpha}(a_i(t - \tau)^\alpha, \alpha), \quad g(t, \tau) = (t - \tau)^{\alpha-1} E_{1/\alpha}(a(t - \tau)^\alpha, \alpha), \\ F_i(t) &= \int_0^t g_i(t, \tau) d\tau, \quad F(t) = \int_0^t g(t, \tau) d\tau, \end{aligned}$$

$\text{Int } A$ , со  $A$  – внутренность и выпуклая оболочка множества  $A$  соответственно. Отметим, что из теоремы 4.1.1 [26] следует, что для всех  $i \in I, t \geq 0, \tau \in [0, t]$  справедливы неравенства  $g_i(t, \tau) \geq 0, g(t, \tau) \geq 0$ .

Пусть  $\gamma_i(t, \tau), i \in I, t \geq 0, \tau \in [0, t]$ , – некоторые ограниченные, измеримые по  $(t, \tau)$ , локально суммируемые по  $\tau$  (при каждом  $t$ ) вектор-функции, которые, следуя [27], будем называть *функциями сдвига*. Зафиксируем некоторый набор функций сдвига  $\gamma(t, \tau) = \{\gamma_i(t, \tau), i \in I\}$  и обозначим

$$\xi_i(t) = f_i(t)x_i^0 - f(t)y^0 + \int_0^t \gamma_i(t, \tau) d\tau.$$

Рассмотрим многозначные отображения

$$\begin{aligned} W_i(t, \tau, v, \gamma_i) &= g_i(t, \tau)U_i - g(t, \tau)v - \gamma_i(t, \tau), \\ W_i(t, \tau, \gamma_i) &= \bigcap_{v \in V} W_i(t, \tau, v, \gamma_i). \end{aligned}$$

**Предположение 1.** Существуют функции сдвига  $\gamma(t, \tau) = \{\gamma_i(t, \tau), i \in I\}$  такие, что для всех  $i \in I, t \geq 0, 0 \leq \tau \leq t$  выполнено условие

$$0 \in W_i(t, \tau, \gamma_i).$$

**Теорема 1.** Пусть выполнено предположение 1 и существуют  $T > 0$ ,  $l \in I$  такие, что  $\xi_l(T) \in M_l$ . Тогда в игре  $G(n + 1)$  происходит поимка.

**Доказательство.** Рассмотрим многозначное отображение

$$U_l(T, \tau, v) = \{u_l \in U_l : g_l(T, \tau)u_l - g(T, \tau)v - \gamma_l(T, \tau) = 0\}.$$

В силу предположения 1  $U_l(T, \tau, v) \neq \emptyset$  для всех  $\tau \in [0, T]$ ,  $v \in V$ . Из теоремы 20.6 [28] следует, что существует измеримый селектор  $u_l^2(\tau, v) \in U_l(T, \tau, v)$ . Полагаем управление преследователя  $P_l$  равным

$$u_l(\tau) = u_l^2(\tau, v(\tau)), \quad \tau \in [0, T].$$

Управления остальных преследователей задаём произвольным образом. Из решений задачи Коши для систем (1), (2) получим (см. [29]) представление

$$x_l(t) - y(t) = \xi_l(t) + \int_0^t (g_l(t, \tau)u_l(\tau) - g(t, \tau)v(\tau) - \gamma_l(t, \tau)) d\tau,$$

поэтому  $x_l(T) - y(T) = \xi_l(T) \in M_l$ . Теорема доказана.

**Следствие 1.** Пусть существует номер  $l \in I$  такой, что:

- 1)  $0 \in g_l(t, \tau)U_l - g(t, \tau)v$  для всех  $v \in V$ ,  $t \geq 0$ ,  $\tau \in [0, t]$ ;
- 2)  $a_l < 0$ ,  $a < 0$ ;
- 3)  $0 \in \text{Int } M_l$ .

Тогда в игре  $G(n + 1)$  происходит поимка.

**Доказательство.** Взяв  $\gamma_l(t, \tau) = 0$  для всех  $t \geq 0$ ,  $\tau \in [0, t]$ , получим, что  $0 \in W_l(t, \tau, \gamma_l)$  для всех  $t \geq 0$ ,  $\tau \in [0, t]$ . Тогда  $\xi_l(t) = f_l(t)x_l^0 - f(t)y^0$ . Из [26, с. 12] следует, что при  $t \rightarrow +\infty$  справедливы асимптотические оценки

$$f(t) = -\frac{1}{at^\alpha \Gamma(1 - \alpha)} + O(1/t^{2\alpha}), \quad f_l(t) = -\frac{1}{a_l t^\alpha \Gamma(1 - \alpha)} + O(1/t^{2\alpha}),$$

поэтому  $\lim_{t \rightarrow +\infty} \xi_l(t) = 0$ . Следовательно, существует  $T > 0$ , для которого  $\xi_l(T) \in M_l$ . Осталось применить теорему 1. Следствие доказано.

В дальнейшем будем считать, что  $\xi_i(t) \notin M_i$  для всех  $i \in I$ ,  $t \geq 0$ . Каждому преследователю  $P_i$ ,  $i \in I$ , поставим в соответствие разрешающую функцию

$$\lambda_i(t, \tau, v) = \sup\{\lambda \geq 0 : \lambda(M_i - \xi_i(t)) \cap W_i(t, \tau, v, \gamma_i) \neq \emptyset\}.$$

**Теорема 2.** Пусть выполнено предположение 1 и существует момент  $T > 0$  такой, что

$$\inf_{v(\cdot)} \max_{i \in I} \int_0^T \lambda_i(T, \tau, v(\tau)) d\tau \geq 1.$$

Тогда в игре  $G(n + 1)$  происходит поимка.

Доказательство данной теоремы проводится аналогично доказательству теоремы 3.1.1 [11].

**Теорема 3.** Пусть  $M_i = \{0\}$  для всех  $i \in I$ , выполнено предположение 1 и существует момент  $T > 0$  такой, что

$$\inf_{v(\cdot)} \max_i \int_0^T \frac{\lambda_i^0(T, \tau, v(\tau))}{\|\xi_i(T)\|} d\tau \geq 1,$$

где

$$\lambda_i^0(t, \tau, v) = \sup\{\lambda \geq 0 : -\lambda \xi_i^0(t) \in g_i(t, \tau)U_i - g(t, \tau)v - \gamma_i(t, \tau)\}, \quad \xi_i^0(t) = \xi_i(t) / \|\xi_i(t)\|.$$

Тогда в игре  $G(n + 1)$  происходит поимка.

**Доказательство.** Пусть  $v(\cdot)$  – произвольная допустимая функция. Определим функции

$$h_i(t, v(\cdot)) = 1 - \int_0^t \frac{\lambda_i^0(T, \tau, v(\tau))}{\|\xi_i(T)\|} d\tau,$$

множества и моменты времени

$$T_i(v(\cdot)) = \{\tau \in [0, T] : h_i(\tau, v(\cdot)) = 0\}, \quad t_i^*(v(\cdot)) = \begin{cases} \inf\{\tau : \tau \in T_i(v(\cdot))\}, & \text{если } T_i(v(\cdot)) \neq \emptyset, \\ +\infty, & \text{если } T_i(v(\cdot)) = \emptyset. \end{cases}$$

Из условия теоремы следует, что существует номер  $l \in I$ , для которого  $t_l^*(v(\cdot)) \leq T$ . Рассмотрим многозначные отображения

$$U_i^1(\tau, v) = \{u_i \in U_i : g_i(T, \tau)u_i - g(T, \tau)v - \gamma_i(T, \tau) = -\lambda_i^0(T, \tau, v)\xi_i^0(T)\},$$

$$U_i^2(\tau, v) = \{u_i \in U_i : g_i(T, \tau)u_i - g(T, \tau)v - \gamma_i(T, \tau) = 0\}.$$

Из теоремы 20.6 работы [28] следует, что существуют измеримые селекторы  $u_i^1(\tau, v) \in U_i^1(\tau, v)$ ,  $u_i^2(\tau, v) \in U_i^2(\tau, v)$ . Задаём управления преследователей  $P_i$ ,  $i \in I$ , следующим образом. Если  $t_i^*(v(\cdot)) < T$ , то полагаем

$$u_i(t) = \begin{cases} u_i^1(t, v(t)), & t \in [0, t_i^*(v(\cdot))], \\ u_i^2(t, v(t)), & t \in (t_i^*(v(\cdot)), T]. \end{cases}$$

Если  $t_i^*(v(\cdot)) \geq T$ , то

$$u_i(t) = u_i^1(t, v(t)), \quad t \in [0, T].$$

Из решений задачи Коши для систем (1), (2) получим (см. [29]) представление

$$x_i(t) - y(t) = \xi_i(t) + \int_0^t (g_i(t, \tau)u_i(\tau) - g(t, \tau)v(\tau) - \gamma_i(t, \tau)) d\tau,$$

поэтому

$$\begin{aligned} x_i(T) - y(T) &= \xi_i(T) - \int_0^{t_i^*(v(\cdot))} \lambda_i^0(T, \tau, v(\tau))\xi_i^0(T) d\tau = \\ &= \xi_i^0(T)\|\xi_i(T)\| \left( 1 - \int_0^{t_i^*(v(\cdot))} \frac{\lambda_i^0(T, \tau, v(\tau))}{\|\xi_i(T)\|} d\tau \right) = 0. \end{aligned}$$

Теорема доказана.

Рассмотрим далее подробнее ситуацию, в которой для всех  $i \in I$

$$M_i = \{0\}, \quad U_i = \{u_i : \|u_i - b_i\| \leq R_i\}, \quad V = \{v : \|v - b\| \leq R\}, \tag{3}$$

где  $b, b_i \in R^k$ ,  $i \in I$ ,  $R_i$ ,  $i \in I$ ,  $R$  – неотрицательные вещественные числа, и в качестве нормы рассматривается евклидова норма.

**Предположение 2.**  $R_i g_i(t, \tau) \geq R g(t, \tau)$  для всех  $i \in I$ ,  $t \geq 0$ ,  $\tau \in [0, t]$ .

Из данного предположения следует, что предположение 1 будет выполнено, если в качестве  $\gamma_i(t, \tau)$  взять  $\gamma_i(t, \tau) = g_i(t, \tau)b_i - g(t, \tau)b$ . Тогда

$$\xi_i(t) = f_i(t)x_i^0 - f(t)y^0 + F_i(t)b_i - F(t)b.$$

Пусть далее  $\delta(t) = \min_{\|p\|=1} \max_i (p, \xi_i^0(t))$ , где  $(a, b)$  – скалярное произведение векторов  $a$  и  $b$ .

**Предположение 3.** Существует момент  $T > 0$  такой, что:

- а)  $0 \in \text{Int co } \{\xi_i(T), i \in I\}$ ;
- б) справедливо неравенство

$$\sum_{i \in I} \|\xi_i(T)\| \leq 2\delta(T)RF(T). \tag{4}$$

**Лемма.** Пусть  $M_i, U_i, V$  определены соотношением (3) и выполнены предположения 2, 3. Тогда для любой допустимой функции  $v(\cdot)$  имеет место неравенство

$$\sum_{i \in I} \left( \|\xi_i(T)\| - \int_0^T \lambda_i^0(T, \tau, v(\tau)) d\tau \right) \leq 0.$$

**Доказательство.** Из определения функций  $\lambda_i^0$  следует, что

$$\lambda_i^0(T, \tau, v) = g(T, \tau)(v - b, \xi_i^0(T)) + \sqrt{g^2(T, \tau)(v - b, \xi_i^0(T))^2 + R_i^2 g_i^2(T, \tau) - \|v - b\|^2 g^2(T, \tau)}.$$

Обозначим  $\hat{v} = v - b$ . Тогда  $\|\hat{v}\| \leq R$ . Из предположения 2 следует, что для всех  $t \in [0, T]$ ,  $i \in I, v \in V$  справедливы неравенства

$$\lambda_i^0(T, t, v) \geq g(T, t)((\hat{v}, \xi_i^0(T)) + |(\hat{v}, \xi_i^0(T))|). \tag{5}$$

Функции  $\lambda_i^0(T, t, v)$  при каждом фиксированном  $t$  являются вогнутыми по  $v$  [12], поэтому существует  $\hat{v}_0, \|\hat{v}_0\| = 1$ , для которого

$$\sum_{i \in I} \lambda_i^0(T, t, \hat{v}) \geq \min_{\|\hat{v}\| \leq R} \sum_{i \in I} \lambda_i^0(T, t, \hat{v}) = \sum_{i \in I} \lambda_i^0(T, t, R\hat{v}_0).$$

Из п. а) предположения 3, работы [30] и определения  $\delta(T)$  следует, что для всякого  $v, \|v\| = 1$ , найдётся номер  $l \in I$ , для которого  $(v, \xi_l^0(T)) \geq \delta(T) > 0$ . Следовательно, из (5) получаем

$$\lambda_l^0(T, t, R\hat{v}_0) \geq 2g(T, t)R(\hat{v}_0, \xi_l^0(T)) \geq 2R\delta(T)g(T, t).$$

Тогда для для каждых  $t \in [0, T]$  и  $v \in V$  справедливы неравенства

$$\sum_{i \in I} \lambda_i^0(T, t, v) \geq \sum_{i \in I} \lambda_i^0(T, t, R\hat{v}_0) \geq \lambda_l^0(T, t, R\hat{v}_0) \geq 2Rg(T, t)\delta(T).$$

Далее имеем

$$\begin{aligned} \sum_{i \in I} \|\xi_i(T)\| - \int_0^T \sum_{i \in I} \lambda_i^0(T, t, v(t)) dt &\leq \sum_{i \in I} \|\xi_i(T)\| - \int_0^T 2\delta(T)Rg(T, t) dt = \\ &= \sum_{i \in I} \|\xi_i(T)\| - 2R\delta(T)F(T) \leq 0. \end{aligned}$$

Лемма доказана.

**Следствие 2.** Пусть выполнены условия леммы 1. Тогда для любой допустимой функции  $v(\cdot)$  найдётся номер  $l \in I$ , для которого имеет место неравенство

$$\|\xi_l(T)\| - \int_0^T \lambda_l^0(T, t, v(t)) dt \leq 0.$$

**Теорема 4.** Пусть  $M_i, U_i, V$  определены соотношением (3) и выполнены предположения 2, 3. Тогда в игре  $G(n + 1)$  происходит поимка.

**Доказательство.** Пусть  $v(\cdot)$  – допустимая функция. В силу следствия 1 существует номер  $l \in I$ , для которого справедливо соотношение

$$\|\xi_l(T)\| - \int_0^T \lambda_l^0(T, t, v(t)) dt \leq 0.$$

Осталось применить теорему 3. Теорема доказана.

**Следствие 3** [20]. Пусть  $a_i = a \leq 0, b_i = b = 0, R_i = R > 0, M_i = \{0\}$  для всех  $i \in I$  и

$$y^0 \in \text{Int co} \{x_1^0, \dots, x_n^0\}.$$

Тогда в игре  $G(n + 1)$  происходит поимка.

**Доказательство.** В данном случае  $f_i(t) = f(t), F_i(t) = F(t) = t^\alpha E_{1/\alpha}(at^\alpha, \alpha + 1)$  для всех  $i \in I, t \geq 0$ . Взяв  $\gamma_i(t, \tau) = 0$ , получаем, что  $\xi_i(t) = f(t)(x_i^0 - y^0)$ . Поэтому для всех  $T > 0$

$$\sum_{i \in I} \|\xi_i(T)\| = |f(T)| \sum_{i \in I} \|\xi_i(0)\| = f(T) \sum_{i \in I} \|\xi_i(0)\|,$$

$\delta(T) = \delta(0) > 0$  в силу [30]. Значит неравенство (4) можно представить в виде

$$\sum_{i \in I} \|\xi_i(0)\| \leq 2\delta(0) \frac{RF(T)}{f(T)}. \tag{6}$$

Если  $a < 0$ , то при  $t \rightarrow +\infty$  справедливы следующие асимптотические оценки [26, с. 12]:

$$f(t) = -\frac{1}{at^\alpha \Gamma(1 - \alpha)} + O(1/t^{2\alpha}), \quad F(t) = -\frac{1}{a} + O(1/t^\alpha).$$

Таким образом,

$$\frac{F(T)}{f(T)} = T^\alpha \Gamma(1 - \alpha) + O(T^\alpha),$$

и поэтому неравенство (6) будет выполнено автоматически при достаточно большом  $T$ .

Если  $a = 0$ , то  $f(t) = 1, F(t) = t^\alpha / \Gamma(1 + \alpha)$ . Следовательно,  $F(T)/f(T) = T^\alpha / \Gamma(1 + \alpha)$ , и поэтому неравенство (6) будет выполнено автоматически при достаточно большом  $T$ . Следствие доказано.

**Замечание.** Пусть  $a \leq a_i < 0$  для всех  $i \in I$ . Тогда в силу [31] для всех  $t \in [0, +\infty), i \in I$  справедливы неравенства  $f(t) \geq f_i(t)$ . Так как ввиду [26, с. 12]

$$\lim_{t \rightarrow +\infty} \frac{F(t)}{F_i(t)} = \frac{a_i}{a} > 0,$$

то для  $\gamma_i = \inf_{t \geq 0} (F(t)/F_i(t))$  выполнено  $\gamma_i > 0$ . Выбирая  $R_i, R$  так, чтобы выполнялись неравенства  $R_i/R \leq \gamma_i$ , получаем, что условия  $RF(t) \geq R_i F_i(t), f_i(t) \geq f(t)$  теоремы 4 будут выполнены.

**3. Достаточные условия уклонения.** В данном пункте считается, что преследователи  $P_i, i \in I$ , могут использовать любую информацию при выборе своих управлений  $u_i(\cdot), i \in I$ , для которых существует решение задачи Коши системы (1), определённое на  $[0, +\infty)$ .

**Определение 4.** В игре  $G(n + 1)$  происходит уклонение от встречи, если существует программное управление  $v(\cdot)$  убегающего  $E$  такое, что для любых траекторий  $x_i(t), i \in I$ ,

преследователей  $P_i$ ,  $i \in I$ , удовлетворяющих системе (1), и для любого  $t > 0$  имеет место  $x_i(t) \neq y(t)$ , где  $y(t)$  – решение задачи Коши (2), порождённое управлением  $v(\cdot)$ .

**Теорема 5.** Пусть  $M_i$ ,  $U_i$ ,  $V$  определены соотношением (3) и существует вектор  $p \in R^k$ ,  $\|p\| = 1$ , такой, что:

1) для всех  $i \in I$ ,  $t \geq 0$  справедливы неравенства

$$\|b + Rp\|F(t) \geq (\|b_i\| + R_i)F_i(t), \quad f(t) \geq f_i(t);$$

2)  $(b + Rp, y^0) \geq 0$ ,  $(b + Rp, y^0) \geq \max_{i \in I}(b + Rp, x_i^0)$ .

Тогда в игре  $G(n + 1)$  происходит уклонение от встречи.

**Доказательство.** Задаём управление убегающего  $E$  следующим образом. Полагаем  $v(t) = b + Rp$  для всех  $t \in [0, +\infty)$ . Пусть  $u_i(t)$ ,  $i \in I$ , – произвольные управления преследователей  $P_i$ ,  $i \in I$ . Определим функции  $\hat{u}_i(t) = u_i(t) - b_i$ ,

$$\tilde{u}_i(t) = \frac{1}{R_i F(t)} \int_0^t g_i(t, s) \hat{u}_i(s) ds, \quad \omega_i(t) = f(t)y^0 - f_i(t)x_i^0.$$

Отметим, что  $F_i(t) \neq 0$  для всех  $t > 0$ . Тогда для всех  $t \geq 0$ ,  $i \in I$  выполняются соотношения

$$\|\hat{u}_i(t)\| \leq R_i, \quad \|\tilde{u}_i(t)\| \leq 1,$$

$$y(t) - x_i(t) = \omega_i(t) + (b + Rp)F(t) - (b_i + R_i \tilde{u}_i(t))F_i(t).$$

Докажем, что  $(b + Rp, \omega_i(t)) \geq 0$  для всех  $i \in I$ ,  $t \geq 0$ . Действительно,

$$(b + Rp, y^0) \geq (b + Rp, x_i^0)$$

для всех  $i \in I$ . Так как  $f_i(t) \geq 0$ ,  $f(t) \geq 0$  и  $f(t) \geq f_i(t)$  для всех  $i \in I$ ,  $t \geq 0$  (см. [31]), то

$$f_i(t)(b + Rp, x_i^0) \leq f_i(t)(b + Rp, y^0) \leq f(t)(b + Rp, y^0) \quad \text{или} \quad (b + Rp, \omega_i(t)) \geq 0.$$

Поэтому для всех  $i \in I$ ,  $t \geq 0$  справедливы неравенства

$$\begin{aligned} \|y(t) - x_i(t)\| &\geq \|\omega_i(t) + (b + Rp)F(t)\| - \|(b_i + R_i \tilde{u}_i(t))F_i(t)\| = \\ &= \sqrt{\|\omega_i(t)\|^2 + 2(b + Rp, \omega_i(t))F(t) + \|b + Rp\|^2 F^2(t)} - F_i(t)(\|b_i\| + R_i) > \\ &> \|b + Rp\|F(t) - (\|b_i\| + R_i)F_i(t) \geq 0. \end{aligned}$$

Теорема доказана.

**Следствие 4.** Пусть  $M_i = \{0\}$ ,  $b_i = b = 0$  для всех  $i \in I$ . Кроме того, для всех  $t \geq 0$ ,  $i \in I$  справедливы неравенства  $RF(t) \geq R_i F_i(t)$ ,  $f(t) \geq f_i(t)$  и существует вектор  $p \in R^k$ ,  $\|p\| = 1$ , такой, что  $(p, y^0) \geq 0$ ,  $(p, y^0) \geq \max_i(p, x_i^0)$ . Тогда в игре  $G(n + 1)$  происходит уклонение от встречи.

**Следствие 5.** Пусть  $a_i = a$ ,  $R_i = R$ ,  $M_i = \{0\}$ ,  $b_i = b = 0$  для всех  $i \in I$  и

$$y^0 \notin \text{Int co} \{x_1^0, \dots, x_n^0\}.$$

Тогда в игре  $G(n + 1)$  происходит уклонение от встречи.

**Доказательство.** Из условия следствия следует, что  $y^0$  и  $\text{co} \{x_1^0, \dots, x_n^0\}$  отделимы. Поэтому существует вектор  $p \in R^k$ ,  $\|p\| = 1$ , такой, что для всех  $i \in I$   $(p, y^0) \geq (p, x_i^0)$ . Задаём управление убегающего  $E$ , полагая  $v(t) = Rp$  для всех  $t \geq 0$ . Тогда

$$\omega_i(t) = f(t)(y^0 - x_i^0), \quad y(t) - x_i(t) = \omega_i(t) + RpF(t) - \tilde{u}_i(t)RF(t),$$

поэтому

$$\begin{aligned} \|y(t) - x_i(t)\| &\geq \|\omega_i(t) + RpF(t)\| - RF(t) = \\ &= \sqrt{\|\omega_i(t)\|^2 + 2R(p, y^0 - x_i^0)f(t)F(t) + R^2F^2(t)} - RF(t) > 0 \end{aligned}$$

для всех  $t > 0$ . Следствие доказано.

**4. Примеры.**

**Пример 1.** В пространстве  $\mathbb{R}^k$  рассматривается игра  $G(3)$ , описываемая системой

$$D^{(\alpha)}x_i = u_i, \quad x_1(0) = (1, 0, \dots, 0), \quad x_2(0) = (-1, 0, \dots, 0),$$

$$D^{(\alpha)}y = v, \quad y(0) = (0, 0, \dots, 0), \quad i \in I = \{1, 2\},$$

причём  $u_i, v \in V = \{(v_1, \dots, v_k) \in \mathbb{R}^k : v_j \in [-1, 1], j = \overline{1, k}\}, M_i = \{0\}$ .

Тогда предположение 1 будет выполнено с нулевыми функциями сдвига. Далее имеем

$$f_i(t) = f(t) = 1, \quad g_i(t, s) = g(t, s) = (t - s)^{\alpha-1} / \Gamma(\alpha), \quad \xi_i(t) = x_i^0 - y^0, \quad \lambda_i(t, \tau, v) = g(t, \tau) \bar{\lambda}_i(v),$$

где  $\bar{\lambda}_i(v) = \sup\{\lambda \geq 0 : -\lambda \xi_i(t) \in V - v\}$ . Поэтому

$$\bar{\lambda}_1(v) = 1 + v_1, \quad \bar{\lambda}_2(v) = 1 - v_1, \quad \max\{\bar{\lambda}_1(v), \bar{\lambda}_2(v)\} \geq 1.$$

Следовательно, для любой допустимой функции  $v(\cdot)$  имеют место соотношения

$$\int_0^t (\lambda_1(t, \tau, v(\tau)) + \lambda_2(t, \tau, v(\tau))) d\tau \geq \int_0^t \max_i \lambda_i(t, \tau, v(\tau)) d\tau \geq \int_0^t g(t, \tau) d\tau = \frac{t^\alpha}{\Gamma(\alpha + 1)}.$$

Выбрав  $T$  таким, чтобы выполнялось неравенство  $T^\alpha \geq 2\Gamma(\alpha + 1)$ , получаем, что для любой допустимой функции  $v(\cdot)$  найдётся номер  $l \in \{1, 2\}$ , для которого

$$\int_0^T \lambda_l(T, \tau, v(\tau)) d\tau \geq 1.$$

Тем самым условие теоремы 2 выполнено. Поэтому в игре  $G(3)$  происходит поимка.

**Пример 2.** В пространстве  $\mathbb{R}^2$  рассматривается игра  $G(4)$ , описываемая системой

$$D^{(\alpha)}x_i = ax_i + u_i, \quad D^{(\alpha)}y = ay + v, \quad i \in I = \{1, 2, 3\},$$

причём  $M_i = \{0\}, U_i = V = \{u : \|u\| \leq R\}$ . Тогда предположение 2 выполнено. Зафиксируем момент  $T > 0$  и выберем  $x_i^0, y^0$  такими, чтобы выполнялись равенства ( $\mu > 0$ )

$$\xi_1(T) = (0, \mu), \quad \xi_2(T) = (-\sqrt{3}\mu/2, -\mu/2), \quad \xi_3(T) = (\sqrt{3}\mu/2, -\mu/2).$$

Получаем

$$0 \in \text{Int co} \{\xi_i^0(T), i \in I\}, \quad \delta(T) = \frac{1}{2}, \quad \sum_{i \in I} \|\xi_i(T)\| = 3\mu.$$

Если  $0 < \mu < RF(T)/3$ , то для игры  $G(4)$  данного примера будут выполнены все условия теоремы 4, т.е. в игре  $G(4)$  будет происходить поимка.

**Пример 3.** В пространстве  $\mathbb{R}^2$  рассматривается игра  $G(4)$ , описываемая системой

$$D^{(\alpha)}x_1 = u_1, \quad \|u_1\| \leq 1, \quad D^{(\alpha)}x_2 = u_2, \quad \|u_2\| \leq 1,$$

$$D^{(\alpha)}x_3 = u_3, \quad \|u_3\| \leq \frac{1}{4}, \quad D^{(\alpha)}y = v, \quad \|v\| \leq 1, \quad I = \{1, 2, 3\},$$

$$x_1^0 = (-\mu/2, -\sqrt{3}\mu/2), \quad x_2^0 = (-\mu/2, \sqrt{3}\mu/2), \quad x_3^0 = (\mu, 0), \quad y^0 = (0, 0), \quad \mu > 0.$$

Тогда для всех  $t \geq 0$

$$f_i(t) = f(t) = 1, \quad F(t) = \frac{t^\alpha}{\Gamma(\alpha + 1)}, \quad \sum_{i \in I} \|\xi_i(t)\| = 3\mu, \quad \delta(t) = \frac{1}{2}.$$

Зададим  $T > 0$  и выберем  $\mu > 0$  так, чтобы выполнялось неравенство  $\mu < F(T)/3$ . Предположение 3 в данном случае будет выполнено. Докажем, что в игре  $G(4)$  происходит уклонение от встречи. Возьмём управление убегающего  $E$  в виде  $v(t) = v_0 = (\sqrt{3}/2, -1/2)$  для всех  $t > 0$ . Имеем

$$y(t) = \frac{v_0 t^\alpha}{\Gamma(\alpha + 1)}, \quad x_i(t) = x_i^0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} u_i(s) ds.$$

Обозначив

$$\hat{u}_i(t) = \int_0^t (t-s)^{\alpha-1} u_i(s) ds \left( \int_0^t (t-s)^{\alpha-1} ds \right)^{-1},$$

получим

$$x_i(t) = x_i^0 + \frac{t^\alpha \hat{u}_i(t)}{\Gamma(\alpha + 1)},$$

причём  $\|\hat{u}_1(t)\| \leq 1$ ,  $\|\hat{u}_2(t)\| \leq 1$ ,  $\|\hat{u}_3(t)\| \leq 1/4$ ,

$$\|x_1(t) - y(t)\| \geq \left\| x_1^0 - \frac{v_0 t^\alpha}{\Gamma(\alpha + 1)} \right\| - \frac{t^\alpha}{\Gamma(\alpha + 1)} \geq \sqrt{\mu^2 + \left( \frac{t^\alpha}{\Gamma(\alpha + 1)} \right)^2} - \frac{t^\alpha}{\Gamma(\alpha + 1)} > 0.$$

Аналогично устанавливается, что  $\|x_2(t) - y(t)\| > 0$  для всех  $t > 0$

$$\begin{aligned} \|x_3(t) - y(t)\| &\geq \left\| x_3^0 - \frac{v_0 t^\alpha}{\Gamma(\alpha + 1)} \right\| - \frac{1}{4} \frac{t^\alpha}{\Gamma(\alpha + 1)} = \\ &= \sqrt{\mu^2 - \sqrt{3}\mu \frac{t^\alpha}{\Gamma(\alpha + 1)} + \left( \frac{t^\alpha}{\Gamma(\alpha + 1)} \right)^2} - \frac{1}{4} \frac{t^\alpha}{\Gamma(\alpha + 1)} = \\ &= \sqrt{\left( \mu - \frac{\sqrt{3}}{2} \frac{t^\alpha}{\Gamma(\alpha + 1)} \right)^2 + \left( \frac{t^\alpha}{2\Gamma(\alpha + 1)} \right)^2} - \frac{1}{4} \frac{t^\alpha}{\Gamma(\alpha + 1)} \geq \frac{1}{4} \frac{t^\alpha}{\Gamma(\alpha + 1)} > 0. \end{aligned}$$

Отметим, что в данном примере не выполнено условие  $R_3 F_3(t) \geq R F(t)$  предположения 2, а все остальные условия теоремы 4 выполнены.

**Пример 4.** Пусть  $M_i, U_i, V$  определены соотношением (3). В пространстве  $\mathbb{R}^2$  рассматривается игра  $G(4)$ , описываемая системой

$$D^{(\alpha)}x_i = u_i, \quad u_i \in U_i, \quad D^{(\alpha)}y = y + v, \quad v \in V, \quad i \in I = \{1, 2, 3\},$$

$$b_1 = b = (0, 0), \quad \|b_2\| = \|b_3\| = \frac{1}{2}R, \quad R_1 = R, \quad R_2 = R_3 = \frac{1}{2}R, \quad R > 0,$$

$$x_1^0 = (0, -1.05), \quad x_2^0 = (-1, -3), \quad x_3^0 = (1, -3), \quad y^0 = (0, -1).$$

Имеем

$$f_i(t) = 1, \quad f(t) = E_{1/\alpha}(t^\alpha, 1) \geq f_i(t), \quad F(t) = t^\alpha E_{1/\alpha}(t^\alpha, \alpha + 1) \geq F_i(t) = t^\alpha / \Gamma(\alpha + 1)$$

для всех  $i \in I$ ,  $t \geq 0$ . Возьмём  $p = (0, 1)$ . Тогда справедливы неравенства

$$\|b + Rp\|F(t) \geq (\|b_i\| + R_i)F_i(t), \quad (y^0, b + Rp) \geq \max_i(x_i^0, b + Rp).$$

Полагаем управление  $v(t)$  убегающего  $E$  равным  $v(t) = v_0 = b + Rp = (0, 1)$  для всех  $t \geq 0$ . Отметим, что управление данного вида гарантировало убегающему уклонение от встречи по теореме 5. Докажем, что преследователь  $P_1$  осуществляет поимку  $E$ , используя управление  $u_1(t) = (0, 1)$ . Имеем

$$y(t) = (0, -E_{1/\alpha}(t^\alpha, 1) + t^\alpha E_{1/\alpha}(t^\alpha, \alpha + 1)), \quad x_1(t) = (0, -1.05 + t^\alpha/\Gamma(\alpha + 1)).$$

Рассмотрим функцию

$$H(t) = -E_{1/\alpha}(t^\alpha, 1) + t^\alpha E_{1/\alpha}(t^\alpha, \alpha + 1) + 1.05 - \frac{t^\alpha}{\Gamma(\alpha + 1)}.$$

Тогда  $H(0) > 0$ . В силу [32, с. 118, формула (4)] справедливо равенство

$$E_{1/\alpha}(z, 1) = \frac{1}{\Gamma(1)} + zE_{1/\alpha}(z, \alpha + 1),$$

поэтому  $H(1) = 0.05 - 1/\Gamma(\alpha + 1) < 0$ . Следовательно, функция  $H$  имеет корень на интервале  $(0, 1)$ , и значит, в игре  $G(4)$  при таком управлении  $v$  убегающего  $E$  преследователь  $P_1$  осуществляет его поимку. Отметим, что в данном случае не выполнено условие  $(b + Rp, y^0) \geq 0$  теоремы 5, а все остальные условия данной теоремы выполнены.

**Заключение.** Получены новые достаточные условия поимки и уклонения от встречи в задаче группового преследования с дробными производными. Для решения задачи использован метод разрешающих функций.

Работа выполнена при финансовой поддержке Российского научного фонда (проект 21-71-10070).

## СПИСОК ЛИТЕРАТУРЫ

1. Айзекс Р. Дифференциальные игры. М., 1967.
2. Понтрягин Л.С. Избранные научные труды. Т. 2. М., 1988.
3. Blaquiere A., Gerard F., Leitmann G. Quantitative and Qualitative Differential Games. New York, 1969.
4. Красовский Н.Н. Игровые задачи в встрече движений. М., 1970.
5. Friedman A. Differential Games. New York, 1971.
6. Красовский Н.Н., Субботин А.И. Позиционные дифференциальные игры. М., 1974.
7. Hajek O. Pursuit Games. New York, 1975.
8. Leitmann G. Cooperative and Noncooperative Many-Player Differential Games. Vienna, 1974.
9. Петросян Л.А. Дифференциальные игры преследования. Л., 1977.
10. Субботин А.И., Ченцов А.И. Оптимизация гарантии в задачах управления. М., 1981.
11. Чикрий А.А. Конфликтно управляемые процессы. Киев, 1992.
12. Григоренко Н.Л. Математические методы управления несколькими динамическими процессами. М., 1990.
13. Благодатских А.И., Петров Н.Н. Конфликтное взаимодействие групп управляемых объектов. Ижевск, 2009.
14. Сатимов Н.Ю., Рихсиев Б.Б. Методы решения задачи уклонения от встречи в математической теории управления. Ташкент, 2000.
15. Эйдельман С.Д., Чикрий А.А. Динамические задачи сближения для уравнений дробного порядка // Укр. мат. журн. 2000. Т. 52. № 11. С. 1566–1583.
16. Чикрий А.А., Матичин И.И. Игровые задачи для линейных систем дробного порядка // Тр. Ин-та математики и механики УрО РАН. 2009. Т. 15. № 3. С. 262–278.

