

СОДЕРЖАНИЕ

Том 57, номер 7, 2021

ЧИСЛЕННЫЕ МЕТОДЫ

Вопросы устойчивости в двумерных математических моделях равновесия плазмы в магнитных ловушках-галатеях <i>К. В. Брушлинский, Е. В. Степин</i>	867
Схемы расщепления решения для эволюционных уравнений второго порядка <i>П. Н. Вабищевич</i>	880
Об одном классе интегро-дифференциальных уравнений в особом случае <i>Н. С. Габбасов</i>	889
Численные методы решения нелинейного операторного уравнения, возникающего в обратной коэффициентной задаче <i>С. В. Гаврилов, А. М. Денисов</i>	900
К вопросу об эквивалентности разностных методов решения задачи Стефана на подвижных и фиксированных сетках <i>А. О. Гусев, О. В. Щерица, О. С. Мажорова</i>	907
Спектральные условия устойчивости явной трёхслойной разностной схемы для многомерного уравнения переноса с возмущениями <i>А. А. Злотник, Б. Н. Четверушкин</i>	922
Квадратурная формула для гармонического потенциала двойного слоя <i>П. А. Крутицкий, И. О. Резниченко</i>	932
Об устойчивых методах Рунге–Кутты для решения гиперболических уравнений разрывным методом Галёркина <i>В. В. Лукин, В. Н. Корчагова, С. М. Сауткина</i>	951
Компактные разностные схемы на трёхточечном шаблоне для гиперболических уравнений второго порядка <i>П. П. Матус, Хоанг Тхи Киеу Ань</i>	963
Алгоритм адаптивной интерполяции на разреженных сетках для численного интегрирования систем обыкновенных дифференциальных уравнений с интервальными неопределённостями <i>А. Ю. Морозов, Д. Л. Ревизников</i>	976
Численное решение уравнений релаксационной модели Баера–Нунциато с помощью разрывного метода Галёркина <i>Р. Р. Тухватуллина, М. В. Алексеев, Е. Б. Савенков</i>	988

ЧИСЛЕННЫЕ МЕТОДЫ

УДК 519.632+517.958:533.9

ВОПРОСЫ УСТОЙЧИВОСТИ В ДВУМЕРНЫХ
МАТЕМАТИЧЕСКИХ МОДЕЛЯХ РАВНОВЕСИЯ ПЛАЗМЫ
В МАГНИТНЫХ ЛОВУШКАХ-ГАЛАТЕЯХ

© 2021 г. К. В. Брушлинский, Е. В. Степин

Ловушки-галатеи для удержания плазмы в магнитном поле, которое создаётся проводниками с током, погружёнными в плазменный объём, составляют перспективный класс объектов разработок в области управляемого термоядерного синтеза. После описания основных свойств и количественных характеристик ловушек центральной проблемой является исследование устойчивости равновесных магнитоплазменных конфигураций. Существенную роль играют здесь математическое моделирование и расчёты, выполненные в терминах дифференциальных уравнений магнитной газодинамики. Они изложены на примере распрямлённой в цилиндр тороидальной ловушки “галатея-пояс”. В статье представлена её плазмостатическая модель и несколько подходов к исследованию устойчивости конфигураций: сходимость итерационных методов установления равновесия в двумерных моделях, краткий обзор исследования магнитогазодинамической устойчивости одномерных конфигураций, окружающих прямой проводник с током, и строгое исследование устойчивости конфигураций относительно двумерных возмущений в линейном приближении. В численных расчётах получены критерии устойчивости в разных приближениях и их соотношения между собой. Указаны пути обобщения результатов на трёхмерные возмущения.

DOI: 10.31857/S0374064121070013

Введение. Тематика, представленная названием нашей статьи, постоянно присутствует в работах по приложениям вычислительной математики к проблемам физики плазмы, начиная с середины прошлого столетия. Интерес к ней обусловлен многообещающими научно-техническими приложениями и, в первую очередь, перспективой получить в большом количестве дешёвую энергию управляемого термоядерного синтеза. Осуществить желаемую реакцию синтеза лёгких элементов таблицы Менделеева предполагается в ловушках, в которых плотная и нагретая до высоких температур плазма должна удерживаться магнитным полем в течение требуемого времени. Особое внимание уделяется преодолению возникающих при этом многочисленных плазменных неустойчивостей. Поэтому основным объектом исследований в данной области науки являются конфигурации плазмы, поля и создающего его электрического тока, находящиеся в равновесии в разнообразных ловушках.

Наиболее распространены ловушки тороидальной формы, позволяющие максимально избежать контактов горячей плазмы с элементами конструкции. В данной работе рассматривается специальный класс тороидальных ловушек, в которых проводники с током, индуцирующим магнитное поле, погружены в плазменный объём, но не соприкасаются с плазмой. Внимание к ним привлечено А.И. Морозовым [1], который назвал их галатеями и инициировал серию конкретных разработок таких ловушек [2–4]. Существенную роль в этих разработках играют математические модели основных физических процессов и большие объёмы относящихся к ним численных исследований. Они облегчают теоретическую часть работы, а также позволяют более экономно вести громоздкие и дорогостоящие эксперименты и анализировать их результаты. С численным моделированием равновесных магнитоплазменных конфигураций в упомянутых выше ловушках-галатеях можно ознакомиться, например, по статьям [5–8], обзору [9] и монографиям [10, с. 65; 11, с. 152]. Математический аппарат моделей достаточно плотной плазмы основывается на приближении механики сплошных сред, т.е. используются дифференциальные уравнения магнитной газодинамики (МГД), а более конкретно – численное решение краевых МГД-задач.

Основные проблемы в исследовании конфигураций в ловушках и соответствующих численных моделей сводятся, во-первых, к описанию их возможных равновесных состояний и,

во-вторых, к их устойчивости относительно хотя бы малых, но достаточно произвольных возмущений.

Равновесия исследуются с помощью краевых задач для уравнений плазмостатики, в общем случае трёхмерных и достаточно сложных. Однако широко распространены ловушки, которые обладают симметрией (плоской, осевой, винтовой) или допускают её в каком-либо приближении. Симметрия позволяет понизить размерность задач о равновесии до двумерных, а в особо простых случаях – даже до одномерных. В двумерных задачах систему уравнений плазмостатики удаётся свести к одному скалярному уравнению Грэда–Шафранова для функции магнитного потока [12, 13]. Одномерные задачи имеют дело с обыкновенными дифференциальными уравнениями, которые несложно решаются аналитически. В исследованиях основных принципиальных вопросов на качественном уровне допускается ещё одно упрощение: тороидальные ловушки “распрямляются” в цилиндр, т.е. в тор бесконечного радиуса. Необходимые для этого количественные поправки на тороидальность оцениваются, например, в работе [7].

Задачи об устойчивости равновесных магнитоплазменных конфигураций в ловушках обсуждаются и рассматриваются во многих научных работах; в общем виде они поставлены в [14–16; 17, с. 82] и решены в ряде конкретных случаев. Из работ последнего времени отметим серию статей С.Ю. Медведева с соавторами (см. статью [18] и приведённую в ней библиографию).

В настоящей работе устойчивость рассматривается на примере двумерных равновесных конфигураций в распрямлённом аналоге тороидальной “галатее-пояса” – цилиндре с двумя погружёнными в него прямыми проводниками. Приведена постановка задачи о МГД-устойчивости относительно произвольных трёхмерных возмущений в линейном приближении. Эта задача сведена к двумерным задачам для отдельных гармоник по координате z , указаны возможные подходы к её решению, получен критерий устойчивости. Кроме того, рассмотрены некоторые упрощения исходной задачи. Во-первых, сходимость итерационных методов установления в численном решении двумерных плазмостатических задач о равновесии трактуется как разновидность устойчивости, но только относительно возмущений магнитного поля той же размерности. Такая устойчивость названа “диффузионной” [19]. Она необходима, но недостаточна для МГД-устойчивости, однако может представлять интерес, поскольку оценивается количественными критериями и указывает на обстоятельства, сопутствующие или препятствующие общепринятой устойчивости. Во-вторых, общим элементом всех ловушек-галатей являются погружённые в плазму проводники с током. Поэтому целесообразно рассмотреть одномерную задачу о плазменной конфигурации, окружающей один прямой проводник и не соприкасающийся с ним. В серии работ последнего времени [20–23] рассмотрены такие задачи с заданным распределением давления по радиусу с различными его вариантами вблизи внешней границы окрестности. Эти конфигурации оказались диффузионно устойчивыми при допустимых значениях давления плазмы, но МГД-устойчивость потребовала более сильных ограничений на давление, зависящих от его распределения у внешней границы.

В целях единообразного представления поставленных вопросов и их решения в п. 1 кратко изложена двумерная математическая модель “галатее-пояса”, даны определение диффузионной устойчивости и её интерпретация в терминах спектральных свойств оператора линеаризованной задачи. В п. 2 кратко перечислены результаты об устойчивости одномерных конфигураций, окружающих один прямой проводник с током. В п. 3 изложена логика МГД-устойчивости двумерных конфигураций в галатеех и результаты численного исследования устойчивости относительно двумерных возмущений, в том числе динамических. Указаны пути исследования МГД-устойчивости относительно произвольных трёхмерных возмущений.

1. Математическая модель равновесия в “галатее-поясе”. Диффузионная устойчивость. Распрямлённый аналог тороидальной ловушки “галатее-пояс” [3] представляет собой бесконечный цилиндр с двумя погружёнными в него прямыми проводниками конечного диаметра. Математическая модель равновесных конфигураций в случаях круглого и квадратного сечений цилиндра плоскостью, перпендикулярной его оси, изложена в работах [6] и [7] соответственно. Конфигурации расположены в приосевой части цилиндра и практически не зависят от формы внешней границы. Поэтому кратко воспроизведём рассматриваемую модель для упрощения в случае квадратной области сечения цилиндра (см. рис. 1, на котором

показаны сечения цилиндра и проводников плоскостью $z = \text{const}$; ось цилиндра совпадает с осью z).

МГД-равновесие описывается уравнениями плазмостатики [10, с. 65; 11, с. 152; 14; 15]

$$\nabla p = \mathbf{j} \times \mathbf{H}, \quad \mathbf{j} = \text{rot } \mathbf{H}, \quad \text{div } \mathbf{H} = 0, \quad \mathbf{j} = \mathbf{j}^{\text{pl}} + \mathbf{j}^{\text{ex}}, \quad (1)$$

связывающими давление p , напряжённость магнитного поля $\mathbf{H} = (H_x, H_y, 0)$ и плотность электрического тока $\mathbf{j} = (0, 0, j)$. Ток складывается из тока j^{pl} в плазме и заданного тока j^{ex} в проводнике. Ток в проводниках вводится в задачу как непрерывная функция, сосредоточенная в основном в области проводников. Такое задание тока позволяет ставить задачу в односвязной области сечения цилиндра, не выделяя из неё “островки” с проводниками [5; 6; 10, с. 69; 11, с. 157]. В наших работах эта функция имеет вид

$$j^{\text{ex}} = j_0 \sum_{k=1}^2 \exp\{-(x - x_k)^2 + y^2/r_c^2\}, \quad j_0 = 2/r_c^2, \quad (2)$$

где $x_1 = 1$, $x_2 = -1$ – абсциссы центров, r_c – условный радиус проводников, а множитель j_0 обеспечивает равенство интеграла от функции (2), взятого по окрестности проводника, величине заданного тока в нём.

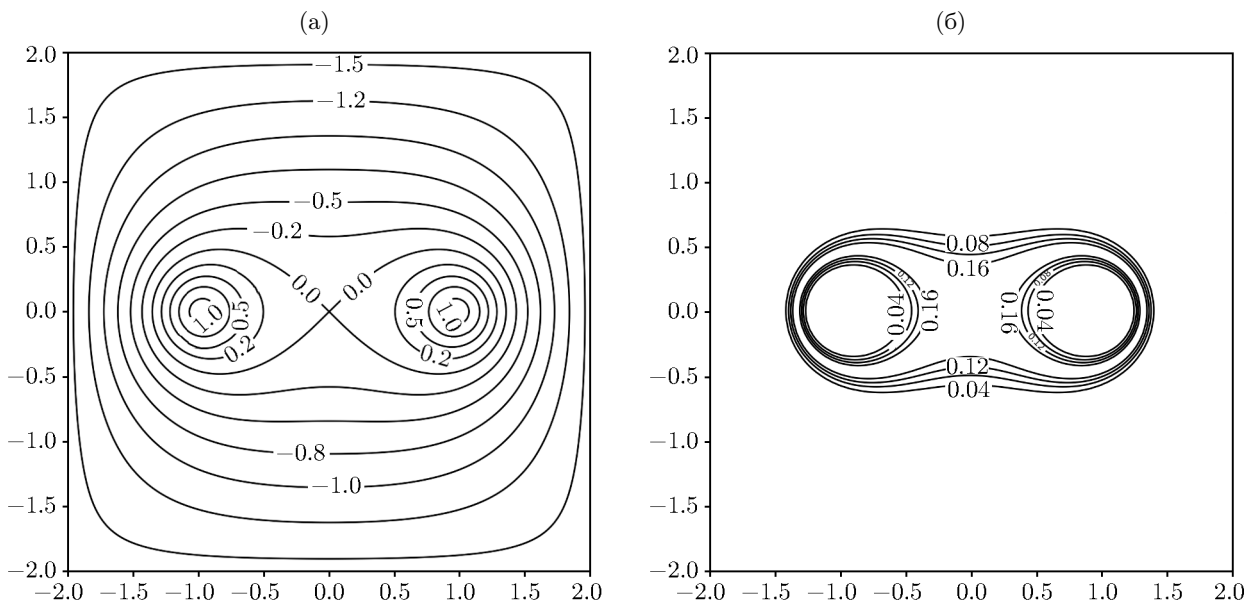


Рис. 1. Равновесные конфигурации магнитного поля (а) и давления плазмы (б) в сечении цилиндра при $r_c = 0.2$, $q = 0.2$, $p_0 = 0.2$.