17. Чикрий А.А., Матичин И.И. О линейных конфликтно управляемых процессах с дробными производными // Тр. Ин-та математики и механики УрО РАН. 2011. Т. 17. № 2. С. 256–270.
18. Chikrii A.A., Matychyn I.I. Game problems for fractional-order systems // *New Trends Innanotechnology and Fractional Calculus Applications*. New York, 2010. P. 233–241.
19. Gotoynov M.I. Solution to a zero-sum differential game with fractional dynamics via approximations // *Dynamic Games and Appl.* 2020. V. 10 (2). P. 417–443.
20. Петров Н.Н. Одна задача группового преследования с дробными производными и фазовыми ограничениями // Вестн. Удмуртского ун-та. Математика. Механика. Компьют. науки. 2017. Т. 27. Вып. 1. С. 54–59.
21. Петров Н.Н. Многократная поимка в одной задаче группового преследования с дробными производными // Тр. Ин-та математики и механики УрО РАН. 2018. Т. 24. № 1. С. 156–164.
22. Петров Н.Н. К задаче группового преследования в дифференциальной игре с дробными производными, фазовыми ограничениями и простой матрицей // Дифференц. уравнения. 2019. Т. 55. № 6. С. 857–864.
23. Банников А.С. Уклонение от группы преследователей в задаче группового преследования с дробными производными и фазовыми ограничениями // Вестн. Удмуртского ун-та. Математика. Механика. Компьют. науки. 2017. Т. 27. Вып. 3. С. 309–314.
24. Petrov N.N. Multiple capture in a group pursuit problem with fractional derivatives and phase restrictions // *Mathematics*. 2021. V. 9 (11). P. 1171.
25. Caputo M. Linear model of dissipation whose  $q$  is almost frequency independent. II // *Geophys. R. Astr. Soc.* 1967. № 13. P. 529–539.
26. Попов А.Ю., Седлецкий А.М. Распределение корней функции Миттаг-Лёффлера // *Соврем. математика. Фунд. направления*. 2011. Т. 40. С. 3–171.
27. Чикрий А.А., Чикрий К.В. Структура образов многозначных отображений в игровых задачах управления движением // *Проблемы управления и информатики*. 2016. № 3. С. 65–78.
28. Половинкин Е.С. Многозначный анализ и дифференциальные включения. М., 1990.
29. Чикрий А.А., Матичин И.И. Об аналоге формулы Коши для линейных систем произвольного дробного порядка // *Доповіді НАН України*. 2007. № 1. С. 50–55.
30. Петров Н.Н. Об управляемости автономных систем // *Дифференц. уравнения*. 1968. Т. 4. № 4. С. 606–617.
31. Pollard H. The completely monotonic character of the Mittag-Leffler function  $E_a(-x)$  // *Bull. of the Amer. Math. Soc.* 1948. V. 54. № 12. P. 1115–1116.
32. Джарбашян М.М. Интегральные преобразования и представления функций в комплексной области. М., 1966.

Удмуртский государственный университет,  
г. Ижевск,  
Институт математики и механики УрО РАН,  
г. Екатеринбург

Поступила в редакцию 24.04.2022 г.  
После доработки 28.03.2023 г.  
Принята к публикации 19.05.2023 г.

УДК 519.63

## РАЗНОСТНЫЕ СХЕМЫ ДЕКОМПОЗИЦИИ НА ОСНОВЕ РАСЩЕПЛЕНИЯ РЕШЕНИЯ И ОПЕРАТОРА ЗАДАЧИ

© 2023 г. П. Н. Вабищевич

Методы декомпозиции области применяются для приближённого решения краевых задач для уравнений с частными производными на параллельных вычислительных системах. Наиболее полно специфика нестационарных задач учитывается при использовании безытерационных схем декомпозиции области. Регионально-аддитивные схемы строятся на основе различных классов схем расщепления. Выделяется новый класс схем декомпозиции области с аддитивным представлением решения на новом слое по времени, который базируется на разделении области на подобласти на основе разбиения единицы. Рассматривается пример задачи Коши для эволюционных уравнений первого порядка с положительным самосопряжённым оператором в конечномерном гильбертовом пространстве. Строятся безусловно устойчивые двух- и трёхслойные схемы расщепления для соответствующей системы уравнений.

DOI: 10.31857/S0374064123070087, EDN: GUXGWW

**Введение.** Итерационное решение многомерных стационарных краевых задач для уравнений с частными производными проводится чаще всего на основе методов декомпозиции области [1, гл. 1]. Специфика нестационарных задач наиболее полно проявляется при использовании безытерационных методов декомпозиции области [2, гл. 8; 3, гл. 9]. Схемы декомпозиции области для решения нестационарных задач можно классифицировать по способу декомпозиции области, по выбору операторов декомпозиции (обменных граничных условий) и по используемой схеме расщепления. Для дифференциальных задач естественно выделять методы декомпозиции области

$$\bar{\Omega} = \bigcup_{\alpha=1}^p \bar{\Omega}_\alpha, \quad \bar{\Omega}_\alpha = \Omega_\alpha \cup \partial\Omega_\alpha, \quad \alpha = 1, 2, \dots, p, \quad (1)$$

с налеганием подобластей ( $\Omega_{\alpha\beta} \equiv \Omega_\alpha \cap \Omega_\beta \neq \emptyset$ ) и без налегания подобластей ( $\Omega_{\alpha\beta} = \emptyset$ ).

С декомпозицией области (1) сопоставляется соответствующее аддитивное представление оператора задачи

$$\mathcal{A} = \sum_{\alpha=1}^p \mathcal{A}_\alpha. \quad (2)$$

В этом случае оператор  $\mathcal{A}_\alpha$  связывается с решением некоторой задачи в подобласти  $\Omega_\alpha$ ,  $\alpha = 1, 2, \dots, p$ . Наиболее общий подход для построения операторов декомпозиции при решении краевых задач для уравнений с частными производными базируется на использовании разбиения единицы для расчётной области. При декомпозиции (1) с отдельной подобластью  $\Omega_\alpha$  свяжем функцию  $\eta_\alpha(\mathbf{x})$ ,  $\alpha = 1, 2, \dots, p$ , такую, что

$$\begin{cases} \eta_\alpha(\mathbf{x}) > 0, & \mathbf{x} \in \Omega_\alpha, \\ \eta_\alpha(\mathbf{x}) = 0, & \mathbf{x} \notin \Omega_\alpha, \end{cases} \quad \alpha = 1, 2, \dots, p, \quad (3)$$

причём

$$\sum_{\alpha=1}^p \eta_\alpha(\mathbf{x}) = 1, \quad \mathbf{x} \in \Omega. \quad (4)$$

Пусть, например, оператор задачи  $A$  есть оператор диффузии:

$$A = -\operatorname{div} k(\mathbf{x}) \operatorname{grad}, \quad \mathbf{x} \in \Omega.$$

Тогда для операторов декомпозиции можно задать три основные конструкции:

$$A_\alpha = \eta_\alpha A, \tag{5}$$

$$A_\alpha = -\operatorname{div} k(\mathbf{x}) \eta_\alpha(\mathbf{x}) \operatorname{grad}, \tag{6}$$

$$A_\alpha = A \eta_\alpha, \quad \alpha = 1, 2, \dots, p. \tag{7}$$

Эта технология используется, начиная со статей [4] (декомпозиция (6)), [5] (декомпозиции (5)–(7)), результаты более поздних работ отражены в обзоре [6] и книге [2, гл. 8].

При решении нестационарных задач для задач с расщеплением (2) используются различные схемы расщепления. В теории аддитивных операторно-разностных схем [7, гл. IX; 8; 9, гл. 4] особо выделим случай простейшего двухкомпонентного расщепления, когда строятся безусловно устойчивые факторизованные схемы расщепления, к которым относятся классические схемы переменных направлений, схемы предиктор-корректор. Двухкомпонентные регионально-аддитивные схемы построены и исследованы в работах [4, 5, 10], в статьях [11, 12] рассмотрены задачи конвекции–диффузии.

Для вычислительной практики, в том числе и при применении методов декомпозиции области, большой интерес представляет расщепление оператора задачи на сумму трёх и более попарно некоммутативных операторов ( $p > 2$  в (2)). Стандартные [8; 9, гл. 5] схемы многокомпонентного расщепления строятся на основе понятия суммарной аппроксимации. Аддитивно-усреднённые схемы суммарной аппроксимации [9, гл. 5; 13] более явно ориентированы на параллельную организацию вычислений. Регионально-аддитивные схемы покомпонентного расщепления исследуются в работе [14]. В настоящее время получили распространение схемы полной аппроксимации для общего многокомпонентного расщепления. В связи с этим отметим регуляризованные аддитивные схемы [15], в которых условия устойчивости достигаются за счёт возмущений операторов разностной схемы. В векторных аддитивных схемах [16, 17] вместо одного уравнения решается система однотипных уравнений. Векторные регионально-аддитивные схемы исследуются в работах [18, 19].

В классических схемах расщепления [8; 9, гл. 4, 5] используется аддитивное представление оператора задачи, при этом решение эволюционных задач для отдельных более простых членов оператора обеспечивает менее затратный с вычислительной точки зрения переход на новый слой во времени. Во многих нестационарных задачах удобнее строить вычислительно приемлемые подзадачи на основе декомпозиции решения, когда более простые задачи формулируются для отдельных компонент решения, в частности, для приближённого решения мультифизических задач можно сгруппировать компоненты решения (моnofизические задачи) для соответствующей системы связанных уравнений. Такие схемы расщепления решений предложены и исследованы в [20, 21] для приближённого решения задачи Коши в конечномерном гильбертовом пространстве для эволюционных уравнений первого и второго порядков.

Идея расщепления решения для построения схем расщепления используется в простейшем варианте при построении аддитивно-усреднённых схем [9, гл. 5; 13]. В этом случае можно рассчитывать на параллельную организацию вычислений для отдельных подзадач. В векторных схемах расщепления [9, гл. 8; 16; 17; 22] осуществляется формальное усложнение задачи с переходом от одного уравнения к системе уравнений. При этом приближённое скалярное решение исходной задачи берётся в виде линейной суперпозиции приближений для отдельных компонент вектора вспомогательной задачи для системы уравнений. Целью нашего исследования является построение схем декомпозиции области на основе аддитивного расщепления как оператора задачи, так и самого приближённого решения. В обоих случаях привлекается разложение единицы для сеточной области. Для системы уравнений строятся безусловно устойчивые двух- и трёхслойные схемы первого и второго порядков аппроксимации по времени. Параллельная реализация обеспечивается выделением диагональной части операторной матрицы системы уравнений и решением задачи в отдельной подобласти. Устойчивость двух-

и трёхслойных схем расщепления исследуется с использованием общих результатов теории устойчивости (корректности) операторно-разностных схем [2, гл. 2; 4; 7, гл. VI; 22].

**1. Краевая задача для параболического уравнения.** Рассматривается модельная краевая задача для параболического уравнения второго порядка. В ограниченной области  $\Omega$  неизвестная функция  $u(\mathbf{x}, t)$  удовлетворяет уравнению

$$\frac{\partial u}{\partial t} - \sum_{\alpha=1}^m \frac{\partial}{\partial x_{\alpha}} \left( k(\mathbf{x}) \frac{\partial u}{\partial x_{\alpha}} \right) = f(\mathbf{x}, t), \quad \mathbf{x} \in \Omega, \quad 0 < t \leq T, \quad (8)$$

где  $k(\mathbf{x}) \geq \kappa > 0$ ,  $\mathbf{x} \in \Omega$ . Уравнение (8) дополним однородным граничным условием Дирихле

$$u(\mathbf{x}, t) = 0, \quad \mathbf{x} \in \partial\Omega, \quad 0 < t \leq T. \quad (9)$$

Кроме того, задаётся начальное условие

$$u(\mathbf{x}, 0) = u^0(\mathbf{x}), \quad \mathbf{x} \in \Omega. \quad (10)$$

Нестационарная задача диффузии (8)–(10) рассматривается на множестве функций  $u(\mathbf{x}, t)$ , удовлетворяющих граничному условию (9). Тогда вместо (8), (9) будем использовать дифференциально-операторное уравнение

$$\frac{du}{dt} + \mathcal{A}u = f(t), \quad 0 < t \leq T. \quad (11)$$

Рассмотрим задачу Коши для эволюционного уравнения (11) с начальным условием

$$u(0) = u^0. \quad (12)$$

Для оператора диффузии положим

$$\mathcal{A}u = - \sum_{\alpha=1}^m \frac{\partial}{\partial x_{\alpha}} \left( k(\mathbf{x}) \frac{\partial u}{\partial x_{\alpha}} \right).$$

На множестве достаточно гладких функций, удовлетворяющих (9), определим гильбертово пространство  $\mathcal{H} = \mathcal{L}_2(\Omega)$  со скалярным произведением и нормой

$$(u, v) = \int_{\Omega} u(\mathbf{x})v(\mathbf{x})d\mathbf{x}, \quad \|u\| = (u, u)^{1/2}.$$

В  $\mathcal{H}$  оператор диффузионного переноса  $\mathcal{A}$  самосопряжён и положительно определён:

$$\mathcal{A} = \mathcal{A}^* \geq \kappa\delta\mathcal{I}, \quad \delta = \delta(\Omega) > 0, \quad (13)$$

где  $\mathcal{I}$  – единичный оператор в пространстве  $\mathcal{H}$ .

Приведём для решения задачи (11)–(13) простейшую априорную оценку, которая будет для нас ориентиром при рассмотрении сеточных задач. С самосопряжённым положительно определённым оператором  $\mathcal{D}$  свяжем гильбертово пространство  $\mathcal{H}_{\mathcal{D}}$ , скалярное произведение и норма в котором определяются формулами

$$(u, v)_{\mathcal{D}} = (\mathcal{D}u, v), \quad \|u\|_{\mathcal{D}} = (u, u)_{\mathcal{D}}^{1/2}$$

соответственно. Домножим скалярно в  $\mathcal{H}$  уравнение (11) на  $\mathcal{A}u$ . С учётом (13) придём к равенству

$$\frac{1}{2} \frac{d}{dt} \|u\|_{\mathcal{A}}^2 + \|\mathcal{A}u\|^2 = (f, \mathcal{A}u). \quad (14)$$

Принимая во внимание, что

$$(f, \mathcal{A}u) \leq \|\mathcal{A}u\|^2 + \frac{1}{4}\|f\|^2,$$

из (14) получаем

$$\frac{d}{dt}\|u\|_{\mathcal{A}}^2 \leq \frac{1}{2}\|f\|^2.$$

Ввиду леммы Гронуолла приходим к искомой оценке

$$\|u(t)\|_{\mathcal{A}}^2 \leq \|u^0\|_{\mathcal{A}}^2 + \frac{1}{2} \int_0^t \|f(\theta)\|^2 d\theta, \tag{15}$$

которая выражает устойчивость решения задачи (11)–(13) по начальным данным и правой части.

**2. Стандартные разностные схемы.** Исследование аппроксимаций по пространству и времени проведём на примере краевых задач в прямоугольнике

$$\Omega = \{\mathbf{x} : \mathbf{x} = (x_1, x_2), \quad 0 < x_\alpha < l_\alpha, \quad \alpha = 1, 2\}.$$

Приближённое решение задаётся на сетке в узлах равномерной прямоугольной сетки в  $\Omega$ :

$$\bar{\omega} = \{\mathbf{x} : \mathbf{x} = (x_1, x_2), \quad x_\alpha = i_\alpha h_\alpha, \quad i_\alpha = 0, 1, \dots, N_\alpha, \quad N_\alpha h_\alpha = l_\alpha\},$$

и пусть  $\omega$  – множество внутренних узлов ( $\bar{\omega} = \omega \cup \partial\omega$ ). Для сеточных функций  $y(\mathbf{x}) = 0$ ,  $\mathbf{x} \in \partial\omega$ , определим гильбертово пространство  $H = L_2(\omega)$  со скалярным произведением и нормой

$$(y, w) = \sum_{\mathbf{x} \in \omega} y(\mathbf{x})w(\mathbf{x})h_1h_2, \quad \|y\| = (y, y)^{1/2}.$$

Считая коэффициент  $k(\mathbf{x})$  в области  $\Omega$  достаточно гладким, сеточный оператор диффузии возьмём в виде

$$\begin{aligned} Ay = & -\frac{1}{h_1^2}k(x_1+0.5h_1, x_2)(y(x_1+h_1, x_2)-y(x_1, x_2))+\frac{1}{h_1^2}k(x_1-0.5h_1, x_2)(y(x_1, x_2)-y(x_1-h_1, x_2))- \\ & -\frac{1}{h_2^2}k(x_1, x_2+0.5h_2)(y(x_1, x_2+h_2)-y(x_1, x_2))+\frac{1}{h_2^2}k(x_1, x_2-0.5h_2)(y(x_1, x_2)-y(x_1, x_2-h_2)). \end{aligned} \tag{16}$$

В  $H$  оператор  $A$  самосопряжён и положительно определён [7, гл. IV]:

$$A = A^* \geq \kappa(\delta_1 + \delta_2)I, \quad \delta_\alpha = \frac{4}{h_\alpha^2} \sin^2 \frac{\pi h_\alpha}{2l_\alpha}, \quad \alpha = 1, 2.$$

После аппроксимации по пространству от (8), (9) приходим к дифференциально-разностному уравнению

$$\frac{dy}{dt} + Ay = f(\mathbf{x}, t), \quad \mathbf{x} \in \omega, \quad 0 < t \leq T. \tag{17}$$

Принимая во внимание (10), уравнение (17) дополним начальным условием

$$y(\mathbf{x}, 0) = u^0(\mathbf{x}), \quad \mathbf{x} \in \omega. \tag{18}$$

Для решения дифференциально-разностной задачи Коши (17), (18) справедлива априорная оценка (см. (15))

$$\|y(t)\|_A^2 \leq \|u^0\|_A^2 + \frac{1}{2} \int_0^t \|f(\theta)\|^2 d\theta. \tag{19}$$

Основное внимание мы уделяем аппроксимации по времени. При построении схем декомпозиции области для задачи (17), (18) отправной точкой для нас являются обычные двухслойные схемы. Пусть  $\tau$  есть шаг равномерной сетки по времени и пусть  $y^n = y(t^n)$ ,  $t^n = n\tau$ ,  $n = 0, 1, \dots, N$ ,  $N\tau = T$ . Уравнение (17) аппроксимируем двухслойной схемой с весами

$$\frac{y^{n+1} - y^n}{\tau} + A(\sigma y^{n+1} + (1 - \sigma)y^n) = f^{n+\sigma}, \quad n = 0, 1, \dots, N - 1, \tag{20}$$

где, например,  $f^{n+\sigma} = f(\sigma t^{n+1} + (1 - \sigma)t^n)$ . Она дополняется начальным условием

$$y^0 = u^0. \tag{21}$$

Разностная схема (20), (21) имеет погрешность аппроксимации  $\mathcal{O}(\tau^2 + (\sigma - 1/2)\tau + h^2)$ , где  $h^2 = (h_1^2 + h_2^2)/2$ .

**Теорема 1.** *Разностная схема (20), (21) безусловно устойчива при  $\sigma \geq 1/2$ , причём для разностного решения справедлива оценка*

$$\|y^{n+1}\|_D^2 \leq \|y^n\|_D^2 + \frac{\tau}{2}\|f^{n+\sigma}\|^2, \quad n = 0, 1, \dots, N - 1, \tag{22}$$

где  $D = A + (\sigma - 1/2)\tau A^2$ .

**Доказательство.** Запишем разностную схему (20) в виде

$$\left( I + \left( \sigma - \frac{1}{2} \right) \tau A \right) \frac{y^{n+1} - y^n}{\tau} + A \frac{y^{n+1} + y^n}{2} = f^{n+\sigma}$$

и домножим её скалярно на  $\tau A(y^{n+1} + y^n)$ . Так как при  $\sigma \geq 1/2$  оператор  $D \geq A$ , то

$$\|y^{n+1}\|_D^2 - \|y^n\|_D^2 + \frac{\tau}{2}\|A(y^{n+1} + y^n)\|^2 = \tau(f^{n+\sigma}, A(y^{n+1} + y^n)).$$

Принимая во внимание равенство

$$(f^{n+\sigma}, A(y^{n+1} + y^n)) \leq \frac{1}{2}\|A(y^{n+1} + y^n)\|^2 + \frac{1}{2}\|f^{n+\sigma}\|^2,$$

приходим к доказываемой оценке (22). Теорема доказана.

Априорная оценка (22) решения разностной задачи (20), (21) является сеточным аналогом априорной оценки (19) для решения дифференциально-разностной задачи (17), (18) ( $D = A + \mathcal{O}(\tau)$ ).

При приближённом решении задачи (17), (18) естественно использовать трёхслойные схемы второго порядка аппроксимации по времени. Пусть

$$\frac{y^{n+1} - y^{n-1}}{2\tau} + A(\sigma y^{n+1} + (1 - 2\sigma)y^n + \sigma y^{n-1}) = f^n, \quad n = 1, 2, \dots, N - 1. \tag{23}$$

Начальные условия аппроксимируем следующим образом:

$$y^0 = u^0, \quad \frac{y^1 - y^0}{\tau} + A \frac{y^1 + y^0}{2} = f^{1/2}. \tag{24}$$

Погрешность разностной схемы (23), (24) есть  $\mathcal{O}(\tau^2 + h^2)$ .

**Теорема 2.** *Разностная схема (23), (24) безусловно устойчива при  $\sigma \geq 1/4$ , причём для разностного решения справедлива оценка*

$$\left( \sigma - \frac{1}{4} \right) \tau^2 \left\| \frac{y^{n+1} - y^n}{\tau} \right\|_A^2 + \left\| \frac{y^{n+1} + y^n}{2} \right\|_A^2 \leq$$

$$\leq \left(\sigma - \frac{1}{4}\right)\tau^2 \left\| \frac{y^n - y^{n-1}}{\tau} \right\|_A^2 + \left\| \frac{y^n + y^{n-1}}{2} \right\|_A^2 + \frac{\tau}{2} \|f^n\|^2, \quad n = 0, 1, \dots, N - 1. \quad (25)$$

**Доказательство.** Введём обозначения

$$\zeta^n = \frac{y^n + y^{n-1}}{2}, \quad \eta^n = \frac{y^n - y^{n-1}}{\tau}.$$

Принимая во внимание тождества

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

$$\sigma y^{n+1} + (1 - 2\sigma)y^n + \sigma y^{n-1} = y^n + \sigma(y^{n+1} - 2y^n + y^{n-1}),$$

перепишем (23) в виде

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

Домножим (26) скалярно в  $H$  на

$$2(\zeta^{n+1} - \zeta^n) = \tau(\eta^{n+1} + \eta^n).$$

С учётом введённых обозначений получим

$$\frac{\tau}{2} \|\eta^{n+1} + \eta^n\|^2 + \left(\sigma - \frac{1}{4}\right)\tau^2 (\|\eta^{n+1}\|_A^2 - \|\eta^n\|_A^2) + \|\zeta^{n+1}\|_A^2 - \|\zeta^n\|_A^2 = \tau(f^n, \eta^{n+1} + \eta^n).$$

Используем для правой части оценку

$$(f^n, \eta^{n+1} + \eta^n) \leq \frac{1}{2} \|f^n\|^2 + \frac{1}{2} \|\eta^{n+1} + \eta^n\|^2$$

и будем иметь

$$\left(\sigma - \frac{1}{4}\right)\tau^2 \|\eta^{n+1}\|_A^2 + \|\zeta^{n+1}\|_A^2 \leq \left(\sigma - \frac{1}{4}\right)\tau^2 \|\eta^n\|_A^2 + \|\zeta^n\|_A^2 + \frac{\tau}{2} \|f^n\|^2. \quad (27)$$

В силу введённых обозначений из (27) получим искомую послынную оценку устойчивости (25) для разностного решения. Теорема доказана.

**3. Декомпозиция области.** При обсуждении методов декомпозиции области для задач типа (8)–(10) обычно выделяют два случая: с наложением и без наложения подобластей. Методы без наложения подобластей связываются с явной формулировкой тех или иных условий на общих границах. В нашем случае самостоятельная задача для интерфейсных границ не формулируется, но в алгоритмах без наложения подобластей соответствующие обменные граничные условия можно получить.

Для методов декомпозиции области принципиальным является вопрос об обменах расчётными данными между отдельными подобластями. В этом плане ориентиром могут быть обычные явные схемы. В данном случае декомпозицию области можно связать с отдельными подмножествами сеточных узлов  $\omega_\alpha$ ,  $\alpha = 1, 2, \dots, p$ :

$$\omega = \bigcup_{\alpha=1}^p \omega_\alpha, \quad \omega_\alpha = \{\mathbf{x} \in \omega, \quad \mathbf{x} \in \Omega_\alpha\}, \quad \alpha = 1, 2, \dots, p.$$

В случае (16) (семиточечный шаблон по пространству) переход на новый слой по времени по явной схеме для нахождения приближённого решения на сетке  $\omega_\alpha$ ,  $\alpha = 1, 2, \dots, p$ , связан с использованием значений решения в приграничных узлах.



Рис. 1. Декомпозиция области.

Для рассматриваемых модельных задач в прямоугольнике простейшая декомпозиция по одной переменной на четыре подобласти представлена на рис. 1 с налеганием подобластей.

Несвязные подобласти можно рассматривать как одну подобласть и разбиение на рис. 1 можно рассматривать как разбиение на две подобласти и описывать двумя функциями:  $\eta_\alpha = \eta_\alpha(x_1)$ ,  $\alpha = 1, 2$ . При приближённом решении задачи (17), (18) можно рассматривать две возможности минимального наложения подобластей. Первая из них связана с тем, что разбиение области проводится по целым узлам – граничные узлы принадлежат нескольким подобластям (двум в нашем случае декомпозиции по одной переменной). Вторая возможность реализуется, когда граница подобластей проходит по полуцелым узлам соответствующей переменной.

Вариант с разбиением по целым узлам отображен на рис. 2. Пусть декомпозиция проводится по переменной  $x_1$ , т.е.  $\theta = x_1$ . Область разбивается по узлу  $\theta = \theta_i$ . Операторы декомпозиции строятся, например, на основе (6). С учётом (16) операторы декомпозиции ( $A_1$  в подобласти  $\Omega_1$  и  $A_2$  в подобласти  $\Omega_2$ ) принимают вид

$$\begin{aligned}
 A_1 y &= \frac{1}{h_1^2} k(x_1 - 0.5h_1, x_2)(y(x_1, x_2) - y(x_1 - h_1, x_2)) - \\
 &- \frac{1}{2h_2^2} k(x_1, x_2 + 0.5h_2)(y(x_1, x_2 + h_2) - y(x_1, x_2)) + \frac{1}{2h_2^2} k(x_1, x_2 - 0.5h_2)(y(x_1, x_2) - y(x_1, x_2 - h_2)), \\
 A_2 y &= -\frac{1}{h_1^2} k(x_1 + 0.5h_1, x_2)(y(x_1 + h_1, x_2) - y(x_1, x_2)) - \\
 &- \frac{1}{2h_2^2} k(x_1, x_2 + 0.5h_2)(y(x_1, x_2 + h_2) - y(x_1, x_2)) + \frac{1}{2h_2^2} k(x_1, x_2 - 0.5h_2)(y(x_1, x_2) - y(x_1, x_2 - h_2)), \\
 &x_1 = \theta_i.
 \end{aligned}$$

Такую декомпозицию мы можем связать с использованием граничных условий Неймана в качестве обменных граничных условий. Связь между отдельными подобластями минимальна и обеспечена обменом данных при  $\theta = \theta_i$ . Значения  $\eta_\alpha(x_1 \pm 0.5h_1, x_2)$ ,  $\eta_\alpha(x_1, x_2 \pm 0.5h_1)$ ,  $\alpha = 1, 2$ , равны 0 или 1.

Вторая возможность с декомпозицией по полуцелым узлам иллюстрируется на рис. 3. В узле  $\theta = \theta_i$  используется разностная аппроксимация с меньшим в два раза потоком.

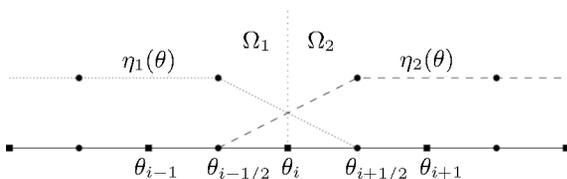


Рис. 2. Декомпозиция по целым узлам.

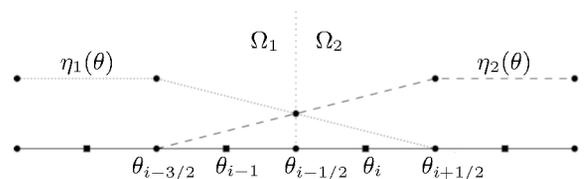


Рис. 3. Декомпозиция по полуцелым узлам.

Применительно к рассматриваемому случаю декомпозиции по переменной  $x_1$  операторы декомпозиции имеют вид

$$\begin{aligned}
 A_1 y &= \frac{1}{2h_1^2} k(x_1 - 0.5h_1, x_2)(y(x_1, x_2) - y(x_1 - h_1, x_2)) - \\
 &- \frac{1}{4h_2^2} k(x_1, x_2 + 0.5h_2)(y(x_1, x_2 + h_2) - y(x_1, x_2)) + \frac{1}{4h_2^2} k(x_1, x_2 - 0.5h_2)(y(x_1, x_2) - y(x_1, x_2 - h_2)),
 \end{aligned}$$

$$\begin{aligned}
 A_2 y = & -\frac{1}{h_1^2} k(x_1 + 0.5h_1, x_2)(y(x_1 + h_1, x_2) - y(x_1, x_2)) + \\
 & + \frac{1}{2h_1^2} k(x_1 - 0.5h_1, x_2)(y(x_1, x_2) - y(x_1 - h_1, x_2)) - \frac{3}{4h_2^2} k(x_1, x_2 + 0.5h_2)(y(x_1, x_2 + h_2) - y(x_1, x_2)) + \\
 & + \frac{3}{4h_2^2} k(x_1, x_2 - 0.5h_2)(y(x_1, x_2) - y(x_1, x_2 - h_2)), \quad x_1 = \theta_i.
 \end{aligned}$$

Для вычислений в области  $\Omega_1$  (см. рис. 3) привлекаются приграничные данные в области  $\Omega_2$  – в узле  $\theta = \theta_i$ . Тем самым при такой декомпозиции области обмены минимальны и совпадают с обменами при реализации явной схемы.

Во втором классе методов декомпозиции области выделяется множество интерфейсных узлов подобластей при решении отдельных задач внутри подобластей. В нашей модельной сеточной задаче в прямоугольнике с квадратной сеткой ( $h = h_1 = h_2$ ) расчётная сетка  $\omega$  разбивается на прямоугольные подобласти грубой сеткой, шаг которой равен  $\hat{h}$ . Границы подобластей (прямые линии) состоят из узлов расчётной сетки. Это множество внутренних граничных узлов обозначим  $\hat{\omega}$ . Фрагмент сетки показан на рис. 4. Подобласть  $\Omega_2$  представляет собой решётку, ширина отдельных рёбер этой решётки равна  $h$ . Область  $\Omega_1$  состоит из несвязанных друг с другом отдельных подобластей.

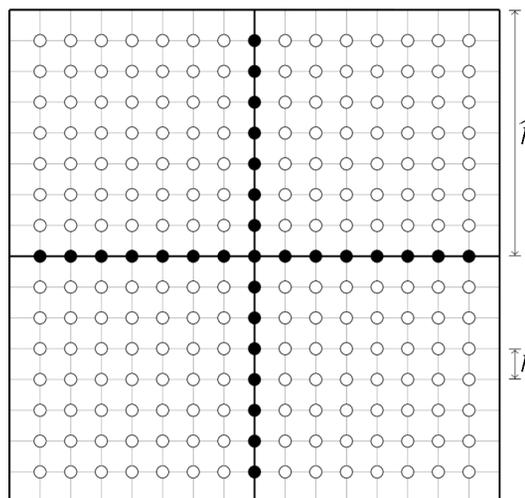


Рис. 4. Сеточная декомпозиция.

Разбиению единицы области (3), (4) сопоставим соответствующее аддитивное представление единицы для сеточных функций, заданных на множестве внутренних узлов  $\omega$ . Положим

$$\sum_{\alpha=1}^p \chi_\alpha = 1, \quad \chi_\alpha \geq 0, \quad \alpha = 1, 2, \dots, p. \tag{28}$$

Аналогично (5) операторы декомпозиции зададим в виде

$$A_\alpha = \chi_\alpha A, \quad \alpha = 1, 2, \dots, p. \tag{29}$$

В силу (28) при таком расщеплении для оператора задачи имеет место аддитивное представление

$$A = \sum_{\alpha=1}^p A_\alpha. \tag{30}$$

Расщепление (30) позволяет от уравнения (17) перейти к уравнению

$$\frac{dy}{dt} + \sum_{\alpha=1}^p A_\alpha y = f(x, t), \quad x \in \omega, \quad 0 < t \leq T. \tag{31}$$

Прямое построение тех или иных схем расщепления для задачи (31), (18) осложнено тем, что отдельные операторные слагаемые  $A_\alpha$ ,  $\alpha = 1, 2, \dots, p$ , не наследуют основные свойства оператора  $A$  – его самосопряжённость и неотрицательность. Однако при использовании операторов декомпозиции (29) уравнение (31) легко симметризуется. Умножив уравнение (31) на самосопряжённый оператор  $A$ , придём к уравнению

$$\tilde{B} \frac{dy}{dt} + \sum_{\alpha=1}^p \tilde{A}_\alpha y = Af(x, t), \quad x \in \omega, \quad 0 < t \leq T, \tag{32}$$

в котором операторы

$$\tilde{B} = A, \quad \tilde{A}_\alpha = A\chi_\alpha A, \quad \alpha = 1, 2, \dots, p,$$

являются самосопряжёнными и неотрицательными. Более того, можно ввести новые переменные  $v = A^{1/2}y$  и вместо (32) рассматривать уравнение

$$\frac{dv}{dt} + \sum_{\alpha=1}^p \tilde{A}_\alpha v = A^{1/2}f(\mathbf{x}, t), \quad \mathbf{x} \in \omega, \quad 0 < t \leq T, \quad (33)$$

с самосопряжёнными и неотрицательными операторами

$$\tilde{A}_\alpha = A^{1/2}\chi_\alpha A^{1/2}, \quad \alpha = 1, 2, \dots, p.$$

Обычные оценки решения уравнения (33) по норме пространства  $H$  (для  $\|v\|$ ) соответствуют использованию оценок в  $H_A$  (для  $\|y\|_A$ ).

Конкретизация операторов декомпозиции типа (28), (29) осуществляется выбором слагаемых  $\chi_\alpha$ ,  $\alpha = 1, 2, \dots, p$ . При рассматриваемой декомпозиции области (см. рис. 4) естественно положить

$$\chi_2(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in \hat{\omega}, \\ 0, & \mathbf{x} \notin \hat{\omega}, \end{cases} \quad \chi_1(\mathbf{x}) = 1 - \chi_2(\mathbf{x}), \quad \mathbf{x} \in \omega.$$

Оператор  $A_2$  связывается с интерфейсными узлами  $\hat{\omega}$ , а  $A_1$  – с внутренними узлами подобластей.

Дифференциальному оператору декомпозиции (7) при разложении единицы для сеточной области (28) сопоставим расщепление (30), в котором

$$A_\alpha = A\chi_\alpha, \quad \alpha = 1, 2, \dots, p. \quad (34)$$

Для построения и исследования схем расщепления для задачи (18), (30), (31), (34) мы можем аналогично (33) перейти к задаче с самосопряжёнными операторами расщепления  $\tilde{A}_\alpha$ ,  $\alpha = 1, 2, \dots, p$ . Введя новую переменную  $v = A^{-1/2}y$ , получим уравнение

$$\frac{dv}{dt} + \sum_{\alpha=1}^p \tilde{A}_\alpha v = A^{-1/2}f(\mathbf{x}, t), \quad \mathbf{x} \in \omega, \quad 0 < t \leq T. \quad (35)$$

Несколько более сложная конструкция связывается с оператором декомпозиции (6).

Мы отметили возможность двухкомпонентного расщепления. Однако во многих случаях приходится ориентироваться на более общую ситуацию многокомпонентного расщепления.

**4. Векторная задача.** Для приближённого решения задачи Коши для уравнений (33) и (35) можно применять различные схемы расщепления [9, гл. 3–8]. Наиболее просто строятся аддитивные схемы при двухкомпонентном расщеплении: операторные аналоги схем Писмена–Рекфорда, Дугласа–Рекфорда, факторизованные схемы. При многокомпонентном расщеплении можем ориентироваться на схемы покомпонентного расщепления, среди которых можно выделить в плане возможности организации параллельных вычислений аддитивно-усреднённые схемы суммарной аппроксимации и аддитивные регуляризованные схемы полной аппроксимации. Новый класс аддитивных схем базируется на переходе от одного скалярного уравнения к системе уравнений – векторные аддитивные схемы. В своём исследовании мы помимо расщепления оператора задачи (30) расщепляем и само приближённое решение. В общем виде схемы расщепления решения в различных вариантах рассматриваются в работах [20, 21]. Для отдельных составляющих решения имеем систему связанных уравнений.