Уравнения (1) и функция (2) представлены в безразмерной форме, т.е. все переменные в них отнесены к единицам измерения, отмеченным индексом u , которые составлены из заданных размерных величин: x_0 – расстояния от оси цилиндра до центров проводников и J_c – электрического тока в каждом из них:

$$x_u = y_u = x_0, \quad H_u = \frac{2J_c}{cx_u}, \quad j_u = \frac{c}{4\pi} \frac{H_u}{x_u}, \quad p_u = \frac{H_u^2}{4\pi}. \quad (3)$$

Рассматриваемая магнитоплазменная конфигурация обладает плоской симметрией ($\partial/\partial z \equiv 0$), т.е. двумерна, что сильно упрощает математический аппарат модели. Магнитное поле $\mathbf{H}$ выражается через z -компоненту вектор-потенциала $\psi = \psi_z$ по формулам

$$\mathbf{H} = \text{rot } \Psi, \quad H_x = \frac{\partial \psi}{\partial y}, \quad H_y = -\frac{\partial \psi}{\partial x}.$$

Линии уровня $\psi(x, y) = \text{const}$ – магнитные силовые линии в сечении цилиндра плоскостью $z = \text{const}$, а физический смысл функции ψ – функция магнитного потока. Система уравнений (1) сводится к одному скалярному уравнению – плоской разновидности уравнения Грэда–Шафранова [12, 13]:

$$\Delta\psi + \frac{dp}{d\psi} + j^{\text{ex}} = 0. \quad (4)$$

Внешняя граница ловушки предполагается непрозрачной для магнитного поля, что отражено в граничном условии $\psi = \psi|_{\Gamma} = \text{const}$. Давление p функционально зависит от ψ , и зависимость $p(\psi)$ должна быть задана при постановке конкретной задачи исходя из дополнительных требований к исследуемой конфигурации.

В задачах о галатеях естественно требовать, чтобы проводники внутри плазменного объёма не контактировали с плазмой. Это достигается, например, заданием $p(\psi)$ в виде

$$p(\psi) = p_0 \exp\{-(\psi - \psi_0)^2/q^2\}. \quad (5)$$

Здесь давление максимально на силовой линии $\psi(x, y) = \psi_0$ – сепаратрисе магнитного поля в “поясе”, проходящей через центр (см. рис. 1), и быстро убывает при удалении от неё, если параметр q достаточно мал. Таким образом, краевая задача для уравнения (4) содержит параметр ψ_0 , определяемый дополнительным условием

$$\psi_0 = \psi(0, 0). \quad (6)$$

На внешней границе без ограничения общности можно положить $\psi|_{\Gamma} = 0$. Задача решается итерационным методом установления. Уравнение (4) заменяется параболическим уравнением

$$\frac{\partial\psi}{\partial t} = \Delta\psi + g(x, y, \psi), \quad (7)$$

где $g(x, y, \psi)$ – младшие члены в (4).

В его разностном аналоге переход от n -го слоя (итерации) к $(n+1)$ -му в главных членах с производными от ψ осуществляется методом продольно-поперечной прогонки [11, с. 246; 24; 25], а нелинейное слагаемое $g(x, y, \psi)$ в (7) и параметр ψ_0 в (6) берутся с предыдущего слоя.

Результаты расчётов “пояса”, полученные в работах [6, 7], проиллюстрированы на рис. 1. Плазменная конфигурация сосредоточена в центре области и имеет форму криволинейного четырёхугольника с выпуклыми внутрь границами и примыкающими к ним узкими полосками вдоль магнитной сепаратрисы. Общий во всех вариантах расчётов результат состоит в том, что итерационный процесс решения задачи сходится лишь при ограничении

$$p_0 < p_0^{\text{diff}} \quad (8)$$

на максимальное значение давления, отнесённое к магнитной единице. Это позволяет считать неравенство (8) условием так называемой диффузионной устойчивости [19] – устойчивости только при возмущениях магнитного поля той же размерности, т.е. двумерных. Она не зависит от численного метода решения задачи, поскольку связана с внутренней природой математической модели – положительной определённой дифференциального оператора

$$L[u] \equiv -\Delta u - \frac{dg}{d\psi}u \quad (9)$$

линеаризованной задачи. Сходимость итераций имеет место, если погрешность решения задачи с уравнением (7) $u = \psi - \psi_{\text{eq}}$, где ψ_{eq} – искомое решение стационарной задачи, стремится к нулю при $t \rightarrow \infty$. Оно строится из экспонент

$$u(t, x, y) = e^{-\lambda t}u(x, y), \quad (10)$$

удовлетворяющих уравнению

$$L[u] = \lambda u. \quad (11)$$

Поэтому $u \rightarrow 0$ в формуле (10) при любых “начальных” данных, если все собственные значения задачи (11) положительны.

Диффузионная устойчивость, очевидно, необходима, но недостаточна для МГД-устойчивости конфигураций, однако может быть полезна при анализе последней, так как содержит грубый, но доступный критерий. Спектр оператора $L[u]$ следует сравнить со спектром оператора $L_0[u] \equiv -\Delta u$, положительно определённого практически при любых однородных граничных условиях. Известно, что если коэффициент при u во втором слагаемом правой части (9) положителен, то собственные значения оператора $L[u]$ смещены вправо по сравнению со спектром оператора $L_0[u]$ [26, с. 86]. Отсюда вытекает достаточное условие диффузионной устойчивости

$$\frac{dg}{d\psi} \equiv \frac{d^2p}{d\psi^2} \leq \mu_1, \quad (12)$$

где μ_1 – старшее (минимальное) собственное значение оператора $L_0[u]$ [21], которое легко находится аналитически и зависит только от геометрии задачи и граничных условий.

В квадратной области $-2 < (x, y) < 2$ на рис. 1 $\mu_1 = \pi^2/8 \approx 1.2321$. Поскольку

$$\frac{d^2p}{d\psi^2} = -\frac{2p_0}{q^2} \left(1 - 2 \left(\frac{\psi - \psi_0}{q} \right)^2 \right) \exp \left(-\frac{(\psi - \psi_0)^2}{q^2} \right) \leq 0.4724 \frac{p_0}{q^2}, \quad (13)$$

то из (12) и (13) следует, что $p_0 \leq 2.117\mu_1 q^2 = 2.61q^2$.

Расчёты показали, что в этом примере со значением $q = 0.2$ итерационный процесс решения задачи сходится при гораздо более слабом ограничении $p_0 \leq 4.5$ [7], что подтверждает лишь достаточность критерия (12) для диффузионной устойчивости. В него коэффициент при u во втором слагаемом оператора (9) входит только посредством максимального значения $d^2p/d\psi^2$, и поэтому без внимания остаётся существенный участок значений ψ , близких к ψ_0 , где $d^2p/d\psi^2 < 0$.

В серии расчётов конфигураций с параметрами $r_c = 0.2$, $q = 0.2$ получены равновесные конфигурации, расположенные на конечном расстоянии от проводников. Пример распределения магнитного поля и давления при $p_0 = 0.2$ приведён на рис. 1. Электрический ток j равен нулю в центре и на сепаратрисе магнитного поля, где давление максимально, и положителен в области между сепаратрисой и внешней границей. Здесь он практически совпадает с $j^{\text{pl}} = j - j^{\text{ex}}$, так как заданный ток j^{ex} сосредоточен вблизи центров проводников. В области между сепаратрисой и проводниками $j^{\text{pl}} = j - j^{\text{ex}} < 0$. Таким образом, сила Ампера $(\mathbf{j} \times \mathbf{H})_z = -jH$ направлена в сторону сепаратрисы и способствует изоляции проводников и внешней границы от плазмы. При значениях $p_0 < 0.8$ границы четырёхугольника выпуклы внутрь, что благоприятствует устойчивости [27]. С возрастанием p_0 конфигурация увеличивается в объёме, становится выпуклой в сторону внешней границы и стремится приблизиться к проводникам. Время установления решения задачи с уравнением (7) возрастает, и по мере приближения $p_0 \rightarrow p_0^{\text{diff}}$ установление прекращается.

Таким образом, двумерная математическая модель равновесия плазмы в ловушке “галатей-пояс” основана на численном решении краевой задачи для уравнения Грэда–Шафранова. Рассмотрены свойства равновесных конфигураций и получено достаточное условие их диффузионной устойчивости в терминах спектра дифференциального оператора линеаризованной задачи.

2. Одномерная модель конфигураций, окружающих проводник с током. Ещё один относительно простой, но достаточно общий подход к исследованию устойчивости равновесия плазмы в галатейях можно рассмотреть в типичном для всех ловушек этого класса элементе – окрестности одного отдельно взятого проводника. Здесь можно ограничиться одномерными задачами в кольцевой окрестности прямого токнесущего проводника и получить определённую количественную информацию об окружающих его, но не соприкасающихся с ним конфигурациях. Эти задачи можно считать обобщением задач о равновесных конфигурациях в круглом цилиндре, в частности, хорошо известного $\mathbb{Z}$ -пинча (см., например, [15; 16; 10, с. 94; 11, с. 168]. Постановка задач и некоторые результаты их решений содержатся в работах [20–23].

Подытожим их результаты и выводы из них кратким обзором, имея в виду, что постановка задач более подробно изложена в работе [21]. Математические модели равновесных конфигураций рассмотрены в кольцевой области $1 < r < R$, содержащейся в сечении плазменного цилиндра плоскостью $z = \text{const}$ и охватывающей прямой круглый проводник радиуса 1. Модели одномерны: единственной независимой пространственной переменной является радиальная координата r . Вектор плотности электрического тока $\mathbf{j}$ в проводнике и окружающей его плазме направлен вдоль оси z , а напряжённость магнитного поля $\mathbf{H}$ – только по азимуту φ . Три скалярные функции $p(r)$, $j(r)$ и $H(r)$ подчиняются двум уравнениям, к которым сводятся уравнения плазмостатики (1):

$$\frac{dp}{dr} = -jH, \quad j = \frac{1}{r} \frac{d(Hr)}{dr}. \quad (14)$$

Они записаны в безразмерной форме. В качестве единицы измерения радиуса выбран радиус проводника: $r_u = r_c$, остальные единицы те же, что и выше (см. (3)). Уравнений (14) недостаточно для определения трёх функций, поэтому в математической модели каждой конкретной конфигурации следует задать одну произвольную функцию в соответствии с дополнительными требованиями или имеющейся информацией. Естественно задать давление $p(r)$ так, чтобы проводник не соприкасался с горячей плазмой, например, в виде квадратичной функции

$$p = p_0 \left(1 - \left(\frac{r_1 - r}{r_1 - 1} \right)^2 \right), \quad 1 < r < r_1, \quad (15)$$

возрастающей от нуля на поверхности проводника до своего максимального значения p_0 на заданном от него расстоянии $r_1 - 1 > 0$. При $r > r_1$ давление остаётся постоянным или убывает в сторону внешней границы различными способами, как показано в работах [20–23]. Примеры рассматривавшихся распределений давления представлены на рис. 2.

Если задано распределение $p(r)$, то магнитное поле $H(r)$ и ток $j(r)$ находятся интегрированием уравнений (14). При этом возникает общее для всех конфигураций ограничение на максимальное значение заданного давления

$$p_0 \leq p_0^{\text{cr}} = 3/(r_1^2 + 2r_1 + 3). \quad (16)$$

Устойчивость этих одномерных конфигураций исследована двумя способами. Первый из них – в терминах рассмотренной выше диффузионной устойчивости, соответствующей сходимости итерационного метода решения краевых задач для уравнения Грэда–Шафранова. Для математической модели одномерных конфигураций в этом уравнении нет необходимости, поскольку она получается аналитически после задания $p(r)$ формулой типа (15), однако их устойчивость может, как уже отмечалось, представлять интерес. Одномерный вариант краевой задачи для уравнения (4) в нашем случае имеет вид

$$\frac{1}{r} \frac{d}{dr} \left(r \frac{d\psi}{dr} \right) + g(\psi) = 0; \quad \frac{d\psi}{dr}(1) = -1, \quad \psi(R) = 0,$$

где

$$g(\psi) = \frac{dp}{d\psi} = \frac{dp/dr}{d\psi/dr} = j, \quad \frac{d\psi}{dr} = -H.$$

Численное решение её методом установления получено во всех рассчитанных вариантах ловушки, достаточный критерий типа (12) также выполнен. Однако участвующий в этом критерии коэффициент

$$\frac{dg}{d\psi} = \frac{1}{H^2} \left(\frac{d^2 p}{dr^2} + \frac{1}{r} \frac{dp}{dr} + \frac{1}{H^2} \left(\frac{dp}{dr} \right)^2 \right)$$

может сильно возрасти при $p_0 \rightarrow p_0^{\text{cr}}$ и $r_1 \rightarrow 1$. Это соответствовало бы неограниченному приближению давления к критическому или конфигурации – к проводнику, что противоречит физическому смыслу ловушек-галатей. С учётом этого замечания можно сделать вывод, что исследованные варианты окружающих проводник конфигураций диффузионно устойчивы.