Для предложенных методов декомпозиции области на основе разбиения единицы сеточной области введём операторы

$$R_\alpha = \chi_\alpha(\mathbf{x})I, \quad \mathbf{x} \in \omega, \quad \alpha = 1, 2, \dots, p.$$

Для линейных операторов  $R_\alpha : H \rightarrow H$ ,  $\alpha = 1, 2, \dots, p$ , имеем

$$R_\alpha = R_\alpha^* \geq 0, \quad \alpha = 1, 2, \dots, p, \quad \sum_{\alpha=1}^p R_\alpha = I. \quad (36)$$

Ориентируемся на расщепление (30), (34). С учётом обозначений  $u(t) = u(\cdot, t)$  ищется приближённое решение задачи Коши

$$\frac{du}{dt} + \sum_{\beta=1}^p AR_\beta u = f(t), \quad 0 < t \leq T, \quad (37)$$

$$u(0) = u^0. \quad (38)$$

Умножив уравнение (37) на  $R_\alpha$ ,  $\alpha = 1, 2, \dots, p$ , получим

$$R_\alpha \frac{du}{dt} + R_\alpha \sum_{\beta=1}^p AR_\beta u = f_\alpha(t), \quad 0 < t \leq T, \quad \alpha = 1, 2, \dots, p, \quad (39)$$

где  $f_\alpha(t) = R_\alpha f(t)$ ,  $\alpha = 1, 2, \dots, p$ . Определим вспомогательные функции  $w_\alpha$ ,  $\alpha = 1, 2, \dots, p$ , из уравнений

$$R_\alpha \frac{dw_\alpha}{dt} + R_\alpha \sum_{\beta=1}^p AR_\beta w_\beta = R_\alpha f(t), \quad 0 < t \leq T, \quad \alpha = 1, 2, \dots, p. \quad (40)$$

Функции  $w_\alpha$  находятся из решения сеточных задач на множестве узлов расчётной сетки  $\omega_\alpha$ , на котором  $\chi_\alpha(x) > 0$ ,  $\alpha = 1, 2, \dots, p$ . С учётом (39) каждая функция  $w_\alpha$ ,  $\alpha = 1, 2, \dots, p$ , сопоставляется с  $u$ . Дополним систему уравнений (40) начальными условиями

$$w_\alpha(0) = u^0, \quad \alpha = 1, 2, \dots, p. \quad (41)$$

**Теорема 3.** Пусть

$$w(t) = \sum_{\alpha=1}^p R_\alpha w_\alpha(t), \quad 0 < t \leq T, \quad (42)$$

где  $w_\alpha(t)$ ,  $\alpha = 1, 2, \dots, p$ , – решение задачи Коши (40), (41). Тогда  $w(t) = u(t)$ , где  $u(t)$  – решение задачи (37), (38).

**Доказательство.** Складывая уравнения (40), получаем

$$\sum_{\alpha=1}^p R_\alpha \frac{dw_\alpha}{dt} + \sum_{\alpha=1}^p R_\alpha A \sum_{\beta=1}^p R_\beta w_\beta = \sum_{\alpha=1}^p R_\alpha f(t).$$

Принимая во внимание (42), для  $w(t)$  имеем уравнение

$$\frac{dw}{dt} + Aw = f(t), \quad 0 < t \leq T.$$

Из (37) и (41) следует начальное условие

$$w(0) = u^0.$$

Эта задача идентична задаче (37), (38) и поэтому  $w(t) = u(t)$  для всех  $0 < t \leq T$ . Теорема доказана.

Удобно ввести вектор искомых величин  $\mathbf{w} = \{w_1, w_2, \dots, w_p\}$ , для начальных данных и правой части положим  $\mathbf{w}^0 = \{u^0, u^0, \dots, u^0\}$ ,  $\mathbf{f} = \{f, f, \dots, f\}$  соответственно. Систему уравнений (40) запишем в виде

$$\mathbf{B} \frac{d\mathbf{w}}{dt} + \mathbf{A}\mathbf{w} = \mathbf{B}\mathbf{f}, \quad 0 < t \leq T, \tag{43}$$

где операторные матрицы

$$\mathbf{B} = \text{diag}\{R_1, R_2, \dots, R_p\}, \quad \mathbf{A} = \{R_\alpha A R_\beta\}, \quad \alpha, \beta = 1, 2, \dots, p. \tag{44}$$

Уравнение (43) дополним начальным условием

$$\mathbf{w}(0) = \mathbf{w}^0. \tag{45}$$

Рассмотрим задачу (43)–(45) на прямой сумме пространств  $\mathbf{H} = H \oplus H \oplus \dots \oplus H$ . Для  $\mathbf{w}, \mathbf{y} \in \mathbf{H}$  скалярное произведение и норма определяются формулами

$$(\mathbf{w}, \mathbf{y}) = \sum_{\alpha=1}^p (w_\alpha, y_\alpha), \quad \|\mathbf{w}\| = (\mathbf{w}, \mathbf{w})^{1/2}.$$

С учётом (44) имеем

$$\mathbf{B} = \mathbf{B}^* \geq 0, \quad \mathbf{A} = \mathbf{A}^* \geq 0.$$

Неотрицательность операторов  $\mathbf{B}$  и  $\mathbf{A}$  для эволюционного уравнения (43) не позволяет рассчитывать на приемлемые априорные оценки для  $\mathbf{w}(t)$ . В нашем случае достаточно ориентироваться на устойчивость  $w(t)$ , которая определяется (см. (42)) линейной комбинацией отдельных компонент вектора  $\mathbf{w}(t)$ .

Умножим уравнение (43) скалярно в  $\mathbf{H}$  на  $d\mathbf{w}/dt$  и получим

$$\left( \mathbf{B} \frac{d\mathbf{w}}{dt}, \frac{d\mathbf{w}}{dt} \right) + \frac{1}{2} \frac{d}{dt} (\mathbf{A}\mathbf{w}, \mathbf{w}) = \left( \mathbf{B}\mathbf{f}, \frac{d\mathbf{w}}{dt} \right). \tag{46}$$

Принимая во внимание неравенство

$$\left( \mathbf{B}\mathbf{f}, \frac{d\mathbf{w}}{dt} \right) \leq \left( \mathbf{B} \frac{d\mathbf{w}}{dt}, \frac{d\mathbf{w}}{dt} \right) + \frac{1}{4} (\mathbf{B}\mathbf{f}, \mathbf{f}),$$

из равенства (46) получаем оценку

$$(\mathbf{A}\mathbf{w}(t), \mathbf{w}(t)) \leq (\mathbf{A}\mathbf{w}^0, \mathbf{w}^0) + \frac{1}{2} \int_0^t (\mathbf{B}\mathbf{f}(\theta), \mathbf{f}(\theta)) d\theta. \tag{47}$$

Для левой части (47) с учётом (42) имеем

$$(\mathbf{A}\mathbf{w}, \mathbf{w}) = \sum_{\alpha=1}^p \left( \sum_{\beta=1}^p R_\alpha A R_\beta w_\beta, w_\alpha \right)_\alpha = \left( A \sum_{\beta=1}^p R_\beta w_\beta, \sum_{\alpha=1}^p R_\alpha w_\alpha \right) = (A\mathbf{w}, \mathbf{w}).$$

Аналогично для правой части (47) с учётом задания начальных данных (45) и правой части уравнения (43) получим

$$(\mathbf{A}\mathbf{w}^0, \mathbf{w}^0) = \|u^0\|_A^2, \quad (\mathbf{B}\mathbf{f}(t), \mathbf{f}(t)) = \|f(t)\|^2.$$

Тем самым приходим к оценке

$$\|\mathbf{w}(t)\|_A^2 \leq \|u^0\|_A^2 + \frac{1}{2} \int_0^t \|f(\theta)\|^2 d\theta$$

для  $w(t)$ , которая совпадает с оценкой решения задачи (36)–(38).

**5. Схемы декомпозиции области.** Построим аппроксимации по времени для задачи Коши (43)–(45), которые обеспечивают переход на новый временной уровень путём решения отдельных задач для компонент вектора решений. Остановимся сначала на двухслойных схемах расщепления [23, гл. 8] для этой векторной задачи, отталкиваясь от стандартных схем с весами.

Для приближённого решения задачи (43), (45) используем схему

$$B \frac{y^{n+1} - y^n}{\tau} + A(\sigma y^{n+1} + (1 - \sigma)y^n) = Bf^{n+\sigma}, \quad n = 0, 1, \dots, N - 1, \quad (48)$$

$$y^0 = w^0. \quad (49)$$

**Теорема 4.** При  $\sigma \geq 1/2$  для векторной схемы (48), (49) с весами имеет место оценка

$$\|y^{n+1}\|_A^2 \leq \|u^0\|_A^2 + \frac{1}{2} \sum_{k=0}^n \tau \|f^{k+\sigma}\|^2, \quad n = 0, 1, \dots, N - 1, \quad (50)$$

в которой

$$y^n = \sum_{\alpha=1}^p R_\alpha y_\alpha^n(t), \quad n = 0, 1, \dots, N. \quad (51)$$

**Доказательство.** Подобно доказательству теоремы 1 запишем (48) в виде

$$(B + \tau D) \frac{y^{n+1} - y^n}{\tau} + A \frac{y^{n+1} + y^n}{2} = Bf^{n+\sigma}, \quad n = 0, 1, \dots, N - 1, \quad (52)$$

где  $D = (\sigma - 1/2)A$ . Умножив (52) на  $2(y^{n+1} - y^n)$  с учётом того, что  $D = D^* \geq 0$  при  $\sigma \geq 1/2$ , получим неравенство

$$\frac{2}{\tau} (B(y^{n+1} - y^n), y^{n+1} - y^n) + (Ay^{n+1}, y^{n+1}) - (Ay^n, y^n) \leq 2(Bf^{n+\sigma}, y^{n+1} - y^n).$$

Принимая во внимание, что

$$2(Bf^{n+\sigma}, y^{n+1} - y^n) \leq \frac{2}{\tau} (B(y^{n+1} - y^n), y^{n+1} - y^n) + \frac{\tau}{2} (Bf^{n+\sigma}, f^{n+\sigma}),$$

имеем

$$(Ay^{n+1}, y^{n+1}) \leq (Ay^n, y^n) + \frac{\tau}{2} (Bf^{n+\sigma}, f^{n+\sigma}).$$

Для  $y^n$ , которое определяется согласно (51), это даёт неравенство

$$\|y^{n+1}\|_A^2 \leq \|y^n\|_A^2 + \frac{\tau}{2} \|f^{n+\sigma}\|^2,$$

из которого следует доказываемая оценка (50). Теорема доказана.

На новом слое по времени решение находится из уравнения

$$(B + \sigma\tau A)y^{n+1} = \psi^n, \quad \psi^n = (B - (1 - \sigma)\tau A)y^{n+1} + \tau Bf^{n+\sigma}, \quad n = 0, 1, \dots, N - 1. \quad (53)$$

Для отдельных компонент решения имеем связанную систему уравнений

$$R_\alpha y_\alpha^{n+1} + \sigma\tau \sum_{\beta=1}^p R_\alpha A R_\beta y_\beta^{n+1} = \psi_\alpha^n, \quad \alpha = 1, 2, \dots, p.$$

Для нахождения решения  $y_\alpha^{n+1}$ ,  $\alpha = 1, 2, \dots, p$ , можем использовать [24, гл. 4] блочные итерационные методы Якоби и блочные треугольные итерационные методы (Seidel, SOR, SSOR).

Такие методы связаны с аддитивным или мультипликативным преобусловливанием Шварца в методах декомпозиции для приближённого решения эллиптических краевых задач. Сеточная задача (53) характеризуется наличием малого параметра – шага по времени  $\tau$ . Кроме того, имеется хорошее начальное приближение  $y_\alpha^n$ ,  $\alpha = 1, 2, \dots, p$ . Эти обстоятельства позволяют ожидать быструю сходимость соответствующих итерационных методов. Мы рассматриваем безытерационные схемы декомпозиции области для нестационарных задач.

Для организации параллельных вычислений наибольший интерес представляют схемы расщепления с выделением диагональной части матрицы оператора  $\mathbf{A}$  на новом слое по времени. В таком варианте аналога блочного метода Якоби положим

$$\mathbf{A}_0 = \text{diag}\{R_1AR_1^*, R_2AR_2^*, \dots, R_pAR_p^*\}.$$

Второй вариант расщепления оператора  $\mathbf{A}$  задачи можно рассматривать как аналог блочного метода Зайделя, когда

$$\mathbf{A}_1 = \begin{pmatrix} \frac{1}{2}R_1AR_1 & 0 & \dots & 0 \\ R_2AR_1 & \frac{1}{2}R_2AR_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ R_pAR_1 & R_pAR_2 & \dots & \frac{1}{2}R_pAR_p \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} \frac{1}{2}R_1AR_1 & R_1AR_2 & \dots & R_1AR_p \\ 0 & \frac{1}{2}R_2AR_2 & \dots & R_2AR_p \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{1}{2}R_pAR_p \end{pmatrix},$$

при этом

$$\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2, \quad \mathbf{A}_1^* = \mathbf{A}_2.$$

Такое расщепление связывается со схемами попеременно-треугольного метода [7, гл. 10; 9, гл. 4; 25, гл. 10]. Работа с блочными треугольными матрицами менее интересна, по сравнению с блочными диагональными матрицами, в плане параллельной реализации. В силу этого ограничимся построением схем декомпозиции области с выделением диагональной части оператора  $\mathbf{A}$ .

Для приближённого решения задачи (43), (45) будем использовать явно-неявную схему первого порядка точности

$$\mathbf{B} \frac{\mathbf{y}^{n+1} - \mathbf{y}^n}{\tau} + \mathbf{A}_0(\sigma \mathbf{y}^{n+1} + (1 - \sigma)\mathbf{y}^n) + (\mathbf{A} - \mathbf{A}_0)\mathbf{y}^n = \mathbf{B}\mathbf{f}^{n+1}, \quad n = 0, 1, \dots, N - 1. \quad (54)$$

**Теорема 5.** При  $\sigma \geq p/2$  для явно-неявной векторной схемы расщепления (49), (54) имеет место априорная оценка (50), (51) для приближённого решения задачи (36)–(38).

**Доказательство.** Запишем схему (54) в виде (52), где теперь

$$\mathbf{D} = \sigma \mathbf{A}_0 - \frac{1}{2}\mathbf{A}.$$

При доказательстве теоремы 4 установлено, что оценка (50), (51) имеет место при  $\mathbf{D} = \mathbf{D}^* \geq 0$ . Сформулируем ограничения на вес  $\sigma$ , которые обеспечивают неотрицательность оператора  $\mathbf{D}$ .

Принимая во внимание неравенство

$$\left(\sum_{\alpha=1}^p a_\alpha\right)^2 \leq p \sum_{\alpha=1}^p a_\alpha^2,$$

с учётом самосопряжённости операторов  $R_\alpha$ ,  $\alpha = 1, 2, \dots, p$ , имеем

$$(\mathbf{A}_0 \mathbf{w}, \mathbf{w}) = \sum_{\alpha=1}^p (R_\alpha AR_\alpha w_\alpha, w_\alpha) = \sum_{\alpha=1}^p ((A^{1/2} R_\alpha w_\alpha)^2, 1) \geq$$

$$\geq \frac{1}{p} \left( \sum_{\alpha=1}^p A^{1/2} R_{\alpha} w_{\alpha}, \sum_{\beta=1}^p A^{1/2} R_{\beta} w_{\beta} \right) = \frac{1}{p} (\mathbf{A} \mathbf{w}, \mathbf{w}).$$

Это приводит к неравенству

$$\mathbf{A}_0 \geq \frac{1}{p} \mathbf{A}. \tag{55}$$

Тем самым при  $\sigma \geq p/2$  оператор  $\mathbf{D}$  неотрицателен. Теорема доказана.

При применении схем декомпозиции области важна не только устойчивость соответствующих схем расщепления. Второй ключевой элемент исследования связан с оценками погрешности приближённого решения, которые получаются при использовании неоднородных по времени приближений (явно-неявных аппроксимаций). Мы имеем условно сходящиеся разностные схемы, точность которых зависит от дискретизации расчётной области по пространству от ширины области налегания. Эти вопросы требуют специального рассмотрения и не затрагиваются в настоящей работе. Ограничимся исследованием аппроксимаций по времени.

Можно улучшить точность схемы декомпозиции области с диагональной матрицей  $\mathbf{A}_0$ , используя трёхслойные аппроксимации по времени. Основной трёхслойной схемой для векторной задачи (43)–(45) является схема второго порядка точности с весом

$$\mathbf{B} \frac{\mathbf{y}^{n+1} - \mathbf{y}^{n-1}}{2\tau} + \mathbf{A}(\sigma \mathbf{y}^{n+1} + (1 - 2\sigma)\mathbf{y}^n + \sigma \mathbf{y}^{n-1}) = \mathbf{B} \mathbf{f}^n, \quad n = 1, 2, \dots, N - 1, \tag{56}$$

$$\mathbf{y}^0 = \mathbf{w}^0, \quad \mathbf{y}^1 = \widetilde{\mathbf{w}}^1. \tag{57}$$

Для аппроксимации второго начального условия (57) со вторым порядком точности привлекается двухслойная схема. Например, положим  $\widetilde{\mathbf{w}}^1 = \{\widetilde{u}^1, \widetilde{u}^1, \dots, \widetilde{u}^1\}$  при использовании симметричной схемы

$$\frac{\widetilde{u}^1 - u^0}{\tau} + A \frac{\widetilde{u}^1 + u^0}{2} = f^{1/2}.$$

Аналогично теореме 4 устанавливается

**Теорема 6.** При  $\sigma \geq 1/4$  для векторной схемы (56), (57) с весами имеет место оценка

$$\begin{aligned} & \left( \mathbf{D} \frac{\mathbf{y}^{n+1} - \mathbf{y}^n}{\tau}, \frac{\mathbf{y}^{n+1} - \mathbf{y}^n}{\tau} \right) + \left\| \frac{\mathbf{y}^{n+1} + \mathbf{y}^n}{2} \right\|_A^2 \leq \\ & \leq \left( \mathbf{D} \frac{\widetilde{u}^1 - u^0}{\tau}, \frac{\widetilde{u}^1 - u^0}{\tau} \right) + \left\| \frac{\widetilde{u}^1 + u^0}{2} \right\|_A^2 + \frac{1}{2} \sum_{k=1}^n \tau \|f^k\|^2, \quad n = 1, 2, \dots, N - 1, \end{aligned} \tag{58}$$

в которой  $\mathbf{D} = (\sigma - 1/4)\tau^2 \mathbf{A}$ , а  $\mathbf{y}^n$ ,  $n = 0, 1, \dots, N$ , определены согласно (51).

**Доказательство.** Введём вспомогательные сеточные функции

$$\boldsymbol{\varsigma}^n = \frac{1}{2}(\mathbf{y}^n + \mathbf{y}^{n-1}), \quad \boldsymbol{\eta}^n = \frac{\mathbf{y}^n - \mathbf{y}^{n-1}}{\tau}.$$

Принимая во внимание равенство

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

запишем схему (56) в виде

$$\mathbf{B} \frac{\boldsymbol{\eta}^{n+1} + \boldsymbol{\eta}^n}{2} + \mathbf{D} \frac{\boldsymbol{\eta}^{n+1} - \boldsymbol{\eta}^n}{\tau} + \mathbf{A} \frac{\boldsymbol{\varsigma}^{n+1} + \boldsymbol{\varsigma}^n}{2} = \mathbf{B} \mathbf{f}^n, \quad n = 1, 2, \dots, N - 1, \tag{59}$$

где  $D = (\sigma - 1/4)\tau^2 A$ . Умножим уравнение (59) скалярно в  $H$  на  $2(\zeta^{n+1} - \zeta^n) = \tau(\eta^{n+1} + \eta^n)$  и получим соотношения

$$\begin{aligned} \frac{\tau}{2}(B(\eta^{n+1} + \eta^n), \eta^{n+1} + \eta^n) &\leq (D\eta^{n+1}, \eta^{n+1}) + (A\zeta^{n+1}, \zeta^{n+1}) \leq \\ &\leq (D\eta^n, \eta^n) + (A\zeta^n, \zeta^n) + \tau(Bf^n, \eta^{n+1} + \eta^n). \end{aligned}$$

С учётом

$$(Bf^n, \eta^{n+1} + \eta^n) \leq \frac{1}{2}(B(\eta^{n+1} + \eta^n), \eta^{n+1} + \eta^n) + \frac{1}{2}(Bf^n, f^n)$$

имеем

$$(D\eta^{n+1}, \eta^{n+1}) + (A\zeta^{n+1}, \zeta^{n+1}) \leq (D\eta^n, \eta^n) + (A\zeta^n, \zeta^n) + \frac{\tau}{2}(Bf^n, f^n).$$

При  $\sigma \geq 1/4$  это неравенство даёт

$$(D\eta^{n+1}, \eta^{n+1}) + (A\zeta^{n+1}, \zeta^{n+1}) \leq (D\eta^1, \eta^1) + (A\zeta^1, \zeta^1) + \frac{1}{2} \sum_{k=1}^n \tau(Bf^k, f^k).$$

Принимая во внимание введённые обозначения и покомпонентные представления для операторов  $A$  и  $D$ , получаем оценку (58). Теорема доказана.

При  $\sigma = 1/4$  имеем априорную оценку для решения задачи (36)–(38) в полупелых узлах  $\tilde{y}^{n+1/2} = (y^{n+1} + y^n)/2$ . В этом случае из (58) получим

$$\left\| \frac{y^{n+1} + y^n}{2} \right\|_A^2 \leq \left\| \frac{\tilde{u}^1 + u^0}{2} \right\|_A^2 + \frac{1}{2} \sum_{k=1}^n \tau \|f^k\|^2, \quad n = 1, 2, \dots, N-1.$$

Отталкиваясь от схемы (56), (57), строится схема декомпозиции области при выделении диагональной части  $A$ . Используем следующую трёхслойную схему второго порядка точности по времени:

$$B \frac{y^{n+1} - y^{n-1}}{2\tau} + A_0(\sigma y^{n+1} + (1-2\sigma)y^n + \sigma y^{n-1}) + (A - A_0)y^n = Bf^n, \quad n = 1, 2, \dots, N-1. \quad (60)$$

**Теорема 7.** При  $\sigma \geq p/4$  для векторной схемы (57), (60) с весами имеет место оценка (58), в которой

$$D = \sigma\tau^2 A_0 - \frac{1}{4}\tau^2 A \geq 0.$$

**Доказательство.** Векторную схему запишем в виде (59) при  $D = \sigma\tau^2 A_0 - \tau^2 A/4$ . Принимая во внимание неравенство (55), при сформулированных ограничениях на  $\sigma$  имеем самосопряжённый неотрицательный оператор  $D$ . Дальнейшие рассуждения проводятся аналогично доказательству теоремы 6. Теорема доказана.

Работа выполнена при финансовой поддержке Российского научного фонда (проекты 23-41-00037 (шп. 1–3) и 23-71-30013 (шп. 4, 5)).

## СПИСОК ЛИТЕРАТУРЫ

1. Toselli A., Widlund O. Domain Decomposition Methods: Algorithms and Theory. Berlin, 2005.
2. Samarskii A.A., Matus P.P., Vabishchevich P.N. Difference Schemes with Operator Factors. Dordrecht, 2002.
3. Mathew T. Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations. Berlin, 2008.
4. Лаевский Ю.М. Методы разбиения области при решении двумерных параболических уравнений // Вариационно-разностные методы в задачах численного анализа. Новосибирск, 1987. № 2. С. 112–128.

5. *Вабищевич П.Н.* Разностные схемы декомпозиции расчётной области при решении нестационарных задач // Журн. вычислит. математики и мат. физики. 1989. Т. 29. С. 1822–1829.
6. *Лаевский М.Ю., Мацокин А.М.* Методы декомпозиции решения эллиптических и параболических краевых задач // Сиб. журн. вычислит. математики. 1999. Т. 2. С. 361–372.
7. *Самарский А.А.* Теория разностных схем. М., 1989.
8. *Marchuk G.I.* Splitting and alternating direction methods // Handbook of Numerical Analysis. V. I. North-Holland, 1990. P. 197–462.
9. *Vabishchevich P.N.* Additive Operator-Difference Schemes: Splitting Schemes. Berlin, 2013.
10. *Вабищевич П.Н.* Регионально-аддитивные разностные схемы стабилизирующей поправки для параболических задач // Журн. вычислит. математики и мат. физики. 1994. Т. 34. С. 1832–1842.
11. *Самарский А.А., Вабищевич П.Н.* Факторизованные разностные схемы декомпозиции области для задач конвекции–диффузии // Дифференц. уравнения. 1997. Т. 33. № 7. С. 967–974.
12. *Вабищевич П.Н.* Разностные схемы декомпозиции области для нестационарных задач конвекции/диффузии // Дифференц. уравнения. 1996. Т. 32. № 7. С. 923–927.
13. *Гордезиани Д.Г., Меладзе Г.В.* О моделировании третьей краевой задачи для многомерных параболических уравнений в произвольной области одномерными уравнениями // Журн. вычислит. математики и мат. физики. 1974. Т. 14. С. 246–250.
14. *Вабищевич П.Н., Вераховский В.А.* Разностные схемы покомпонентного расщепления–декомпозиции области // Вестн. Моск. ун-та. Вычислит. математика и кибернетика. 1994. № 3. С. 17–22.
15. *Самарский А.А., Вабищевич П.Н.* Регуляризованные аддитивные схемы полной аппроксимации // Докл. РАН. 1998. Т. 358. С. 461–464.
16. *Абрашин В.Н.* Об одном варианте метода переменных направлений решения многомерных задач математической физики. I // Дифференц. уравнения. 1990. Т. 26. № 2. С. 314–323.
17. *Вабищевич П.Н.* Векторные аддитивные разностные схемы для эволюционных уравнений первого порядка // Журн. вычислит. математики и мат. физики. 1996. Т. 36. С. 44–51.
18. *Самарский А.А., Вабищевич П.Н.* Векторные аддитивные схемы декомпозиции области для параболических задач // Дифференц. уравнения. 1995. Т. 31. № 9. С. 1563–1569.
19. *Vabishchevich P.N.* Domain decomposition methods with overlapping subdomains for the time-dependent problems of mathematical physics // Comput. Methods in Appl. Math. 2008. V. 8. P. 393–405.
20. *Efendiev Y., Vabishchevich P.N.* Splitting methods for solution decomposition in nonstationary problems // Appl. Math. and Comput. 2021. V. 397. P. 125785.
21. *Вабищевич П.Н.* Схемы расщепления решения для эволюционных уравнений второго порядка // Дифференц. уравнения. 2021. Т. 57. № 7. С. 880–888.
22. *Абрашин В.Н., Вабищевич П.Н.* Векторные аддитивные схемы для эволюционных уравнений второго порядка // Дифференц. уравнения. 1998. Т. 34. № 12. С. 1666–1674.
23. *Вабищевич П.Н.* Численные методы решения нестационарных задач. М., 2021.
24. *Saad Y.* Iterative Methods for Sparse Linear Systems. Philadelphia, 2003.
25. *Самарский А.А., Николаев Е.С.* Методы решения сеточных уравнений. М., 1978.

Институт проблем безопасного развития  
атомной энергетики РАН, г. Москва,  
Северо-Восточный федеральный университет  
имени М.К. Аммосова, г. Якутск

Поступила в редакцию 08.01.2023 г.  
После доработки 08.01.2023 г.  
Принята к публикации 18.04.2023 г.

УДК 519.63+517.957

## ГЛОБАЛЬНО УСТОЙЧИВЫЕ РАЗНОСТНЫЕ СХЕМЫ ДЛЯ УРАВНЕНИЯ ФИШЕРА

© 2023 г. П. П. Матус, Д. Пылак

Построены и исследованы безусловно монотонные и глобально устойчивые разностные схемы для уравнения Фишера. Показано, что при определённом выборе входных данных эти схемы наследуют главное свойство устойчивого решения дифференциальной задачи. Доказана безусловная монотонность рассматриваемых разностных схем и получена априорная оценка разностного решения в равномерной норме. Доказано устойчивое поведение разностного решения в нелинейном случае при жёстких ограничениях на входные данные. Полученные результаты обобщены на многомерные уравнения, для аппроксимации которых применяются экономичные разностные схемы.

DOI: 10.31857/S0374064123070099, EDN: GUZVVS

**Введение.** Популяционная модель Фишера или Колмогорова–Петровского–Пискунова [1, 2] встречается в различных задачах, например, в теории горения, в теории фазовых переходов, в физике плазмы и др. Дифференциальные свойства решения задачи Коши для уравнения данной модели изучаются в фундаментальной монографии [3, с. 465]. Указывается класс входных данных, для которых точное неотрицательное решение не превосходит единицы для любого  $0 \leq t < +\infty$ . Естественно для этой модели строить такие вычислительные методы, которые наследуют это свойство. В данной работе строятся и исследуются безусловно монотонные и глобально устойчивые разностные схемы для уравнения Фишера. При определённом выборе входных данных задачи эти схемы наследуют главное свойство устойчивого решения дифференциальной задачи:  $0 \leq u(x, t) \leq 1$  при всех  $(x, t) \in \bar{Q}_T = \{(x, t) : 0 \leq x \leq l, 0 \leq t < +\infty\}$ . Устойчивое поведение разностного решения в нелинейном случае доказывается при ограничениях на входные данные  $0.5 \leq u_0(x), \mu_1(t), \mu_2(t) \leq 1$ . В п. 3 полученные результаты обобщаются на многомерное уравнение Фишера, для численного решения которого используются экономичные локально-одномерные схемы.

### 1. Вспомогательные результаты.

**1.1. Двусторонние оценки.** В дальнейшем нам понадобится нестандартный принцип максимума для общей канонической формы записи разностных схем [4, гл. IV, § 2], позволяющий устанавливать двусторонние оценки разностного решения для нелинейных задач [5]. Пусть задано начальное число точек – сетка  $\bar{\omega}_h = \omega_h \cup \gamma_h$ , где  $\omega_h$  – множество внутренних точек,  $\gamma_h$  – множество граничных узлов. *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 любых  $x \in \omega_h$  и принимающие вещественные значения. Для каждой точки  $x \in \omega_h$  сопоставим одно и только одно уравнение вида

$$A(x)y(x) = \sum_{\xi \in M'(x)} B(x, \xi)y(\xi) + F(x), \quad x \in \omega_h, \quad (1)$$

называемое *канонической формой записи* разностной схемы. Для уравнения (1) в граничных узлах зададим условие Дирихле

$$y(x) = \mu(x), \quad x \in \gamma_h. \quad (2)$$

Отметим, что при аппроксимации граничных условий второго или третьего рода сетка может не содержать граничных узлов.

Будем предполагать выполнение обычных условий положительности коэффициентов

$$A(x) > 0, \quad B(x, \xi) > 0 \quad \text{для всех } \xi \in M'(x), \quad x \in \omega_h, \quad (3)$$

$$D(x) = A(x) - \sum_{\xi \in M'(x)} B(x, \xi) > 0 \quad \text{для всех } \xi \in M'(x), \quad x \in \omega_h, \quad (4)$$

гарантирующих однозначную разрешимость схемы в равномерной норме.

Сформулируем утверждение, позволяющее установить двусторонние оценки сеточного решения при незнакоопределённых входных данных задачи  $\mu(x)$ ,  $F(x)$ .

**Лемма.** Пусть выполнены условия положительности коэффициентов (3), (4). Тогда максимальное и минимальное значения решения разностной схемы (1), (2) принадлежат интервалу изменения входных данных

$$m_1 \leq y(x) \leq m_2, \quad x \in \bar{\omega}_h, \quad (5)$$

где

$$m_1 = \min \left\{ \min_{x \in \gamma_h} \mu(x), \min_{x \in \omega_h} \frac{F(x)}{D(x)} \right\}, \quad m_2 = \max \left\{ \max_{x \in \gamma_h} \mu(x), \max_{x \in \omega_h} \frac{F(x)}{D(x)} \right\}.$$

Доказательство этого утверждения можно найти в работе [5].