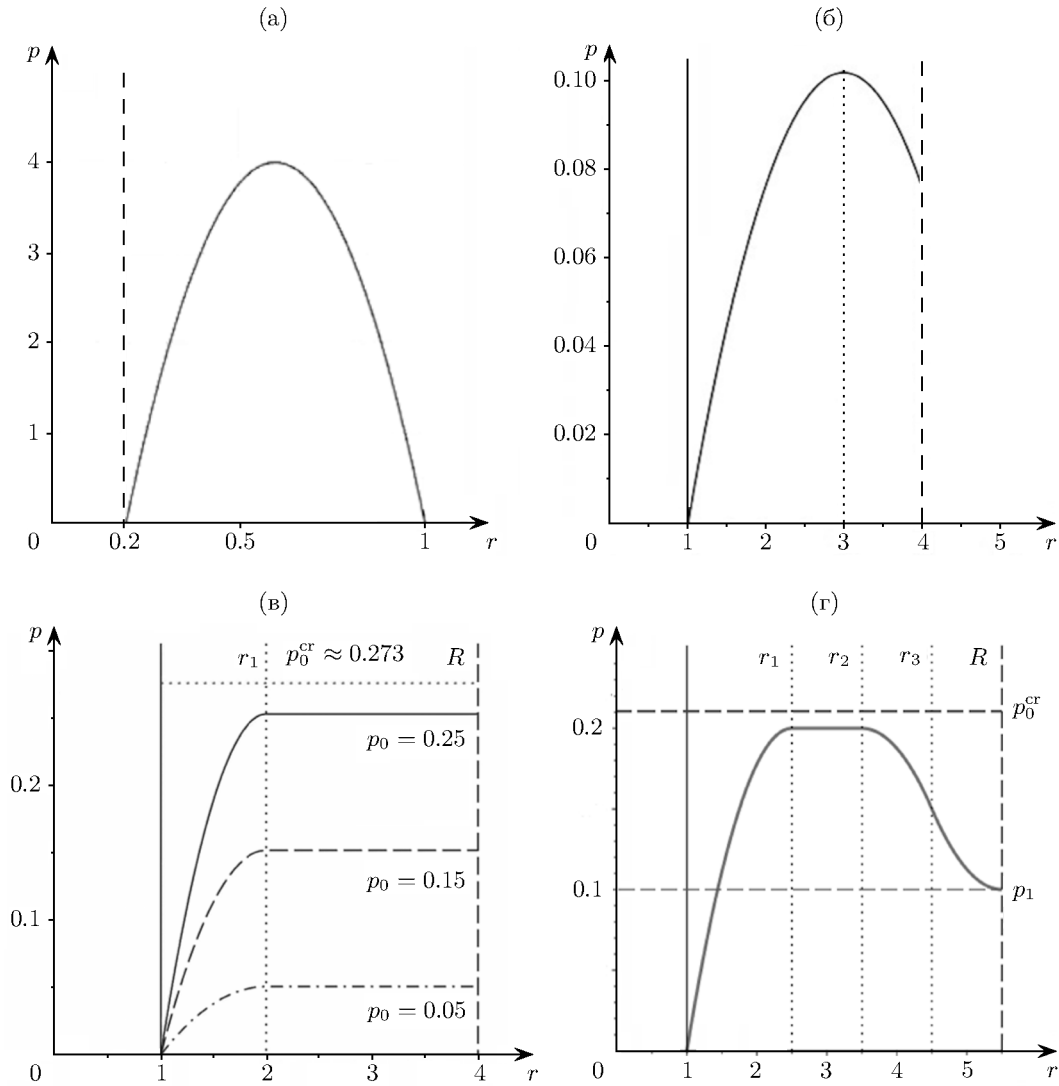


Рис. 2. Распределения давления в кольцевых конфигурациях плазмы вокруг проводника с током [20–23].

Второй способ – строгое исследование МГД-устойчивости равновесных магнитоплазменных конфигураций относительно произвольных трёхмерных малых возмущений в линейном приближении. Его общая логика в приложении к одномерным конфигурациям в цилиндре и, в первую очередь, к $\mathbb{Z}$ -пинчу, известная из ранних работ [15; 16; 17, с. 71], изложена в монографиях [10, с. 90; 11, с. 165]. Результаты расчётов, относящиеся к рассматриваемым здесь конфигурациям, окружающим прямой проводник, также представлены в работах последнего времени [20–23].

Воспроизведём вкратце схему исследований [21] и полученные результаты. Состояние покоя (1), (14) возмущается малыми величинами p_1 , $\mathbf{H}_1$ и $\mathbf{v}_1$, зависящими от времени и всех трёх пространственных координат. Уравнения магнитной газодинамики линеаризуются относительно этих величин. Из возникшей при этом системы семи линейных уравнений первого порядка удаётся исключить p_1 и $\mathbf{H}_1$ и привести её к следующему трёхмерному уравнению второго порядка для вектора скорости $\mathbf{v}_1$:

$$\rho \frac{\partial^2 \mathbf{v}_1}{\partial t^2} = \nabla(\gamma p \operatorname{div} \mathbf{v}_1 + \nabla p \cdot \mathbf{v}_1) + \operatorname{rot} \operatorname{rot} (\mathbf{v}_1 \times \mathbf{H}) \times \mathbf{H} + \mathbf{j} \times \operatorname{rot} (\mathbf{v}_1 \times \mathbf{H}) \equiv -\mathbb{K}[\mathbf{v}_1]. \quad (17)$$

Решения линейного однородного уравнения с не зависящими от времени t коэффициентами формируются из экспонент

$$\mathbf{v}_1(t, r, \varphi, z) = e^{i\omega t} \mathbf{v}(r, \varphi, z), \quad (18)$$

а параметр ω в показателе степени определяется спектральной задачей на собственные значения

$$\rho\omega^2\mathbf{v} = \mathbb{K}[\mathbf{v}]. \quad (19)$$

Дифференциальный оператор $\mathbb{K}[\mathbf{v}]$ с однородными краевыми условиями, например, $v_n = 0$ на границе области, является самосопряжённым, следовательно, его собственные значения ω^2 действительны. Отсюда следует, что решения (18) не возрастают со временем, если $\omega^2 > 0$, и экспоненциально возрастают, если $\omega^2 < 0$, т.е. условием МГД-устойчивости конфигураций является положительная определённость оператора $\mathbb{K}$. В одномерных задачах его коэффициенты, составленные из функций $p(r)$, $H(r)$ и $j(r)$ в состоянии покоя, не зависят также от φ и z , поэтому возмущения (18) зависят от этих переменных также экспоненциально:

$$\mathbf{v}_1(t, r, \varphi, z) = e^{i\omega t + im\varphi - ikz} \mathbf{v}(r), \quad (20)$$

где m – целое, k – действительное, так как решения (20) должны быть периодическими по φ и ограниченными при $z \rightarrow \pm\infty$.

Спектральная задача (19) распадается на бесконечную серию одномерных задач для отдельных гармоник возмущений с параметрами m и k :

$$\rho\omega_{m,k}^2 \mathbf{v}_{m,k} = \mathbb{K}_{m,k}[\mathbf{v}_{m,k}]. \quad (21)$$

Из векторных уравнений (21) удаётся исключить компоненты v_φ и v_z и свести поиск решений к задачам для одного скалярного уравнения относительно функции $u = v_r$:

$$-\frac{d}{dr} \left(F_{m,k} \frac{du}{dr} \right) + G_{m,k} u = 0; \quad u|_\Gamma = 0, \quad (22)$$

в котором коэффициенты $F_{m,k}$ и $G_{m,k}$ нелинейно зависят от собственного значения $\omega_{m,k}^2$ (уравнение Соловьёва) [16]. Численно решая задачу (22), например, методом “стрельбы”, можно найти старшие (минимальные) собственные значения $\omega_{m,k}^2$, отвечающие за устойчивость конфигураций относительно соответствующих гармоник возмущений.

Дальнейшие упрощения связаны с тем, что для качественного исследования устойчивости нет необходимости в вычислении собственных значений операторов $\mathbb{K}_{m,k}$ – достаточно установить их положительность. Для этого достаточно лишь найти “границу устойчивости” в области параметров равновесной конфигурации – условие на параметры, при котором старшее собственное значение обращается в нуль, т.е. условие, при котором краевая задача для уравнений (21), (22) при $\omega^2 = 0$ имеет нетривиальное решение.

В частном случае $m = 0$ (т.е. когда возмущения не зависят от φ) уравнение (22) при $\omega^2 = 0$ вырождается, и устойчивость определяется с помощью энергетического принципа [16; 28; 17, с. 75]: квадратичная форма $\mathbb{K}[\mathbf{v}]\mathbf{v}$ положительна при всех значениях k , если выполнено условие

$$-\frac{r}{p} \frac{dp}{dr} < \frac{2\gamma H^2}{p + H^2}, \quad (23)$$

где $\gamma = 5/3$ – показатель адиабаты.

Этому условию удовлетворяют все конфигурации на рис. 2 при $1 < r \leq r_1$, т.е. на возрастающей ветви $p(r)$, и заведомо не удовлетворяют конфигурации на рис. 2, а, где левая часть неравенства (23) неограниченно возрастает при $r \rightarrow R$, а правая остаётся ограниченной. В связи с этим рассмотрена конфигурация на рис. 2, б, где внешняя граница $r = R_1$ не допускает обращения $p(r)$ в нуль. В расчётах получена таблица значений R_1 , допускающих неравенство (23), в зависимости от параметров r_1 и p_0 [21]. Из неё следует, что попытки отодвинуть границу $r = R$ от $r = r_1$ заметно усиливают ограничение на максимум давления p_0 .

В остальных случаях $m \geq 1$ уравнение (22) при $\omega^2 = 0$ имеет вид

$$-\frac{d}{dr} \left(\frac{H^2}{\eta r} \frac{du}{dr} \right) + \left[\frac{m^2 H^2}{r^3} - \frac{4\alpha^2 H^2}{\eta r} + \frac{d}{dr} \left(\frac{(2-\eta)H^2}{\eta r^2} \right) \right] u = 0, \quad (24)$$

где $\alpha = k/m$, $\eta = 1 + \alpha^2 r^2$. В численном решении краевой задачи для этого уравнения показано, что конфигурации на рис. 2, б устойчивы при $R < R_1$ для гармоник с низкими значениями волнового числа $|k| \lesssim 2$. При возрастании $|k|$ устойчивость требует более строгих ограничений на p_0 при каждом значении $R = R_1$ по сравнению с упомянутой выше таблицей [21].

Приведённые результаты показывают, что причины обнаруженной неустойчивости сосредоточены в основном у внешней границы кольцевой окрестности проводника, где плотность плазмы уменьшается с ростом радиуса. Они же подтверждены анализом конфигурации на рис. 2, в, где граница проведена непосредственно через область максимального давления. Здесь все варианты устойчивы в пределах естественного ограничения давления (16) [22]. Отдельное внимание уделено периферии конфигурации на рис. 2, г. Здесь давление ограничено снизу $p \geq p_1 > 0$, и переход от $p(r_2) = p_0$ к $p(R) = p_1$ осуществляется более плавно и разделён значением $r = r_3$, где $d^2 p/dr^2 = 0$ и абсолютная величина dp/dr максимальна. Условие устойчивости (23) при $m = 0$ выполнено везде, если оно выполнено при $r = r_3$. При $m \geq 1$ в численном решении краевой задачи для уравнения (24) установлено, что неустойчивость провоцируется скоростью убывания $p(r)$ при $r > r_2$, т.е. сокращением области $r_2 < r < R$ и уменьшением граничного значения p_1 [23]. Это соответствует известной тенденции неустойчивости Z-пинча [15].

Изложенные результаты представляют интерес в первую очередь тем, что показывают влияние распределения давления $p(r)$ вблизи проводника на устойчивость. Здесь рассмотренная одномерная модель наиболее оправдана. Результаты, относящиеся к периферии, следует считать ориентировочными, поскольку территории любых ловушек с двумя и более проводниками, примыкающие к внешней границе, по существу неоднородны и нуждаются в исследованиях в, как минимум, двумерных моделях.

Таким образом, в рассмотренных математических моделях равновесных магнитоплазменных конфигураций в окрестности прямого проводника с током исследована зависимость их диффузионной и МГД-устойчивости от распределения давления плазмы вдоль радиуса ловушки.

3. Двумерные задачи о МГД-устойчивости. Исследование устойчивости равновесных конфигураций в ловушках, допускающих плоскую симметрию, в частности, в цилиндрическом аналоге “галатей-пояса”, ведётся в соответствии с упомянутой выше общей логикой линейной теории. В основе такого исследования лежит анализ решений начально-краевой задачи для уравнения

$$\rho \frac{\partial^2 \mathbf{v}_1}{\partial t^2} = \nabla(\gamma p \operatorname{div} \mathbf{v}_1 + \nabla p \cdot \mathbf{v}_1) + \operatorname{rot} \operatorname{rot} (\mathbf{v}_1 \times \mathbf{H}) \times \mathbf{H} + (\mathbf{j} - \mathbf{j}^{\text{ex}}) \times \operatorname{rot} (\mathbf{v}_1 \times \mathbf{H}) \equiv -\tilde{\mathbb{K}}[\mathbf{v}_1] \quad (25)$$

с граничным условием $v_{1n} = 0$ на внешней границе или, что то же самое, с условием положительной определённости линейного оператора $\tilde{\mathbb{K}}[\mathbf{v}_1]$. Уравнение (25) отличается от (17) заменой тока $\mathbf{j}$ на плазменный ток $\mathbf{j}^{\text{pl}} = \mathbf{j} - \mathbf{j}^{\text{ex}}$, что позволяет рассматривать задачи в односвязных областях, не исключая из них территории проводников. Коэффициенты уравнения составлены из функций $p(x, y)$, $H(x, y)$ и $j(x, y)$, т.е. не зависят от переменных t и z , поэтому его решения зависят от них экспоненциально $\mathbf{v}_1(t, x, y, z) = e^{i\omega t - ikz} \mathbf{v}(x, y)$ с действительными значениями параметра k , чтобы избежать неограниченного роста при $z \rightarrow \pm\infty$. Задача с уравнением (25) становится спектральной задачей типа (19), которая превращается в однопараметрическую серию двумерных векторных задач $\rho \omega_k^2 \mathbf{v}_k = \tilde{\mathbb{K}}_k[\mathbf{v}_k]$ для фурье-гармоник возмущений скорости. Здесь оператор $\tilde{\mathbb{K}}_k$ означает соответствующую гармонику оператора $\tilde{\mathbb{K}}$.

Собственные значения ω_k^2 действительны, а вектор-функции $\mathbf{v}_k$, вообще говоря, комплексны. Действительные части этих вектор-функций удобно представить в терминах функций $u(x, y) = \operatorname{Re} v_{xk}$, $v(x, y) = \operatorname{Re} v_{yk}$, $w(x, y) = \operatorname{Im} v_{zk}$, опустив в дальнейшем индекс k :

$$\rho \omega^2 u = L_1[u, v] + k M_1[w], \quad \rho \omega^2 v = L_2[u, v] + k M_2[w], \quad \rho \omega^2 w = M_3[w] + k L_3[u, v] + k^2 M_4[w], \quad (26)$$

где

$$L_1[u, v] = -\frac{\partial}{\partial x}(\gamma p \operatorname{div} \mathbf{v}) + H_y \Delta \xi - \frac{\partial j^{\text{pl}}}{\partial x} \xi, \quad L_2[u, v] = -\frac{\partial}{\partial y}(\gamma p \operatorname{div} \mathbf{v}) - H_x \Delta \xi - \frac{\partial j^{\text{pl}}}{\partial y} \xi,$$

$$L_3[u, v] = \gamma p \operatorname{div} \mathbf{v} + H_x \frac{\partial \xi}{\partial y} - H_y \frac{\partial \xi}{\partial x} + j^{\text{pl}} \xi,$$

$$M_1[w] = -\frac{\partial}{\partial x}(\gamma p w) - H_y \left(\frac{\partial H_y w}{\partial x} - \frac{\partial H_x w}{\partial y} \right) - j^{\text{pl}} H_y w,$$

$$M_2[w] = -\frac{\partial}{\partial y}(\gamma p w) + H_x \left(\frac{\partial H_y w}{\partial x} - \frac{\partial H_x w}{\partial y} \right) + j^{\text{pl}} H_x w,$$

$$M_3[w] = -(\mathbf{H} \cdot \nabla) \operatorname{div} \mathbf{H} w, \quad M_4[w] = (\gamma p + H_x^2 + H_y^2) w, \quad \xi = H_x v - H_y u.$$

Заметим, что упростить уравнения (26) дальше, сведя их, например, к системе двух уравнений для функций u и v , аналогичной уравнению Соловьёва (22), здесь не удаётся, поскольку третье из них не позволяет выразить функцию w аналитически через функции u , v и их производные. Однако из неё очевидно выделяется система уравнений, если положить в ней $k = 0$, которая распадается на систему из двух уравнений

$$\rho \omega^2 u = L_1[u, v], \quad \rho \omega^2 v = L_2[u, v] \quad (27)$$

для u и v и одно уравнение

$$\rho \omega^2 w = M_3[w] \quad (28)$$

для w .

С ними естественно связать ещё один промежуточный вариант устойчивости – устойчивость относительно двумерных возмущений, не зависящих от координаты z . Этот вариант устойчивости шире рассмотренной в п. 1 диффузионной устойчивости, которая ограничена возмущениями только магнитного поля, а здесь учитывается также возможность динамики выведенной из равновесия плазмы.

Собственные значения задачи (28) всегда положительны, в чём легко убедиться с помощью интеграла по квадрату $|x| < 2$, $|y| < 2$:

$$\begin{aligned} \omega^2 \iint \rho w^2 dx dy &= \iint M_3[w] w dx dy = - \iint \mathbf{H} w \cdot \nabla \operatorname{div} (\mathbf{H} w) dx dy = \\ &= \iint (\operatorname{div} (\mathbf{H} w))^2 dx dy - \oint_{\Gamma} H_n w \operatorname{div} (\mathbf{H} w) ds, \end{aligned} \quad (29)$$

в котором граница Γ непрозрачна для магнитного поля ($H_n = 0$). Отсюда следует, что осевая компонента возмущений $w(x, y)$, не зависящих от z , не может оказать влияние на неустойчивость двумерных конфигураций. Нетривиальным объектом на рассматриваемом двумерном этапе исследования остаётся краевая задача (27) с граничным условием $v_n = 0$:

$$u(\pm 2, y) = 0, \quad v(x, \pm 2) = 0.$$

Здесь уместно обратить также внимание на квадратичную форму

$$\rho \omega^2 (u^2 + v^2) = L_1[u, v] u + L_2[u, v] v = -\mathbf{v} \cdot \nabla (\gamma p \operatorname{div} \mathbf{v}) - \xi \Delta \xi - \xi \mathbf{v} \cdot \nabla j^{\text{pl}} \quad (30)$$

и её интеграл по упомянутому выше квадрату

$$\iint (L_1 u + L_2 v) dx dy = \iint (\gamma p (\operatorname{div} \mathbf{v})^2 + (\nabla \xi)^2 + j^{\text{pl}} \operatorname{div} (\xi \mathbf{v})) dx dy. \quad (31)$$

Если эта форма неотрицательна на множестве всех собственных функций (u, v) задачи (27), то соответствующая равновесная конфигурация устойчива относительно двумерных возмущений. Более сильное условие состоит в следующем: её неотрицательность при любых функциях (u, v) достаточна для устойчивости. Поскольку

$$j^{\text{pl}} \operatorname{div}(\xi \mathbf{v}) = \operatorname{div}(j^{\text{pl}} \xi \mathbf{v}) - \xi(\mathbf{v} \cdot \nabla) j^{\text{pl}} \quad \text{и} \quad \iint \operatorname{div}(j^{\text{pl}} \xi \mathbf{v}) dx dy = \oint_{\Gamma} j^{\text{pl}} \xi v_n ds = 0,$$

это условие принимает вид

$$\iint \xi \mathbf{v} \cdot \nabla j^{\text{pl}} \leq \iint (\gamma p (\operatorname{div} \mathbf{v})^2 + (\nabla \xi)^2) dx dy. \quad (32)$$

Устойчивость относительно двумерных возмущений скорости (u, v) можно исследовать, как и в п. 2, с помощью отыскания “границы устойчивости” в области параметров, т.е. условия, при котором $\omega^2 = 0$ является собственным значением задачи (27), или, иначе говоря, условия, при котором система линейных однородных уравнений $L_1[u, v] = 0$, $L_2[u, v] = 0$ имеет нетривиальное решение. Оно может быть найдено в процессе численного решения этих уравнений итерационным методом установления:

$$\frac{\partial u}{\partial t} = -L_1[u, v] = 0, \quad \frac{\partial v}{\partial t} = -L_2[u, v] = 0. \quad (33)$$

Здесь t – итерационный параметр, играющий роль “времени”, а знак минус в правых частях обеспечивает корректность задачи Коши с уравнениями (33) параболического типа. “Начальными условиями” (нулевой итерацией) могут быть любые функции $u(x, y)$ и $v(x, y)$.

Равновесные конфигурации плазмы описываются решениями краевой задачи для уравнения Грэда–Шафранова с фиксированными геометрией и параметрами $r_c = 0.2$, $q = 0.2$ в формулах (2), (5). Искомой границей устойчивости является значение максимального давления p_0 в интервале $0 < p_0 < p_0^{\text{diff}}$, допускающем диффузионную устойчивость относительно двумерных возмущений магнитного поля.

В результате численного решения задачи (33) при небольших значениях параметра $p_0 < p_0^* \sim 3$ устанавливается нулевая скорость в центральной части квадрата, где давление плазмы заметно отличается от нуля (см. рис. 1). Время установления растёт вместе со значениями p_0 и естественно зависит от выбора начальных условий. Вне конфигурации, где давление практически нулевое, скорость может долго сохранять начальное значение, что можно отнести к выбранным начальным условиям. При $p_0 \rightarrow p_0^*$ в области конфигурации появляются ненулевые значения скорости, близкой по направлению к магнитному полю $\mathbf{H}$. При этом обращаются в нуль величины ξ , $\operatorname{div} \mathbf{v}$ и квадратичная форма (30), что соответствует собственному значению $\omega^2 = 0$, т.е. границей устойчивости является значение $p_0 = p_0^* < p_0^{\text{diff}} \sim 5$. При значениях $p_0 > p_0^*$ решения краевой задачи с уравнениями (33) неограниченно возрастают, что соответствует неустойчивости конфигураций относительно двумерных возмущений равновесия, включающих ненулевые значения скорости. Полученный результат конкретизирует тенденцию, установленную выше при исследовании одномерных моделей конфигураций, окружающих прямой проводник: устойчивость конфигураций в цилиндрическом аналоге “пояса” относительно двумерных возмущений, включая динамические, требует приблизительно более сильного ограничения на максимальное давление $p_0 < p_0^*$ по сравнению с диффузионной устойчивостью $p_0 < p_0^{\text{diff}}$ относительно возмущений магнитного поля. Таким образом, справедливо следующее

Утверждение. Для МГД-устойчивости двумерных конфигураций плазмы в ловушке “га-латея-пояс” произвольные малые двумерные возмущения, включая динамические, требуют более сильного ограничения на давление плазмы, измеренного в магнитных единицах, по сравнению с диффузионной устойчивостью относительно возмущений магнитного поля. Параметры ограничений находятся с помощью численного решения соответствующих задач.

Наконец, развитием и завершением изложенных исследований устойчивости в ловушках-галатеях должно стать их распространение на трёхмерные малые возмущения равновесных двумерных конфигураций. В основе лежит краевая задача с уравнениями (26) для фурье-компонент произвольных возмущений с произвольными значениями действительного параметра k . “Граница устойчивости” в области $0 < p_0 < p_0^{\text{diff}}$ соответствует существованию нетривиального решения уравнений (26) при $\omega^2 = 0$, которое следует искать численными методами. С другой стороны, устойчивость конфигураций равносильна положительной определённости квадратичной формы

$$F = F_0[u, v, w] + kF_1[u, v, w] + k^2F_2[w], \quad (34)$$

где $F_0 = L_1u + L_2v + M_3w$, $F_1 = M_1u + M_2w + L_3w$, $F_2 = M_4w$, или её интеграла по сечению цилиндра. При этом F_0 можно упростить до рассмотренной выше формы (30), (31), поскольку её последнее слагаемое положительно (см. (29)). Квадратный полином (34) с положительными коэффициентами F_0 (что требуется для двумерной устойчивости) и F_2 , очевидно, положителен при любых значениях $k \in \mathbb{R}$, если он не имеет действительных нулей, т.е. его дискриминант отрицателен. Отсюда следует условие, ограничивающее абсолютную величину формы F_1 или её интеграла,

$$F_1^2 \leq 4F_0F_2. \quad (35)$$

В трёхмерных задачах оно играет ту же роль, что и неравенство (32) в двумерных.

Таким образом, МГД-устойчивость обсуждаемых конфигураций относительно произвольных малых трёхмерных возмущений имеет место при выполнении условия, которое можно сформулировать любым из двух эквивалентных между собой способов: 1) отсутствие нетривиальных решений краевой задачи с уравнениями (26) при $\omega^2 = 0$; 2) выполнение неравенства (35) для достаточно представительного набора функций (u, v, w) . Его проверка потребует большого объёма вычислений.

Заключение. В статье представлена математическая модель равновесных магнитоплазменных конфигураций в цилиндре с погружёнными в плазму двумя прямыми проводниками с током – распрямлённом аналоге тороидальной магнитной ловушки “галатея-пояс”. Эта модель основана на численном решении двумерных краевых задач для полулинейного дифференциального уравнения Грэда–Шафранова эллиптического типа. Приведены примеры расчёта геометрии конфигураций и их параметров, а также их зависимости от условий задачи. Проведено достаточно полное исследование устойчивости конфигураций: одномерных – в окрестности одного отдельного проводника и двумерных – сначала относительно двумерных возмущений магнитного поля (диффузионная устойчивость), а затем МГД-устойчивости относительно двумерных возмущений всех переменных, включая скорость. На всех этапах получены ограничения на безразмерную величину давления плазмы, требуемые для устойчивости, и иерархия этих ограничений. Указаны пути исследования устойчивости двумерных равновесных конфигураций относительно произвольных трёхмерных возмущений.

Работа выполнена при поддержке Московского центра фундаментальной и прикладной математики (соглашение с Министерством науки и высшего образования Российской Федерации № 075-15-2019-1623).

СПИСОК ЛИТЕРАТУРЫ

1. Морозов А.И. О галатеях – плазменных ловушках с омываемыми плазмой проводниками // Физика плазмы. 1992. Т. 18. Вып. 3. С. 305–316.
2. Морозов А.И., Пустовитов В.Д. О стеллараторе с левитирующими обмотками // Физика плазмы. 1991. Т. 17. Вып. 10. С. 1276.
3. Морозов А.И., Франк А.Г. Тороидальная магнитная ловушка-галатея с азимутальным током // Физика плазмы. 1994. Т. 20. № 11. С. 982–989.
4. Морозов А.И., Бугрова А.И., Бишаев А.М., Липатов А.С., Козинцева М.В. Параметры плазмы в модернизированной ловушке-галатее “Тримикс-М” // Журн. техн. физики. 2007. Т. 77. № 12. С. 15–20.