**1.2. Определение монотонности разностной схемы в нелинейном случае.** В теории вычислительных методов понятие монотонности имеет большое значение, поскольку монотонность обеспечивает отсутствие нефизических осцилляций в численных расчётах. Сохранение свойства монотонности разностной схемы важно для задач с сильно меняющимися входными данными [6]. В линейном случае монотонность разностной схемы следует из положительности коэффициентов разностной схемы [6] или сеточного принципа максимума [7]. В нелинейном случае понятие монотонности несколько сложнее [8, 9]. Самое правильное и естественное определение дано в монографии [10, с. 123]. Для абстрактной задачи (1), (2) оно формулируется следующим образом. Возмущая входные данные  $F(x)$ ,  $\mu(x)$ , получаем возмущённую задачу

$$A(x)\tilde{y}(x) = \sum_{\xi \in M'(x)} B(x, \xi)\tilde{y}(\xi) + \tilde{F}(x), \quad x \in \omega_h,$$

$$\tilde{y}(x) = \tilde{\mu}(x), \quad x \in \gamma_h.$$

В соответствии с работами [8; 9; 10, с. 123] разностная схема (1), (2) называется *монотонной*, если из условий

$$\tilde{F}(x) - F(x) \geq 0, \quad x \in \omega_h, \quad \tilde{\mu}(x) - \mu(x) \geq 0, \quad x \in \gamma_h, \quad (6)$$

$$\tilde{F}(x) - F(x) \leq 0, \quad x \in \omega_h, \quad \tilde{\mu}(x) - \mu(x) \leq 0, \quad x \in \gamma_h,$$

следует выполнение соответствующих неравенств

$$\tilde{y}(x) - y(x) \geq 0, \quad (\tilde{y}(x) - y(x) \leq 0), \quad x \in \bar{\omega}_h. \quad (7)$$

**1.3. Точная разностная схема.** При аппроксимации нелинейного члена в уравнении Фишера будем ориентироваться на точную разностную схему [11; 12, с. 16]

$$\frac{y_{n+1} - y_n}{\tau} = \left[ \frac{1}{y_{n+1} - y_n} \int_{y_n}^{y_{n+1}} \frac{dy}{f(y)} \right]^{-1}, \quad y_0 = u_0,$$

$$y_n = y(t_n), \quad t_n \in \bar{\omega}_\tau = \{t_n = n\tau, \quad n = 0, 1, \dots\},$$

аппроксимирующую задачу Коши для обыкновенного дифференциального уравнения

$$\frac{du}{dt} = f(u), \quad u = u(t), \quad u(0) = u_0.$$

**Следствие.** Для квадратичной нелинейности  $f(u) = u^2$  точная разностная схема имеет безытерационный вид

$$\frac{y_{n+1} - y_n}{\tau} = y_n y_{n+1}, \quad y_0 = u_0. \tag{8}$$

**2. Разностная схема для уравнения Фишера.**

**2.1. Постановка задачи.** Уравнение Фишера, известное также как уравнение Колмогорова–Петрова–Пискунова [1], названо в честь статистика и биолога Рональда Фишера, предложившего его в 1937 г. для описания процессов популяционной динамики [2]. Поставим для этого уравнения начальную задачу с краевыми условиями Дирихле:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \lambda u(1 - u), \quad \lambda = \text{const} > 0, \tag{9}$$

$$u(x, 0) = u_0(x), \quad u(0, t) = \mu_1(t), \quad u(l, t) = \mu_2(t). \tag{10}$$

В монографии [3, с. 465] изучаются вопросы существования неотрицательных и устойчивых решений при  $t \rightarrow \infty$  в виде бегущих волн. В частности, доказывается, что если входные данные задачи удовлетворяют условиями

$$0 \leq u_0(x), \mu_1(t), \mu_2(t) \leq 1, \quad (x, t) \in \overline{Q}_T, \tag{11}$$

то и для точного решения имеет место аналогичная двусторонняя оценка

$$0 \leq u(x, t) \leq 1, \quad (x, t) \in \overline{Q}_T. \tag{12}$$

**2.2. Разностная схема и априорная оценка.** На равномерной пространственно-временной сетке с постоянными по пространству и времени шагами  $h$  и  $\tau$  соответственно

$$\overline{\omega} = \overline{\omega}_h \times \overline{\omega}_\tau, \quad \overline{\omega}_h = \omega_h \cup \{x_0 = 0, x_N = l\}, \quad \omega_h = \{x_i = ih, i = \overline{1, N-1}, hN = l\},$$

$$\overline{\omega}_\tau = \{t_n = n\tau, n = 0, 1, \dots\} = \omega_\tau \cup \{t_0 = 0\}$$

аппроксимируем дифференциальную задачу (9), (10) неявной разностной схемой с учётом формулы (8):

$$y_t = \hat{y}_{\overline{x}x} + \lambda y(1 - \hat{y}), \tag{13}$$

$$y(x, 0) = u_0(x), \quad x \in \overline{\omega}_h, \quad \hat{y}_0 = \hat{\mu}_1, \quad \hat{y}_N = \hat{\mu}_2, \quad t \in \omega_\tau. \tag{14}$$

В работе используются стандартные обозначения теории разностных схем [4, гл. V, § 1]:

$$v = v_i^n = v(x_i, t_n), \quad \hat{v} = v_i^{n+1}, \quad v_t = (\hat{v} - v)/\tau, \quad v_{\overline{x}x} = (v_{i+1}^n - 2v_i^n + v_{i-1}^n)/h^2.$$

Запишем схему (13), (14) в каноническом виде (1), (2):

$$C_i^n y_i^{n+1} = A y_{i-1}^{n+1} + B y_{i+1}^{n+1} + F_i^n, \quad y_0^{n+1} = \mu_1^{n+1}, \quad y_N^{n+1} = \mu_2^{n+1},$$

где

$$A = B = \gamma, \quad \gamma = \tau/h^2, \quad C_i^n = 1 + 2\gamma + \lambda\tau y_i^n, \quad F_i^n = (1 + \lambda\tau)y_i^n, \quad D_i^n = 1 + \lambda\tau y_i^n.$$

Относительно входных данных предполагаем выполненными условия (11). Докажем, что и для решения разностной схемы (13), (14) при всех  $(x, t) \in \overline{\omega}$  выполнен разностный аналог

дифференциального свойства решения непрерывной задачи (12). Следуя методу математической индукции, предположим, что  $0 \leq y_i^n \leq 1$  для всех  $i = 0, \bar{N}$ . При таком предположении выполнены все условия леммы, и на основании неравенства (5) находим двустороннюю оценку

$$m_1^n \leq y_i^{n+1} \leq m_2^n, \tag{15}$$

где

$$m_1^n = \min \left\{ \min_{x \in \gamma_h} \left\{ \mu_1^{n+1}, \mu_2^{n+1} \right\}, \min_{x \in \omega_h} \frac{(1 + \lambda\tau)y_i^n}{1 + \lambda\tau y_i^n} \right\} \geq 0,$$

$$m_2^n = \max \left\{ \max_{x \in \gamma_h} \left\{ \mu_1^{n+1}, \mu_2^{n+1} \right\}, \max_{x \in \omega_h} \frac{(1 + \lambda\tau)y_i^n}{1 + \lambda\tau y_i^n} \right\} \leq \max \left\{ 1, \frac{1 + \lambda\tau y_i^n}{1 + \lambda\tau y_i^n} \right\} \leq 1.$$

Следовательно, оценка (15) выполнена для произвольного  $n = 0, 1, \dots$ . На основании данного неравенства заключаем, что

$$\max_{t \in \bar{\omega}_\tau} \|y(t)\|_{\bar{C}} \leq 1, \quad \|\cdot\|_{\bar{C}} = \max_{x \in \bar{\omega}_h} |\cdot|.$$

**2.3. Безусловная монотонность.** Разностную схему (13), (14) будем называть *безусловно монотонной*, если она монотонна при произвольных соотношениях сеточных шагов  $\tau$  и  $h$ .

Рассмотрим разностную задачу с возмущёнными входными данными

$$\tilde{y}_t = \hat{y}_{\bar{x}x} + \lambda\tilde{y}(1 - \hat{y}), \tag{16}$$

$$\tilde{y}(x, 0) = \tilde{u}_0(x), \quad x \in \bar{\omega}_h, \quad \hat{y}_0 = \hat{\mu}_1, \quad \hat{y}_N = \hat{\mu}_2. \tag{17}$$

Относительно возмущённых входных данных предполагаем выполненными условия (6) и (11):

$$\tilde{u}_0(x) - u_0(x) \geq 0, \quad x \in \bar{\omega}_h, \quad \tilde{\mu}_k(t_{n+1}) - \mu_k(t_{n+1}) \geq 0, \quad k = 1, 2, \quad t_{n+1} \in \omega_\tau.$$

Вычитая из уравнений (16), (17) соответствующие уравнения (13), (14), получаем задачу для возмущения  $\bar{y} = \tilde{y} - y$ :

$$\bar{y}_t = \hat{y}_{\bar{x}x} + \lambda\bar{y} - \lambda\hat{y}\bar{y} + \lambda\hat{y}\bar{y}, \tag{18}$$

$$\hat{y}_0 = \hat{\mu}_1, \quad \hat{y}_N = \hat{\mu}_2, \quad \bar{y}(x, 0) = \bar{u}_0(x). \tag{19}$$

Определим сеточные функции

$$F = (1 + \lambda\tau)\bar{y} - \lambda\tau\hat{y}\bar{y}, \quad D = 1 + \lambda\tau\hat{y} \geq 0.$$

Пусть на  $n$ -м временном слое разностное решение является монотонным в смысле (7). Докажем, что оно является монотонным и на  $(n + 1)$ -м временном слое. Так как сеточная функция  $\hat{y}$  удовлетворяет неравенству  $y^{n+1} \leq 1$ , то функция  $F = (1 + \lambda\tau - \lambda\tau\hat{y})\bar{y} \geq \bar{y} \geq 0$  является неотрицательной. На основании сформулированной леммы для  $\bar{y}^{n+1}$  получаем неравенство

$$\bar{y}^{n+1} \geq \min \left\{ \min_{x \in \gamma_h} \left\{ \bar{\mu}_1^{n+1}, \bar{\mu}_2^{n+1} \right\}, \min_{x \in \omega_h} \frac{F}{D} \right\} \geq 0.$$

Итак, мы доказали монотонность разностной схемы при произвольных соотношениях на сеточные шаги  $\tau$  и  $h$ .

**2.4. Глобальная устойчивость.** Понятие устойчивости является составной частью корректности постановки дифференциальной или разностной задачи и в общем случае означает непрерывную зависимость решения задачи  $u$  от входных данных  $\varphi$ , т.е. существует такая неотрицательная функция  $\rho(t)$ , не зависящая от решения и входных данных, которая удовлетворяет неравенству

$$\|\tilde{u} - u\|_1 \leq \rho(t)\|\tilde{\varphi} - \varphi\|_2, \tag{20}$$

где  $\|\cdot\|_1$  и  $\|\cdot\|_2$  – некоторые нормы,  $\tilde{u}$  – решение той же дифференциальной задачи с возмущёнными входными данными. Определение глобальной устойчивости даётся в работе [13].

**Определение.** Предположим, что решение дифференциальной или разностной задачи удовлетворяет оценке (20). Тогда задача называется *глобально устойчивой*, если  $\rho(t) \rightarrow \text{const}$  при  $t \rightarrow +\infty$ , и *асимптотически устойчивой*, если  $\rho(t) \rightarrow 0$  при  $t \rightarrow +\infty$ .

Если входные данные дифференциальной задачи (9), (10) удовлетворяют условиям (11), то возникают решения типа бегущей волны [13, с. 442]

$$u(x, t) = U(z), \quad z = x - ct, \quad c = 2,$$

удовлетворяющие регулярным состояниям

$$\lim_{z \rightarrow -\infty} U(z) = 0, \quad \lim_{z \rightarrow +\infty} U(z) = 1.$$

В этой же монографии отмечается [13, с. 447], что малые возмущения могут быть неустойчивы в некоторой конечной области.

**Теорема.** Пусть выполнены следующие условия на входные данные:

$$0.5 \leq u_0(x), \quad \mu_k(t) \leq 1, \quad k = 1, 2, \quad (x, t) \in \overline{Q}_T. \tag{21}$$

Тогда решение разностной задачи (13), (14) является глобально устойчивым и имеет место априорная оценка в равномерной норме для любого  $t \in \overline{\omega}_\tau$ :

$$\|\tilde{y}(t) - y(t)\|_{\overline{C}} \leq \max_{(x,t) \in \overline{Q}_T} \left\{ \max_{k=1,2} |\tilde{\mu}_k(t) - \mu_k(t)|, \|\tilde{u}_0 - u_0\|_{\overline{C}} \right\}. \tag{22}$$

**Доказательство.** Покажем, что при предположениях (21) имеют место соотношения

$$0.5 \leq y_i^n, \quad \tilde{y}_i^n \leq 1 \quad \text{для всех } i = \overline{0, N}, \quad n = 0, 1, \dots \tag{23}$$

Докажем сначала, что справедливы нижние оценки для решения разностных задач (13), (14) и (16), (17). Действительно,

$$\frac{F_i^n}{D_i^n} = \frac{1 + \lambda\tau}{1 + \lambda\tau y_i^n} \geq \frac{1}{2}.$$

Следовательно, если  $0.5 \leq y_i^k \leq 1$  для всех  $k = \overline{0, n}$ , то на основании леммы для решения  $y_i^{n+1}$  справедливо неравенство

$$y_i^{n+1} \geq \min \left\{ \min_{k=1,2} \mu_k(t_{n+1}), \min_{i=1, N-1} \frac{F_i^n}{D_i^n} \right\} \geq \frac{1}{2}.$$

Дальнейшие рассуждения повторяют вывод априорной оценки (15).

Обратимся теперь к доказательству требуемого неравенства (22). Максимум модуля сеточной функции  $\tilde{y} = \tilde{y} - y$  и решения задачи (18), (19) может достигаться на границе области, и тогда оценка (22) справедлива. Если же максимум достигается во внутренней точке, то на основании леммы для решения задачи (18), (19) имеет место неравенство

$$\tilde{y}_i^{n+1} \leq \max_{x \in \omega_h} \left\{ \frac{(1 + \lambda\tau) - \lambda\tau \tilde{y}}{1 + \lambda\tau \tilde{y}} \right\} \tilde{y}_i^n \leq \max_{x \in \omega_h} \frac{1 + \lambda\tau(1 - \tilde{y})}{1 + \lambda\tau \tilde{y}}.$$

В силу (23) для всех  $i$  и  $n$  выполняется условие  $1 - y_i^{n+1} \leq \tilde{y}_i^n$ . Следовательно, из последнего неравенства находим

$$\tilde{y}_i^{n+1} \leq \max_{1 \leq i \leq N-1} \frac{1 + \lambda\tau(1 - y_i^{n+1})}{1 + \lambda\tau \tilde{y}} \tilde{y}_i^n \leq \|y_i^n\|_C.$$

Итак,

$$\|\tilde{y}^{n+1} - y^{n+1}\|_C \leq \|\tilde{y}^n - y^n\|_C.$$

Объединив полученные оценки для  $\bar{y}^{n+1}$ , приходим к неравенству (22). Теорема доказана.

**Замечание.** Полученные выше результаты обобщаются на обобщённое уравнение Фишера [3, с. 454]

$$\frac{\partial u}{\partial t} = \sum_{\alpha=1}^2 \frac{\partial}{\partial x_\alpha} \left( k_\alpha(x, t, u) \frac{\partial u}{\partial x_\alpha} \right) + \sum_{\alpha=1}^2 r_\alpha(x) \frac{\partial u}{\partial x_\alpha} + \lambda(1 - u)u, \quad x = (x_1, x_2).$$

Вывод основывается на том, что минимальное и максимальное значения решения не зависят как от конвективных коэффициентов  $r_\alpha(x_1, x_2)$ , так и от диффузионных  $k_\alpha(x_1, x_2, t, u)$  [5].

**3. Многомерное уравнение Фишера.** Используя метод суммарной аппроксимации, можно построить экономичные аддитивные схемы для многомерного уравнения Фишера в  $p$ -мерной области  $\bar{G} = G \cup \Gamma$ . В данной работе ограничимся рассмотрением локально одномерной схемы [4, с. 486]. Пусть  $x = (x_1, x_2, \dots, x_p)$  – точка  $p$ -мерного евклидова пространства  $\mathbb{R}^p$ . Рассмотрим в цилиндре  $\bar{Q}_T = \bar{G} \times [0 \leq t \leq T]$ ,  $\bar{G} = \{x = (x_1, x_2, \dots, x_p) : 0 \leq x_\alpha \leq l_\alpha, \alpha = \overline{1, p}\}$ , первую краевую задачу для  $p$ -мерного уравнения Фишера

$$\frac{\partial u}{\partial t} = Lu + \lambda u(1 - u), \quad L = \sum_{\alpha=1}^p L_\alpha, \quad L_\alpha = \frac{\partial^2}{\partial x_\alpha^2}, \quad (x, t) \in Q_T, \tag{24}$$

$$u|_\Gamma = \mu(x, t), \quad t \geq 0, \quad u(x, 0) = u_0(x), \quad x \in \bar{G}. \tag{25}$$

Для аппроксимации дифференциальной задачи (24), (25) будем использовать чисто неявную локально-одномерную схему вида

$$\frac{y_{i_\alpha}^{n+\alpha/p} - y_{i_\alpha}^{n+(\alpha-1)/p}}{\tau} = \Lambda_\alpha y_{i_\alpha}^{n+\alpha/p}, \quad \alpha = \overline{1, p-1}, \quad i_\alpha = \overline{1, N_\alpha - 1}, \tag{26}$$

$$\frac{y_{i_p}^{n+1} - y_{i_p}^{n+(p-1)/p}}{\tau} = \Lambda_p y^{n+1} + \lambda y^{n+(p-1)/p} y^{n+1}, \quad \alpha = p, \tag{27}$$

с граничными и начальными условиями

$$y^{n+\alpha/p} = \mu^{n+\alpha/p} \quad \text{при} \quad x \in \gamma_{h,\alpha}, \quad y(x, 0) = u_0(x), \quad x \in \bar{\omega}_h. \tag{28}$$

Здесь также используются обычные обозначения теории разностных схем

$$\omega_h = \{x = (i_1 h_1, i_2 h_2, \dots, i_p h_p), \quad 0 \leq i_\alpha \leq N_\alpha - 1, \quad N_\alpha h_\alpha = l_\alpha, \quad \alpha = \overline{1, p}\},$$

$$\bar{\omega}_h = \omega_h \cup \gamma_h, \quad \gamma_{h,\alpha} = \omega_h \cup \{x_\alpha = 0, l_\alpha\},$$

$$\Lambda_\alpha y = y_{\bar{x}_\alpha x_\alpha} = \frac{1}{h_\alpha^2} (y^{(+1_\alpha)} - 2y + y^{(-1_\alpha)}).$$

Кратко опишем вычислительный алгоритм, реализующий схему (26)–(28). Пусть известно  $y^n$ . Чтобы найти  $y^{n+1}$  на новом слое, мы должны решить  $p$  уравнений с краевыми условиями (28), последовательно полагая  $\alpha = \overline{1, p}$ . Для определения  $y^{n+\alpha/p}$  получаем краевую задачу

$$A_\alpha y_{i_{\alpha-1}}^{n+\alpha/p} - C_{i_\alpha} y_{i_\alpha}^{n+\alpha/p} + B_\alpha y_{i_{\alpha+1}}^{n+\alpha/p} = -F_{i_\alpha}^{n+(\alpha-1)/p}, \quad \alpha = \overline{1, p},$$

$$y^{n+\alpha/p} = \mu^{n+\alpha/p} \quad \text{при} \quad x \in \gamma_{h,\alpha}, \quad \alpha = \overline{1, p},$$

где

$$A_\alpha = B_\alpha = \gamma_\alpha, \quad \gamma_\alpha = \tau/h_\alpha^2, \quad \alpha = \overline{1, p}, \quad C_{i_\alpha} = \begin{cases} 1 + A_\alpha + B_\alpha, & \alpha = \overline{1, p-1}, \\ 1 + A_p + B_p + \lambda\tau y_{i_\alpha}^{n+(p-1)/p}, & \alpha = p, \end{cases}$$

$$F_{i_\alpha}^{n+(\alpha-1)/p} = \begin{cases} y_{i_\alpha}^{n+(\alpha-1)/p}, & \alpha = \overline{1, p-1}, \\ (1 + \lambda\tau) y_{i_\alpha}^{n+(p-1)/p}, & \alpha = p. \end{cases}$$

Полагая последовательно  $\alpha = \overline{1, p}$  и меняя направления прогонок, определяем  $y^{n+1/p}$ ,  $y^{n+2/p}$ , ...,  $y^{n+1}$ , затратив при этом  $O(1)$  операций на узел действий. Таким образом, локально-одномерная схема является экономичной. Для получения глобальных априорных оценок разностного решения, как и в одномерном случае, предположим выполнение условий вида (11), (12)

$$0 \leq \mu(x, t), u_0(x) \leq 1 \quad \text{при всех } (x, t) \in \overline{Q}_T. \tag{29}$$

Применив теперь лемму к разностным уравнениям (26) и (28), получим оценки

$$m_1^{n+\alpha/2} \leq y_i^{n+\alpha/p} \leq m_2^{n+\alpha/p},$$

где

$$m_1^{\alpha/p} = \min \left\{ \min_{x \in \gamma_{h,\alpha}} \mu^{n+\alpha/p}, \min_{x \in \omega_h} y^{n+(\alpha-1)/p} \right\}, \quad \alpha = \overline{1, p-1},$$

$$m_2^{\alpha/p} = \max \left\{ \max_{x \in \gamma_{h,\alpha}} \mu^{n+\alpha/p}, \max_{x \in \omega_h} y^{n+(\alpha-1)/p} \right\}, \quad \alpha = \overline{1, p-1}.$$

Следуя методу математической индукции, предположим, что  $\|y^k\|_{\overline{C}} \leq 1$  для всех  $k = \overline{0, n}$ . Здесь  $\|y^n\|_{\overline{C}} = \max_{x \in \overline{\omega}_h} |y(x, t_n)|$ . Тогда из неравенств (29) находим  $0 \leq y_{i_\alpha}^{n+(p-1)/p} \leq 1$  для всех  $i_\alpha$ .

Обратимся теперь к разностному уравнению (20). Снова применяя лемму, получаем неравенство

$$y_i^{n+1} \leq \max \left\{ \max_{\alpha=\overline{1,p}} \max_{x \in \gamma_h} \mu^{n+\alpha/p}, \max_{x \in \omega_h} \frac{(1 + \lambda\tau) y_{i_\alpha}^{n+(p-1)/p}}{1 + \lambda\tau y_{i_\alpha}^{n+(p-1)/p}} \right\} \leq \max \left\{ 1, \frac{1 + \lambda\tau y_{i_\alpha}^{n+(p-1)/p}}{1 + \lambda\tau y_{i_\alpha}^{n+(p-1)/p}} \right\} \leq 1.$$

Это завершает доказательство того факта, что при выполнении условий (29) решение разностной задачи унаследует свойство глобальной ограниченности решения дифференциальной задачи и при всех  $0 \leq t < +\infty$  имеет место неравенство  $\max_{t \in \overline{\omega}_\tau} \|y(t)\| \leq 1$ . Как и в одномерном случае, решение разностной задачи при выполнении условий

$$0.5 \leq \mu(x, t), u_0(x) \leq 1 \quad \text{при всех } (x, t) \in \overline{Q}_T$$

является глобально устойчивым.

### СПИСОК ЛИТЕРАТУРЫ

1. Колмогоров А.Н., Петровский И.Г., Пискунов Н.С. Исследование уравнения диффузии, соединённой с возрастанием количества вещества и его применение к одной биологической проблеме // Бюлл. Московского гос. ун-та. Секция А. 1937. Т. 1. № 6. С. 1–25.
2. Fisher R.A. The wave of advance of advantageous genes // Ann. Hum. Genetic. 1937. V. 7. № 4. P. 353–369.
3. Murray J.D. Mathematical Biology: an Introduction. Berlin, 2001.
4. Самарский А.А. Теория разностных схем. М., 1989.

5. *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.
6. *Годунов С.К.* Разностный метод численного расчёта разрывных решений уравнений гидродинамики // *Мат. сб.* 1959. Т. 47 (89). № 3. С. 271–306.
7. *Matus P.* The maximum principle and some of its applications // *Comput. Methods Appl. Math.* 2002. V. 2. P. 50–91.
8. *Matus P., Lemeshevsky S.* Stability and monotonicity of difference schemes for nonlinear scalar conservation laws and multidimensional quasi-linear parabolic equations // *Comput. Methods Appl. Math.* 2009. V. 9. № 3. P. 253–280.
9. *Матус П.П., Утебаев Б.Д.* Компактные и монотонные разностные схемы для параболических уравнений // *Мат. моделирование.* 2021. Т. 33. № 4. С. 60–78.
10. *Godlewski E., Raviart P.-A.* *Hyperbolic Systems of Conservation Law.* Paris, 1991.
11. *Матус П.П., Ирхин В.А., Лапиньска-Хицоневич М., Лемешевский С.В.* О точных разностных схемах для гиперболических и параболических уравнений // *Дифференц. уравнения.* 2007. Т. 43. № 7. С. 978–986.
12. *Lemeshevsky S., Matus P., Poliakov D.* *Exact Finite-Difference Schemes.* Berlin, 2016.
13. *Jovanovic B., Lemeshevsky S., Matus P.* On the stability of differential-operator equations and operator-difference schemes as  $t \rightarrow \infty$  // *Comput. Methods Appl. Math.* 2002. V. 2. № 2. P. 153–170.

Институт математики НАН Беларуси,  
г. Минск,  
Католический университет имени Иоанна-Павла II,  
г. Люблин, Польша

Поступила в редакцию 24.03.2023 г.  
После доработки 24.03.2023 г.  
Принята к публикации 19.05.2023 г.

УДК 519.63+517.958+532.5

## К ВОПРОСУ О ЧИСЛЕННОМ РЕШЕНИИ НЕКОНСЕРВАТИВНЫХ ГИПЕРБОЛИЧЕСКИХ СИСТЕМ УРАВНЕНИЙ

© 2023 г. Р. Р. Полехина, М. В. Алексеев, Е. Б. Савенков

Численно исследованы вопросы, связанные с отсутствием сходимости при применении формально консервативных по пути разностных схем для решения неконсервативных гиперболических систем уравнений. Эта проблема является центральной при построении корректных разностных схем для решения указанного класса задач. Изложены базовые понятия теории неконсервативных гиперболических систем уравнений и соответствующие проблемы построения разностных схем для их решения. Предложен вариант метода HLL, позволяющий использовать произвольный явно заданный путь. Для модельной системы уравнений Бюргерса явно вычислены ударные адиабаты и пути, соответствующие вязкой регуляризации системы заданного вида. Проанализированы причины отсутствия сходимости численных решений к точным при некорректном применении соответствующих алгоритмов. Показано, что по крайней мере в частном рассмотренном случае формально консервативный по пути вариант метода HLL даёт правильное решение задачи.

DOI: 10.31857/S0374064123070105, EDN: GVFCXZ

**Введение.** Настоящая работа посвящена практическому анализу ряда вопросов сходимости численных решений неконсервативных гиперболических систем уравнений. Основной сложностью при построении алгоритмов для решения задач этого класса является то, что численные решения зачастую сходятся, но не к тому точному решению, которое предполагается. Эта проблема является центральной при построении численных методов решения неконсервативных гиперболических систем уравнений. Причина её появления состоит в том, что определение обобщённого (разрывного) решения в неконсервативном случае существенно сложнее, чем в случае консервативных гиперболических систем уравнений. Эти вопросы подробно обсуждаются, например, в работе [1].

Принципиальная теоретическая причина отсутствия сходимости считается известной [1] и связанной с особенностью определения так называемых “неконсервативных произведений” в общей математической теории неконсервативных гиперболических систем уравнений. Качественно эта проблема заключается в следующем. Одним из способов определения корректных разрывных решений гиперболических систем является использование регуляризованной задачи, которая имеет вид диссипативного (“вязкого”) возмущения исходной системы уравнений. В этом случае разрывные решения определяются как пределы решений возмущённой задачи при исчезающе малой вязкости. В случае консервативных гиперболических систем вид возмущённой системы (т.е. вид соответствующего диссипативного оператора) до известной степени не имеет значения – все “разумные” регуляризации определяют один и тот же класс обобщённых решений. В неконсервативных системах это не так – различные вязкие возмущения определяют различные классы обобщённых решений исходной задачи. Далее любая устойчивая разностная схема неявно определяет особый вид диссипативного возмущения исходной системы – её первое дифференциальное приближение [2, с. 17]. Численные решения сходятся к тем разрывным решениям задачи, которые выделяются диссипативным оператором разностной схемы (точнее, её первым дифференциальным приближением). В случае если этот оператор отличается от использованного при определении классов разрывных решений в исходной задаче, будет наблюдаться отсутствие сеточной сходимости.

Задача построения разностных схем, обеспечивающих сходимость численных решений к априорно заданному классу обобщённых решений, в настоящее время до конца не решена и является одной из центральных проблем построения разностных схем для неконсервативных гиперболических систем уравнений.

Описанный выше механизм потери сходимости принимается общим – считается, что “правильная” разностная схема должна иметь в точности те же самые диссипативные механизмы, что и использованные при постановке исходной задачи. Таким образом, решение задачи лежит в области построения разностных схем с априорным и “тонким” контролем её диссипативных свойств. Известно лишь небольшое количество работ, посвящённых этому вопросу. Для примера отметим работы Ю.И. Шокина и его коллег (см. [2, 3]), где был развит метод дифференциального приближения и рассматривался вопрос построения алгоритмов для решения уравнений Эйлера газовой динамики в простых (не консервативных) переменных; а также работы, посвящённые построению энтропийно-устойчивых схем для консервативных [4, 5] и неконсервативных [6] задач. Отметим также статью [7], где конкретным способом построения схем с контролируемыми диссипативными свойствами является метод дифференциального приближения.

Вместе с тем понятно, что вопрос о способах построения корректных в указанном выше смысле разностных схем неотъемлемо связан с вопросом о конкретных причинах отсутствия сходимости к нужному классу обобщённых решений. Настоящая работа посвящена исследованию ряда возникающих при этом частных вопросов. В п. 1 кратко и на неформальном уровне строгости описаны основные аспекты математической теории неконсервативных гиперболических систем, включая основные положения теории неконсервативных произведений Г. Даля Мазо, Ф. Лефлоха и Ф. Мурата (так называемая “теория DLM”). В п. 2 рассмотрены основные проблемы, возникающие при построении численных аппроксимаций неконсервативных систем. В п. 3 рассмотрен пример так называемой консервативной по пути разностной схемы типа HLL. Модельная задача для неконсервативной системы уравнений Бюргерса описана в п. 4. Наконец, в п. 5 представлены результаты численного исследования, а именно, для указанной модельной задачи строятся численные решения, определённые как с помощью вязкой диссипации, так и в рамках теории DLM. Показано, что качественное отличие решений задач, полученных различными способами, в конкретном рассматриваемом случае связано с причинами качественного характера: задача Римана с одними и теми же начальными данными может иметь различные решения. Состояния на скачке в задаче Римана в одном случае принадлежат ударной адиабате, соответствующей использованному способу регуляризации и учитываемому разностной схемой, и, таким образом, начальный разрыв сохраняется с течением времени, а в другом – нет.

**1. Неконсервативные гиперболические системы уравнений.** Основным объектом обсуждения в настоящей работе являются неконсервативные гиперболические системы уравнений первого порядка, т.е. системы вида

$$\frac{\partial \mathbf{q}}{\partial t} + \mathbf{A}(\mathbf{q}) \frac{\partial \mathbf{q}}{\partial x} = 0, \quad (1)$$

где  $\mathbf{q} \in \Omega \subset \mathbb{R}^n$  и  $\mathbf{A}(\mathbf{q}) \in \mathbb{R}^{n \times n}$  гладко зависит от  $\mathbf{q}$ . “Неконсервативное произведение”

$$\mathbf{A} \frac{\partial \mathbf{q}}{\partial x} \neq \frac{\partial \mathbf{F}}{\partial x}$$

для какого-либо вектора плотности потока  $\mathbf{F}$ . Здесь и далее будем считать, что решение  $\mathbf{q} = \mathbf{q}(x, t)$  зависит от пространственной координаты  $x$  и времени  $t$ . В случае если речь будет идти о задаче Коши, будем считать, что она рассматривается в области  $x \in \mathbb{R}$  при  $t \geq 0$ , причём при  $t > 0$  её решение  $\mathbf{q}$  определяется уравнением (1), а при  $t = 0$  – начальным условием

$$\mathbf{q}(x, 0) = \mathbf{q}_0(x).$$

Основной проблемой при построении математической теории систем указанного вида является невозможность определить обобщённое (разрывное) решение традиционным способом, как это делается для консервативных систем уравнений: в случае разрывного решения  $\mathbf{q}$  неконсервативное произведение  $\mathbf{A} \partial \mathbf{q} / \partial x$  теряет смысл даже в смысле классической теории обобщённых функций. Корректное определение разрывного решения может быть дано различными

способами. Простейший из них основан на подходящей регуляризации уравнения (1) (см., например, [8]). В соответствии с этой идеей далее будем рассматривать только такие разрывные решения, которые являются пределом (в том или ином смысле) гладких решений расширенной системы вида

$$\frac{\partial \mathbf{q}_\varepsilon}{\partial t} + \mathbf{A}(\mathbf{q}_\varepsilon) \frac{\partial \mathbf{q}_\varepsilon}{\partial x} = \varepsilon \frac{\partial}{\partial x} \left[ \mathbf{B} \frac{\partial \mathbf{q}_\varepsilon}{\partial x} \right] \quad (2)$$

при  $\varepsilon \rightarrow 0$ . Конкретный вид такого возмущения обосновывается тем, что он часто встречается в физических и механических приложениях. Естественно считать, что  $\varepsilon > 0$  и матрица  $\mathbf{B}$  такова, что задача Коши для уравнения (2) является корректной [9].

Очевидно, что решения задачи Коши для (2) зависят от вида матрицы  $\mathbf{B}$  (и, вообще, от вида регуляризирующего оператора в правой части (2)). Возникает вопрос: для каких систем (1) предел решений (2) при  $\varepsilon \rightarrow 0$  не будет зависеть от вида матрицы  $\mathbf{B}$ ? Для ответа рассмотрим малую окрестность точки  $x$ , лежащей на линии разрыва. В этой окрестности можно считать в первом приближении, что решение задачи (1) зависит только от величины  $\xi = x - \sigma t$ , где уравнение  $x - \sigma t = \text{const}$  определяет касательную к линии разрыва на плоскости  $Oxt$ . Для гладких решений системы (2) также будем считать, что они зависят от переменной  $\xi$ . Такие решения называются *решениями типа бегущей волны*.

Положим  $\mathbf{q}_\varepsilon = \mathbf{q}_\varepsilon(\xi)$ , где  $\mathbf{q}_\varepsilon$  – решение (2),  $\xi = x - \sigma t$ . Подставив его в (2), получим

$$-\sigma \frac{\partial \mathbf{q}_\varepsilon}{\partial \xi} + \mathbf{A} \frac{\partial \mathbf{q}_\varepsilon}{\partial \xi} = \varepsilon \frac{\partial}{\partial \xi} \left[ \mathbf{B} \frac{\partial \mathbf{q}_\varepsilon}{\partial \xi} \right]. \quad (3)$$

Непосредственной проверкой можно убедиться в том, что решение этой системы будет иметь вид

$$\mathbf{q}_\varepsilon(\xi) = \mathbf{q}_1(\xi/\varepsilon). \quad (4)$$

Таким образом, решения задачи при различных  $\varepsilon$  являются перепараметризацией друг друга.

Предположим, что система (3) имеет решение, удовлетворяющее граничным условиям

$$\mathbf{q}_\varepsilon(\xi) = \mathbf{q}^\pm, \quad d\mathbf{q}_\varepsilon/d\xi \rightarrow 0 \quad \text{при} \quad \xi \rightarrow \pm\infty. \quad (5)$$

Отметим, что вопрос о существовании решения типа бегущей волны является нетривиальным (см. [10]).

Из предыдущих построений видно, что предельное при  $\varepsilon \rightarrow 0$  решение  $\mathbf{q}$  будет разрывным и удовлетворяет условиям

$$\mathbf{q}(\xi) = \begin{cases} \mathbf{q}^-, & \xi < 0, \\ \mathbf{q}^+, & \xi > 0. \end{cases} \quad (6)$$

Далее, проинтегрировав систему (3) по переменной  $\xi$  от  $-\infty$  до  $+\infty$  с учётом граничных условий (5), получим

$$-\sigma(\mathbf{q}^+ - \mathbf{q}^-) = - \int_{-\infty}^{+\infty} \mathbf{A}(\mathbf{q}_\varepsilon) \frac{d\mathbf{q}_\varepsilon}{d\xi} d\xi. \quad (7)$$

Иначе (7) можно записать в виде

$$-\sigma(\mathbf{q}^+ - \mathbf{q}^-) = - \int_{\Gamma} \mathbf{A}(\mathbf{q}_\varepsilon) d\mathbf{q}_\varepsilon, \quad (8)$$

где  $\Gamma$  – зависящий от матрицы  $\mathbf{B}$  путь в фазовом пространстве решений, “соединяющий” точки  $\mathbf{q}^-$  и  $\mathbf{q}^+$  в смысле выполнения граничных условий (5).

Выражения (7) и (8) – тождества, верные для решения системы (3). Их правые части зависят как от  $\varepsilon$ , так и от  $\mathbf{B}$ . Для того чтобы всё выражение не зависело от  $\mathbf{B}$  при  $\varepsilon \rightarrow 0$ ,

необходимо, чтобы это было справедливо для его правой части, что возможно в случае, когда подынтегральное выражение является полным дифференциалом, т.е.

$$\mathbf{A}(\mathbf{q}_\varepsilon) d\mathbf{q}_\varepsilon \equiv d\mathbf{F}(\mathbf{q}_\varepsilon), \quad \mathbf{A}(\mathbf{q}_\varepsilon) \frac{d\mathbf{q}_\varepsilon}{d\xi} \equiv \frac{d\mathbf{F}}{d\mathbf{q}_\varepsilon} \frac{d\mathbf{q}_\varepsilon}{d\xi} = \frac{d\mathbf{F}}{d\xi}$$

для подходящего вектора  $\mathbf{F}$ . Тогда условия (7) и (8) могут быть записаны как

$$-\sigma(\mathbf{q}^+ - \mathbf{q}^-) = - \int_{\Gamma} \mathbf{A}(\mathbf{q}_\varepsilon) d\mathbf{q}_\varepsilon = - \int_{\Gamma} d\mathbf{F} = -(\mathbf{F}(\mathbf{q}^+) - \mathbf{F}(\mathbf{q}^-)),$$

или просто

$$-\sigma(\mathbf{q}^+ - \mathbf{q}^-) = -(\mathbf{F}(\mathbf{q}^+) - \mathbf{F}(\mathbf{q}^-)), \tag{9}$$

а квазилинейная система (1) примет вид

$$\frac{\partial \mathbf{q}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = 0. \tag{10}$$

Система (10) имеет вид закона сохранения, при этом величины  $\mathbf{q}$  называются консервативными переменными,  $\mathbf{F}$  – вектором плотности потока; соотношение (9) выражает стандартного вида условие Ранкина–Гюгонио для консервативных систем.

Таким образом, консервативные гиперболические системы – это такие квазилинейные системы, для которых обобщённые решения, построенные как вязкие пределы, не зависят от вида регуляризующего оператора (по крайней мере, если он имеет использованный выше вид). Эта характеристика гиперболических законов сохранения (консервативных гиперболических систем уравнений) является крайне важной. Благодаря ей отпадает вопрос, какой именно регуляризующий оператор используется для построения решения задачи. С точки зрения разностных схем это позволяет не заботиться, до известной степени, о виде искусственной вязкости при построении разностной схемы – не важно, вводится ли она явно или неявно (и “видна” только в первом дифференциальном приближении).