5. Брушлинский К.В., Зуева Н.М., Михайлова М.С., Морозов А.И., Пустовитов В.Д., Тузова Н.Б. Численное моделирование прямых винтовых шнуров с проводниками, погруженными в плазму // Физика плазмы. 1994. Т. 20. № 3. С. 284–292.
6. Брушлинский К.В., Игнатов П.А. Плазмостатическая модель магнитной ловушки “галатея-пояс” // Журн. вычислит. математики и мат. физики. 2010. Т. 50. № 12. С. 2184–2194.
7. Брушлинский К.В., Кондратьев И.А. Сравнительный анализ расчётов равновесия плазмы в тороидальных и цилиндрических магнитных ловушках // Мат. моделирование. 2018. Т. 30. № 6. С. 76–94.
8. Tao B., Jin X., Li Z., Tong W. Equilibrium configuration reconstruction of multipole galatea magnetic trap based on magnetic measurement // IEEE Trans. on Plasma Sci. 2019. V. 47. № 7. P. 3114–3123.
9. Морозов А.И., Савельев В.В. О галатеех – ловушках с погруженными в плазму проводниками // Успехи физ. наук. 1998. Т. 168. № 11. С. 1153–1194.
10. Брушлинский К.В. Математические и вычислительные задачи магнитной газодинамики. М., 2009.
11. Брушлинский К.В. Математические основы вычислительной механики жидкости, газа и плазмы. М., 2017.
12. Шафранов В.Д. О равновесных магнитогидродинамических конфигурациях // Журн. эксп. и теор. физики. 1957. Т. 33. Вып. 3 (9). С. 710–722.
13. Grad H., Rubin H. Hydrodynamic equilibria and force-free fields // Proc. 2nd United Nations Int. Conf. on the Peaceful Uses of Atomic Energy. Geneva, 1958. V. 31. P. 190–197.
14. Шафранов В.Д. Равновесие плазмы в магнитном поле // Вопросы теории плазмы / Под ред. М.А. Леонтовича. М., 1963. Вып. 2. С. 92–131.
15. Кадомцев Б.Б. Гидромагнитная устойчивость плазмы // Вопросы теории плазмы / Под ред. М.А. Леонтовича. М., 1963. Вып. 2. С. 132–176.
16. Соловьев Л.С. Гидромагнитная устойчивость замкнутых плазменных конфигураций // Вопросы теории плазмы / Под ред. М.А. Леонтовича. М., 1972. Вып. 6. С. 210–290.
17. Бейтман Г. МГД-неустойчивости. М., 1982.
18. Медведев С.Ю., Мартынов А.А., Дроздов В.В., Иванов А.А., Пошехонов Ю.Ю., Коновалов С.В., Виллард Л. МГД-устойчивость и энергетический принцип без предположения о вложенности магнитных поверхностей двумерных равновесий // Физика плазмы. 2019. Т. 45. № 2. С. 120–132.
19. Брушлинский К.В. Два подхода к задаче об устойчивости равновесия плазмы в цилиндре // Прикл. математика и механика. 2001. Т. 65. Вып. 2. С. 235–243.
20. Брушлинский К.В., Кривцов С.А., Степин Е.В. Об устойчивости равновесия плазмы в окрестности прямого проводника с током // Журн. вычислит. математики и мат. физики. 2020. Т. 60. № 4. С. 153–163.
21. Брушлинский К.В., Степин Е.В. Математические модели равновесных конфигураций плазмы, окружающей проводники с током // Дифференц. уравнения. 2020. Т. 56. № 7. С. 901–909.
22. Brushlinskii K.V., Stepin E.V. Mathematical model and stability investigation of plasma equilibrium around a current-carrying conductor // J. Phys.: Conf. Ser. 2020. V. 1686. P. 012030.
23. Brushlinskii K.V., Stepin E.V. Plasma equilibrium and stability in a current-carrying conductor vicinity // J. Phys.: Conf. Ser. 2020. V. 1640. P. 012018.
24. Peaceman D.W., Rachford H.H. The numerical solution of parabolic and elliptic differential equations // J. Soc. Industr. Appl. Math. 1955. V. 3. № 1. P. 28–42.
25. Douglas J. On the numerical integration of $\partial^2 u / \partial x^2 + \partial^2 u / \partial y^2 = \partial u / \partial t$ by implicit method // J. Soc. Industr. Appl. Math. 1955. V. 3. № 1. P. 42–65.
26. Арсенин В.Я. Методы математической физики и специальные функции. М., 1984.
27. Веденов А.А., Велихов Е.П., Сагдеев Р.З. Устойчивость плазмы // Успехи физ. наук. 1961. Т. 73. Вып. 4. С. 701–766.
28. Bernstein I.B., Frieman E.A., Kruskal M.D., Kulsrud R.M. Energy principle for the hydromagnetic stability problem // Proc. Roy. Soc. 1958. V. 244. P. 17–50. (Перевод: Бернштейн А.Б. и др. Энергетический принцип для проблемы гидродинамической устойчивости // Управляемые термоядерные реакции (сб. перев. материалов). М., 1960. Вып. 26. С. 226–260).

Институт прикладной математики
им. М.В. Келдыша РАН, г. Москва,
Национальный исследовательский
ядерный университет “МИФИ”, г. Москва,
Московский государственный университет
им. М.В. Ломоносова

Поступила в редакцию 01.03.2021 г.
После доработки 01.03.2021 г.
Принята к публикации 27.04.2021 г.

ЧИСЛЕННЫЕ МЕТОДЫ

УДК 519.63

СХЕМЫ РАСЩЕПЛЕНИЯ РЕШЕНИЯ
ДЛЯ ЭВОЛЮЦИОННЫХ УРАВНЕНИЙ
ВТОРОГО ПОРЯДКА

© 2021 г. П. Н. Вабищевич

Для нестационарных задач приближённое решение на новом слое по времени часто получают из ряда более простых задач. В стандартных схемах расщепления используют аддитивное расщепление оператора задачи на более удобные для вычислительной реализации операторы и явно-неявные аппроксимации по времени. В работе рассматривается новый класс схем расщепления, который связывается с аддитивным представлением не оператора задачи, а самого решения. Предложена новая общая процедура расщепления решения на основе аддитивного представления единичного оператора через операторы ограничения и продолжения для вспомогательных пространств. Построены и исследуются безусловно устойчивые схемы расщепления при приближённом решении задачи Коши для эволюционного уравнения второго порядка, которое рассматривается в конечномерном гильбертовом пространстве.

DOI: 10.31857/S0374064121070025

Введение. Численное решение нестационарных краевых задач для уравнений с частными производными базируется на использовании, чаще всего, неявных аппроксимаций по времени, которые обеспечивают для соответствующих разностных схем их безусловную устойчивость по начальным данным и правой части [1, гл. 6; 2, гл. II]. Явные схемы более просты для нахождения приближённого решения на новом слое, но имеют жёсткие ограничения на шаги по времени. Целью многих исследований является построение схем, которые бы наследовали, в той или иной мере, достоинства явных и неявных аппроксимаций: сохраняли бы устойчивость, но были бы более удобными для вычислительной реализации.

Упрощение задачи на новом слое без потери устойчивости достигается, в частности, использованием неоднородных аппроксимаций по времени. Широкое распространение получили явно-неявные схемы (IMEX-методы) (см., например, [3, гл. IV]). В этом случае оператор задачи расщепляется на два операторных слагаемых с выделением приемлемого для вычислительной реализации слагаемого, которое берётся с верхнего слоя, а другое слагаемое – с нижнего слоя по времени. В недавней работе [4] нами проведено исследование явно-неявных двух- и трёх-слойных операторно-разностных схем для эволюционного уравнения первого порядка как для стандартного расщепления основного оператора задачи, так и для расщепления оператора при производной решения по времени.

В схемах расщепления [5, гл. II; 6, гл. 3] используется аддитивное представление оператора задачи. Переход на новый слой по времени осуществляется при помощи решения эволюционных задач для отдельных операторных слагаемых. Во многих нестационарных задачах вычислительно приемлемые подзадачи имеет смысл строить на основе декомпозиции решения, когда более простые задачи формулируются для отдельных составляющих решения. Такие схемы расщепления решения построены в работе [7] для приближённого решения задачи Коши в конечномерном гильбертовом пространстве для эволюционного уравнения первого порядка.

В настоящей работе схемы расщепления решения строятся для эволюционных уравнений второго порядка. Декомпозиция решения проводится как на прямой сумме подпространств, так и на новом варианте, который связывается с аддитивным представлением единичного оператора на основе операторов ограничения и продолжения. Схемы расщепления основываются на использовании явно-неявных аппроксимаций по времени при выделении диагональной части или треугольного расщепления соответствующей операторной матрицы. Устойчивость предложенных трёхслойных схем расщепления решения исследуется с привлечением общих

результатов теории устойчивости (корректности) операторно-разностных схем [2, гл. IV; 8, гл. 4].

1. Постановка задачи. Рассматривается задача Коши для эволюционного уравнения второго порядка в конечномерном гильбертовом пространстве V . Для того чтобы не усложнять текст несущественными техническими деталями, ограничимся однородным уравнением

$$\frac{d^2 u}{dt^2} + Au = 0, \quad 0 < t \leq T. \quad (1)$$

Дополним уравнение (1) начальными условиями

$$u(0) = u^0, \quad \frac{du}{dt}(0) = \tilde{u}^0. \quad (2)$$

Будем считать, для упрощения, что линейный оператор $A : V \rightarrow V$ является постоянным (не зависит от t), самосопряжённым и положительным:

$$\frac{d}{dt}A = A\frac{d}{dt}, \quad A = A^* > 0. \quad (3)$$

К задаче (1)–(3) мы приходим, например, при численном решении начально-краевых задач для гиперболических уравнений второго порядка после дискретизации по пространственным переменным.

Скалярное произведение векторов $u, v \in V$ обозначим через (u, v) , а норму вектора u – через $\|u\| = (u, u)^{1/2}$. Для самосопряжённого и положительного оператора D определяется гильбертово пространство V_D со скалярным произведением $(u, v)_D = (Du, v)$ и нормой $\|u\|_D = (u, v)_D^{1/2}$.

При приближённом решении задачи Коши (1), (2) чаще всего ориентируются на неявные аппроксимации по времени, которые дают безусловно устойчивые схемы. Будем использовать равномерную сетку по времени с шагом τ , и пусть

$$y^n = y(t^n), \quad t^n = n\tau, \quad n = \overline{0, N}, \quad N\tau = T.$$

Наше рассмотрение естественно начать с трёхслойной схемы с весом $\sigma = \text{const}$:

$$\frac{y^{n+1} - 2y^n + y^{n-1}}{\tau^2} + A(\sigma y^{n+1} + (1 - 2\sigma)y^n + \sigma y^{n-1}) = 0, \quad n = \overline{1, N-1}, \quad (4)$$

при заданных начальных условиях

$$y^0 = u^0, \quad y^1 = \bar{u}^1. \quad (5)$$

Второе начальное условие на решениях уравнениях (1) определим со вторым порядком, например, из уравнения

$$\left(I + \frac{\tau^2}{2}A\right)\bar{u}^1 = u^0 + \tau\tilde{u}^0,$$

где I – единичный оператор в V . Разностная схема (4), (5) аппроксимирует задачу (1), (2) со вторым порядком по τ при достаточной гладкости решения $u(t)$.

При формулировке условий устойчивости трёхслойных схем мы можем ориентироваться на общие результаты устойчивости операторно-разностных схем [2, гл. IV; 8, гл. 4]. Наша цель состоит в получении аналогов априорных оценок, которые имеют место для дифференциальной задачи. Простейшую оценку для решения задачи (1)–(3) можно получить, скалярно домножая уравнение (1) на du/dt , что даёт

$$\frac{d}{dt} \left(\left\| \frac{du}{dt} \right\|^2 + \|u\|_A^2 \right) = 0.$$

В результате приходим к равенству

$$\left\| \frac{du}{dt}(t) \right\|^2 + \|u(t)\|_A^2 = \|\tilde{u}^0\|^2 + \|u^0\|_A^2, \quad (6)$$

которое обеспечивает устойчивость решения по начальным данным и консервативность.

Исследование устойчивости операторно-разностных схем базируется на следующем вспомогательном утверждении.

Лемма. Пусть в трёхслойной схеме

$$C \frac{y^{n+1} - 2y^n + y^{n-1}}{\tau^2} + Ay^n = 0, \quad n = \overline{1, N-1}, \quad (7)$$

C и A – постоянные самосопряжённые операторы. Тогда для решения имеет место равенство

$$\mathcal{E}^{n+1} = \mathcal{E}^n, \quad n = \overline{1, N-1}, \quad (8)$$

в котором

$$\mathcal{E}^{n+1} = \left(A \frac{y^{n+1} + y^n}{2}, \frac{y^{n+1} + y^n}{2} \right) + \left(G \frac{y^{n+1} - y^n}{\tau}, \frac{y^{n+1} - y^n}{\tau} \right),$$

где $G = C - \tau^2 A/4$.

Доказательство. Принимая во внимание равенство

$$y^n = \frac{y^{n+1} + 2y^n + y^{n-1}}{4} - \frac{\tau^2}{4} \frac{y^{n+1} - 2y^n + y^{n-1}}{\tau^2},$$

запишем схему (7) в виде

$$G \frac{y^{n+1} - 2y^n + y^{n-1}}{\tau^2} + A \frac{y^{n+1} + 2y^n + y^{n-1}}{4} = 0. \quad (9)$$

Введём новые переменные $s^n = (y^n + y^{n-1})/2$, $r^n = (y^n - y^{n-1})/\tau$, в которых равенство (9) принимает вид

$$G \frac{r^{n+1} - r^n}{\tau} + A \frac{s^{n+1} + s^n}{2} = 0.$$

Домножая его скалярно на $2(s^{n+1} - s^n) = \tau(r^{n+1} + r^n)$, получаем

$$(Gr^{n+1}, r^{n+1}) + (As^{n+1}, s^{n+1}) = (Gr^n, r^n) + (As^n, s^n).$$

Возвращаясь к старым переменным, приходим к равенству (8). Лемма доказана.

Следствие 1. При выполнении неравенств

$$A > 0, \quad C - \frac{\tau^2}{4} A > 0 \quad (10)$$

$\mathcal{E}^{n+1}$ определяет норму решения, а (8) является оценкой устойчивости

$$\left\| \frac{y^{n+1} - y^n}{\tau} \right\|_G^2 + \left\| \frac{y^{n+1} + y^n}{2} \right\|_A^2 = \left\| \frac{\bar{u}^1 - u^0}{\tau} \right\|_G^2 + \left\| \frac{\bar{u}^1 + u^0}{2} \right\|_A^2, \quad n = \overline{1, N-1}, \quad (11)$$

для разностной схемы (5), (7).

Следствие 2. При менее жёстких, чем (10), ограничениях

$$A > 0, \quad C - \frac{\tau^2}{4} A \geq 0$$

оценка устойчивости имеет место в более слабой норме: вместо равенства (11) выполняется неравенство

$$\left\| \frac{y^{n+1} + y^n}{2} \right\|_A^2 \leq \left\| \frac{\bar{u}^1 - u^0}{\tau} \right\|_G^2 + \left\| \frac{\bar{u}^1 + u^0}{2} \right\|_A^2, \quad n = \overline{1, N-1},$$

для $0.5(y^{n+1} + y^n)$.

Теперь мы можем сформулировать условия устойчивости схемы (4), (5). Уравнение (4) записывается в виде (7) при $C = I + \sigma\tau^2 A$. Тем самым

$$G = I + \left(\sigma - \frac{1}{4} \right) \tau^2 A$$

и условия устойчивости (10) будут выполнены при стандартных [2, гл. IV; 8, гл. 4] ограничениях на вес: $\sigma \geq 0.25$.

Основной недостаток неявных схем связан с вычислительной сложностью задачи для определения решения на новом слое по времени. При использовании схемы (4), (5) приближённое решение y^{n+1} находится как решение задачи

$$(I + \sigma\tau^2 A)y^{n+1} = \varphi^n$$

при известной правой части

$$\varphi^n = (I + \sigma\tau^2 A)(2y^n - y^{n-1}) - \tau^2 Ay^n.$$

Актуальной является задача построения таких аппроксимаций по времени, которые приводили бы к более простым алгоритмам для нахождения приближённого решения на новом слое по времени.

В методах расщепления [5, гл. II; 6, гл. 3] упрощение задачи достигается за счёт того или иного представления оператора (операторов) задачи в виде суммы более простых операторов. Если оператор задачи A имеет аддитивное представление

$$A = \sum_{\alpha=1}^p A_\alpha,$$

то для приближённого решения задачи (1), (2) применяются явно-неявные аппроксимации. Переход на новый слой по времени обеспечивается решением ряда более простых задач, которые формируются отдельными операторами A_α , $\alpha = \overline{1, p}$:

$$(I + \sigma\tau^2 A_\alpha)y_\alpha^{n+1} = \varphi_\alpha^n, \quad \sigma_\alpha = \text{const}, \quad \alpha = \overline{1, p}.$$

Мы рассматриваем новый класс численных методов приближённого решения нестационарных задач. Эти методы основываются не на расщеплении операторов, а на расщеплении самого решения [7]. Выделим общие конструкции аддитивного представления приближённого решения и построим соответствующие явно-неявные схемы для задачи Коши (1), (2).

2. Расщепление решения. Будем работать с семейством некоторых конечномерных гильбертовых пространств V_α , $\alpha = \overline{1, p}$, которые имеют, например, существенно меньшую размерность, чем пространство V . Из решений в пространствах V_α строится аддитивное представление решения в исходном пространстве V .

Для каждого из этих пространств определены линейные операторы – оператор R_α ограничения с пространства V на пространство V_α и оператор R_α^* продолжения с V_α на V :

$$R_\alpha : V \rightarrow V_\alpha, \quad R_\alpha^* : V_\alpha \rightarrow V, \quad \alpha = \overline{1, p}.$$

Мы предполагаем, что для V имеет место разложение

$$V = \sum_{\alpha=1}^p R_{\alpha}^* V_{\alpha},$$

так что для каждого $u \in V$ справедливо равенство

$$u = \sum_{\alpha=1}^p R_{\alpha}^* u_{\alpha}, \quad u_{\alpha} \in V_{\alpha}, \quad \alpha = \overline{1, p}. \quad (12)$$

Тем самым имеет место аддитивное представление для $u \in V$:

$$u = \sum_{\alpha=1}^p w_{\alpha}, \quad w_{\alpha} = R_{\alpha}^* u_{\alpha}, \quad w_{\alpha} \in V, \quad \alpha = \overline{1, p}.$$

Чтобы получить систему уравнений для отдельных слагаемых u_{α} , $\alpha = \overline{1, p}$, подставим представление (12) в уравнение (1) и поочерёдно домножим его на R_{α} , $\alpha = \overline{1, p}$. Это даёт

$$\sum_{\beta=1}^p R_{\alpha} R_{\beta}^* \frac{d^2 u_{\beta}}{dt^2} + \sum_{\beta=1}^p R_{\alpha} A R_{\beta}^* u_{\beta} = 0, \quad \alpha = \overline{1, p}. \quad (13)$$

Полученная система уравнений дополняется начальными условиями

$$u_{\alpha}(0) = u_{\alpha}^0, \quad \frac{du_{\alpha}}{dt}(0) = \tilde{u}_{\alpha}^0, \quad \alpha = \overline{1, p}, \quad (14)$$

которые вытекают из условий (2). В (14) векторы u_{α}^0 , $\alpha = \overline{1, p}$, и $\tilde{u}_{\alpha}^0$, $\alpha = \overline{1, p}$, – решения, возможно не единственные, систем уравнений

$$\sum_{\beta=1}^p R_{\alpha} R_{\beta}^* u_{\beta}^0 = R_{\alpha} u^0, \quad \sum_{\beta=1}^p R_{\alpha} R_{\beta}^* \tilde{u}_{\beta}^0 = R_{\alpha} \tilde{u}^0, \quad \alpha = \overline{1, p}.$$

Наша цель состоит в том, чтобы построить такие аппроксимации по времени для задачи Коши (13), (14), которые обеспечивали бы переход на новый слой по времени с помощью решения отдельных задач для u_{α} , $\alpha = \overline{1, p}$. В этом случае, принимая во внимание представление (12), речь идёт о схемах расщепления решения для задачи (1), (2).

Запишем систему (13) в виде одного уравнения второго порядка для векторных величин. Определим вектор $\mathbf{u} = \{u_1, \dots, u_p\}$ и от задачи (13), (14) перейдём к задаче Коши

$$\mathbf{B} \frac{d^2 \mathbf{u}}{dt^2} + \mathbf{A} \mathbf{u} = 0, \quad (15)$$

$$\mathbf{u}(t) = \mathbf{u}^0, \quad \frac{d\mathbf{u}}{dt}(0) = \tilde{\mathbf{u}}^0, \quad (16)$$

где $\mathbf{B}$ и $\mathbf{A}$ – следующие операторные матрицы: $\mathbf{B} = (R_{\alpha} R_{\beta}^*)$, $\mathbf{A} = (R_{\alpha} A R_{\beta}^*)$, $\alpha, \beta = \overline{1, p}$.

Задачу (15), (16) естественно рассматривать на прямой сумме $\mathbf{V} = V_1 \oplus \dots \oplus V_p$ пространств $V_1, \dots, V_p$; при этом для $\mathbf{u}, \mathbf{v} \in \mathbf{V}$ скалярное произведение и норма определяются выражениями

$$(\mathbf{u}, \mathbf{v}) = \sum_{\alpha=1}^p (u_{\alpha}, v_{\alpha})_{\alpha} \quad \text{и} \quad \|\mathbf{u}\| = (\mathbf{u}, \mathbf{u})^{1/2},$$

здесь и далее $(\cdot, \cdot)_{\alpha}$ – скалярное произведение в пространстве V_{α} , $\alpha = \overline{1, p}$.

В силу того, что $(R_\alpha R_\beta^*)^* = R_\beta R_\alpha^*$, $\alpha, \beta = \overline{1, p}$, оператор $\mathbf{B}$ самосопряжён. Кроме того, непосредственно убеждаемся в том, что справедливо равенство

$$(\mathbf{B}\mathbf{u}, \mathbf{u}) = \left(\left(\sum_{\alpha=1}^p R_\alpha^* u_\alpha \right)^2, 1 \right), \quad (17)$$

из которого вытекает неотрицательность оператора $\mathbf{B}$. В силу равенств (17) и (12) имеем

$$(\mathbf{B}\mathbf{u}, \mathbf{u}) = \|\mathbf{u}\|^2. \quad (18)$$

Аналогичные свойства при выполнении условий (3) имеют место для оператора $\mathbf{A}$:

$$(\mathbf{A}\mathbf{u}, \mathbf{u}) = \left(\left(\sum_{\alpha=1}^p R_\alpha^* u_\alpha \right)^2, 1 \right)_A = \|\mathbf{u}\|_A^2. \quad (19)$$

Таким образом, приходим к векторной задаче (15), (16), в которой $\mathbf{B} = \mathbf{B}^* \geq 0$, $\mathbf{A} = \mathbf{A}^* \geq 0$. Домножая уравнение (15) скалярно в $\mathbf{V}$ на $d\mathbf{u}/dt$, получаем

$$\frac{d}{dt} \left(\left(\mathbf{B} \frac{d\mathbf{u}}{dt}, \frac{d\mathbf{u}}{dt} \right) + (\mathbf{A}\mathbf{u}, \mathbf{u}) \right) = 0.$$

Тем самым имеет место равенство

$$\left(\mathbf{B} \frac{d\mathbf{u}}{dt}, \frac{d\mathbf{u}}{dt} \right) + (\mathbf{A}\mathbf{u}, \mathbf{u}) = (\mathbf{B}\tilde{\mathbf{u}}^0, \tilde{\mathbf{u}}^0) + (\mathbf{A}\mathbf{u}^0, \mathbf{u}^0),$$

которое с учётом равенств (18), (19) равносильно оценке (6) для решения задачи (1), (2).

После декомпозиции (12) решения задачи (1), (2) приходим к векторной задаче (13), (14) для отдельных слагаемых решения u_α , $\alpha = \overline{1, p}$. Они связаны друг с другом компонентами не только операторной матрицы $\mathbf{A}$, но и матрицы $\mathbf{B}$. Чтобы получить распадающиеся задачи, нужно строить те или иные расщепления обеих операторных матриц $\mathbf{A}$ и $\mathbf{B}$. Особенно проблематично это сделать для матрицы $\mathbf{B}$ – приходится ориентироваться на четырёхслойные разностные аппроксимации по времени. Явно-неявные аппроксимации при расщеплении оператора при производной по времени для эволюционных уравнений первого порядка рассмотрены в статье [4] и использовались в [7] при построении схем расщепления решения. В настоящей работе рассматриваются варианты расщепления решения в случае, когда операторная матрица $\mathbf{B}$ является диагональной и необходимо расщеплять только операторную матрицу $\mathbf{A}$. Выделим два типа такого расщепления, которые связаны с дополнительными ограничениями на пространства V_α , $\alpha = \overline{1, p}$.

Будем искать решение $\mathbf{u}$ на прямой сумме подпространств, когда для $\mathbf{V}$ используется декомпозиция

$$\mathbf{V} = \bigoplus_{\alpha=1}^p R_\alpha^* V_\alpha.$$

В этом случае для операторов ограничения имеет место соотношение

$$R_\alpha R_\beta^* = \begin{cases} > 0, & \alpha = \beta, \\ 0, & \alpha \neq \beta, \end{cases} \quad \alpha, \beta = \overline{1, p}. \quad (20)$$

Поэтому для отдельных слагаемых решения в представлении (12) получаем $R_\alpha R_\alpha^* u_\alpha = R_\alpha u$, $\alpha = \overline{1, p}$. Для векторной задачи (15), (16) в случае (20) операторная матрица $\mathbf{B}$ является диагональной и положительной, а матрица $\mathbf{A}$ та же, что и выше:

$$\mathbf{B} = \text{diag}(R_1 R_1^*, \dots, R_p R_p^*), \quad \mathbf{A} = (R_\alpha A R_\beta^*), \quad \alpha, \beta = \overline{1, p}. \quad (21)$$

Отметим второй подход к аддитивному представлению решения (12), который также приводит к диагональной операторной матрице $\mathbf{B}$. Эта новая конструкция предложена в работах [9, 10] для построения схем декомпозиции области для нестационарных задач. Будем считать, что для операторов ограничения справедливо равенство

$$\sum_{\alpha=1}^p R_{\alpha}^* R_{\alpha} = I. \quad (22)$$

Тем самым в (12) имеем $u_{\alpha} = R_{\alpha} u$, $\alpha = \overline{1, p}$. Тогда при расщеплении решения (12), (22) в векторной задаче (15), (16) для операторных матриц $\mathbf{B}$ и $\mathbf{A}$ получаем

$$\mathbf{B} = \mathbf{I}, \quad \mathbf{A} = (R_{\alpha} A R_{\beta}^*), \quad \alpha, \beta = \overline{1, p}, \quad (23)$$

где $\mathbf{I}$ – единичная операторная матрица в V .

3. Аппроксимация по времени. Мы ориентируемся на схемы расщепления решения, в которых отдельные компоненты v_{α} , $\alpha = \overline{1, p}$, вектора $\mathbf{v}$ определяются из независимых задач. Благодаря выбору подпространств (20) или (22) эти компоненты (см. (21) и (23)) связываются только операторной матрицей $\mathbf{A}$.

Схемы расщепления решения мы можем строить на выделении диагональной части операторной матрицы $\mathbf{A}$, тогда приходим к аналогам блочных методов Якоби. Второй вариант базируется на треугольном разложении операторной матрицы $\mathbf{A}$ – приходим к аналогам блочных методов Зейделя. Подобные классы аддитивных схем используются, например, при численном решении краевых задач для систем параболических уравнений [6, гл. 10].