С другой стороны, для неконсервативных систем корректность постановки задачи о поиске разрывных решений уравнения (1) возможна, только если для заданной квазилинейной системы уравнений явно указывается регуляризующий оператор. При этом пределы гладких решений для разных матриц  $\mathbf{B}$  при  $\varepsilon \rightarrow 0$  порождают различные классы обобщённых решений. Не задавая конкретный вид  $\mathbf{B}$ , невозможно ни формально определить обобщённые решения, ни указать их конкретный класс.

**1.1. Условия Ранкина–Гюгонио.** Фундаментальным следствием консервативной постановки (10) являются так называемые условия Ранкина–Гюгонио на разрывах. В частности, они позволяют провести классификацию разрывов и используются при построении решения задачи Римана. С физической точки зрения они выражают условия непрерывности консервативных величин на разрывах. Для системы (10) их можно записать в точности в виде (9). Для неконсервативной системы (1) их роль играет соотношение (8).

Обратим внимание на следующий факт. Как уже отмечалось, при различных  $\varepsilon$  решения возмущённой задачи отличаются лишь параметризацией, (см. (4)). Тогда

$$\frac{d\mathbf{q}_\varepsilon}{d\xi} = \left. \frac{d\mathbf{q}_1(\alpha)}{d\alpha} \right|_{\alpha=\xi/\varepsilon}$$

и значение в правой части интегралов (7) или (8) *не зависит* от конкретного значения параметра  $\varepsilon$ . Это явно говорит о том, что условия (7) или (8) являются свойством квазилинейной системы (1) (при заданном выборе способа регуляризации, т.е. оператора  $\mathbf{B}$ ), а не системы (3) при заданном  $\varepsilon$ .

Идеи, изложенные выше, обеспечили развитие теории DLM, которая в настоящее время является широко распространённым формализмом для анализа разрывных решений квазилинейных неконсервативных систем гиперболических уравнений.

Необходимые для дальнейшего элементы этой теории изложены далее.

**1.2. Теория DLM.** Основы современной теории квазилинейных неконсервативных гиперболических систем уравнений были заложены в фундаментальных работах А.И. Вольперта (см. [11]) и в дальнейшем приобрели законченный вид в цикле работ Г. Даля Мазо, Ф. Лефлонга и Ф. Мурата [12, 13].

С абстрактной математической точки зрения теория DLM позволяет корректно построить элементы анализа в пространствах BV функций с ограниченными вариациями (под которыми далее можно понимать кусочно-дифференцируемые в классическом смысле функции с конечным числом точек разрыва). Её основным приложением является строгая теория обобщённых решений квазилинейных неконсервативных систем гиперболических уравнений, а основным теоретическим результатом – корректное определение “неконсервативных произведений”, что позволяет найти значения недивергентных квазилинейных дифференциальных операторов на разрывных функциях из пространства BV.

Поскольку нас интересуют в основном приложения теории DLM к анализу квазилинейных неконсервативных гиперболических уравнений, строгое изложение теории ниже даваться не будет. Вместо этого покажем, каким образом её основные определения могут быть получены как простые обобщения построений, записанных выше.

Используем построения п. 1.1 и обратим внимание, что выражение  $A(q_\varepsilon) \times \partial q_\varepsilon / \partial x$  при  $\varepsilon \rightarrow 0$  можно формально рассматривать как определение неконсервативного произведения  $A(q) \partial q / \partial x$ . При этом первое выражение корректно определено в силу того, что решение  $q_\varepsilon$  гладкое, и приближает второе в силу соответствия между граничными условиями (5) и условием (6). Более того, как было показано, значение выражения

$$-\int_{\Gamma} A(q_\varepsilon) dq_\varepsilon \equiv -\int_{\Gamma} A(q_1) dq_1$$

не зависит явно от величины  $\varepsilon$  и регуляризующего оператора  $B$  (хотя само  $q_\varepsilon$ , конечно же, зависит от обоих). Эти соображения позволяют “забыть” о том, что  $q_\varepsilon$  – решение задачи (3), и написать

$$A(q) \frac{\partial q}{\partial x} = \lim_{\varepsilon \rightarrow 0} A(q_\varepsilon) \frac{\partial q_\varepsilon}{\partial x},$$

где  $q_\varepsilon$  – какой-либо путь в пространстве решений, удовлетворяющий условиям (5), но не обязательно удовлетворяющий уравнению (3). Строгая теория говорит, что сходимость в последнем пределе следует понимать в смысле сходимости мер. Далее такой путь будем обозначать как  $\psi = \psi(s; q^-, q^+)$ . Соответствующее ему  $q_\varepsilon$  определим как

$$q_\varepsilon(x) = \begin{cases} q^-, & x \in (a, x_0 - \varepsilon), \\ \psi\left(\frac{x - x_0 + \varepsilon}{2\varepsilon}; q^-, q^+\right), & x \in (x_0 - \varepsilon, x_0 + \varepsilon), \\ q^+, & x \in (x_0 + \varepsilon, b), \end{cases}$$

где  $x_0 \in (a, b)$  – координата точки разрыва,  $[a, b]$  – отрезок, содержащий разрыв. Таким образом определённое  $q_\varepsilon$  не является решением (3), но обладает всеми требуемыми свойствами, а именно,  $q_\varepsilon$  интерполирует между значениями  $q^-$  и  $q^+$  на отрезке длины  $2\varepsilon$  и является достаточно гладким.

Для такого  $q_\varepsilon$  можно показать, что при  $\varepsilon \rightarrow 0$

$$A(q_\varepsilon) \frac{\partial q_\varepsilon}{\partial x} \rightarrow C \delta(x - x_0), \quad C = \int_{\Gamma} A(\psi(s; q^-, q^+)) \frac{\partial \psi(s; q^-, q^+)}{\partial s} ds, \tag{11}$$

где сходимость понимается в смысле сходимости мер,  $\delta(x - x_0)$  – мера Дирака в точке  $x_0$ . Соотношение (11) является определением неконсервативного произведения в теории DLM. Отметим, что в качестве  $\psi$  естественно, но не обязательно, можно взять, при нужном выборе

параметризации, решение  $q_1(\xi)$  задачи (3). Так же легко видеть, что конкретная параметризация  $\psi$  не важна, так как интеграл в (11) от неё не зависит. Поэтому обычно говорят о *семействе* путей  $\psi$ , а сам путь  $\Gamma$  параметризуют отрезком  $[0, 1]$ .

Пусть  $\psi$  является отображением  $\psi : [0, 1] \times \Omega \times \Omega \mapsto \Omega$ , удовлетворяющим условию интерполяции  $\psi(0; q^-, q^+) = q^-$ ,  $\psi(1; q^-, q^+) = q^+$  и ряду условий гладкости, которые здесь не формулируем (см., например, [12]). Отметим, что если задать путь линейным, т.е. положить

$$\psi(s; q^-, q^+) = q^- + s(q^+ - q^-), \tag{12}$$

то теория DLM воспроизведёт теорию А.И. Вольперта.

Можно показать, что в рамках теории DLM могут быть получены соотношения Ранкина–Гюгиони, полностью аналогичные описанным в п. 1.1, а именно (ср. с (7) и (8))

$$-\sigma(q^+ - q^-) = - \int_0^1 A(\psi(s; q^-, q^+)) \frac{\partial \psi(s; q^-, q^+)}{\partial s} ds. \tag{13}$$

Подчеркнём ещё раз, что задание конкретного пути либо конкретного регуляризующего оператора является необходимым для корректной постановки задачи (1). Более того, информация о пути не содержится в том или ином виде непосредственно в системе уравнений (1); она является внешней по отношению к ней и должна задаваться из сторонних соображений.

Понятно, что задав определённое  $B$ , мы тем самым задаём вполне определённое  $\psi$ . Вместе с тем, задав определённое  $\psi$ , не очевиден способ указать вид оператора  $B$  такой, чтобы заданное  $\psi$  удовлетворяло бы уравнениям (3) (и (2)) (и непонятно, можно ли сделать это в общем случае). Считать (для целей дальнейшего прикладного анализа) первичной величиной  $\psi$  или  $B$  – вопрос удобства. По крайней мере можно сказать, что при построении теории часть свойств для  $\psi$ , если это необходимо, могут быть *постулированы* – и это проще, чем задать свойства  $B$ , которые бы обеспечивали выполнение нужных свойств для  $\psi$ . С другой стороны, при анализе, например, разностных схем методом дифференциальных приближений удобнее сравнивать свойства регуляризующего оператора в (2) и в первом дифференциальном приближении разностной схемы, применяемой для решения уравнения (1).

Однако теория DLM не является единственным способом определения неконсервативных произведений. Среди них отметим подход работы [14], посвящённой общим вопросам определения произведений обобщённых функций, и изложенную в [15] теорию, основанную на аппроксимации специальных ударных кривых.

Обратим внимание на следующее. В классической теории консервативных гиперболических систем известны два подхода, широко используемых для их анализа и построения численных алгоритмов: первый основан на концепции вязкого приближения, энтропийных функций и энтропийных неравенств; второй изучает свойства таких систем в терминах характеристик системы уравнений. Эти два подхода не являются альтернативными (в исключаяющем смысле), а согласованы и непротиворечиво дополняют друг друга (см. [16]). Теория DLM является “удобной” в том смысле, что позволяет сделать то же самое для неконсервативных систем уравнений.

В работе [13] было получено, что для строго гиперболической системы уравнений с линейно-вырожденными или истинно нелинейными полями задача Римана для (1) при достаточно малом скачке в начальных данных имеет единственное решение со стандартной “лаксонской” структурой, т.е. представляет собой  $n + 1$  постоянное состояние, разделённое центрированными волнами разрежения, контактными разрывами и ударными волнами при некоторых дополнительных условиях, накладываемых на путь. Отличие от консервативного случая заключается только в том, что состояния на ударной волне связаны условиями Ранкина–Гюгиони в виде (13), а не (9).

**2. Разностные схемы.** В настоящее время известно множество приложений, в которых необходимо применение математических моделей, представляющих собой неконсервативные квазилинейные гиперболические системы уравнений. В качестве примера назовём широко распространённые уравнения типа Баера–Нунциато для описания динамики многофазных сред

самой различной природы [17]. Эта система уравнений содержит в себе уравнения для описания эволюции объёмных долей фаз, которые необходимо являются неконсервативными. По этой причине построению эффективных вычислительных алгоритмов для решения систем уравнений (1) уделяется значительное внимание.

Понятно, что разработка такого алгоритма является нетривиальной задачей – если, например, использовать формальные разностные аппроксимации уравнения (1), то такие аппроксимации не будут ничего “знать” о пути или виде диссипативного оператора, выделяющих единственный класс разрывных решений задачи. На практике в этом случае может отсутствовать сеточная сходимость численных решений разностной схемы к “точному” решению задачи (см., например, [1]). Более того, указанная проблема существует даже в том случае, если для построения аппроксимаций используются подходы, формально учитывающие заданный вид пути. К последним относятся, в частности, так называемые *формально консервативные по пути схемы* (“formally path-conservative schemes”), которые обобщают классический метод Годунова для рассматриваемого класса задач. Для системы уравнений (1) консервативные по пути схемы более подробно рассматриваются в п. 3. Здесь лишь отметим, что в рамках указанного формализма были предложены, например, схемы типа Лакса–Фридрихса [1], Рое [18, 19], HLL [19] и др.

Несмотря на то, что конкретный вид пути явно учитывается при построении разностных схем указанного класса, они всё равно не свободны от указанных выше проблем – во многих случаях их решение не сходится к заданному точному решению задачи. Возможное объяснение этому эффекту было дано в работе [1]. Его суть заключается в следующем: первое дифференциальное приближение разностной схемы, в том числе консервативной по пути, включает в себя диссипативные слагаемые, отличающиеся от использованных при определении обобщённого решения задачи. Это приводит к тому, что при измельчении шага расчётной сетки численные решения сходятся к такому классу обобщённых решений, который определяется чисто схемным диссипативным оператором (или порождённым им путём). Таким образом, для построения корректной разностной схемы мало формально учитывать заданный вид пути, необходимо ещё обеспечить заданный вид первого дифференциального приближения схемы.

Отметим, что такой проблемы принципиально не возникает при построении аппроксимаций консервативных гиперболических систем (как указано в п. 1), для консервативных систем вид диссипативного оператора до известной степени не важен. По этой причине одним из возможных направлений в построении разностных схем для неконсервативных систем гиперболических уравнений является построение разностных схем с априорно заданным схемным диссипативным оператором. Примеры построения таких схем (в частности, в классе энтропийно-устойчивых схем) предложены в работах [4–7].

Тем не менее, как будет показано далее, вопросы, связанные со сходимостью разностных схем для неконсервативных гиперболических систем уравнений, могут являться более тонкими. Для этого ниже рассмотрен конкретный пример. В п. 3 опишем конкретный вариант консервативной по пути схемы, используемой в дальнейших численных расчётах.

**3. Вариант метода HLL.** Рассмотрим вариант метода HLL, который относится к классу формально консервативных по пути схем. Он позволяет формально использовать произвольный заданный аналитически путь и обобщает вариант метода, предложенный в работе [19], где рассматривается только случай линейного пути.

Рассмотрим, как и ранее, систему (1). Пусть  $q \in \Omega$ ,  $\Omega \subset \mathbb{R}^n$  – область значений решения задачи. Для упрощения будем считать, что для задачи (1) рассматривается задача Коши на всей вещественной прямой, т.е.  $x \in \mathbb{R}$ .

Введём разностную сетку с узлами  $x_i$ ,  $i = 0, \pm 1, \pm 2, \dots$ . Узлы сетки будем относить к центрам расчётных ячеек  $\omega_i$ ,  $\omega_i = [x_{i-1/2}, x_{i+1/2}]$ ,  $x_{i+1/2} = (x_i + x_{i+1})/2$ . Далее величины, отнесённые к ячейкам или к их центрам, будем обозначать с использованием целых индексов  $i$ ; к границам ячеек – с использованием полуцелых индексов  $i + 1/2$ . Для удобства шаги  $\Delta x = x_{i+1} - x_i$  и  $\Delta t$  по пространству и по времени, соответственно, считаем постоянными.

Для численного решения задачи (1) будем использовать разностную схему вида

$$q_i^{n+1} = q_i^n - \frac{\Delta t}{\Delta x} (D_{i-1/2}^- + D_{i+1/2}^+), \quad D_{i+1/2}^\pm = D^\pm(q_{i-s}^n, \dots, q_{i+p}^n), \quad (14)$$

где верхний индекс  $n$  обозначает номер временного шага,  $D^\pm$  – непрерывные функции из  $\Omega^{p+q+1}$  в  $\Omega$ , которые удовлетворяют условию  $D^\pm(\mathbf{q}, \dots, \mathbf{q}) = 0$  для всех  $\mathbf{q} \in \Omega$ . Последнее условие гарантирует, что рассмотренная схема точно аппроксимирует постоянные решения.

Схема вида (14) называется *формально консервативной по пути  $\psi$*  при условии, что выполняется соотношение [20]

$$D^-(\mathbf{q}_{-s}, \dots, \mathbf{q}_p) + D^+(\mathbf{q}_{-s}, \dots, \mathbf{q}_p) = \int_0^1 \mathbf{A}(\psi(s; \mathbf{q}_0, \mathbf{q}_1)) \frac{\partial \psi(s; \mathbf{q}_0, \mathbf{q}_1)}{\partial s} ds \tag{15}$$

для всех  $\mathbf{q}_i \in \Omega$  и  $i = -s, \dots, p$ .

Построим теперь консервативную по пути схему типа HLL. Представленный ниже вывод в целом повторяет рассмотренный в [19], однако не делает предположений о линейности пути.

Будем считать, что неконсервативное произведение записано в виде

$$\mathbf{A} \frac{\partial \mathbf{q}}{\partial x} = \frac{\partial \mathbf{F}}{\partial x} + \mathbf{C} \frac{\partial \mathbf{q}}{\partial x}.$$

Рассмотрим задачу Римана для системы (1) с состояниями  $\mathbf{q}_L$  и  $\mathbf{q}_R$  слева и справа от разрыва соответственно. Будем считать, что её решение имеет автомодельный вид, т.е.  $\mathbf{q}(x, t) = \mathbf{q}(\xi)$ , где  $\xi = x/t$ . Тогда систему уравнений (1) можно записать как

$$\mathbf{q} - \frac{\partial(\xi \mathbf{q})}{\partial \xi} + \frac{\partial \mathbf{F}}{\partial \xi} + \mathbf{C}(\mathbf{q}) \frac{\partial \mathbf{q}}{\partial \xi} = 0. \tag{16}$$

Приближённое решение задачи Римана будем представлять в виде

$$\mathbf{q}(\xi) = \begin{cases} \mathbf{q}_L, & \xi \leq s_L, \\ \mathbf{q}_*, & s_L < \xi < s_R, \\ \mathbf{q}_R, & \xi \geq s_R, \end{cases} \tag{17}$$

где  $\mathbf{q}_{L,R}$  – левое и правое состояния решения на разрыве,  $\mathbf{q}_*$  – некоторое промежуточное состояние, которое будет определено в дальнейшем,  $s_L = s_L(\mathbf{q}_L, \mathbf{q}_R)$  и  $s_R = s_R(\mathbf{q}_L, \mathbf{q}_R)$  – оценки минимального и максимального собственных чисел матрицы  $\mathbf{A}$ , причём считается, что  $s_L < 0 < s_R$  по допущению.

Проинтегрировав равенство (16) на отрезке  $\xi \in [s_L, s_R]$  с учётом (17), получим

$$\mathbf{q}_*(s_R - s_L) - (\mathbf{q}_R s_R - \mathbf{q}_L s_L) + (\mathbf{F}_R - \mathbf{F}_L) + \int_{\mathbf{q}_L}^{\mathbf{q}_R} \mathbf{C}(\mathbf{q}) d\mathbf{q} = 0, \tag{18}$$

где

$$\int_{\mathbf{q}_L}^{\mathbf{q}_R} \mathbf{C}(\mathbf{q}) d\mathbf{q} = \int_0^1 \mathbf{C}(\psi(s; \mathbf{q}_L, \mathbf{q}_*)) \frac{\partial \psi(s; \mathbf{q}_L, \mathbf{q}_*)}{\partial s} ds + \int_0^1 \mathbf{C}(\psi(s; \mathbf{q}_*, \mathbf{q}_R)) \frac{\partial \psi(s; \mathbf{q}_*, \mathbf{q}_R)}{\partial s} ds.$$

Введём обозначение

$$\tilde{\mathbf{V}}(\mathbf{q}_a, \mathbf{q}_b) = \int_0^1 \mathbf{C}(\psi(s; \mathbf{q}_a, \mathbf{q}_b)) \frac{\partial \psi(s; \mathbf{q}_a, \mathbf{q}_b)}{\partial s} ds.$$

Тогда (18) можно записать в виде

$$\mathbf{q}_*(s_R - s_L) - (\mathbf{q}_R s_R - \mathbf{q}_L s_L) + (\mathbf{F}_R - \mathbf{F}_L) + \tilde{\mathbf{V}}(\mathbf{q}_L, \mathbf{q}_*) + \tilde{\mathbf{V}}(\mathbf{q}_*, \mathbf{q}_R) = 0. \tag{19}$$

В силу того, что в последнем соотношении величины  $\mathbf{q}_{L,R}$  и путь  $\psi$  заданы и известны, его можно рассматривать как нелинейное алгебраическое уравнение относительно неизвестного значения  $\mathbf{q}_*$ . При реализации метода оно решается подходящим способом численно. В качестве начального приближения удобно использовать оценку  $\mathbf{q}_*$ , полученную по формуле

$$\tilde{\mathbf{q}}^* = \frac{1}{s_R - s_L} [\mathbf{q}_R s_R - \mathbf{q}_L s_L - (\mathbf{F}_R - \mathbf{F}_L) - \tilde{\mathbf{V}}(\mathbf{q}_R, \mathbf{q}_L)].$$

Это уравнение соответствует приближению, когда промежуточное состояние  $\mathbf{q}$  не учитывается при вычислении последнего интеграла в (18).

Далее, проинтегрировав (16) по  $\xi$  на промежутках  $[s_L, 0]$  и  $[0, s_R]$ , получим

$$-\mathbf{q}_* s_L + \mathbf{q}_L s_L + \int_{s_L}^0 \mathbf{A}(\mathbf{q}) \frac{\partial \mathbf{q}}{\partial \xi} d\xi = 0, \quad \mathbf{q}_* s_R - \mathbf{q}_R s_R + \int_0^{s_R} \mathbf{A}(\mathbf{q}) \frac{\partial \mathbf{q}}{\partial \xi} d\xi = 0.$$

Интегралы в последних выражениях естественно принять за “потоки”  $D_{HLL}^-$  и  $D_{HLL}^+$  соответственно. Выразив их явно, получим

$$D_{HLL}^- = +\mathbf{q}_* s_L - \mathbf{q}_L s_L, \quad D_{HLL}^+ = -\mathbf{q}_* s_R + \mathbf{q}_R s_R.$$

Далее необходимо в выражения для  $D_{HLL}^\pm$  подставить значения  $\mathbf{q}_*$ , определяемые решением уравнения (19). В результате придём к соотношениям

$$D_{HLL}^-(\mathbf{q}_L, \mathbf{q}_R) = -\frac{s_L}{s_R - s_L} (\mathbf{F}_R - \mathbf{F}_L + \tilde{\mathbf{V}}(\mathbf{q}_L, \mathbf{q}_*) + \tilde{\mathbf{V}}(\mathbf{q}_*, \mathbf{q}_R)) + \frac{s_L s_R}{s_R - s_L} (\mathbf{q}_R - \mathbf{q}_L),$$

$$D_{HLL}^+(\mathbf{q}_L, \mathbf{q}_R) = \frac{s_R}{s_R - s_L} (\mathbf{F}_R - \mathbf{F}_L + \tilde{\mathbf{V}}(\mathbf{q}_L, \mathbf{q}_*) + \tilde{\mathbf{V}}(\mathbf{q}_*, \mathbf{q}_R)) - \frac{s_L s_R}{s_R - s_L} (\mathbf{q}_R - \mathbf{q}_L). \quad (20)$$

Значение  $\mathbf{q}_*$  в этом случае определяется как решение уравнения (19). Однако на практике оказывается, что в большинстве рассмотренных авторами случаев оценка  $\tilde{\mathbf{q}}_*$  является достаточно точной для того, чтобы использовать её в выражениях (20) непосредственно, без последующего уточнения. Такой вариант метода является существенно более экономичным, так как для большинства ячеек расчётной сетки решения уравнения (19) не требуется.

Наконец, положим в схеме (14)

$$D_{i-1/2}^- = D_{HLL}^-(\mathbf{q}_{i-1}, \mathbf{q}_i), \quad D_{i+1/2}^+ = D_{HLL}^+(\mathbf{q}_i, \mathbf{q}_{i+1}). \quad (21)$$

Непосредственной проверкой можно убедиться, что справедливо следующее

**Утверждение.** *Разностная схема (14) с “потоками”, определёнными в соответствии с (21), является формально консервативной по пути в смысле справедливости соотношения (15).*

**Доказательство.** Просуммировав величины  $D_{HLL}^\pm$  исходя из их определения, получим

$$D_{HLL}^- + D_{HLL}^+ = \mathbf{F}_R - \mathbf{F}_L + \int_0^1 \mathbf{C}(\psi(s; \mathbf{q}_L, \mathbf{q}_R)) \frac{\partial \psi(s; \mathbf{q}_L, \mathbf{q}_R)}{\partial s} ds = \int_0^1 \mathbf{A}(\psi(s; \mathbf{q}_L, \mathbf{q}_R)) \frac{\partial \psi(s; \mathbf{q}_L, \mathbf{q}_R)}{\partial s} ds.$$

#### 4. Неконсервативная система уравнений Бюргерса. Диссипативное возмущение.

В качестве модельной задачи, на примере которой часто иллюстрируются вопросы сходимости разностных схем для неконсервативных гиперболических систем уравнений, рассмотрим модельную систему уравнений Бюргерса [21, 6]

$$\frac{\partial \mathbf{q}}{\partial t} + \mathbf{A}(\mathbf{q}) \frac{\partial \mathbf{q}}{\partial x} = 0, \quad \mathbf{q} = \begin{pmatrix} u \\ v \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} u & u \\ v & v \end{pmatrix}. \quad (22)$$

В [21] разрывные решения системы (22) определены путём введения диссипативного слагаемого вида

$$\frac{\partial \mathbf{q}}{\partial t} + \mathbf{A} \frac{\partial \mathbf{q}}{\partial x} = \mathbf{D} \frac{\partial^2 \mathbf{q}}{\partial x^2}, \quad \mathbf{D} = \begin{pmatrix} \varepsilon_1 & \varepsilon_1 \\ \varepsilon_2 & \varepsilon_2 \end{pmatrix}, \quad (23)$$

где  $\varepsilon_{1,2} > 0$  – малые параметры. Разрывные решения системы (22) определяются как пределы решений системы (23) при  $\varepsilon_{1,2} \rightarrow 0$ .

В работе [21] показано, что решения уравнения (23) типа бегущей волны с начальными данными (6) существуют только в том случае, если переменные  $\mathbf{q}^\pm$  удовлетворяют соотношениям

$$v_+ = \frac{\varepsilon_2}{\varepsilon_1 + \varepsilon_2} (2\sigma - u_- - v_-) + \frac{\varepsilon_1 v_- - \varepsilon_2 u_-}{\varepsilon_1 + \varepsilon_2} \exp\left(2 - 2 \frac{u_- + v_-}{\sigma}\right), \quad u_+ = 2\sigma - u_- - (v_- + v_+), \quad (24)$$

где  $\sigma$  – скорость бегущей волны, которая, в свою очередь, не может быть произвольной и удовлетворяет условию

$$\frac{u_- + v_-}{2} \leq \sigma \leq u_- + v_-.$$

Непосредственно профиль решения типа бегущей волны в рассматриваемой задаче не может быть получен аналитически. Для его определения поступим следующим образом. Исследуем структуру разрыва, которая удовлетворяет соотношениям (24). Запишем систему уравнений (23) в эквивалентной форме:

$$\begin{aligned} \frac{\partial}{\partial t} (u + v) + \frac{\partial}{\partial x} \left( \frac{[u + v]^2}{2} \right) &= (\varepsilon_1 + \varepsilon_2) \frac{\partial^2 (u + v)}{\partial x^2}, \\ \frac{\partial}{\partial t} \left( \frac{v}{u + v} \right) &= \left( \frac{\varepsilon_1}{\varepsilon_1 + \varepsilon_2} - \frac{v}{u + v} \right) \frac{\varepsilon_1 + \varepsilon_2}{u + v} \frac{\partial^2 (u + v)}{\partial x^2}. \end{aligned} \quad (25)$$

Будем искать решение (25) вида  $u = u(\xi)$  и  $v = v(\xi)$ , где  $\xi = x - \sigma t$ , с краевыми условиями

$$\mathbf{q} \rightarrow \mathbf{q}^\pm, \quad \frac{d\mathbf{q}}{d\xi} \rightarrow 0 \quad \text{при} \quad \xi \rightarrow \pm\infty. \quad (26)$$

В этом случае система уравнений (25) принимает вид

$$\begin{aligned} -\sigma \frac{d(u + v)}{d\xi} + \frac{d}{d\xi} \left( \frac{[u + v]^2}{2} \right) &= (\varepsilon_1 + \varepsilon_2) \frac{d^2(u + v)}{d\xi^2}, \\ -\sigma \frac{d}{d\xi} \left( \frac{v}{u + v} \right) &= \left( \frac{\varepsilon_1}{\varepsilon_1 + \varepsilon_2} - \frac{v}{u + v} \right) \frac{\varepsilon_1 + \varepsilon_2}{u + v} \frac{d^2(u + v)}{d\xi^2}. \end{aligned} \quad (27)$$

Сделаем замену переменных  $w = u + v$  в (27) и введём переменную  $\tilde{w} = dw/d\xi$ . Тогда из (27) получим систему обыкновенных дифференциальных уравнений первого порядка относительно неизвестных  $(w, \tilde{w}, v)$  при заданном значении скорости разрыва  $\sigma$ :

$$\frac{dw}{d\xi} = \tilde{w}, \quad \frac{d\tilde{w}}{d\xi} = \tilde{w} \frac{w - \sigma}{\varepsilon_1 + \varepsilon_2}, \quad \frac{dv}{d\xi} = \frac{v}{w} \tilde{w} - \frac{w - \sigma}{\sigma} \left( \frac{\varepsilon_1}{\varepsilon_1 + \varepsilon_2} - \frac{v}{w} \right) \tilde{w}. \quad (28)$$

Эта система уравнений должна быть дополнена граничными условиями, которые являются следствием условий (26). Заметим, что состояния  $\mathbf{Q}_+ = (u_+ + v_+, 0, v_+)$  и  $\mathbf{Q}_- = (u_- + v_-, 0, v_-)$  являются её положениями равновесия. Тогда граничные условия для системы (28) имеют вид

$$\mathbf{Q} \rightarrow \mathbf{Q}_\pm \quad \text{при} \quad \xi \rightarrow \pm\infty, \quad (29)$$

где  $\mathbf{Q} = (w, \tilde{w}, v)$ .

Чтобы задача о структуре разрыва имела решение, необходимо существование интегральной кривой, удовлетворяющей (28) и краевым условиям (29), т.е. соединяющей два положения равновесия.

Описанную выше задачу для системы уравнений (28) с краевыми условиями (29) будем называть *задачей о структуре разрыва* [22, с. 84]. Используемый в работе способ её численного решения представляет собой вариант метода стрельбы и предполагает следующую последовательность шагов.

1. Выбрать “опорное” положение равновесия  $Q_0$ , например,  $Q_0 = Q_+$ .
2. Выбрать направление “стрельбы” как единичный вектор  $n \in \mathbb{R}^3$  в пространстве  $Ow\tilde{w}v$ .
3. Построить численно (удобным способом) траекторию системы (28), проходящую через точку  $Q_\delta = Q_0 + n\delta$ , где  $\delta$  – малый параметр.
4. Проверить “попадание” построенной траектории в  $\delta$ -окрестность второго положения равновесия, т.е. точки  $Q_-$ . Если это условие выполнено, то считается, что требуемая интегральная кривая определена, в противном случае необходимо скорректировать направление вектора  $n$ .

Так как вычислительные затраты для решения системы (28) минимальны, решение задачи легко может быть найдено перебором траекторий, определяемых сравнительно небольшим числом векторов  $n$ .

На практике численное решение системы уравнений (28) с краевыми условиями (29) реализуется следующим образом. Определим векторы, по направлению которых будем отступать от положения равновесия  $Q_0$  следующим образом:

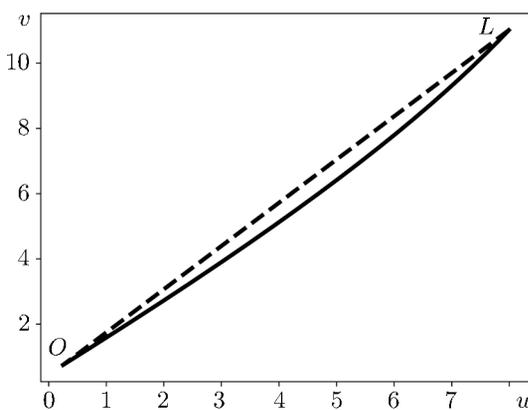
$$n_k = (n_k^1, n_k^2, n_k^3) = (\sin \theta_k \cos \phi_k, \sin \theta_k \sin \phi_k, \cos \theta_k),$$

где  $\theta_k = 2\pi k/k_{\max}$ ,  $\phi_k = 2\pi k/k_{\max}$ ,  $k = \overline{0, k_{\max}}$ , и положим  $k_{\max}$  в диапазоне от 6 до 20,  $\delta = 10^{-4}$ . Перебор этих направлений в рассматриваемых ниже задачах позволяет определить интегральную кривую (28) с заданной точностью  $\delta$ .

**5. Численные результаты.**

**5.1. Задача о структуре разрыва.** Рассмотрим задачу о структуре разрыва для неконсервативной системы уравнений Бюргера (23). Аналогично работе [6] выберем набор параметров

$$q^- = (7.99, 11.01), \quad q^+ = (0.25, 0.75), \quad \sigma = 10, \tag{30}$$



**Рис. 1.** Решение задачи (28) о структуре разрыва (сплошная линия) и прямая, соединяющая точки  $O$  и  $L$  (штриховая линия).

который удовлетворяет соотношениям (24).

В уравнении (23) положим  $\varepsilon_1 = \varepsilon_2 = 10^{-4}$ .

Численные эксперименты показывают, что в этом случае существует только одна такая интегральная кривая уравнения (28) с краевыми условиями (29), она показана на рис. 1 в плоскости переменных  $(u, v)$  (точка  $O$  соответствует значению  $q_+$ , точка  $L$  –  $q_-$ ). Напомним, что построенная структура разрыва определяет профиль решения задачи (23) и, как следствие, соответствующий путь.

**5.2. Путь.** Для построения в явном виде выражения для пути аппроксимируем интегральную кривую, представленную на рис. 1, полиномом

$$v = c_3u^3 + c_2u^2 + c_1u + c_0,$$

где  $c_3 = 0.00465794$ ,  $c_2 = -0.01583278$ ,  $c_1 = 1.14905787$ ,  $c_0 = 0.46234259$ . Тогда путь можно параметризовать следующим образом:

$$\psi_u(s) = u_+ + (1 - s)(u_- - u_+), \quad \psi_v(s) = c_3\psi_u(s)^3 + c_2\psi_u(s)^2 + c_1\psi_u(s) + c_0, \tag{31}$$

где параметр  $s \in [0, 1]$ . Уточним, что конкретная параметризация пути не имеет значения (см. п. 1).

**5.3. Соотношения Ранкина–Гюгоньо.** В общем случае, когда решение системы (22) определяется в рамках теории DLM заданным путём, условия Гюгоньо имеют вид (13), которые запишем здесь в виде

$$\sigma[\mathbf{q}] = \int_0^1 \mathbf{A}(\boldsymbol{\psi}(s; \mathbf{q}^-, \mathbf{q}^+)) \frac{\partial \boldsymbol{\psi}(s; \mathbf{q}^-, \mathbf{q}^+)}{\partial s} ds, \quad \boldsymbol{\psi} = \begin{pmatrix} \psi_u \\ \psi_v \end{pmatrix}. \quad (32)$$

В случае, если путь отождествляется с профилем решения типа бегущей волны для системы (23), последнее уравнение сводится в точности к системе соотношений (25).

Выбрав путь  $\boldsymbol{\psi}$ , подставим его в (32), проинтегрируем выражение (32) и исключим переменную  $\sigma$ . В итоге получим ударную адиабату – соотношение вида  $f(\mathbf{q}^-, \mathbf{q}^+) = 0$ , которая для фиксированного состояния  $\mathbf{q}^+$  до разрыва описывает возможные состояния  $\mathbf{q}^-$  после разрыва.

Рассмотрим два случая задания пути: линейный, определяемый соотношением (12), и построенный выше (31), соответствующий вязкой регуляризации вида (23). В качестве исходного выберем состояние  $\mathbf{q}^+ = (0.25, 0.75)$  (см. (30)).

На рис. 2 представлены ударные адиабаты. Точка  $O$  соответствует состоянию перед разрывом, точки  $L$  и  $M$  – состояния за разрывом для соответствующих ударных адиабат.

Как и следовало ожидать, состояния  $\mathbf{q}_{\pm}$ , определённые в соответствии с (30), лежат на одной ударной адиабате, соответствующей регуляризованной задаче (23). Обратим внимание, что ударная адиабата, соответствующая *линейному* пути, не проходит через точку  $L$ .

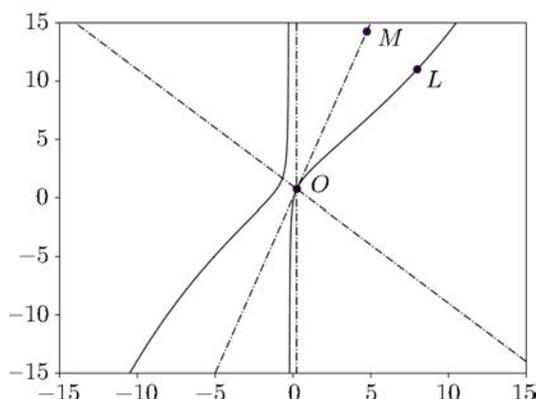
Вообще говоря, это означает следующее.