Выделим диагональную часть $\mathbf{A}_0$ операторной матрицы $\mathbf{A}$:

$$\mathbf{A}_0 = \text{diag} (R_1 A R_1^*, \dots, R_p A R_p^*).$$

Для приближённого решения задачи (15), (16) будем использовать трёхслойную схему

$$\mathbf{B} \frac{\mathbf{y}^{n+1} - 2\mathbf{y}^n + \mathbf{y}^{n-1}}{\tau^2} + \mathbf{A}_0(\sigma \mathbf{y}^{n+1} + (1 - 2\sigma)\mathbf{y}^n + \sigma \mathbf{y}^{n-1}) + (\mathbf{A} - \mathbf{A}_0)\mathbf{y}^n = 0, \quad n = \overline{1, N-1}, \quad (24)$$

$$\mathbf{y}^0 = \mathbf{u}^0, \quad \mathbf{y}^1 = \bar{\mathbf{u}}^1, \quad (25)$$

с весовым параметром $\sigma = \text{const} > 0$. Второе начальное условие (25) рассчитывается, например, с использованием двухслойной схемы

$$\left(\mathbf{I} + \frac{\tau^2}{2} \mathbf{A}_0 \right) \bar{\mathbf{u}}^1 = \mathbf{u}^0 + \tau \tilde{\mathbf{u}}^0 + \frac{\tau^2}{2} (\mathbf{A}_0 - \mathbf{A}) \mathbf{u}^0.$$

Для нахождения приближённого решения на новом слое в рассматриваемых нами случаях (21) или (23) имеем систему распадающихся уравнений

$$(\mathbf{B} + \sigma \tau^2 \mathbf{A}_0) \mathbf{w}^{n+1} = \boldsymbol{\varphi}^n,$$

которая однозначно разрешима. Условия устойчивости даются следующим утверждением с учётом того, что $\mathbf{A} \geq 0$.

Теорема 1. Для трёхслойной явно-неявной схемы (24), (25) при выполнении условий (21) или (23) имеет место равенство (8), в котором

$$\mathcal{E}^{n+1} = \left(\mathbf{A} \frac{\mathbf{y}^{n+1} + \mathbf{y}^n}{2}, \frac{\mathbf{y}^{n+1} + \mathbf{y}^n}{2} \right) + \left(\mathbf{G} \frac{\mathbf{y}^{n+1} - \mathbf{y}^n}{\tau}, \frac{\mathbf{y}^{n+1} - \mathbf{y}^n}{\tau} \right), \quad (26)$$

где $\mathbf{G} = \mathbf{B} + \sigma \tau^2 \mathbf{A}_0 - \tau^2 \mathbf{A}/4$.

При $\sigma \geq 0.25p$ справедливо неравенство $\mathbf{G} > 0$, которое обеспечивает безусловную устойчивость схемы с выполнением априорной оценки

$$\left\| \frac{\mathbf{y}^{n+1} - \mathbf{y}^n}{\tau} \right\|_{\mathbf{G}}^2 \leq \left\| \frac{\bar{\mathbf{u}}^1 - \mathbf{u}^0}{\tau} \right\|_{\mathbf{G}}^2 + \left\| \frac{\bar{\mathbf{u}}^1 + \mathbf{u}^0}{2} \right\|_{\mathbf{A}}^2, \quad n = \overline{1, N-1}. \quad (27)$$

Доказательство. Запишем схему (24) в виде

$$\mathbf{C} \frac{\mathbf{y}^{n+1} - 2\mathbf{y}^n + \mathbf{y}^{n-1}}{\tau^2} + \mathbf{A}\mathbf{y}^n = 0, \quad n = \overline{1, N-1}, \quad (28)$$

где $\mathbf{C} = \mathbf{B} + \sigma\tau^2\mathbf{A}_0$. На основании доказанной леммы справедливо равенство (8) при задании $\mathcal{E}^{n+1}$ равенством (26).

Очевидно, что $\mathbf{G} = \mathbf{C} - \tau^2\mathbf{A}/4$. Поэтому с учётом (19) и неравенства

$$\left(\sum_{\alpha=1}^p a_{\alpha} \right)^2 \leq p \sum_{\alpha=1}^p a_{\alpha}^2$$

получаем

$$(\mathbf{A}_0\mathbf{v}, \mathbf{v}) = \sum_{\alpha=1}^p ((R_{\alpha}^*v_{\alpha})^2, 1)_A \geq \frac{1}{p}(\mathbf{A}\mathbf{v}, \mathbf{v}).$$

Тем самым $\mathbf{G} > 0$ при $\sigma \geq 0.25p$. Это даёт нам возможность из равенства (8), (26) вывести оценку устойчивости по начальным данным (27). Теорема доказана.

Можно ориентироваться на оценки устойчивости непосредственно для исходной задачи (1), (2). Например, в случае (21) вследствие (18), (19) при $\sigma \geq 0.25p$ имеем

$$\left(\mathbf{A} \frac{\mathbf{y}^{n+1} + \mathbf{y}^n}{2}, \frac{\mathbf{y}^{n+1} + \mathbf{y}^n}{2} \right) + \left(\mathbf{G} \frac{\mathbf{y}^{n+1} - \mathbf{y}^n}{\tau}, \frac{\mathbf{y}^{n+1} - \mathbf{y}^n}{\tau} \right) \geq \left\| \frac{\mathbf{y}^{n+1} - \mathbf{y}^n}{\tau} \right\|^2 + \left\| \frac{\mathbf{y}^{n+1} + \mathbf{y}^n}{2} \right\|_{\mathbf{A}}^2$$

с учётом того, что

$$\mathbf{y}^{n+1} = \sum_{\alpha=1}^p R_{\alpha}^* y_{\alpha}^{n+1}, \quad \mathbf{y}^{n+1} = \{y_1^{n+1}, \dots, y_p^{n+1}\}.$$

Равенство (8), (26) приводит нас к аналогу оценки (6):

$$\left\| \frac{\mathbf{y}^{n+1} - \mathbf{y}^n}{\tau} \right\|^2 + \left\| \frac{\mathbf{y}^{n+1} + \mathbf{y}^n}{2} \right\|_{\mathbf{A}}^2 \leq \left\| \frac{\bar{\mathbf{u}}^1 - \mathbf{u}^0}{\tau} \right\|^2 + \left\| \frac{\bar{\mathbf{u}}^1 + \mathbf{u}^0}{2} \right\|_{\mathbf{A}}^2 + \left((\mathbf{G} - \mathbf{B}) \frac{\bar{\mathbf{u}}^1 - \mathbf{u}^0}{\tau}, \frac{\bar{\mathbf{u}}^1 - \mathbf{u}^0}{\tau} \right), \quad (29)$$

где $n = \overline{1, N-1}$.

Любое расщепление операторов задачи приводит к дополнительным погрешностям при приближённом решении нестационарных задач. Повышения точности схем расщепления для задачи (15), (16) можно достичь использованием треугольного расщепления операторной матрицы $\mathbf{A}$. Определим треугольные матрицы

$$\mathbf{A}_1 = \begin{pmatrix} \frac{1}{2}R_1AR_1^* & 0 & \cdots & 0 \\ R_2AR_1^* & \frac{1}{2}R_2AR_2^* & \cdots & 0 \\ \cdots & \cdots & \cdots & 0 \\ R_pAR_1^* & R_pAR_2^* & \cdots & \frac{1}{2}R_pAR_p^* \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} \frac{1}{2}R_1AR_1^* & R_1AR_2^* & \cdots & R_1AR_p^* \\ 0 & \frac{1}{2}R_2AR_2^* & \cdots & R_2AR_p^* \\ 0 & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \frac{1}{2}R_pRR_p^* \end{pmatrix}.$$

Тогда очевидны равенства

$$\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2, \quad \mathbf{A}_1^* = \mathbf{A}_2. \quad (30)$$

Расщепления (30) связываются со схемами попеременно-треугольного метода Самарского [11].

При двухкомпонентном расщеплении (30) наибольший интерес представляют аналоги классических схем переменных направлений. Такие аддитивные схемы для эволюционных уравнений второго порядка рассмотрены в [6, гл. 4]. Для приближённого решения задачи (15), (16), (30) будем использовать схему

$$(B + \sigma\tau^2 A_1)B^{-1}(B + \sigma\tau^2 A_2)\frac{y^{n+1} - 2y^n + y^{n-1}}{\tau^2} + Ay^n = 0, \quad n = \overline{1, N-1}. \quad (31)$$

Вычислительная реализация (31) обеспечивается последовательным определением компонент решения на новом слое по времени. В плане адаптации в архитектуре параллельных вычислительных систем схема (24) более предпочтительна, так как компоненты решения находятся независимо друг от друга.

Теорема 2. Для трёхслойной явно-неявной схемы (25), (31) при выполнении условий (21) или (23) имеет место равенство (8), (26), в котором

$$G = C - \frac{\tau^2}{4}A, \quad C = B + \sigma\tau^2 A + \sigma^2\tau^4 A_1 B^{-1} A_2.$$

При $\sigma \geq 0.25$ справедливы соотношения $G = G^* > 0$, что обеспечивает безусловную устойчивость схемы с выполнением априорной оценки (27).

Доказательство. Исследование устойчивости проводится по схеме доказательства теоремы 1. Схема (31) записывается в виде (28) при $C = C^* \geq B + \sigma\tau^2 A$. Устойчивость в силу доказанной в работе леммы имеет место при стандартных ограничениях на вес $\sigma \geq 0.25$. Теорема доказана.

Работа выполнена при финансовой поддержке Правительства Российской Федерации (соглашение № 14.Y26.31.0013).

СПИСОК ЛИТЕРАТУРЫ

1. Samarskii A.A. The Theory of Difference Schemes. New York, 2001.
2. Самарский А.А., Гулин А.В. Устойчивость разностных схем. М., 1973.
3. Hundsdorfer W., Verwer J. Numerical Solution of Time-Dependent Advection–Diffusion–Reaction Equations. Berlin, 2003.
4. Вабищевич П.Н. Явно-неявные схемы для эволюционных уравнений первого порядка // Дифференц. уравнения. 2020. Т. 56. № 7. С. 910–917.
5. Marchuk G.I. Splitting and alternating direction methods // Handbook of Numerical Analysis / Eds. P.G. Ciarlet, J.-L. Lions. V. I. North-Holland, 1990. P. 197–462.
6. Vabishchevich P.N. Additive Operator-Difference Schemes: Splitting Schemes. Berlin, 2014.
7. Efendiev Y., Vabishchevich P.N. Splitting methods for solution decomposition in nonstationary problems // Appl. Math. and Comput. 2021. V. 397. P. 125785.
8. Samarskii A.A., Matus P.P., Vabishchevich P.N. Difference Schemes with Operator Factors. Dordrecht, 2002.
9. Vabishchevich P.N. Vector domain decomposition schemes for parabolic equations // Comput. Math. Math. Phys. 2017. V. 57. № 9. P. 1511–1527.
10. Vabishchevich P.N. Two classes of vector domain decomposition schemes for time-dependent problems with overlapping subdomains // In: I. Lirkov, S. Margenov Eds. Large-Scale Scientific Computing. LSSC 2017. Lecture Notes in Computer Science. V. 10665. Springer, 2018. P. 85–92.
11. Samarskii A.A. An economical algorithm for the numerical solution of systems of differential and algebraic equations // U.S.S.R. Comput. Math. Math. Phys. 1964. V. 4. № 3. P. 263–271.

Институт проблем безопасного развития
атомной энергетики РАН, г. Москва,
Северо-Восточный федеральный университет
им. М.К. Аммосова, г. Якутск

Поступила в редакцию 22.01.2021 г.
После доработки 22.01.2021 г.
Принята к публикации 27.04.2021 г.

УДК 519.642.2

ОБ ОДНОМ КЛАССЕ ИНТЕГРО-ДИФФЕРЕНЦИАЛЬНЫХ
УРАВНЕНИЙ В ОСОБОМ СЛУЧАЕ

© 2021 г. Н. С. Габбасов

Исследовано линейное интегро-дифференциальное уравнение с особым дифференциальным оператором в главной части. Для его приближённого решения в пространстве обобщённых функций предложен и обоснован специальный обобщённый вариант метода коллокации.

DOI: 10.31857/S0374064121070037

Объектом исследования в работе является линейное интегро-дифференциальное уравнение (ИДУ) вида

$$Ax \equiv x^{(p)}(t) \prod_{j=1}^q (t - t_j)^{m_j} + \int_{-1}^1 K(t, s)x(s) ds = y(t), \quad (1)$$

в котором $t \in I \equiv [-1, 1]$, числа $t_j \in (-1, 1)$, $m_j \in \mathbb{N}$ ($j = \overline{1, q}$) и $p \in \mathbb{Z}_+$ фиксированы; K и y – известные непрерывные функции, обладающие определёнными свойствами “гладкости” точечного характера, а x – искомая функция. Очевидно, что задача о нахождении решения уравнения (1) в классе обычных гладких функций является некорректно поставленной. Следовательно, возникает важный вопрос о построении основных пространств, обеспечивающих корректность этой задачи. При изучении этого вопроса естественно учитывать то, что при $p = 0$ ИДУ (1) представляет собой интегральное уравнение третьего рода (УТР) (т.е. в этом смысле эти уравнения являются “родственными”). Последнее уравнение встречается в ряде задач теорий переноса нейтронов, упругости, рассеяния частиц (см., например, [1 и библиографию в ней; 2, с. 121–129]), теории уравнений с частными производными смешанного типа [3], а также теории сингулярных интегральных уравнений с вырождающимся символом [4]. При этом, как правило, естественными классами решений УТР являются специальные пространства обобщённых функций типа D или V . Под D (соответственно, V) понимается пространство обобщённых функций, построенных при помощи функционала “дельта-функция Дирака” (соответственно, “конечная часть интеграла по Адамару”). Подробный обзор полученных результатов и обширную библиографию по УТР можно найти в монографии [5, с. 3–11, 168–173] и в диссертации [6, с. 3–6, 106–114].

ИДУ (1) при $q = 1$, $t_1 = 0$ исследовано в работе [7, с. 25–43], в которой с использованием известных результатов по УТР построена теория Нётера для такого уравнения в классах гладких и обобщённых функций типа D . При этом вопросы фредгольмовости и обратимости оператора A , поиска точного решения ИДУ (1) в пространстве обобщённых функций остались не решёнными. Следует отметить, что исследуемые ИДУ точно решаются лишь в очень редких частных случаях. Поэтому особенно актуальна разработка эффективных методов их приближённого решения в пространствах обобщённых функций с соответствующим теоретическим обоснованием. Построение таких методов до настоящего времени в литературе не рассматривалось.