1. В решении задачи Римана о распаде разрыва для системы (22) в случае, если её обобщённые решения определены с помощью регуляризации (23) (или, эквивалентно, соответствующим путём), разрыв с параметрами (30) не “распадается”, а существует неограниченное время. Другими словами, в этом случае параметры (30) удовлетворяют условиям Ранкина–Гюгоньо для системы (22), и состояния на разрыве при  $t > 0$  совпадают с таковыми при  $t = 0$ .

2. В то же время если обобщённые решения (22) определены с помощью линейного пути тот же самый разрыв с параметрами (30) не удовлетворяет условиям Гюгоньо, распад начального разрыва действительно происходит и состояния на разрыве при  $t > 0$  становятся отличными от таковых при  $t = 0$ .

С точки зрения теории неконсервативных систем гиперболических уравнений это наблюдение является тривиальным фактом, демонстрирующим её основные особенности.

**5.4. Численное решение системы (22).** Как уже отмечалось, центральной и не до конца решённой проблемой построения разностных схем для неконсервативных систем уравнений является отсутствие сходимости построенных аппроксимаций к точному решению задачи. Традиционно это связывают с тем, что первое дифференциальное приближение разностной схемы включает в себя диссипативные слагаемые, которые порождают путь, отличающийся от использованного при постановке задачи (см. п. 1, а также [1, 6, 20, 23–25]). В п. 3 был приведён пример разностной схемы, которая относится к классу *формально* консервативных по пути (см. также библиографию в п. 2). Термин “формальный” обозначает, что использование при конструировании схемы явно заданного пути вовсе не гарантирует, что первое дифференциальное приближение схемы *фактически* реализует тот же самый путь, который был заложен при её конструировании.



**Рис. 2.** Ударная адиабата, соответствующая (31) (сплошная линия), и ударная адиабата, соответствующая линейному пути (штрихпунктирная линия).

Численно решим систему уравнений (22) с помощью консервативного вдоль пути численного метода HLL, представленного в п. 3. Рассмотрим задачу Римана о распаде начального разрыва с начальными данными (30). В начальный момент времени  $t = 0$  имеются два однородных состояния:  $q = O(u_+, v_+)$  при  $x > 2$  и  $q = L(u_-, v_-)$  при  $x < 2$ .

На рис. 3, а представлены результаты численных расчётов с линейным путём. Проанализируем соответствующую ударную адиабату, представленную на рис. 2 штрихпунктирной кривой. Точка  $O$  соответствует состоянию перед разрывом, точка  $L$  – состоянию за разрывом. Видно, что точка  $L$  не принадлежит ударной адиабате для линейного пути, поэтому начальный разрыв распался на два разрыва:  $O \rightarrow M$  и  $M \rightarrow L$ . Точка  $M$ , в свою очередь, принадлежит ударной адиабате для линейного пути. Таким образом, численный расчёт сошёлся к верному решению.

На рис. 3, б показаны результаты аналогичных расчётов, полученных по той же схеме, но с путём, определённым в соответствии с (31). Видно, что теперь решение совпадает с решением, соответствующим (24). Это согласовано с тем, что точка  $L$  принадлежит ударной адиабате для пути (31) (сплошная кривая на рис. 2).

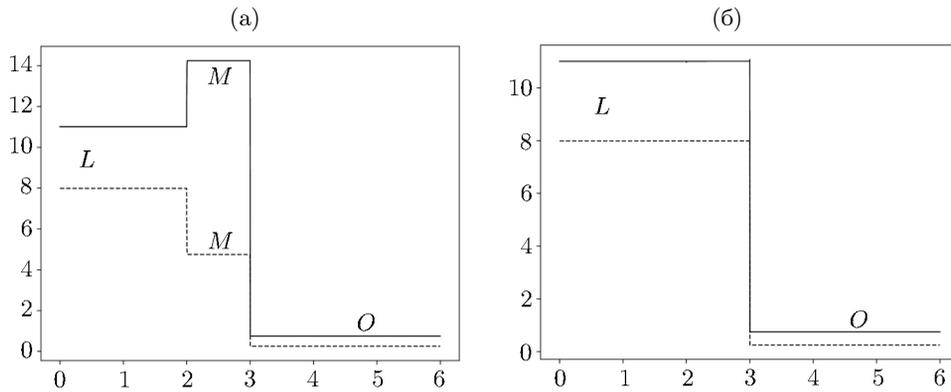


Рис. 3. Результаты численных расчётов для  $u$  (штриховая линия) и  $v$  (сплошная линия) на момент времени  $t = 0.1$  с.

Таким образом, первый вывод, который можно сделать, заключается в том, что для рассмотренного варианта метода HLL и по крайней мере конкретной постановки задач формально согласованная по пути схема из п. 3 является фактически согласованной.

Второй вывод, не менее важный, состоит в том, что несоответствие численных решений задач с линейным путём и путём (31) в рассматриваемом случае связано с тем, что для одной из них данные (30) удовлетворяют условиям Ранкина–Гюгонно, а для другой – нет, в силу чего происходит распад разрыва. Более того, параметры скачков после разрыва в численном решении оказываются правильными.

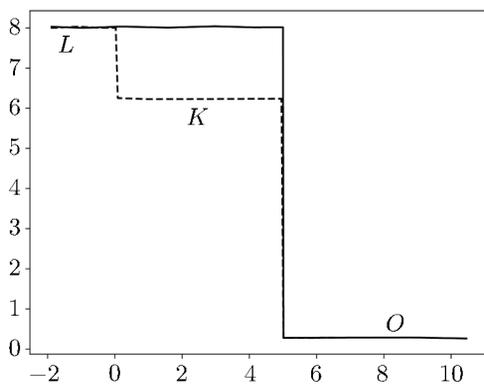


Рис. 4. Результаты численных расчётов из статьи [6] при  $t = 0.5$  с.

разрыв распался на два разрыва:  $O \rightarrow K$  и  $K \rightarrow L$ . Там указывается, что задача решалась численно консервативным по пути методом Годунова, а путь определялся вязким профилем расширенной системы уравнений.

Однако способ учёта такого пути, равно как и детали реализации, в этой статье не приводятся. При этом несоответствие численного и аналитического решений в [6] объясняется

Для примера рассмотрим иную интерпретацию полученных результатов. Так, рассматривая описанный выше тест для неконсервативной системы уравнений Бюргерса, авторы работ [6, 23] утверждают, что им не удалось добиться того, чтобы численное решение (22) сошло к решению, определяемому соотношениями (24). Результаты этих расчётов (конкретно из работы [6]) приведены на рис. 4. Видно, что начальный

наличием первого дифференциального приближения, которое выделяет другой класс решений. Результаты настоящей работы показывают, что в рассмотренном авторами [6] примере проблема со сходимостью связана с недопустимостью рассматриваемого разрыва для пути, заданного в численном методе.

Более того, эта проблема бы сохранилась, если бы авторы использовали полностью “правильные” схемы, которые являлись бы не формально, а *фактически* консервативными по пути (т.е. имели в своём первом дифференциальном приближении такой же диссипативный оператор, как и в дифференциальной задаче).

**Заключение.** В работе неконсервативная гиперболическая система уравнений Бюргерса исследовалась аналитически и численно консервативным по пути методом HLL. Показано, что при его корректном применении, по крайней мере для рассмотренной модельной задачи, он даёт правильное решение. По результатам решения задачи о структуре разрыва получены явные формулы для пути, отвечающего вязкому профилю расширенной неконсервативной системы Бюргерса [21].

В результате решения задачи Римана о распаде разрыва для модельной задачи в дифференциальной и разностной постановках было получено, что решения имеют качественно отличный вид. Наблюдаемое отсутствие сходимости к точному решению связано с тем, что исходной постановке задачи и разностному решению соответствуют принципиально различные виды ударных адиабат и формы условий Ранкина–Гюгонио.

Результаты пп. 4, 5 получены Полехиной Р.Р. при поддержке Министерства науки и высшего образования Российской Федерации в рамках реализации программы Московского центра фундаментальной и прикладной математики по соглашению № 075-15-2022-283. Результаты пп. 1–3 получены Алексеевым М.В. и Савенковым Е.Б. при поддержке Российского научного фонда (проект 22-11-00203).

#### СПИСОК ЛИТЕРАТУРЫ

1. *LeFloch P.G., Mohammadian M.* Why many theories of shock waves are necessary: kinetic functions, equivalent equations, and fourth-order models // *J. of Comput. Phys.* 2008. V. 227. № 8. P. 4162–4189.
2. *Шокин Ю.И.* Метод дифференциального приближения. Новосибирск, 1979.
3. *Karni S.* Viscous shock profiles and primitive formulations // *SIAM J. on Numer. Anal.* 1992. V. 29. № 6. P. 1592–1609.
4. *Tadmor E.* The numerical viscosity of entropy stable schemes for systems of conservation laws. I // *Math. Comp.* 1987. V. 49. P. 91–103.
5. *Tadmor E., Zhong W.* Novel entropy stable schemes for 1D and 2D fluid equations // *Hyperbolic Problems: Theory, Numerics, Applications.* / Eds. S. Benzoni-Gavage, D. Serre. Berlin; Heidelberg, 2008.
6. *Castro M.J. et al.* Entropy conservative and entropy stable schemes for nonconservative hyperbolic systems // *SIAM J. on Numer. Anal.* 2013. V. 51. № 3. P. 1371–1391.
7. *LeFloch P., Mishra S.* Numerical methods with controlled dissipation for small-scale dependent shocks // *Acta Numerica.* 2014. V. 23. P. 743–816.
8. *Гельфанд И.М.* Некоторые задачи теории квазилинейных уравнений // *Успехи мат. наук.* 1959. Т. 14. Вып. 2 (86). С. 87–158.
9. *Петровский И.Г.* О проблеме Cauchy для систем линейных уравнений с частными производными в области неаналитических функций // *Бюлл. Московского гос. ун-та. Секц. А. Математика и механика.* 1938. Т. 1. Вып. 7.
10. *Majda A., Pego L.* Stable viscosity matrices for systems of conservation laws // *J. of Differ. Equat.* 1985. V. 56. № 2. P. 229–262.
11. *Вольперт А.И.* Пространства  $BV$  и квазилинейные уравнения // *Мат. сб.* 1967. Т. 73 (115). № 2. С. 255–302.
12. *LeFloch P., Liu T.-P.* Existence theory for nonlinear hyperbolic systems in nonconservative form. Berlin; New York, 1993.
13. *Dal Maso G., Lefloch P.G., Murat F.* Definition and weak stability of nonconservative products // *J. de Math. Pur. et Appl.* 1995. V. 74. № 6. P. 483–548.
14. *Colombeau J.-F.* Multiplication of distributions // *Bull. of the Amer. Math. Soc.* 1990. V. 23. № 2. P. 251–268.

15. *Alouges F., Merlet B.* Approximate shock curves for non-conservative hyperbolic systems in one space dimension // *J. of Hyperbolic Differ. Equat.* 2004. V. 1. № 4. P. 769–788.
16. *Lax P.* Hyperbolic systems of conservation laws and the mathematical theory of shock waves // *CBMS-NSF Regional Conference Series in Applied Mathematics*. V. 11. SIAM, 1973.
17. *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. 2. № 6. P. 861–889.
18. *Toumi I.* A weak formulation of Roe's approximate Riemann solver // *J. of Comput. Phys.* 1992. V. 102. № 2. P. 360–373.
19. *Dumbser M., Balsara D.S.* A new efficient formulation of the HLLEM Riemann solver for general conservative and non-conservative hyperbolic systems // *J. of Comput. Phys.* 2016. V. 304. P. 275–319.
20. *Parés C.* Numerical methods for nonconservative hyperbolic systems: atheoretical framework // *SIAM J. on Numer. Anal.* 2006. V. 44. № 1. P. 300–321.
21. *Berthon C.* Schéma nonlinéaire pour l'approximation numérique d'un système hyperbolique non conservatif // *Comptes Rendus Mathématique*. 2002. V. 335. № 12. P. 1069–1072.
22. *Куликовский А.Г., Свешникова Е.И.* Нелинейные волны в упругих средах. М., 1998.
23. *Pimentel-García E., Castro M.J., Chalons C., de Luna T.M., Parés C.* In-cell discontinuous reconstruction path-conservative methods for non conservative hyperbolic systems-Second-order extension // *J. of Comput. Phys.* 2022. V. 459. P. 111152.
24. *Chalons C.* Path-conservative in-cell discontinuous reconstruction schemes for non conservative hyperbolic systems // *Commun. Math. Sci.* 2020. V. 18. № 1. P. 1–30.
25. *Polekhina R.R., Korneev B.A., Savenkov E.B.* Numerical study of multiphase hyperbolic models // *J. of Comput. and Appl. Math.* 2023. V. 423. P. 114925.

Институт прикладной математики  
имени М.В. Келдыша РАН, г. Москва

Поступила в редакцию 03.03.2023 г.  
После доработки 03.03.2023 г.  
Принята к публикации 19.05.2023 г.

УДК 519.633

## СЕТОЧНО-ХАРАКТЕРИСТИЧЕСКИЙ МЕТОД ПОВЫШЕННОГО ПОРЯДКА ДЛЯ СИСТЕМ ГИПЕРБОЛИЧЕСКИХ УРАВНЕНИЙ С КУСОЧНО-ПОСТОЯННЫМИ КОЭФФИЦИЕНТАМИ

© 2023 г. Н. И. Хохлов, И. Б. Петров

Рассмотрен новый подход для повышения порядка точности сеточно-характеристического метода в области скачка коэффициентов, основанный на кусочно-полиномиальной интерполяции для схем второго и третьего порядков точности, для случая, когда граница раздела сред согласована с конечно-разностной сеткой. Метод предназначен для численного моделирования распространения динамических волновых возмущений в гетерогенных средах. Для описания рассматриваемых физических процессов использованы системы гиперболических уравнений с переменными коэффициентами. Приведены описание численного метода и результаты его тестирования.

DOI: 10.31857/S0374064123070117, EDN: GVISAO

**Введение.** Распространение динамических возмущений наблюдается в различных областях физики, таких как акустика, теория упругости и электромагнетизм. Для математического описания распространения волновых возмущений в гетерогенных средах применяются системы гиперболических уравнений с переменными коэффициентами. При переходе от среды с одними параметрами к среде с другими параметрами изменяется скорость распространения волн и другие характеристики среды. При численном моделировании необходимо корректно учитывать границу перехода от одной среды к другой. В данной работе рассматривается случай разрывного изменения коэффициентов системы и предлагается построение конечно-разностного метода, обладающего повышенным порядком точности в области разрыва коэффициентов, а именно рассматривается кусочно-постоянное распределение параметров среды, причём разрыв её параметров привязан к узлам конечно-разностной сетки, т.е. параметры среды привязаны к ячейкам. Значения искомого физических параметров привязаны к узлам конечно-разностной сетки (сдвинутая сетка).

Рассматриваем одномерную систему уравнений, однако несмотря на то, что решение одномерного уравнения на первый взгляд имеет слабое практическое значение, оно позволяет получить ряд важных результатов, в том числе имеющих прикладное значение. Прежде всего, методы, использующие расщепление по пространственным координатам, сводятся к решению одномерных систем гиперболических уравнений вдоль направлений расщепления [1]. Используя их, решение, полученное для одномерного случая, может быть перенесено на многомерный случай. Также в задачах сейсмологии или акустики океана зачастую геологическая модель может быть описана в виде многослойной среды [2]. В задачах медицины, ультразвуковой фильтрации, оптики также возникают такого рода среды [3].

Изучим подробнее вопрос решения рассматриваемых систем уравнений в области разрыва коэффициентов. Данная задача может рассматриваться как задача погружённого интерфейса и является проблемой, возникающей в большом количестве практических приложений. Для решения такого рода задач уже предложен ряд численных методов. Опишем их подробнее. Постановки можно разделить на два типа: когда граница раздела сред согласована с сеткой, т.е. скачок среды приходится на ячейку или узел, и противоположный случай, когда граница может быть не согласована с сеткой.

Следует заметить, что метод погружённого интерфейса достаточно полно представлен для задач параболического и эллиптического типа (см. [4–8] и др.). Для задач гиперболического типа он представлен не так широко.

Для случая, когда граница раздела сред согласована с конечно-разностной сеткой, существует ряд подходов для повышения порядка в области скачка параметров. Так, в работе [1] рассматривается модификация конечно-объёмного метода типа Годунова на примере уравнений акустики и линейной теории упругости, приводятся точное решение задачи распада-разрыва на границе интерфейса и метод, учитывающий это. Рассматриваются TVD методы, обладающие монотонным поведением. Однако, как описано в ней, порядок схемы в области скачка коэффициентов падает. В статье [9] предложен подход повышения порядка точности конечно-разностного метода в области скачка параметров для конечно-разностных методов на сдвинутых сетках типа FDTD [3], изучен подход усреднения (гомогенизации) коэффициентов конечно-разностной схемы для получения второго порядка на границе раздела сред. В простейшей реализации методы данного класса дают первый порядок точности в области скачка параметров. Также для решения такого рода задач на неструктурных сетках широко применяется разрывный метод Галёркина, описанный в [10] в двумерном случае на неструктурированных (или нерегулярных) сетках, там же предложен подход, позволяющий строить методы произвольного порядка. В частности, данный метод даёт возможность рассчитывать границы раздела сред и между различными физическими средами, типа акустическая-упругая среда [11].

Что касается методов погружённого интерфейса с несогласованной сеткой, то одной из первых работ в данном направлении для двумерного уравнения акустики была статья [12]. В ней рассматривается подход, позволяющий обсчитывать границу произвольной формы для двумерного уравнения акустики. Подход, построенный на методе неопределённых коэффициентов, используется в [13]. Позже он был распространён на уравнение упругости [12]. Предложенные авторами методы требуют предварительного решения систем линейных алгебраических уравнений для получения коэффициентов в конечно-разностной схеме.

В работе [14] авторами предложен так называемый явный упрощённый метод погружённого интерфейса (ESIM), отличие которого от классического метода погружённого интерфейса заключается в том, что на всей сетке используется один и тот же метод повышенного порядка точности и только в точках вблизи интерфейса применяются некоторые корректировки, не зависящие от используемого метода. В [14] авторы рассматривают одномерный случай, а в [15] предложили обобщение данного подхода на двумерный случай. Метод позволяет производить расчёт границы произвольной формы на структурированной прямоугольной сетке. В дальнейшем данный подход был применён и для других систем гиперболических уравнений [16]. Метод выглядит достаточно универсальным, однако требует препроцессинга для вычисления матриц коэффициентов на границе и хранения дополнительных данных для границы. Для задач, где параметры среды меняются плавно или достаточно много границ раздела сред, применение такого рода методов может быть затруднительно.

В статье [17] предложен метод CFM четвёртого порядка точности для двумерного уравнения акустики, согласно которому к узлам вблизи границы раздела сред прибавляется компенсирующая добавка после каждого шага, чтобы учесть скачок параметров. Потенциально метод может быть расширен на более высокие порядки и, в том числе, на случай подвижной границы. Однако он также требует препроцессинга и хранения дополнительных параметров для описания границы.

В данной работе рассматривается новый подход для повышения порядка точности в области скачка коэффициентов для случая, когда граница раздела сред согласована с конечно-разностной сеткой.

Ранее для сеточно-характеристического метода уже разрабатывались подходы для явного выделения разрыва матрицы коэффициентов [18]. Использовались компактные сеточно-характеристические схемы [19] второго и третьего порядков точности. Следует отметить, что в отличие от методов с несогласованной сеткой, таких как ESIM, CFM и др., предложенный в данной работе метод более простой и не требует предварительного расчёта коэффициентов схемы или хранения дополнительных данных. Он скорее ближе к методу [9] и больше подходит для задач, где существует достаточно много границ раздела сред, таких как задачи геофизики, сейсмологии и т.д.

Идея предложенного метода построена на сеточно-характеристическом методе [20] и на кусочно-полиномиальной интерполяции [21]. Сеточно-характеристический метод применяется для достаточно широкого круга задач, например, для расчёта различных геофизических задач [22, 23], в том числе с наличием трещиноватых неоднородностей [24, 25]. Также сеточно-характеристический метод применяется в задачах движения поездов [26], в задачах сейсмоки и сейсмостойкости [27] и др.

Работа построена следующим образом. Вначале рассматривается базовый подход решения систем уравнений гиперболического типа на примере уравнения акустики, описывается метод для случая разрывных коэффициентов системы. Проводится тестирование и исследование метода, показывается пониженный порядок в области скачка. На основе одномерного уравнения переноса предлагается модифицированный метод, имеющий повышенный порядок точности в области скачка скорости переноса. Затем данный подход обобщается на уравнение акустики и проводится тестирование метода. В конце работы формулируются выводы.

**1. Сеточно-характеристический метод для одномерного уравнения акустики.** В одномерном случае система линейных уравнений акустики имеет вид [1]

$$v_t + \frac{1}{\rho} p_x = 0, \quad p_t + K v_x = 0, \tag{1}$$

где  $p(x, t)$  – возмущение давления от некоторого постоянного значения  $p_0$ ,  $v(x, t)$  – скорость,  $\rho(x)$  – плотность упругого материала или газа,  $K(x)$  – модуль объёмной упругости (в случае акустической среды он равен первому параметру Ламе). Дополнительно введём такие параметры как акустический импеданс  $Z = \sqrt{K\rho}$  и скорость распространения продольных волн (скорость звука)  $c = \sqrt{K/\rho}$ .

Систему уравнений (1) можно записать в матричном виде как систему дифференциальных уравнений в частных производных первого порядка

$$\mathbf{U}_t + \mathbf{A}\mathbf{U}_x = 0,$$

где  $\mathbf{U} = \begin{pmatrix} p(x, t) \\ v(x, t) \end{pmatrix}$ ,

$$\mathbf{A} = \begin{pmatrix} 0 & K \\ 1/\rho & 0 \end{pmatrix} = \begin{pmatrix} 0 & c/Z \\ cZ & 0 \end{pmatrix}. \tag{2}$$

Будем рассматривать случай, когда параметры среды  $Z$  и  $c$  кусочно-постоянны и имеют разрыв в некоторой точке  $x = x_\alpha$ :

$$(Z, c) = \begin{cases} (Z_l, c_l), & \text{если } x < x_\alpha, \\ (Z_r, c_r), & \text{если } x > x_\alpha. \end{cases}$$

Условия сшивки в точке разрыва параметров среды для уравнений акустики имеют вид [12]

$$[p] = 0, \quad [v] = 0. \tag{3}$$

Система (1) является гиперболической, поэтому матрица (2) имеет полный набор действительных собственных чисел и векторов [1]. Матрица  $\mathbf{A}$  представима в виде  $\mathbf{A} = \mathbf{R}\mathbf{\Lambda}\mathbf{R}^{-1}$ , где  $\mathbf{R}$  – матрица собственных векторов,  $\mathbf{R}^{-1}$  – обратная к ней и  $\mathbf{\Lambda}$  – матрица собственных чисел:

$$\mathbf{\Lambda} = \begin{pmatrix} -c & 0 \\ 0 & c \end{pmatrix}.$$

Введём характеристические переменные (инварианты Римана)

$$\mathbf{W} = \begin{pmatrix} \omega^- \\ \omega^+ \end{pmatrix} = \mathbf{R}^{-1}\mathbf{U} = \frac{1}{2Z} \begin{pmatrix} -p + Zv \\ p + Zv \end{pmatrix}, \tag{4}$$

тогда

$$\mathbf{U} = \mathbf{R}\mathbf{W} = \begin{pmatrix} -Z & Z \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \omega^- \\ \omega^+ \end{pmatrix} = \begin{pmatrix} Z(\omega^+ - \omega^-) \\ \omega^+ + \omega^- \end{pmatrix}, \tag{5}$$

или  $p = Z(\omega^+ - \omega^-)$ ,  $v = \omega^+ + \omega^-$ .

Дискретизацию будем рассматривать с постоянными шагами по времени и пространству соответственно  $\tau$  и  $h$ . Индекс по времени будем обозначать  $n$ , по координате –  $m$ . Без ограничения общности рассмотрим трёхточечный шаблон по координате  $(t^n, x_{m-1})$ ,  $(t^n, x_m)$ ,  $(t^n, x_{m+1})$ ,  $(t^{n+1}, x_m)$ .

В дальнейшем, где смысл понятен из контекста, индексы будем опускать. Как уже оговаривалось ранее, в работе рассматривается согласованная сетка, которую будем располагать таким образом, чтобы точка скачка коэффициентов  $x = x_\alpha$  попадала в узел расчётной сетки. Не ограничивая общности, будем рассматривать случай  $x_\alpha = 0$ . Характеристики  $\omega^+$  и  $\omega^-$  попадают, соответственно, в левую и правую стороны.

Рассмотрим алгоритм для случая постоянных коэффициентов, т.е.  $(Z_l, c_l) = (Z_r, c_r) = (Z, c)$ . В этом случае можно использовать следующий алгоритм вычисления значений  $\mathbf{U}^{n+1}$ .

*Шаг 1.* Переходим к характеристическим переменным,  $\mathbf{W}^n = \mathbf{R}^{-1}\mathbf{U}^n$ .

*Шаг 2.* Решаем одномерные уравнения переноса для характеристических переменных  $\omega^+$  и  $\omega^-$  и находим  $\mathbf{W}^{n+1}$ .

*Шаг 3.* Делаем обратное преобразование  $\mathbf{U}^{n+1} = \mathbf{R}\mathbf{W}^{n+1}$ .

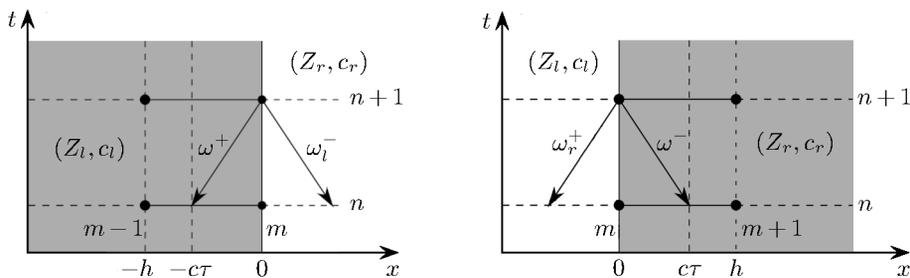
Следует заметить, что данный алгоритм может быть достаточно эффективно реализован. Это связано с наличием аналитического выражения для матриц  $\mathbf{R}$  и  $\mathbf{R}^{-1}$ , а также тем, что фактически хранить переменные  $\mathbf{W}$  на всей расчётной сетке не нужно, а можно использовать только несколько значений в зависимости от ширины шаблона разностной схемы по пространству. На шаге 2 алгоритма для восстановления решения на следующем временном слое можно использовать различные методы, в частности, конечно-объёмные [9] или сеточно-характеристические [20, 28], что и предлагается в данной работе. Также следует отметить, что данный алгоритм можно записать в матричном виде, минуя переход к характеристическим переменным, как, например, выполняется в [1].

**2. Системы с переменными коэффициентами.** Для переменных коэффициентов алгоритм, описанный ранее, не может быть применим. Это связано с тем, что преобразование переменных подразумевает возможность внесения  $\mathbf{R}^{-1}$  под знак дифференциала, что сделать нельзя, если матрица  $\mathbf{R}^{-1}$  зависит от координаты  $x$ .

**Утверждение 1.** Для случая с кусочно-постоянным распределением коэффициентов переход к следующему временному слою возможен по формулам

$$p^{n+1} = 2 \frac{Z_l Z_r}{Z_l + Z_r} (\omega^+ - \omega^-), \quad v^{n+1} = \frac{2}{Z_l + Z_r} (Z_l \omega^+ + Z_r \omega^-). \tag{6}$$

Чтобы решить данную задачу, при вычислении в узле расчётной сетки, который попадает на границу раздела сред, разделим среду на два полупространства и будем их рассматривать независимо, соответственно слева и справа, как показано на рис. 1.



**Рис. 1.** Разделение расчётной сетки для различных коэффициентов матрицы системы.

Вначале рассмотрим левое полупространство. Будем рассматривать точку  $\mathbf{U}_m$  как граничную, соответственно справа находится полупространство с параметрами среды  $(Z_r, c_r)$ , а слева – с параметрами  $(Z_l, c_l)$ . Обозначим характеристическую переменную, выходящую за границу области интегрирования, через  $\omega_l^-$ . Она неизвестна и её необходимо получить из граничного условия. Пусть  $\omega_l^{-n+1} = \alpha_l \omega^{+n+1}$ , где  $\alpha_l$  – некий скалярный коэффициент.

Аналогично для правого полупространства. Характеристика, выходящая за границу области интегрирования, обозначена как  $\omega_r^{+n+1} = \alpha_r \omega^{-n+1}$ , где  $\alpha_r$  – некий скалярный коэффициент. Далее для нахождения  $\alpha_l$  и  $\alpha_r$  будем использовать условие (3)  $[\mathbf{U}] = 0$ , т.е. на временном шаге  $n + 1$  необходимо, чтобы  $\mathbf{U}_l^{n+1} = \mathbf{U}_r^{n+1}$ . Отсюда находим

$$\alpha_r = 2 \frac{\omega^+}{\omega^-} \frac{Z_l}{Z_l + Z_r} - \frac{Z_l - Z_r}{Z_l + Z_r}, \quad \alpha_l = 2 \frac{\omega^-}{\omega^+} \frac{Z_r}{Z_l + Z_r} + \frac{Z_l - Z_r}{Z_l + Z_r},$$

и значит, выражение (6).

В результате получаем следующий алгоритм вычисления в области разрыва коэффициентов.

*Шаг 1.* Переходим к характеристическим переменным независимо слева и справа, используя (4): слева  $\omega^+ = (p + Z_l v)/(2Z_l)$  и справа  $\omega^- = (-p + Z_r v)/(2Z_r)$ .

*Шаг 2.* Используя какой-либо метод, решаем одномерные уравнения переноса для характеристических переменных  $\omega^+$  и  $\omega^-$ .

*Шаг 3.* Делаем обратное преобразование, используя (6).

**Замечание 1.** Можно проверить, что равенства (6) при  $Z_l = Z_r = Z$  переходят в (5). Алгоритм по вычислительной сложности практически не отличается от алгоритма для постоянных коэффициентов.

**Замечание 2.** Для кусочно-постоянных параметров среды значения материала нужно привязывать к ячейкам. Для всех точек сетки можно применять метод с переменными коэффициентами, в случае одинаковых значений параметров среды слева и справа от узла он вырождается в метод для постоянных коэффициентов, а по вычислительной сложности практически такой же. Функции  $p$  и  $v$  непрерывны в точке  $x = x_\alpha$ , однако производная имеет разрыв. Условия на производные легко получить, исходя из соображения, что поскольку  $\mathbf{U}$  непрерывно в любой момент времени  $t$ , то и производная по времени  $\mathbf{U}_t$  также непрерывна, отсюда получаем условие на производную [12]

$$[\mathbf{A}\mathbf{U}_x] = 0$$

или

$$\left[ \frac{1}{\rho} p_x \right] = 0, \quad [K v_x] = 0.$$

При применении конечно-разностных схем с шаблоном, использующим точки слева и справа от центральной, будет возникать интерполяция через точку разрыва производных, что следует учитывать при расчёте.

**Замечание 3.** Коэффициенты  $\alpha_r$  и  $\alpha_l$  можно разложить на два – коэффициент для прошедшей и для отражённой волн. Тогда можно построить граничное условие с заданным импедансом, так как за границей области находится бесконечное полупространство с заданными параметрами среды. Сделать это можно следующим образом. Рассмотрим на примере  $\alpha_l$ . Нетрудно заметить, что коэффициент состоит из двух частей, и если занулить в нём параметр  $\omega^-$  (из правого полупространства ничего не приходит), то получим  $\alpha_l^R = (Z_l - Z_r)/(Z_l + Z_r)$  – коэффициент отражения. Тогда другая часть отвечает за прошедшую волну:

$$\alpha_l^T = 2 \frac{\omega^-}{\omega^+} \frac{Z_r}{Z_l + Z_r}, \quad \alpha_l = \alpha_l^T + \alpha_l^R.$$

Для задания граничного условия нужно занулить  $\alpha_l^T$ , тогда вычисление характеристических переменных, уходящих вправо ( $\omega^-$ ), не требуется. Аналогично для правой части получим  $\alpha_r = \alpha_r^T + \alpha_r^R$ , где  $\alpha_r^R = (Z_r - Z_l)/(Z_l + Z_r)$  и  $\alpha_r^T = 2(\omega^+/\omega^-)(Z_r/(Z_l + Z_r))$ .

**3. Тестирование метода.** Прежде чем начать переход к повышению порядка точности в области скачка коэффициентов, рассмотрим работу предложенного метода. Как уже упоминалось ранее, для решения одномерных уравнений переноса, которым удовлетворяют характеристические переменные, можно использовать различные методы. Рассмотрим для функции  $u = u(x, t)$  простейшее однородное линейное уравнение переноса

$$u_t + \lambda u_x = 0. \quad (7)$$

Будем полагать, что  $\lambda = \text{const} > 0$ , поскольку для отрицательных  $\lambda$  все построения аналогичны и могут быть выполнены заменой  $\lambda$  на  $-\lambda$  и сеточного шаблона на симметричный по  $x$  относительно рассчитываемой точки  $(t^n, x_m)$ . Число Куранта обозначим как  $\sigma = \lambda\tau/h$ . Простейшая схема первого порядка точности – это схема типа Куранта–Изаксона–Риса [29]:

$$u^{n+1} = u_m - \sigma(u_m - u_{m-1}). \quad (8)$$

Схема Лакса–Вендрофа [30] второго порядка точности по времени и по координате имеет вид

$$u^{n+1} = u_m - \frac{1}{2}\sigma((u_{m+1} - u_{m-1}) - \sigma(u_{m-1} - 2u_m + u_{m+1})), \quad (9)$$

а схема третьего порядка точности или схема Рундана [31] –

$$u^{n+1} = u_m + \frac{1}{2}\sigma(\Delta_m + \Delta_{m+1}) + \frac{1}{2}\sigma^2(\Delta_m - \Delta_{m+1}) + \frac{1}{6}\sigma(\sigma^2 - 1)(\Delta_{m-1} - 2\Delta_m + \Delta_{m+1}), \quad (10)$$

где  $\Delta_i = u_{i-1} - u_i$ .

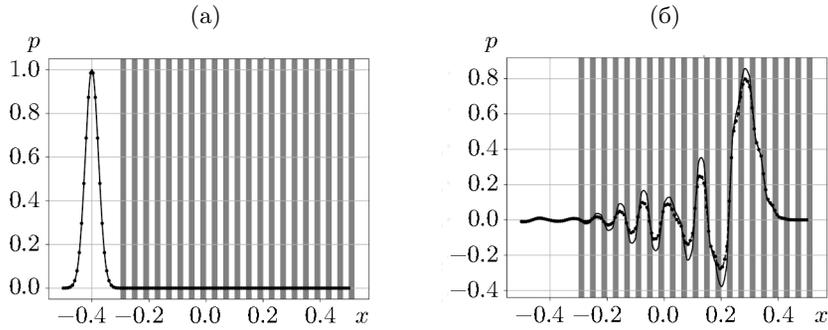
Следует отметить, что приведённые разностные схемы немонотонны. Как уже упоминалось ранее, основная область применения, которая предлагается для данных методов – геофизика. В ней при решении задач распространения волновых возмущений в условиях малых деформаций отсутствуют нелинейные эффекты и все изменения в переносе изначальной формы импульса возникают при переходе через границу раздела сред. При точном решении на границе раздела сред нефизичные осцилляции при переходе через границу минимальны и на практике применяются, в основном, немонотонные методы [3, 9, 10]. В задачах такого рода более важен перенос импульса на большие расстояния и сохранение его амплитуды. В случае потенциально-го расширения предложенного метода на другой класс задач, когда монотонность разностных схем важна, возможно применение монотонных сеточно-характеристических схем [32]. Однако следует учесть, что по норме  $L_\infty$  они будут давать первый порядок точности, что показано, например, в работе [19].

**3.1. Многослойная среда с переменным импедансом.** Рассмотрим пример с многослойной средой, аналогичный разобранным в [1]. Расчётная область  $[-0.5, 0.5]$ . Параметры среды слева:  $c = 1$ ,  $\rho = 1$ . Затем, начиная с точки  $-0.3$ , начинают чередоваться слои толщиной  $0.02$  с и плотностью  $1$  и  $3$ , скорость звука везде постоянная и равна  $1$ . Шаг сетки  $h = 0.001$ , шаг по времени  $\tau = 0.0004$ . Использовалась схема третьего порядка точности (10). Начальные данные заданы в виде импульса, движущегося вправо. Импульс имеет форму

$$f(x) = e^{-625(x+0.25)^2/4}. \quad (11)$$

Результат моделирования для тонкослойной среды приведён на рис. 2. Значения  $p$  (сплошная линия) получены схемой третьего порядка (10),  $\sigma = 0.4$ , значения  $p_{\text{exact}}$  (точки) посчитаны схемой первого порядка (8),  $\sigma = 1$ . Решение  $p_{\text{exact}}$  названо “точным”, поскольку при  $\sigma = 1$  схема точно переносит решение, при этом в ней нет интерполяции через разрыв. Хорошо видно прохождение импульса через сложную слоистую структуру.

Данный пример демонстрирует возможности метода по учёту множества границ раздела сред, показано отсутствие нефизичных осцилляций при переходе импульса через границы. Несмотря на то, что схема третьего порядка даёт некоторую ошибку, она всё равно сохраняет основную структуру решения и скорость распространения импульса.



**Рис. 2.** Примеры расчёта распространения гладкого импульса через слоистую структуру с переменными значениями импеданса и постоянной скоростью звука: а -  $t = 0.0$ ,  $\tau = 0$ ; б -  $t = 0.8$ ,  $\tau = 2000$ .

**3.2. Сеточная сходимость.** Исследуем рассматриваемый метод на сеточную сходимость. Для этого возьмём двухслойную среду с параметрами  $c_l = 1$ ,  $\rho_l = 1$ ,  $c_r = 2$ ,  $\rho_r = 2$ , соответственно импедансы  $Z_l = 1$ ,  $Z_r = 4$ , граница раздела сред  $x_\alpha = 0$ . Расчётная область  $[-0.5, 1]$ . Импульс (11) переходит из левой области в правую. Расчёт ведётся до времени  $t = 0.6$ . Ошибка измеряется в промежутке  $[0, 1]$ .

Расчёт проводится с числом Куранта  $\sigma = 0.4$  для скорости звука  $c_l$  и соответственно  $\sigma = 0.8$  для скорости звука  $c_r$ . Число узлов в сетке  $N = 150, 300, 600, 1200, 2400, 4800$ . Для расчёта порядка сходимости использовались следующие нормы:  $L_1 = \sum_i |x_i| h$ ,  $L_\infty = \max |x_i|$ .

В табл. 1 представлены ошибки и порядки сходимости для схем второго (9) и третьего (10) порядков соответственно. Видно, что численный порядок схем отличается от теоретических и близок к первому или равен ему. Это связано с тем, что искомая функция имеет точку излома в области скачка матрицы коэффициентов, что приводит к понижению порядка до первого.

**Таблица 1.** Сеточная сходимость для метода Лакса–Вендрофа (9) и метода Рундана (10)

$h$	Метод Лакса–Вендрофа				Метод Рундана			
	$L_1$	Порядок	$L_\infty$	Порядок	$L_1$	Порядок	$L_\infty$	Порядок
0.01	1.201E-01	–	1.000E-02	–	4.185E-02	–	1.000E-02	–
0.005	5.699E-02	1.076	5.728E-01	0.419	1.198E-02	1.804	1.487E-01	1.555
0.0025	1.741E-02	1.711	2.351E-01	1.285	3.414E-03	1.811	4.593E-02	1.694
0.00125	5.009E-03	1.797	7.061E-02	1.735	1.519E-03	1.169	1.773E-02	1.373
0.000625	1.608E-03	1.639	2.223E-02	1.667	7.515E-04	1.015	8.237E-03	1.106
0.0003125	6.059E-04	1.408	7.953E-03	1.483	3.755E-04	1.001	4.048E-03	1.025
0.00015625	2.631E-04	1.203	3.209E-03	1.309	1.878E-04	1.000	2.016E-03	1.006

**4. Повышение порядка в области скачка коэффициентов.** Далее на примере линейного уравнения переноса (7) рассмотрим вопрос повышения порядка точности метода для случая, когда скорость переноса  $\lambda$  кусочно-постоянна и имеет разрыв в некоторой точке  $x = x_\alpha$ :

$$\lambda = \begin{cases} \lambda_l, & \text{если } x < x_\alpha, \\ \lambda_r, & \text{если } x > x_\alpha. \end{cases}$$

Условие сшивки в точке разрыва параметров среды для уравнений переноса имеет вид

$$[u] = 0. \tag{12}$$

Рассмотрим решение уравнения (7), используя разностные схемы (9) и (10), удовлетворяющее начальному условию

$$u(0, x) = \sin^4(\pi x) \tag{13}$$

в области  $[-0.9, 0.1]$  и периодическим граничным условиям, размер всей области интегрирования  $[-1, 2]$ . Число Куранта  $\sigma = 0.4$ . Расчёт проводился до момента времени  $t = 2.5$ .

Скорость переноса  $\lambda = 2$  в интервале  $[0.5, 1.5]$  и  $\lambda = 1$  в остальной расчётной области. Нетрудно посчитать, что за время  $t = 2.5$  импульс должен пройти один период и вернуться в начальное положение.

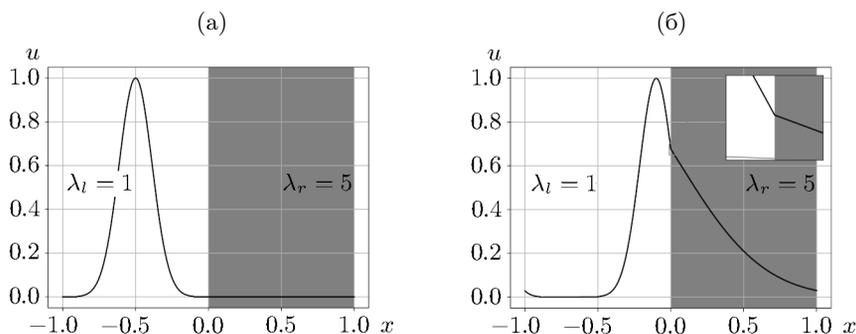
Результат тестирования приведён в табл. 2 для методов (9) и (10) соответственно. Как видно из таблицы, нет существенного падения порядка точности до первого, как это было с уравнением акустики, однако порядок всё равно не дотягивает до теоретического. Это связано с той же проблемой. Сама функция не имеет разрыва (12) в точке  $x_\alpha$ , однако её производная будет иметь скачок [12]

$$[\lambda u_x] = 0. \tag{14}$$

**Таблица 2.** Сеточная сходимость для уравнения переноса для метода Лакса–Вендрофа (9) и метода Рунанова (10)

$h$	Метод Лакса–Вендрофа				Метод Рунанова			
	$L_1$	Порядок	$L_\infty$	Порядок	$L_1$	Порядок	$L_\infty$	Порядок
0.04	0.35505	–	0.76520	–	0.12593	–	0.29088	–
0.02	0.17547	1.01681	0.43151	0.82646	0.04354	1.53214	0.10808	1.42829
0.01	0.06069	1.53172	0.15813	1.44828	0.00753	2.53094	0.01937	2.47988
0.005	0.01658	1.87222	0.04303	1.87770	0.00112	2.75251	0.00294	2.71970
0.0025	0.00454	1.86971	0.01199	1.84312	0.00017	2.70403	0.00047	2.63913
0.00125	0.00133	1.76548	0.00363	1.72465	0.00003	2.54744	0.00008	2.48088
0.000625	0.00044	1.59786	0.00122	1.57067	0.00001	2.36544	0.00002	2.31922

Рассмотрим подробнее распространение гладкого импульса через границу раздела сред. На рис. 3 приведён пример расчёта распространения гладкого импульса через границу раздела сред. Расчёт проводился в области  $[-1, 1]$ , значения  $x_\alpha = 0$ ,  $\lambda_l = 1$ ,  $\lambda_r = 2$ ,  $h = 0.002$  (шаг по координате),  $\tau = 0.00025$  (шаг по времени). Из графика хорошо видно, что в точке  $x = x_\alpha$  искомая функция имеет точку излома, т.е. выполняются условия непрерывности функции (12) и скачка производной (14).



**Рис. 3.** Пример расчёта распространения гладкого импульса через границу раздела сред: а – по схеме третьего порядка (10),  $\sigma = 0.4$ ; б – для уравнения переноса (7). На врезке показан начальный импульс и момент прохождения через границу раздела поверхностей.

При использовании разностных схем возникает интерполяция через точку излома, в результате чего падает порядок. Для устранения точки излома в работе предлагается новый метод, основанный на кусочно-полиномиальной интерполяции. Похожий подход описан, например, в работе [21].

Не нарушая общности, рассмотрим случай  $x_\alpha = 0$ . Далее аргумент  $t$ , где это возможно, будем опускать. Также предполагаем, что скорость переноса  $\lambda(x) > 0$  во всей рассматриваемой области.

**Утверждение 2.** Функция

$$u^*(x) = u\left(\frac{\lambda(x)}{\lambda_l}x\right) \tag{15}$$

в случае двух полупространств

$$u^*(x) = \begin{cases} u(x), & \text{если } x < x_\alpha, \\ u\left(\frac{\lambda_r}{\lambda_l}x\right), & \text{если } x > x_\alpha, \end{cases} \quad (16)$$

удовлетворяет условиям непрерывности  $[u^*] = 0$  и  $[u_x^*] = 0$ .

Действительно, условие непрерывности для (15) будет такое же, как и для функции  $u$  (12). Далее для производной из (16) следует

$$u_x^* = \frac{\partial u^*(x)}{\partial x} = \begin{cases} \frac{\lambda_l}{\lambda_l} \frac{\partial u(x)}{\partial x}, & \text{если } x < x_\alpha, \\ \frac{\lambda_r}{\lambda_l} \frac{\partial u(\lambda_r \lambda_l^{-1}x)}{\partial x}, & \text{если } x > x_\alpha. \end{cases}$$

Из (14) получим равенства

$$[\lambda u_x] = \left[ \frac{\lambda}{\lambda_l} u_x \right] = [u_x^*] = 0,$$

т.е. у функции  $u^*$  отсутствует скачок производной в точке  $x = x_\alpha = 0$ . На рис. 4 приведены графики функций  $u$  (штриховая кривая) и  $u^*$  (15) (сплошная). Стрелкой показано фактическое преобразование координат в правой полуплоскости. Видно, что устранена точка излома и, соответственно, скачок производной.

**Утверждение 3.** Для функций  $u$  и  $u^*$  верно равенство  $u(x_\alpha, t + \tau) = u^*(x_\alpha, t + \tau)$ .

Действительно, из (16) следует, что для  $x < x_\alpha$  данные функции совпадают и

$$u(x_\alpha, t + \tau) = u(x_\alpha - \lambda_l \tau, t) = u^*(x_\alpha - \lambda_l \tau, t).$$

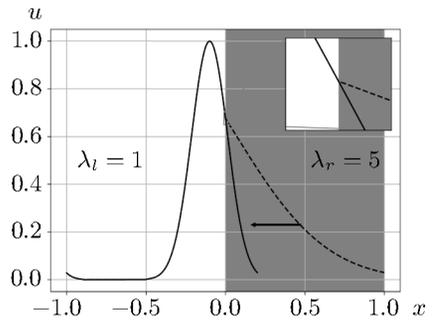


Рис. 4. Графики функции  $u^*$  и  $u$ .

**5. Построение сеточно-характеристической схемы повышенного порядка для кусочно-постоянной скорости переноса.** Рассмотрим метод повышенного порядка, построенный на аппроксимации решения полиномом заданной степени. Строить сеточно-характеристическую схему будем на четырёхточечном шаблоне по координатам

$$(t^n, x_{m-2}), \quad (t^n, x_{m-1}), \quad (t^n, x_m), \quad (t^n, x_{m+1}), \quad (t^{n+1}, x_m). \quad (17)$$

Схема максимального порядка, которая может быть построена на данном шаблоне, имеет третий порядок [28]. Схематично шаблон (17) (по координате) приведён на рис. 5, значения функции  $u$  показаны сплошной линией,  $u^*$  – штриховой. Шаг сетки постоянен и равен  $h$ , значение скорости переноса  $\lambda(x)$  кусочно-постоянное и постоянно в пределах ячейки, имеет полуцелый индекс.

Исходя из характеристических свойств уравнения, можно записать

$$u(x_m, t^{n+1}) = u(x_m - \lambda_{m-1/2} \tau, t^n) = u^*(x_m - \lambda_{m-1/2} \tau, t^n).$$

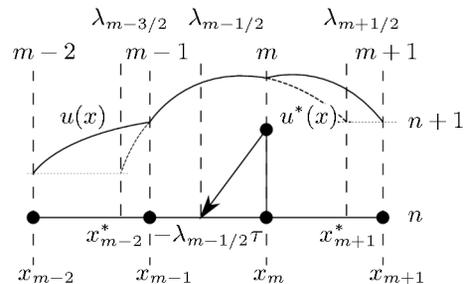


Рис. 5. Четырёхточечный шаблон для построения разностной схемы.

Необходимо построить интерполяционный полином для функции (15), однако значения этой функции для построения полинома известны только в точках

$$x_{m-2}^* = x_{m-1} - h \frac{\lambda_{m-3/2}}{\lambda_{m-1/2}}, \quad x_{m-1}, \quad x_m, \quad x_{m+1}^* = x_m + h \frac{\lambda_{m+1/2}}{\lambda_{m-1/2}},$$

т.е. фактически нужно строить конечно-разностный метод для уравнения переноса с постоянной скоростью переноса  $\lambda = \lambda_{m-1/2}$  на шаблоне с переменным шагом  $(h_1, h_2, h_3)$ , где  $h_1 = h\lambda_{m-3/2}/\lambda_{m-1/2}$ ,  $h_2 = h$ ,  $h_3 = h\lambda_{m+1/2}/\lambda_{m-1/2}$ .

Для построения такого рода схем можно использовать разложение в ряд Тейлора около точки  $x_m$  либо нахождение коэффициентов интерполяционного полинома. Обозначим  $u_i = u_{m+i}$ ,  $i = -2, -1, 0, 1$ . Тогда для схемы второго порядка можно построить полином вида

$$f_2(x) = ax^2 + bx + c.$$

Если  $x_m = 0$ , то  $u_m^{n+1} = f(-\lambda_{m-1/2}\tau)$ ,

$$a = \frac{u_{-1}h_3 - u_0(h_2 + h_3) + u_1h_2}{h_3h_2(h_3 + h_2)}, \quad d = \frac{-u_{-1}h_3^2 + u_0(h_3^2 - h_2^2) + u_1h_2^2}{h_3h_2(h_3 + h_2)}, \quad c = u_0. \quad (18)$$

Для схемы третьего порядка имеем

$$f_3(x) = (h_1 + h_2 + x) \left( (h_2 + x) \left( \frac{x(\theta_2 - \theta_3)}{h_1 + h_2 + h_3} + \theta_1 \right) + \theta_0 \right) + u_{-2}, \quad (19)$$

где

$$\theta_0 = \frac{u_{-1} - u_{-2}}{h_1}, \quad \theta_1 = \frac{h_1(u_0 - u_{-1}) - h_2(u_{-1} - u_{-2})}{h_1h_2(h_1 + h_2)},$$

$$\theta_2 = \frac{h_2(u_1 - u_0) - h_3(u_0 - u_{-1})}{h_2h_3(h_2 + h_3)}, \quad \theta_3 = \frac{h_1(u_0 - u_1) - h_2(u_1 - u_2)}{h_1h_2(h_1 + h_2)}.$$

**6. Тестирование.** В табл. 3 приведены результаты тестирования порядка сходимости методов (18) и (19) соответственно. Параметры тестирования такие же, что и для табл. 2. Хорошо видно, что порядок сходимости методов совпадает с теоретическим порядком схем.

**Таблица 3.** Сеточная сходимость для уравнения переноса для метода второго порядка с переменными коэффициентами (18) и метода третьего порядка с переменными коэффициентами (19)

$h$	Метод (18)				Метод (19)			
	$L_1$	Порядок	$L_\infty$	Порядок	$L_1$	Порядок	$L_\infty$	Порядок
0.04	0.35689	–	0.76195	–	0.12035	–	0.27664	–
0.02	0.17176	1.05503	0.42513	0.84179	0.04147	1.53706	0.10246	1.43291
0.01	0.05815	1.56268	0.14958	1.50701	0.00671	2.62698	0.01632	2.65075
0.005	0.01514	1.94130	0.03822	1.96841	0.00088	2.92413	0.00213	2.93525
0.0025	0.00378	2.00068	0.00954	2.00243	0.00011	2.98906	0.00027	2.98990
0.00125	0.00094	2.00293	0.00238	2.00141	0.00001	2.99940	0.00003	2.99806
0.000625	0.00024	2.00157	0.00060	2.00039	0.00000	3.00032	0.00000	2.99948

**7. Схема для уравнения акустики с переменными коэффициентами.** Для системы уравнений линейной теории акустики (1) можно поступить аналогичным образом. Однако характеристические переменные имеют скачок не только производной, но и самой функции, так как из (3) и (5) следует, что  $[\mathbf{R}\omega] = [\mathbf{U}] = 0$ .

Сами искомые функции не имеют разрыва в точке  $x_\alpha$  (3). Для сохранения повышенного порядка точности интерполяцию стоит применять в искомым функциях. Следует учитывать, что скачок производных для уравнения акустики отличается от скачка уравнения переноса (14), а именно [12]

$$[\mathbf{U}] = [\mathbf{A}\mathbf{U}_x] = 0,$$

или отдельно для  $p_x$  и  $v_x$

$$\left[ \frac{1}{\rho} p_x \right] = 0, \quad [K v_x] = 0.$$

Исходя из изложенного выше получен следующий алгоритм расчёта.

*Шаг 1.* Рассчитываем  $p_-^{n+1} = p^n(x_\alpha - c_l\tau)$ ,  $v_-^{n+1} = v^n(x_\alpha - c_l\tau)$  и  $p_+^{n+1} = p^n(x_\alpha + c_r\tau)$ ,  $v_+^{n+1} = v^n(x_\alpha + c_r\tau)$ , используя (18) или (19).

*Шаг 2.* Переходим к характеристическим переменным  $\omega^+ = (p_+ + Z_l v_+)/ (2Z_l)$  и  $\omega^- = (-p_- + Z_r v_-)/ (2Z_r)$ , используя (4).

*Шаг 3.* Делаем обратное преобразование (6).

**8. Тестирование нового алгоритма для уравнения акустики.** Проведём тестирование сходимости, аналогичное тому, что приведено в табл. 1.

Результаты тестирования приведены в табл. 4 для методов второго и третьего порядков соответственно. Видно, что абсолютное значение норм ошибок значительно уменьшилось и порядок стал ближе к теоретическому. Для схемы второго порядка он равен 2 и для схемы третьего порядка 2.8.

**Таблица 4.** Сеточная сходимость для метода второго порядка (18) и метода третьего порядка (19) для одномерного уравнения акустики

$h$	Метод второго порядка				Метод третьего порядка			
	$L_1$	Порядок	$L_\infty$	Порядок	$L_1$	Порядок	$L_\infty$	Порядок
0.01	1.137E-01	–	1.000E-02	–	4.011E-02	–	1.000E-02	–
0.005	5.237E-02	1.118	5.117E-01	0.492	1.086E-02	1.885	1.383E-01	1.616
0.0025	1.505E-02	1.799	1.958E-01	1.386	1.760E-03	2.625	2.563E-02	2.432
0.00125	3.810E-03	1.981	5.107E-02	1.939	2.340E-04	2.911	3.554E-03	2.851
0.000625	9.532E-04	1.999	1.262E-02	2.016	3.031E-05	2.949	4.677E-04	2.926
0.0003125	2.382E-04	2.000	3.140E-03	2.008	4.010E-06	2.918	6.297E-05	2.893
0.00015625	5.955E-05	2.000	7.837E-04	2.002	5.585E-07	2.844	8.973E-06	2.811

На рис. 2 значения  $p^*$  (штриховая линия) получены схемой третьего порядка (19) и модифицированным методом с переменными коэффициентами,  $\sigma = 0.4$ . Видно практически полное совпадение с “точным” решением.

**Заключение.** В работе предложен новый метод построения сеточно-характеристических схем для систем гиперболических уравнений с переменными коэффициентами. На примере одномерного уравнения акустики показано уменьшение абсолютного значения ошибки и повышение порядка точности метода. Метод основан на сеточно-характеристическом подходе [20] и кусочно-полиномиальной интерполяции [21], имеет повышенный порядок в области скачка параметров среды. В отличие от методов, построенных по принципу погружённого интерфейса, он имеет аналитическое выражение и вычислительно менее затратный (в области скачка). По своему типу он скорее подходит для расчёта задач геофизики и им подобных. Метод может быть расширен на двумерный и трёхмерный случаи. Для этого можно использовать методы, использующие расщепление по пространственным координатам, которые сводятся к решению одномерных систем гиперболических уравнений вдоль направлений расщепления [1]. К его преимуществам можно отнести повышенный порядок точности в области скачка параметров среды, меньшую интегральную ошибку вычислений, аналитические выражения для коэффициентов схемы, отсутствие необходимости хранения и вычисления дополнительных параметров, возможность задания параметров среды в каждой ячейке; относительную простоту реализации (не требуется предварительного вычисления и хранения дополнительных данных) и скорость работы. К недостаткам можно отнести требование согласованности параметров среды с сеткой.

В ряде случаев учёт нормали к границе раздела поверхностей может быть критичным. В качестве дальнейшего направления работы предполагается доработка метода для случая несогласованной сетки и учёт нормали к поверхности раздела поверхностей в многомерном случае.

Авторы выражают благодарность Ю.И. Скалько за полезные консультации по работе.

Работа выполнена при финансовой поддержке Российского научного фонда (проект 21-11-00139).

## СПИСОК ЛИТЕРАТУРЫ

1. *LeVeque R.J.* Finite Volume Methods for Hyperbolic Problems. Cambridge, 2002.
2. *Brekhovskikh L.M., Godin O.A.* Acoustics of Layered Media I. Berlin; Heidelberg, 1990.
3. *Moczo P., Kristek J., Galis M.* The Finite-Difference Modelling of Earthquake Motions: Waves and Ruptures. Cambridge, 2014.
4. *Li Z., Ito K.* The Immersed Interface Method: Numerical Solutions of PDEs Involving Interfaces and Irregular Domains. Siam, 2006.
5. *Xu J.* Estimate of the convergence rate of finite element solutions to elliptic equations of second order with discontinuous coefficients // arXiv Prepr. arXiv1311.4178. 2013.
6. *Adjerid S., Ben-Romdhane M., Lin T.* Higher degree immersed finite element methods for second-order elliptic interface problems // Int. J. Numer. Anal. & Model. 2014. V. 11. № 3. P. 541–566.
7. *He X., Lin T., Lin Y., Zhang X.* Immersed finite element methods for parabolic equations with moving interface // Numer. Methods Partial Differ. Equat. 2013. V. 29. № 2. P. 619–646.
8. *Tong F., Wang W., Feng X., Zhao J., Li Z.* How to obtain an accurate gradient for interface problems? // J. Comput. Phys. 2020. V. 405. P. 109070.
9. *Lisitsa V., Podgornova O., Tcheverda V.* On the interface error analysis for finite difference wave simulation // Comput. Geosci. 2010. V. 14. № 4. P. 769–778.
10. *Kaser M., Dumbser M.* An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes-I. The two-dimensional isotropic case with external source terms // Geophys. J. Int. 2006. V. 166. № 2. P. 855–877.
11. *Wilcox L., Stadler G., Burstedde C., Ghattas O.* A high-order discontinuous Galerkin method for wave propagation through coupled elastic-acoustic media // J. Comput. Phys. 2010. V. 229. № 24. P. 9373–9396.
12. *Zhang C., LeVeque R.J.* The immersed interface method for acoustic wave equations with discontinuous coefficients // Wave Motion. 1997. V. 25. № 3. P. 237–263.
13. *Тихонов А.Н., Самарский А.А.* Однородные разностные схемы // Журн. вычислит. математики и мат. физики. 1961. Т. 1. № 1. С. 5–67.
14. *Piraux J., Lombard B.* A new interface method for hyperbolic problems with discontinuous coefficients: one-dimensional acoustic example // J. Comput. Phys. 2001. V. 168. № 1. P. 227–248.
15. *Lombard B., Piraux J.* Numerical treatment of two-dimensional interfaces for acoustic and elastic waves // J. Comput. Phys. 2004. V. 195. № 1. P. 90–116.
16. *Chiavassa G., Lombard B.* Time domain numerical modeling of wave propagation in 2D heterogeneous porous media // J. Comput. Phys. 2011. V. 230. № 13. P. 5288–5309.
17. *Abraham D.S., Marques A.N., Nave J.C.* A correction function method for the wave equation with interface jump conditions // J. Comput. Phys. 2018. V. 353. P. 281–299.
18. *Golubev V., Shevchenko A., Khokhlov N., Petrov I., Malovichko M.* Compact grid-characteristic scheme for the acoustic system with the piece-wise constant coefficients // Int. J. Appl. Mech. 2022. V. 14. № 2. P. 2250002.
19. *Khokhlov N.I., Petrov I.B.* On one class of high-order compact grid-characteristic schemes for linear advection // Russ. J. Numer. Anal. Math. Model. 2016. V. 31. № 6. P. 355–368.
20. *Favorskaya A.V., Zhdanov M.S., Khokhlov N.I., Petrov I.B.* Modelling the wave phenomena in acoustic and elastic media with sharp variations of physical properties using the grid-characteristic method // Geophys. Prospect. 2018. V. 66. № 8. P. 1485–1502.
21. *Ito K., Takeuchi T.* Immersed interface CIP for one dimensional hyperbolic equations // Commun. Comput. Phys. 2014. V. 16. № 1. P. 96–114.
22. *Stognii P.V., Khokhlov N.I., Petrov I.B.* The numerical solution of the problem of the contact interaction in models with gas pockets // J. Phys. Conf. Ser. 2021. V. 1715. № 1. P. 012058.
23. *Golubev V.I., Khokhlov N.I., Nikitin I.S., Churyakov M.A.* Application of compact grid-characteristic schemes for acoustic problems // J. Phys. Conf. Ser. 2020. V. 1479. № 1. P. 012058.
24. *Khokhlov N.I., Favorskaya A., Furgailo V.* Grid-characteristic method on overlapping curvilinear meshes for modeling elastic waves scattering on geological fractures // Minerals. 2022. V. 12. № 12. P. 1597.
25. *Khokhlov N., Favorskaya A., Mitkovets I., Stetsyuk V.* Grid-characteristic method using Chimera meshes for simulation of elastic waves scattering on geological fractured zones // J. Comput. Phys. 2021. V. 446. P. 110637.
26. *Kozhemyachenko A.A., Petrov I.B., Favorskaya A.V., Khokhlov N.I.* Boundary conditions for modeling the impact of wheels on railway track // Comput. Math. Math. Phys. 2020. V. 60. № 9. P. 1539–1554.

27. *Favorskaya A.V., Khokhlov N.I., Petrov I.B.* Grid-characteristic method on joint structured regular and curved grids for modeling coupled elastic and acoustic wave phenomena in objects of complex shape // *Lobachevskii J. Math.* 2020. V. 41. № 4. P. 512–525.
28. *Kholodov Y.A., Kholodov A.S., Tsybulin I. V.* Construction of monotone difference schemes for systems of hyperbolic equations // *Comput. Math. Math. Phys.* 2018. V. 58. № 8. P. 1226–1246.
29. *Courant R., Isaacson E., Rees M.* On the solution of nonlinear hyperbolic differential equations by finite differences // *Commun. Pure Appl. Math.* 1952. V. 5. № 3. P. 243–255.
30. *Lax P., Wendroff B.* Systems of conservation laws // *Commun. Pure Appl. Math.* 1960. V. 13. № 2. P. 217–237.
31. *Русанов В.В.* Разностные схемы третьего порядка точности для прямого вычисления разрывных решений // *Докл. АН СССР.* 1968. Т. 180. № 6. С. 1303–1305.
32. *Холодов А.С., Холодов Я.А.* О критериях монотонности разностных схем для уравнений гиперболического типа // *Журн. вычислит. математики и мат. физики.* 2006. Т. 46. № 9. С. 1638–1667.

Московский физико-технический институт  
(национальный исследовательский университет)

Поступила в редакцию 17.02.2023 г.  
После доработки 29.03.2023 г.  
Принята к публикации 19.05.2023 г.

УДК 517.925+517.93

## СУЩЕСТВОВАНИЕ ЕДИНСТВЕННОЙ НЕПОДВИЖНОЙ ТОЧКИ ОТОБРАЖЕНИЙ, ПОРОЖДЁННЫХ МНОГОМЕРНОЙ СИСТЕМОЙ С РЕЛЕЙНЫМ ГИСТЕРЕЗИСОМ

© 2023 г. А. М. Камачкин, В. В. Евстафьева, Д. К. Потапов

Рассматривается многомерная система обыкновенных дифференциальных уравнений с релейным гистерезисом. Параметры системы полагаются такими, что существует семейство непрерывных операторов, каждый из которых отображает некоторое связное компактное множество в себя. При этом оператору соответствует периодическая орбита с чётным числом точек переключения в фазовом пространстве системы. Для семейства операторов получено необходимое и достаточное условие существования единственной неподвижной точки.

DOI: 10.31857/S0374064123070129, EDN: GWRHOJ

**Введение. Постановка задачи.** Исследуем многомерную нелинейную динамическую систему со сложным разбиением фазового пространства траекториями разных конфигураций, в том числе периодическими орбитами с различным числом точек переключения.

Рассмотрим систему обыкновенных дифференциальных уравнений вида

$$\dot{X} = AX + BF(\sigma). \quad (1)$$

Здесь  $X$  – вектор состояний системы,  $X \in \mathbb{R}^n$ ,  $A$  – постоянная невырожденная квадратная матрица  $n$ -го порядка с вещественными элементами,  $B$  – постоянный ненулевой вектор из  $\mathbb{R}^n$  с вещественными элементами,  $F(\sigma)$  – многозначная функция, описывающая нелинейность типа неидеального (гистерезисного) двухпозиционного реле с пороговыми значениями  $l_1$ ,  $l_2$  ( $l_1 < l_2$ ) и значениями выхода  $m_1$ ,  $m_2$  ( $m_1 < m_2$ ),  $\sigma = (\Gamma, X)$  – скалярное произведение векторов  $\Gamma$  и  $X$ , где  $\Gamma$  – постоянный ненулевой вектор из  $\mathbb{R}^n$  с вещественными элементами.

Функция  $F(\sigma(t))$  определена при непрерывном входе  $\sigma(t)$  для  $t \geq 0$  в классе кусочно-непрерывных функций и задаётся следующим образом: из неравенства  $\sigma(t) \leq l_1$  следует равенство  $F(\sigma) = m_1$ , из неравенства  $\sigma(t) \geq l_2$  – равенство  $F(\sigma) = m_2$ , а из неравенств  $l_1 < \sigma(t) < l_2$  ( $t_1 < t \leq t_2$ ) – равенство  $F(\sigma(t_1)) = F(\sigma(t_2))$ , т.е.  $F(\sigma(t))$  принимает постоянное значение на отрезке  $[t_1, t_2]$ , если  $F(\sigma(t_1)) = m_1$  и  $\sigma(t) < l_2$  или  $F(\sigma(t_1)) = m_2$  и  $\sigma(t) > l_1$ . Обход по петле гистерезиса на плоскости  $(\sigma, F(\sigma))$  совершается против хода часовой стрелки. В приложениях функцию  $F(\sigma(t))$  называют *нелинейной характеристикой системы*.

Из последних исследований нелинейных систем с обратной связью в форме двухпозиционного реле с гистерезисом отметим работы [1–9]. В данной статье продолжены эти исследования.

В пространстве  $\mathbb{R}^n$  уравнение  $(\Gamma, X) = l_i$ ,  $i = 1, 2$ , определяет гиперплоскость, которую называют *поверхностью переключения*, поскольку в её точках происходит переключение реле. Эти точки называют *точками переключения*. Поверхность переключения обозначим через  $L_i$ ,  $i = 1, 2$ .

Траектория любого непрерывного решения системы (1) с точками переключения в фазовом пространстве состоит из кусков траекторий в силу систем

$$\dot{X} = AX + Bm_1, \quad \dot{X} = AX + Bm_2, \quad (2)$$

склеивание которых происходит в точке переключения. Непрерывному периодическому решению системы (1) соответствует замкнутая фазовая траектория (периодическая орбита) с

точкой переключения  $X^*$ , удовлетворяющей равенствам  $X^* = X(t_0) = X(t_0 + T)$ ,  $(\Gamma, X^*) = l_i$  ( $i = 1, 2$ ), где  $t_0$  – начальный момент времени,  $T$  – период, за который изображающая точка решения возвращается в точку  $X^*$ .

Пусть система (1) не имеет положений равновесия. Будем рассматривать непрерывные периодические решения с траекториями, составленными из конечного числа кусков траекторий ввиду систем (2), число этих кусков совпадает с числом точек переключения за период  $T$ . В силу решения системы (1) поверхность  $L_i$  (или её подмножество) отображается в себя.

Непрерывный оператор  $P$ , отображающий некоторое связное множество  $S_i \subset L_i$ ,  $i = 1, 2$ , в себя в силу решения системы (1), представим в виде

$$P(X_0, T(X_0)) = e^{A(T(X_0)-t_0)} \left( X_0 + \int_{t_0}^{T(X_0)} e^{-A(\tau-t_0)} BF(\sigma) d\tau \right), \tag{3}$$

где  $X_0 = X(t_0)$  – начальная точка отображения,  $X_0 \in S_i$ ,  $T(X_0)$  – время возврата изображающей точки по траектории, задаваемой системой (1), в множество  $S_i$ . Точка  $X_0$  произвольная, т.е. может не принадлежать периодической орбите.

Пусть существует периодическое решение  $X(\cdot)$  системы (1) с  $2k$ ,  $k \in \mathbb{N}$ , точками переключения  $X^1, X^2, \dots, X^{2k}$ . Изображающая точка решения начинает движение в точке  $X^1$  на поверхности  $L_1$  в момент времени  $t_0$  и достигает поверхности  $L_2$  в точке  $X^2$  в момент времени  $t_1$  в силу системы (1) при  $m_i = m_1$ , затем переходит на  $L_1$  в точку  $X^3$  в момент времени  $t_2$  при  $m_i = m_2$ , далее достигает  $L_2$  в точке  $X^4$  в момент времени  $t_3$  при  $m_i = m_1$  и так продолжает движение от  $L_1$  к  $L_2$  и обратно  $k$  раз. В момент времени  $t_{2k}$  изображающая точка возвращается на поверхность  $L_1$  в начальную точку  $X^1$  в силу системы (1) при  $m_i = m_2$ . Значит  $t_{2k} = T$ , где  $T$  – период решения.

Пусть имеют место неравенства

$$-(\Gamma, A^{-1}Bm_2) < l_1, \quad -(\Gamma, A^{-1}Bm_1) > l_2. \tag{4}$$

Заметим, что матрица  $A^{-1}$  существует в силу невырожденности  $A$ . Неравенства (4) дают условие, при котором для систем (2) выполняется равенство  $\dot{X} = 0$  в точке  $X_i = -A^{-1}Bm_i$ ,  $i = 1, 2$ , лежащей в фазовом пространстве вне области между поверхностями переключения (т.е. вне зоны неоднозначности  $F(\sigma)$ ).

Далее запишем систему относительно точек и моментов времени переключения в соответствии с описанным выше поведением движения изображающей точки решения. Итак, имеем

$$\begin{aligned} (\Gamma, X^1) = l_1, \quad X^2 &= e^{A(t_1-t_0)} X^1 + \int_{t_0}^{t_1} e^{A(t_1-\tau)} Bm_1 d\tau, \\ (\Gamma, X^2) = l_2, \quad X^3 &= e^{A(t_2-t_1)} X^2 + \int_{t_1}^{t_2} e^{A(t_2-\tau)} Bm_2 d\tau, \\ &\dots \\ (\Gamma, X^{2k-1}) = l_1, \quad X^{2k} &= e^{A(t_{2k-1}-t_{2k-2})} X^{2k-1} + \int_{t_{2k-2}}^{t_{2k-1}} e^{A(t_{2k-1}-\tau)} Bm_1 d\tau, \\ (\Gamma, X^{2k}) = l_2, \quad X^1 &= e^{A(t_{2k}-t_{2k-1})} X^{2k} + \int_{t_{2k-1}}^{t_{2k}} e^{A(t_{2k}-\tau)} Bm_2 d\tau. \end{aligned} \tag{5}$$

Система (5) описывает процесс отображения множества  $S_1 \subset L_1$  в себя. Аналогично можно построить процесс отображения  $S_2 \subset L_2$  в себя. При этом точка  $X^j \in S_i$ ,  $j = \overline{1, 2k}$ ,  $i = 1, 2$ , принадлежит траектории периодического решения.

**Замечание 1.** В общем случае, когда в последнем уравнении вместо точки  $X^1$  берём  $X^{2k+1}$ , система вида (5) описывает процесс отображения множества  $S_i$ ,  $i = 1, 2$ , в себя в силу систем (2). Если  $X^{2k+1} \neq X^1$ , то  $X^j \in S_i$ ,  $j = \overline{1, 2k+1}$ , но не принадлежит траектории периодического решения.

Таким образом, система (5) для любого  $k \in \mathbb{N}$  задаёт отображение множества  $S_1$  в себя с оператором  $P$ , определённым формулой (3). Оператор  $P$  обозначим через  $P_2$ , если отображение множества  $S_1$  в себя имеет два перехода в силу систем (2), через  $P_4$ , если четыре перехода и т.д. В результате получим семейство  $\{P_{2k}\}_{k=1}^{\infty}$ , каждый оператор которого соответствует отображению множества  $S_1$  в себя. Если оператор  $P_{2k}$  имеет неподвижную точку, то существует периодическое решение системы (1) с  $2k$  точками переключения в фазовом пространстве.

Задача состоит в том, чтобы установить условия, при которых система (1) имеет периодические решения, соответствующие операторам из семейства  $\{P_{2k}\}_{k=1}^{\infty}$ .

**1. Свойства множества и семейства операторов.** Рассмотрим свойства множества  $S_i$  ( $i = 1, 2$ ), которое отображается в себя оператором  $P_{2k}$  ( $k \in \mathbb{N}$ ) вида (5).

Пусть собственные числа матрицы  $A$  имеют отрицательные вещественные части. Тогда на поверхности  $L_i$  существует выпуклое компактное множество  $S_i$ , которое в силу системы (1) отображается в себя оператором  $P_{2k}$  для любого  $k \in \mathbb{N}$  (см. подробнее [7, 8]). Множество  $S_i$  является пересечением ограниченного замкнутого множества из  $\mathbb{R}^n$  с поверхностью  $L_i$ . Значит, множество  $S_i$  принадлежит пространству  $\mathbb{R}^{n-1}$ .

В статье [7] доказано, что изображающая точка любого решения системы (1) с точками переключения за конечное время попадает и остаётся в ограниченной области фазового пространства, т.е. оператор  $P_{2k}$  отображает множество  $S_i$  в себя для любого  $k$ . В случае существования периодического решения с  $2k$  изолированными точками переключения в фазовом пространстве установлено, что при  $(\Gamma, B) \neq 0$  эти точки локально непрерывно зависят от параметров  $A$  и  $B$ , а также от  $\Gamma$ ,  $l_1$ ,  $l_2$ ,  $m_1$  и  $m_2$ . Точка переключения является изолированной в том смысле, что в её окрестности не существует точек переключения других периодических орбит. Кроме того, если  $(\Gamma, B) \neq 0$ , то в точках переключения траектория решения одновременно не касается поверхностей переключения.

В работе [8] установлено, что если точка касания принадлежит множеству  $S_i$ , то оно может быть заменено его подмножеством, не содержащим точку касания. Это подмножество сохраняет свойства связности, выпуклости и компактности множества  $S_i$ . Поэтому далее полагаем, что  $S_i$  не содержит точку касания и выполняется условие  $(\Gamma, B) \neq 0$ .

Пусть  $\mathcal{P} = \{P_{2k}\}_{k=1}^{\infty}$ . Тогда согласно (5) имеем, что  $\mathcal{P}$  – это семейство отображений, определённых на множестве  $S_i$  ( $i = 1, 2$ ) из пространства  $\mathbb{R}^{n-1}$  со значениями в этом множестве. Если  $k = 1$ , то  $P_2(X)$  является непрерывным отображением множества  $S_i$  в себя с двумя переходами. Для любых двух различных точек  $X'$  и  $X''$  выполняется условие

$$P_2(X') \neq P_2(X'').$$

Отображение  $P_2(X)$  взаимно однозначное. При  $k \geq 2$  получаем последовательность  $\{P_{2k}\}_{k=2}^{\infty}$ , каждый член которой является суперпозицией непрерывных отображений вида  $P_2(X)$ . Поэтому семейство  $\mathcal{P}$  также различает точки. Кроме того, семейство  $\mathcal{P}$  различает замкнутые множества (см. [8]).

**2. Неподвижные точки семейства операторов.** Поставленная задача сводится к решению проблемы существования неподвижной точки оператора  $P_{2k}$ ,  $k \in \mathbb{N}$ , и установлению связи неподвижных точек семейства операторов  $\mathcal{P}$ .

Теорема Какутани (см., например, [10, с. 638]) даёт достаточное условие существования неподвижной точки замкнутого многозначного отображения  $f: K \rightarrow 2^K$ . Здесь  $K$  – непустое компактное выпуклое множество в банаховом пространстве,  $2^K$  – множество всех подмножеств

множества  $K$  и для каждой точки  $x \in K$  множество  $f(x)$  является непустым выпуклым подмножеством множества  $K$ .

Воспользуемся модификацией теоремы Какутани, которую называют *теоремой Маркова–Какутани*. Приведём формулировку этой теоремы, сохраняя обозначения из монографии [11].

**Теорема 1** [11, с. 220]. Пусть  $E$  – отделимое топологическое векторное пространство,  $K$  – непустое компактное выпуклое множество в  $E$  и  $\Gamma$  – некоторое семейство непрерывных отображений множества  $K$  в себя, удовлетворяющее следующим двум условиям:

а) если  $u \in \Gamma$ ,  $x, y \in K$  и  $\alpha, \beta$  – такие положительные числа, что  $\alpha + \beta = 1$ , то  $u(\alpha x + \beta y) = \alpha u(x) + \beta u(y)$ ;

б) существуют такое натуральное число  $n$  и такие подсемейства  $\Gamma_i$  семейства  $\Gamma$ ,  $0 \leq i \leq n-1$ , что  $\{1\} = \Gamma_n \subset \Gamma_{n-1} \subset \dots \subset \Gamma_0 = \Gamma$ , где  $1$  – тождественное отображение на  $K$ , и каждой паре  $u', u'' \in \Gamma_{i-1}$  соответствует такой элемент  $s \in \Gamma_i$ , что  $u'u'' = u''u's$ .

Тогда существует точка  $x_0 \in K$  такая, что  $u(x_0) = x_0$  для всех отображений  $u \in \Gamma$ .

По условию б) теоремы 1 элементы семейства  $\Gamma$  попарно коммутируют. Рассмотрим частный случай, когда  $\Gamma$  является абелевой полугруппой непрерывных отображений множества  $K$  в себя, которая содержит тождественное отображение  $1$ . Если  $1 \notin \Gamma$ , то тождественное отображение всегда можно присоединить к семейству  $\Gamma$ , не нарушая при этом другие условия теоремы 1 [11, с. 222]. Следовательно, в этом случае теорема Маркова–Какутани утверждает существование неподвижной точки одновременно для всех отображений из некоторого семейства непрерывных отображений непустого выпуклого компактного подмножества топологического векторного пространства в себя.

Покажем, что семейство  $\mathcal{P}$  удовлетворяет условиям теоремы Маркова–Какутани. Имеет место

**Теорема 2.** Пусть в системе (1) матрица  $A$  имеет собственные числа с отрицательными вещественными частями,  $(\Gamma, B) \neq 0$  и выполнены неравенства (4). Семейство операторов  $\mathcal{P}$  имеет единственную неподвижную точку тогда и только тогда, когда для  $n \geq 3$  выполняется условие

$$\Lambda(P_2) = 1 + (-1)^{n-2} \operatorname{sign} J_{P_2}(X_{\text{fp}}) = 0, \quad (6)$$

где  $X_{\text{fp}}$  – неподвижная точка отображения  $P_2(X)$ ,  $J_{P_2}(X_{\text{fp}})$  – якобиан отображения  $P_2(X)$ , вычисленный в точке  $X_{\text{fp}}$ .

**Доказательство. Необходимость.** Пусть семейство операторов  $\mathcal{P}$  имеет единственную неподвижную точку. Доказательство проведём от противного. Согласно теореме 2 из [8] при условиях, наложенных на параметры системы (1), существует по крайней мере одна неподвижная точка  $X_{\text{fp}}$  отображения  $P_2(X)$ . Рассмотрим, например, множество  $S_1 \setminus \{X_{\text{fp}}\}$ , которое является гомотопически эквивалентным сфере  $S^{n-2}$ . Здесь и далее  $n \geq 3$ . Тогда по теореме Лефшеца о неподвижной точке (см. [12, с. 419]) для отображения  $P_2(X)$  число Лефшеца вычисляется по формуле  $\Lambda(P_2) = 1 + (-1)^{n-2} \operatorname{deg} P_2(X)$ , где  $\operatorname{deg} P_2(X)$  – топологическая степень непрерывного отображения  $P_2(X)$ , и если  $\Lambda(P_2) \neq 0$ , то существует ещё одна неподвижная точка отображения  $P_2(X)$  на множестве  $S_1 \setminus \{X_{\text{fp}}\}$ . Множество  $S_1$  не только компактно, но и связно. Поэтому согласно [13, с. 95] имеем  $\Lambda(P_2) = 1 + (-1)^{n-2} \operatorname{sign} J_{P_2}(X_{\text{fp}})$ . Если для отображения  $P_2(X)$  существует неподвижная точка, отличная от  $X_{\text{fp}}$ , то семейство  $\mathcal{P}$  имеет две неподвижные точки. Получили противоречие, которое доказывает необходимость условия (6).

**Достаточность.** Пусть выполнено условие (6) для  $n \geq 3$ . Покажем, что семейство операторов  $\mathcal{P}$  имеет единственную неподвижную точку. Для семейства  $\mathcal{P}$  выполняются условия теоремы 1. Во-первых, отображение  $P_{2k}$  удовлетворяет свойству, близкому к линейности, что следует из формализации вида (5). Во-вторых, операторы семейства  $\mathcal{P}$  попарно коммутируют. Действительно, оператор  $P_{2k} \in \mathcal{P}$  для любого  $k$  взаимно однозначно отображает  $S_i$ ,  $i = 1, 2$ , в себя, при этом обратный к нему оператор  $P_{2k}^{-1}$  тоже принадлежит  $\mathcal{P}$ . Если  $P_{2l} \in \mathcal{P}$  и  $P_{2m} \in \mathcal{P}$ , то  $P_{2l}P_{2m} \in \mathcal{P}$ ,  $l, m \in \mathbb{N}$ , т.е. имеет место равенство  $(P_{2l}P_{2m})X = P_{2l}(P_{2m}X)$ . Как отмечалось выше, к семейству  $\mathcal{P}$  можно присоединить тождественное отображение. Поэтому существует точка  $X_{\text{fp}} \in S_i$  такая, что  $P_{2k}(X_{\text{fp}}) = X_{\text{fp}}$  для любого  $k \in \mathbb{N}$ , в том числе и для оператора  $P_2(X)$ , для которого справедливо условие (6), т.е.

$$\Lambda(P_2) = 1 + (-1)^{n-2} \operatorname{sign} J_{P_2}(X_{\text{fp}}) = 0.$$

Следовательно, на множестве  $S_i \setminus \{X_{\text{fp}}\}$  не существует неподвижной точки, отличной от  $X_{\text{fp}}$ . Таким образом,  $X_{\text{fp}}$  – единственная неподвижная точка семейства  $\mathcal{P}$ . Теорема доказана.

**Замечание 2.** При выполнении условий теоремы 2 оператор  $P_{2k}$ ,  $k \geq 2$ , может иметь неподвижную точку, отличную от  $X_{\text{fp}}$ . В работе [8] получена формула для якобиана  $J_{P_2}(X)$  отображения  $P_2(X)$ . Аналогично можно рассмотреть семейства вида  $\mathcal{P}$ , где “младшими” операторами являются операторы  $P_4$ ,  $P_6$  и т.д. Тогда возникает проблема получения якобианов этих операторов, что является отдельной непростой задачей.

**Замечание 3.** Теорема 2 для семейства операторов  $\{P_{2k}\}_{k=1}^{\infty}$  получена с помощью теоремы Маркова–Какутани. На самом деле можно применить теорему Шаудера, согласно которой у каждого оператора  $P_{2k}$  существует по крайней мере одна неподвижная точка, и ограничиться рассмотрением оператора  $P_2$ , поскольку в силу построения операторов  $P_{2k}$  неподвижная точка у  $P_2$  является неподвижной точкой у всех операторов  $P_{2k}$ .

**Заключение.** В статье сделан ещё один шаг в решении проблемы существования и единственности периодического решения системы (1), а именно, установлено, что в фазовом пространстве системы (1) одновременно могут существовать периодические орбиты с различным числом точек переключения, причём как с выходом из зоны неоднозначности нелинейной характеристики, так и без выхода. Теорема 2 даёт необходимое и достаточное условие существования единственной неподвижной точки одновременно у всех возможных отображений, порождаемых системой (1). В частности, отсюда следует существование единственной унимодальной орбиты (периодической орбиты с двумя точками переключения), которую в инженерной практике удобно принимать за рабочий режим колебательной системы, описываемой системой вида (1).

Работа выполнена при финансовой поддержке Российского научного фонда (проект 23-21-00069).

## СПИСОК ЛИТЕРАТУРЫ

1. Камачкин А.М., Потапов Д.К., Евстафьева В.В. Динамика и синхронизация циклических структур осцилляторов с гистерезисной обратной связью // Вестн. С.-Петерб. ун-та. Прикл. математика. Информатика. Процессы управления. 2020. Т. 16. Вып. 2. С. 186–199.
2. Фурсов А.С., Тодоров Т.С., Крылов П.А., Митрев Р.П. О существовании колебательных режимов в одной нелинейной системе с гистерезисами // Дифференц. уравнения. 2020. Т. 56. № 8. С. 1103–1121.
3. Евстафьева В.В. О существовании двухточечно-колебательных решений возмущённой релейной системы с гистерезисом // Дифференц. уравнения. 2021. Т. 57. № 2. С. 169–178.
4. Евстафьева В.В. Существование  $T/k$ -периодических решений нелинейной неавтономной системы с кратным собственным числом матрицы // Мат. заметки. 2021. Т. 109. № 4. С. 529–543.
5. Євстаф'єва В.В. Існування двоточково-коливних розв'язків релейної неавтономної системи з кратним власним числом дійсної симетричної матриці // Укр. мат. журн. 2021. Т. 73. № 5. С. 640–650.
6. Фурсов А.С., Митрев Р.П., Крылов П.А., Тодоров Т.С. О существовании периодического режима в одной нелинейной системе // Дифференц. уравнения. 2021. Т. 57. № 8. С. 1104–1115.
7. Kamachkin A.M., Potapov D.K., Yevstafyeva V.V. Continuous dependence on parameters and boundedness of solutions to a hysteresis system // Appl. Math. 2022. V. 67. № 1. P. 65–80.
8. Камачкин А.М., Потапов Д.К., Евстафьева В.В. Неподвижные точки отображения, порождённого системой обыкновенных дифференциальных уравнений с релейным гистерезисом // Дифференц. уравнения. 2022. Т. 58. № 4. С. 456–469.
9. Евстафьева В.В. Синтез управления возмущённой системой с неоднозначной нелинейностью // Автоматика и телемеханика. 2023. № 3. С. 44–64.
10. Канторович Л.В., Акилов Г.П. Функциональный анализ. М., 1984.
11. Эдвардс Р. Функциональный анализ. Теория и приложения. М., 1969.
12. Лefшец С. Алгебраическая топология. М., 1949.
13. Понтрягин Л.С. Гладкие многообразия и их применения в теории гомотопий. М., 1976.

Санкт-Петербургский государственный университет

Поступила в редакцию 17.02.2023 г.  
После доработки 29.05.2023 г.  
Принята к публикации 14.06.2023 г.

УДК 517.927.4+517.988.63

## ОБ ОГРАНИЧЕННЫХ ТРАЕКТОРИЯХ АВТОНОМНОЙ СИСТЕМЫ С ВЫДЕЛЕННОЙ ПОЛОЖИТЕЛЬНО ОДНОРОДНОЙ НЕЛИНЕЙНОСТЬЮ

© 2023 г. Э. Мухамадиев, А. Н. Наимов, М. В. Быстрецкий

Исследованы ограниченные траектории автономной системы с выделенной положительно однородной нелинейностью, являющейся градиентом гладкой функции. Доказано существование нестационарных ограниченных траекторий, лежащих в связных компонентах множества точек, где положительно однородная функция отрицательна, и ненулевых стационарных точек в тех связных компонентах, эйлерова характеристика замыкания которых отлична от нуля. Существование нестационарных ограниченных траекторий обосновано с использованием метода Важевского, а существование стационарных точек – с помощью методов вычисления вращения конечномерных векторных полей.

DOI: 10.31857/S0374064123070130, EDN: GXBVQD

**Введение.** Рассмотрим автономную систему вида

$$x'(t) = \nabla v(x(t)) + \eta(|x(t)|)(x(t) - \nabla v(x(t))), \quad x(t) \in \mathbb{R}^n, \quad (1)$$

где  $\mathbb{R}^n$  – евклидово пространство  $n$ -мерных векторов с вещественными координатами,  $n \geq 2$ ,  $\nabla v$  – градиент функции  $v \in C^2(\mathbb{R}^n \setminus \{0\}; \mathbb{R}^1)$ , являющейся положительно однородной порядка  $m > 0$ , т.е.  $v(\lambda y) \equiv \lambda^m v(y)$  для любых  $\lambda > 0$ , и  $\nabla v(y) \neq 0$  при всех  $y \in \mathbb{R}^n \setminus \{0\}$ ,

$$\eta(s) = \begin{cases} 1, & s \leq 0.5, \\ 2(1-s), & 0.5 < s < 1, \\ 0, & s \geq 1. \end{cases}$$

При любом  $y \in \mathbb{R}^n$  существует единственное решение  $\varphi(t, y)$  автономной системы (1), удовлетворяющее начальному условию  $\varphi(0, y) = y$  и определённое на наибольшем интервале  $(\tau_-(y), \tau_+(y))$  изменения  $t$ . Множество  $\{\varphi(t, y) : t \in (\tau_-(y), \tau_+(y))\}$  называют *траекторией решения*  $\varphi(t, y)$  или *траекторией автономной системы* (1), выпущенной из точки  $y$ . Решение  $\varphi(t, y)$  и его траекторию назовём *ограниченными вправо (влево)*, если скалярная функция  $|\varphi(t, y)|$  ограничена на промежутке  $(0, \tau_+(y))$  ( $(\tau_-(y), 0)$ ); в этом случае  $\tau_+(y) = +\infty$  (соответственно  $\tau_-(y) = -\infty$ ). Ограниченное в обе стороны решение и его траекторию назовём просто *ограниченными*.

В статье исследованы ограниченные траектории автономной системы (1), что связано с работами [1, 2], в которых доказаны теоремы существования ограниченных решений для систем нелинейных обыкновенных дифференциальных уравнений вида

$$x'(t) = \nabla v(x(t)) + f(t, x(t)), \quad x(t) \in \mathbb{R}^n. \quad (2)$$

Здесь  $\nabla v$  – главная нелинейная часть, являющаяся градиентом положительно однородной функции порядка  $m > 2$ ,  $f$  – возмущение, удовлетворяющее условию

$$|y|^{1-m} \sup_{t \in \mathbb{R}} |f(t, y)| \rightarrow 0 \quad \text{при} \quad |y| \rightarrow \infty. \quad (3)$$

Из результатов работы [1, теорема 5] следует, что если  $\nabla v(y) \neq 0$  при всех  $y \in \mathbb{R}^n \setminus \{0\}$  и отлично от нуля вращение  $\gamma(\nabla v, S^{n-1})$  векторного поля  $\nabla v : S^{n-1} \mapsto \mathbb{R}^n$  на единичной сфере

$S^{n-1} = \{y \in \mathbb{R}^n : |y| = 1\}$ , то система уравнений (2) имеет по крайней мере одно ограниченное решение при любом возмущении  $f$ , удовлетворяющем условию (3). В статье [2, с. 162] построен пример системы уравнений вида (2), имеющей нестационарные ограниченные решения при  $\gamma(\nabla v, S^{n-1}) = 0$ . В связи с этими результатами авторам представляется актуальным исследование ограниченных решений системы уравнений (2) при конкретных возмущениях  $f$ , удовлетворяющих условию (3).

В настоящей работе доказано существование нестационарных ограниченных траекторий автономной системы (1), лежащих в связных компонентах непустого множества  $\{y \in \mathbb{R}^n : v(y) < 0\}$ . Кроме того, доказано существование ненулевых стационарных точек в тех связных компонентах, эйлерова характеристика замыкания которых отлична от нуля. Существование нестационарных ограниченных траекторий доказано с использованием метода Важевского [3, гл. 10, § 3], а существование стационарных точек обосновано с применением методов вычисления вращения конечномерных векторных полей [4, гл. 1].

**1. Основные результаты.** Предположим, что выполнены следующие условия:

- 1) функция  $v \in C^2(\mathbb{R}^n \setminus \{0\}; \mathbb{R}^1)$ ;
- 2) существует число  $m > 0$  такое, что  $v(\lambda y) \equiv \lambda^m v(y)$  для любого  $\lambda > 0$ ;
- 3)  $\nabla v(y) \neq 0$  для всех  $y \in \mathbb{R}^n \setminus \{0\}$ ;
- 4) множество  $\{y \in \mathbb{R}^n : v(y) < 0\}$  не пусто.

Если выполнены условия 1)–4), то имеем:

а) для любой точки  $y \in \mathbb{R}^n$  существует единственное решение  $\varphi(t, y)$ ,  $t \in (\tau_-(y), \tau_+(y))$ , автономной системы (1), удовлетворяющее начальному условию  $\varphi(0, y) = y$ ;

б) для функции  $v$  верна формула Эйлера  $\langle \nabla v(y), y \rangle \equiv m v(y)$ , где  $\langle z, y \rangle = z_1 y_1 + \dots + z_n y_n$  – скалярное произведение в пространстве  $\mathbb{R}^n$ ;

в) определено вращение  $\gamma(\nabla v, S^{n-1})$  векторного поля  $\nabla v : S^{n-1} \mapsto \mathbb{R}^n$  на  $S^{n-1}$  (другими словами, определена степень отображения  $\nabla v / |\nabla v| : S^{n-1} \mapsto S^{n-1}$ ) (см., например, [4, с. 16]);

д) множество  $\Omega_-(v) = \{y \in S^{n-1} : v(y) < 0\}$  либо совпадает с  $S^{n-1}$ , либо состоит из конечного числа связных компонент  $\Omega_-^j(v)$ ,  $j = 1, p_-(v)$ , с гладкими границами, и в каждой граничной точке вектор  $\nabla v$  сонаправлен с вектором внешней нормали;

е) для каждой связной компоненты  $\Omega_-^j(v)$  определена эйлерова характеристика  $\chi(\overline{\Omega_-^j(v)})$  её замыкания [5, с. 35–36, 410];

ф) справедливо равенство  $\chi(\overline{D_-^j(v)}) = \chi(\overline{\Omega_-^j(v)})$ , где  $\overline{D_-^j(v)}$  – замыкание области

$$D_-^j(v) = \{y \in \mathbb{R}^n : 0.5 < |y| < 1, \quad y/|y| \in \Omega_-^j(v)\}.$$

Далее обозначим через  $\Omega$  произвольную связную компоненту множества  $\Omega_-(v)$  и определим

$$D = \{y \in \mathbb{R}^n : 0.5 < |y| < 1, \quad y/|y| \in \Omega\}, \quad \Omega_1 = \Omega, \quad \Omega_2 = \{y : |y| = 0.5, \quad y/|y| \in \Omega\}.$$

Имеют место следующие теоремы.

**Теорема 1.** Пусть выполнены условия 1)–4). Тогда существуют множества  $M_i \subset \Omega_i$ ,  $i = 1, 2$ , такие, что для любой точки  $y \in M_1 \cup M_2$  выпущенная из неё траектория автономной системы (1) лежит в области  $D$  при  $t \in (0, +\infty)$ .

**Теорема 2.** Если выполнены условия теоремы 1 и  $\chi(\overline{\Omega}) \neq 0$ , то существует стационарная точка автономной системы (1), лежащая в области  $D$ .

Таким образом, если обозначим через  $p_-(v)$  число связных компонент множества  $\Omega_-(v)$ , то, согласно теореме 1, у автономной системы (1) ограничены по крайней мере  $p_-(v)$  траекторий, выпущенных из  $\Omega_2$ , и ограничены вправо по крайней мере  $p_-(v)$  траекторий, выпущенных из  $\Omega_1$ . Кроме того, согласно теореме 2, в тех областях  $D_-^j(v)$ , где  $\chi(\overline{D_-^j(v)}) \neq 0$ , имеется стационарная точка автономной системы (1).

Теорема 1 доказана с использованием метода Важевского [3, гл. 10, § 3]. Суть данного метода состоит в том, что с помощью набора функций выделяется область в фазовом пространстве, где остаются некоторые решения  $x(t)$  системы дифференциальных уравнений при

возрастании  $t$ . В случае автономной системы (1) посредством функции  $v$  выделена область  $D$  и доказано существование траекторий, лежащих в области  $D$  при  $t > 0$ . В теореме 2 существование стационарной точки доказано с применением методов вычисления вращения конечномерных векторных полей [4, гл. 1], а именно, доказана формула

$$\gamma(\nabla v + f, \partial D) = -\chi(\bar{\Omega}),$$

где  $f(y) = \eta(|y|)(y - \nabla v(y))$ , откуда в силу условия  $\chi(\bar{\Omega}) \neq 0$  и принципа ненулевого вращения [4, с. 20] следует существование стационарной точки, лежащей в области  $D$ .

**2. Доказательство теоремы 1.** Очевидно, что если  $\Omega_-(v) = S^{n-1}$ , то для любой точки  $y \in \Omega_1 \cup \Omega_2$  выпущенная из неё траектория автономной системы (1) ограничена вправо.

Пусть  $\Omega_-(v) \neq S^{n-1}$ . Для любого решения  $\varphi(t, y)$ ,  $t \in (\tau_-(y), \tau_+(y))$ , автономной системы (1) имеем

$$\frac{d}{dt}v(\varphi(t, y)) = (1 - \eta(|\varphi(t, y)|))|\nabla v(\varphi(t, y))|^2 + \eta(|\varphi(t, y)|)mv(\varphi(t, y)),$$

$$\frac{1}{2} \frac{d}{dt}|\varphi(t, y)|^2 = (1 - \eta(|\varphi(t, y)|))mv(\varphi(t, y)) + \eta(|\varphi(t, y)|)|\varphi(t, y)|^2.$$

Отсюда следует:

1) если  $y \in \Gamma$ , где  $\Gamma = \partial D \setminus (\Omega_1 \cup \Omega_2)$ , то  $\varphi(t, y) \notin \bar{D}$  при  $t \in (0, \tau_+(y))$ ;

2) если  $y \in \Omega_1 \cup \Omega_2$ , то существует наибольший интервал  $(0, \tau(y))$ , где имеет место включение  $\varphi(t, y) \in D$ ,  $t \in (0, \tau(y))$ . При этом если  $\tau(y) < +\infty$ , то  $\varphi(\tau(y), y) \in \Gamma$ . В силу непрерывной зависимости решения  $\varphi(t, y)$  от  $y$  для  $y \in \Omega_1 \cup \Omega_2$  из малой окрестности  $\partial\Omega_1 \cup \partial\Omega_2$  имеем  $\tau(y) < +\infty$ . Следовательно, каждое из множеств  $\Omega_i^0 = \{y \in \Omega_i : \tau(y) < +\infty\}$ ,  $i = 1, 2$ , не пусто и состоит, по крайней мере, из  $p$  связных компонент, где  $p$  – число связных компонент границы  $\partial\Omega$ . Функция  $\tau$  непрерывна на множестве  $\Omega_1^0 \cup \Omega_2^0$ .

Определим  $M_i = \partial\Omega_i^0 \setminus \partial\Omega_i$ ,  $i = 1, 2$ . Проверим, что  $M_i \neq \emptyset$ ,  $i = 1, 2$ ; тем самым теорема 1 будет доказана. Для определённости предположим, что  $M_1 = \emptyset$ . Тогда  $\Omega_1^0 = \Omega_1$  и  $\varphi(\tau(y), y) \in \Gamma$  для любого  $y \in \Omega_1$ . Функция  $|\varphi(\tau(y), y)|$  непрерывна на множестве  $\Omega_1$  и принимает наименьшее значение в какой-то точке  $y_0 \in \Omega_1$ , так как в точках  $y$ , близких к  $\partial\Omega_1$ , значение  $|\varphi(\tau(y), y)|$  близко к единице. Тогда, в силу непрерывной зависимости решения от начальных данных, в шаре малого радиуса с центром в точке  $y_0$  найдётся точка  $z$ , для которой выпущенная из неё траектория, пересекая множество  $\Omega_1$ , через конечное время  $t_1$  окажется в точке  $\varphi(t_1, z) \in \Gamma$ , где  $|\varphi(t_1, z)| < |\varphi(\tau(y_0), y_0)|$ . Пришли к противоречию. Теорема 1 доказана.

**3. Доказательство теоремы 2.** Аналогично работе [6, лемма 2] можно построить положительно однородную порядка  $m$  функцию  $v_1 \in C^2(\mathbb{R}^n \setminus \{0\}; \mathbb{R}^1)$  такую, что  $v_1$  совпадает с  $v$  на  $\Omega$ ,  $\Omega_-(v_1) = \Omega$  и  $\nabla v_1(y) \neq 0$ ,  $y \in \mathbb{R}^n \setminus \{0\}$ . Тогда имеем  $\gamma(\nabla v + f, \partial D) = \gamma(\nabla v_1 + f_1, \partial D)$ , где  $f(y) = \eta(|y|)(y - \nabla v(y))$ ,  $f_1(y) = \eta(|y|)(y - \nabla v_1(y))$ .

Векторное поле  $\nabla v_1 + f_1$  внутри сферы  $S^{n-1}$  обращается в нуль в точке 0 и, возможно, в точках области  $D$ . Поэтому, согласно свойству вращения [4, с. 16], верно равенство

$$\gamma(\nabla v_1, S^{n-1}) = \gamma(\nabla v_1 + f_1, \partial(|y| < 0.5)) + \gamma(\nabla v_1 + f_1, \partial D) = 1 + \gamma(\nabla v_1 + f_1, \partial D).$$

Следовательно,

$$\gamma(\nabla v_1, S^{n-1}) = 1 + \gamma(\nabla v + f, \partial D).$$

Применяя к функции  $v_1$  формулу из работы [7], находим

$$\gamma(\nabla v_1, S^{n-1}) = 1 - \chi(\bar{\Omega}_-(v_1)) = 1 - \chi(\bar{\Omega}).$$

Значит,  $\gamma(\nabla v + f, \partial D) = -\chi(\bar{\Omega})$ . Отсюда, в силу условия  $\chi(\bar{\Omega}) \neq 0$  и принципа ненулевого вращения [4, с. 20], следует существование точки из  $D$ , где векторное поле  $\nabla v + f$  обращается в нуль. Данная точка является стационарной для автономной системы (1). Теорема 2 доказана.

**4. Пример функции.** Рассмотрим функцию вида

$$v(y) = |y|^{m+1-q} \prod_{i=1}^q (\langle c_i, y \rangle - d_i |y|), \quad (4)$$

где  $m > 1$ ,  $c_i \in \mathbb{R}^n$ ,  $d_i \in \mathbb{R}^1$ ,  $i = \overline{1, q}$ ,  $q \geq 1$ . В статье [8] доказано, что для положительно однородной функции  $v$ , заданной формулой (4), условие  $\nabla v(y) \neq 0$  при всех  $y \in \mathbb{R}^n \setminus \{0\}$  равносильно двум условиям:

$$A_1: |c_i| \neq |d_i|, \quad i = \overline{1, q};$$

$A_2$ : при любых  $i, j \in \{1, \dots, q\}$ ,  $i \neq j$ , верна импликация

$$(|c_i| > |d_i|, |c_j| > |d_j|) \implies |d_i c_j - d_j c_i| > (|c_i|^2 |c_j|^2 - \langle c_i, c_j \rangle^2)^{1/2}.$$

Множество  $\Omega_-(v)$  не пусто лишь в том случае, когда не пусто множество  $\{i : |c_i| > |d_i|\}$ . В этом случае граница произвольной связной компоненты  $\Omega$  множества  $\Omega_-(v)$  состоит из конечного числа  $(n-2)$ -мерных ориентируемых гладких многообразий без края, диффеоморфных  $S^{n-2}$ . Другими словами, связную компоненту  $\Omega$  можно получить из  $S^{n-1}$ , удаляя какие-то  $k$  множеств, каждое из которых диффеоморфно шару  $B^{n-1} = \{y \in \mathbb{R}^{n-1} : |y| \leq 1\}$ . Поэтому, зная  $k$  и воспользовавшись известными свойствами эйлеровой характеристики [5, с. 35–36, 410], находим

$$\chi(\overline{\Omega}) = \chi(S^{n-1}) - k\chi(B^{n-1}) + k\chi(S^{n-2}),$$

$$\chi(\overline{\Omega}) = 1 + (-1)^{n-1} - k + k(1 + (-1)^{n-2}) = 1 + (-1)^n(k-1).$$

Отсюда, согласно теореме 2, можно получить ответ на вопрос о существовании стационарной точки в соответствующей области  $D$ .

Работа выполнена при финансовой поддержке Российского научного фонда (проект 23-21-00032).

#### СПИСОК ЛИТЕРАТУРЫ

1. Мухамадиев Э. О построении правильной направляющей функции для системы дифференциальных уравнений // Докл. АН СССР. 1970. Т. 190. № 4. С. 777–779.
2. Мухамадиев Э., Наимов А.Н. Критерии существования периодических и ограниченных решений для трёхмерных систем дифференциальных уравнений // Тр. Ин-та математики и механики УрО РАН. 2021. Т. 27. № 1. С. 157–172.
3. Хартман Ф. Обыкновенные дифференциальные уравнения. М., 1970.
4. Красносельский М.А., Забрейко П.П. Геометрические методы нелинейного анализа. М., 1975.
5. Борисович Ю.Г., Близняков Н.М., Израилевич Я.А., Фоменко Т.Н. Введение в топологию. М., 2014.
6. Mukhamadiev E., Naimov A.N. On the homotopy classification of positively homogeneous functions of three variables // Iss. Anal. 2021. V. 10. № 2. P. 67–78.
7. Мухамадиев Э. Ограниченные решения и гомотопические инварианты систем нелинейных дифференциальных уравнений // Докл. РАН. 1996. Т. 351. № 5. С. 596–598.
8. Мухамадиев Э., Наимов А.Н. Об априорной оценке и существовании периодических решений для одного класса систем нелинейных обыкновенных дифференциальных уравнений // Изв. вузов. Математика. 2022. № 4. С. 37–48.

Вологодский государственный университет

Поступила в редакцию 06.02.2023 г.

После доработки 06.02.2023 г.

Принята к публикации 19.05.2023 г.