В данной работе построена полная теория разрешимости общего ИДУ (1) в некотором пространстве типа D обобщённых функций (фредгольмовость уравнения, условия разрешимости, алгоритм отыскания точного решения, достаточные условия непрерывной обратимости оператора A). Более того, разработан полиномиальный прямой проекционный метод, специально приспособленный к приближённому решению ИДУ (1) в классе обобщённых функций, и дано его обоснование в смысле [8, гл. 1]. Именно, доказана теорема существования и

единственности решения соответствующего приближённого уравнения, установлены оценки погрешности этого решения и доказана сходимость последовательности приближённых решений к точному решению в пространстве обобщённых функций. Исследованы также вопросы устойчивости и обусловленности аппроксимирующих уравнений.

1. Пространства основных и обобщённых функций. Пусть $C \equiv C(I)$ – банахово пространство всех непрерывных на I функций с обычной макс-нормой и $m \in \mathbb{N}$. Следуя [9], скажем, что функция $f \in C$ принадлежит классу $C\{m; t_0\} \equiv C_{t_0}^{\{m\}}(I)$, если в точке $t_0 \in (-1, 1)$ существует тейлоровская производная $f^{\{m\}}(t_0)$ порядка m (естественно считаем, что $C\{0; t_0\} \equiv C$).

Далее, пусть $t_1, t_2, \dots, t_q$ – произвольно фиксированные попарно различные точки интервала $(-1, 1)$. Каждой точке t_j поставим в соответствие некоторое натуральное число m_j ($j = \overline{1, q}$). Введём векторное пространство

$$C\{\overline{m}; \overline{\tau}\} \equiv C_{\overline{\tau}}^{\{\overline{m}\}}(I) \equiv \bigcap_{j=1}^q C\{m_j; t_j\},$$

где $\overline{m} \equiv (m_1, m_2, \dots, m_q)$ и $\overline{\tau} \equiv (t_1, t_2, \dots, t_q)$ – конечные наборы соответствующих величин. Построим теперь основное в наших исследованиях пространство

$$Y \equiv \{y \in C\{\overline{m}; \overline{\tau}\} : y^{\{i\}}(t_j) = 0 \quad (i = \overline{0, p-1}, \quad j = \overline{1, q})\},$$

где $p \in \mathbb{Z}_+$ таково, что $p < \mu$, $\mu \equiv \min\{m_1, m_2, \dots, m_q\}$. Снабдим его нормой

$$\|y\|_Y \equiv \|Ty\|_C + \sum_{j=1}^q \sum_{i=p}^{m_j-1} |y^{\{i\}}(t_j)|, \quad (2)$$

где $T : Y \rightarrow C$ – “характеристический” оператор класса Y , определяемый следующим образом:

$$(Ty)(t) \equiv \left[y(t) - \sum_{j=1}^q \sum_{i=p}^{m_j-1} y^{\{i\}}(t_j) H_{ji}(t) \right] (u(t))^{-1} \equiv H(t) \in C,$$

$$u(t) \equiv \prod_{j=1}^q (t - t_j)^{m_j}, \quad H(t_j) \equiv \lim_{t \rightarrow t_j} H(t) \quad (j = \overline{1, q}),$$

H_{ji} – фундаментальные полиномы Эрмита степени $m - 1$ по системе $\{t_j\}$, $m \equiv \sum_{j=1}^q m_j$.

Лемма 1. *i) Включение $y \in Y$ равносильно равенству*

$$y(t) = (u \cdot H)(t) + \sum_{j=1}^q \sum_{i=p}^{m_j-1} \alpha_{ji} H_{ji}(t), \quad (3)$$

причём $Ty = H \in C$ с точностью до устранимого разрыва в точках t_j ($j = \overline{1, q}$), а $y^{\{i\}}(t_j) = \alpha_{ji} \in \mathbb{R}$ ($i = \overline{p, m_j-1}$, $j = \overline{1, q}$).

ii) Пространство Y по норме (2) полно и вложено в пространство C .

Справедливость леммы 1 легко следует из леммы 1.4.5 [5, с. 23] для класса $C\{\overline{m}; \overline{\tau}\}$ и определения пространства Y основных функций.

Критерий компактности множеств в пространстве Y устанавливает

Лемма 2. *Множество $M \subset Y$ относительно компактно в Y тогда и только тогда, когда: i) M ограничено; ii) семейство $T(M)$ непрерывных на I функций равностепенно непрерывно.*

Доказательство проводится так же, как и доказательство теоремы 1.2.2 из [5, с. 15]. Отличие заключается лишь в том, что роль пространства $C\{m; 0\}$ и его “характеристического” оператора, фигурирующих в теореме 1.2.2, играют соответственно Y и T .

Обозначим через $C^{(p)} \equiv C^{(p)}(I)$ векторное пространство p раз непрерывно дифференцируемых на I функций и наделим его специальной нормой

$$\|z\|_{(p)} \equiv \|Dz\|_C + \sum_{i=0}^{p-1} |z^{(i)}(-1)| \quad (z \in C^{(p)}), \quad (4)$$

где $Dz \equiv z^{(p)}(t) \in C$. В силу формулы Тейлора (см., например, [10, с. 20]) очевидно, что включение $z \in C^{(p)}$ равносильно равенству

$$z(t) = (Jg)(t) + \sum_{i=0}^{p-1} \alpha_i(t+1)^i, \quad (5)$$

здесь $Dz = g \in C$, $z^{(i)}(-1) = \alpha_i i!$ ($i = \overline{0, p-1}$);

$$Jg \equiv (J_{p-1}g)(t) \equiv ((p-1)!)^{-1} \int_{-1}^t (t-s)^{p-1} g(s) ds;$$

при этом $J : C \rightarrow C^{(p)}$, $(Jg)^{(i)} = J_{p-i-1}g$ ($i = \overline{0, p-1}$), $DJg = g$.

Из соотношений (5) и (4) легко следует

Лемма 3. *Пространство $C^{(p)}$ с нормой (4) полно и нормально вложено в пространство C .*

Следствие 1. *Обычная норма $\|\cdot\|_{C^{(p)}}$ в $C^{(p)}$ и норма (4) эквивалентны, т.е. существует постоянная $d \geq 1$ такая, что $\|z\|_{(p)} \leq \|z\|_{C^{(p)}} \leq d\|z\|_{(p)}$ для любой функции $z \in C^{(p)}$, где $\|z\|_{C^{(p)}} \equiv \sum_{i=0}^p \|z^{(i)}\|_C$.*

Пусть $C_{-1}^{(p)} \equiv C_{-1}^{(p)}(I) \equiv \{z \in C^{(p)} | z^{(i)}(-1) = 0 \text{ } (i = \overline{0, p-1})\}$ – банахово пространство гладких функций с нормой $\|z\|_{(p)} \equiv \|Dz\|_C$.

Теперь над пространством Y основных функций построим семейство $X \equiv D_{-1}^{(p)}\{\overline{m}; \overline{\tau}\}$ обобщённых функций $x(t)$ вида

$$x(t) \equiv z(t) + \sum_{j=1}^q \sum_{i=0}^{m_j-p-1} \gamma_{ji} \delta^{\{i\}}(t-t_j), \quad (6)$$

где $t \in I$, $z \in C_{-1}^{(p)}$, $\gamma_{ji} \in \mathbb{R}$ – произвольные постоянные, а δ и $\delta^{\{i\}}$ – соответственно дельта-функция Дирака и её “тейлоровские” производные, действующие на пространстве Y основных функций по следующему правилу:

$$(\delta^{\{i\}}, y) \equiv \int_{-1}^1 \delta^{\{i\}}(t-t_j) y(t) dt \equiv (-1)^i y^{\{i\}}(t_j) \quad (y \in Y, \quad i = \overline{0, m_j-p-1}, \quad j = \overline{1, q}). \quad (7)$$

Очевидно, что векторное пространство X по норме

$$\|x\|_X \equiv \|z\|_{(p)} + \sum_{j=1}^q \sum_{i=0}^{m_j-p-1} |\gamma_{ji}| \quad (8)$$

является банаховым.

2. Фредгольмовость ИДУ (1). Пусть задано ИДУ (1):

$$(Ax)(t) \equiv (Vx)(t) + (Kx)(t) = y(t) \quad (t \in I), \quad (9)$$

$$V \equiv UD, \quad Uf \equiv (u \cdot f)(t), \quad Kx \equiv \int_{-1}^1 K(t, s)x(s) ds,$$

где $y \in Y$, ядро K удовлетворяет следующим условиям:

$$K(\cdot, s) \in C, \quad K(t, \cdot) \in Y, \quad \varphi_{ji}(s) \equiv K_t^{\{i\}}(t_j, s) \in C \quad (i = \overline{p, m_j - 1}, j = \overline{1, q}),$$

$$\psi_{ji}(t) \equiv K_s^{\{i\}}(t, t_j) \in Y \quad (i = \overline{0, m_j - p - 1}, j = \overline{1, q}); \quad (10)$$

а $x \in X$ – искомая обобщённая функция.

Теорема 1. При выполнении условий (10) оператор $A : X \rightarrow Y$ фредгольмов.

Доказательство. Предварительно изучим уравнение

$$Vx \equiv x^{(p)}(t) \prod_{j=1}^q (t - t_j)^{m_j} = y(t), \quad (11)$$

в котором $p < \mu$, $\mu \equiv \min\{m_1, m_2, \dots, m_q\}$; $y \in Y$. Покажем, что оператор $V : X \rightarrow Y$ ограничен. В силу равенств (6) и (11) имеем

$$(Dx)(t) = (Dz)(t) + \sum_{j=1}^q \sum_{i=0}^{m_j-p-1} \gamma_{ji} \delta^{\{i+p\}}(t - t_j) = (Dz)(t) + \sum_{j=1}^q \sum_{k=p}^{m_j-1} \gamma_{j, k-p} \delta^{\{k\}}(t - t_j). \quad (12)$$

Тогда, учитывая свойство

$$(u(t) \cdot \delta^{\{k\}}(t - t_j), \varphi(t)) \equiv (\delta^{\{k\}}, (u\varphi)(t)) \equiv (-1)^k (u\varphi)^{\{k\}}(t_j) = 0$$

$$(k = \overline{0, m_j - 1}, j = \overline{1, q}; \quad \varphi \in C\{\overline{m}; \overline{\tau}\}), \quad (13)$$

получаем $Vx \equiv U Dx = U Dz$, откуда на основании соотношений (2), (3) и (8) следует, что

$$\|Vx\|_Y = \|UDz\|_Y \equiv \|TUDz\|_C = \|Dz\|_C \equiv \|z\|_{(p)} \leq \|x\|_X,$$

т.е. $\|V\| \equiv \|V\|_{X \rightarrow Y} \leq 1$.

Теперь найдём в пространстве $X \equiv D_{-1}^{(p)}\{\overline{m}; \overline{\tau}\}$ решение уравнения (11) и индекс оператора V . Из равенств (12) и (13) вытекает, что в пространстве X общее решение однородного уравнения $Vx = 0$ имеет вид

$$\tilde{x}(t) = \sum_{j=1}^q \sum_{i=0}^{m_j-p-1} \gamma_{ji} \delta^{\{i\}}(t - t_j) \quad (\gamma_{ji} \in \mathbb{R}).$$

Следовательно, $\alpha(V) \equiv \dim \ker V = m - pq$. С другой стороны, неоднородное уравнение (11) разрешимо в пространстве X тогда и только тогда, когда выполнены дополнительные условия $(\delta^{\{i\}}(t - t_j), y) = 0$ ($i = \overline{p, m_j - 1}, j = \overline{1, q}$). При их выполнении общее решение уравнения (11) представляется формулой

$$x(t) = (JT y)(t) + \sum_{j=1}^q \sum_{i=0}^{m_j-p-1} \gamma_{ji} \delta^{\{i\}}(t - t_j) \quad (\gamma_{ji} \in \mathbb{R}).$$

Это означает, что $\beta(V) \equiv \dim \operatorname{coker} V = m - pq$. Таким образом, $\operatorname{ind} V \equiv \alpha(V) - \beta(V) = 0$, т.е. оператор $V : X \rightarrow Y$ фредгольмов.

Далее обсудим свойства интегрального оператора $\mathcal{K}$. В силу соотношений (9), (6) и (7) имеем

$$(\mathcal{K}x)(t) = (\mathcal{K}z)(t) + \sum_{j=1}^q \sum_{i=0}^{m_j-p-1} (-1)^i \gamma_{ji} \psi_{ji}(t). \quad (14)$$

Отсюда с учётом условий (10) видим, что $Kx \in Y$ ($x \in X$). Используя определение (2), следствие 1 и определение (8), последовательно находим, что

$$\|\mathcal{K}x\|_Y \leq \|\mathcal{K}z\|_Y + \sum_{j=1}^q \sum_{i=0}^{m_j-p-1} |\gamma_{ji}| \|\psi_{ji}\|_Y \leq 2dd_0 \|x\|_X,$$

где $d_0 \equiv \|T_t \mathcal{K}\|_C + (m - pq) \max_{s \in I; i, j} |\varphi_{ji}(s)|$. Следовательно, оператор $\mathcal{K}$ действует из X в Y ограниченно, причём $\|\mathcal{K}\| \equiv \|\mathcal{K}\|_{X \rightarrow Y} \leq 2dd_0$.

Пусть $L \equiv \{x\} \subset X$ – произвольное ограниченное множество. Рассуждая аналогично случаю интегральных уравнений третьего рода (см. [5, с. 52, 53]), с использованием леммы 2 несложно показать, что множество $M \equiv \mathcal{K}(L)$ относительно компактно в Y . Другими словами, оператор $\mathcal{K} : X \rightarrow Y$ вполне непрерывен. Тогда утверждение теоремы 1 непосредственно следует из того, что возмущение нётерова оператора вполне непрерывным оператором сохраняет нётеровость и не изменяет его индекса.

3. Непрерывная обратимость интегро-дифференциального оператора. Рассмотрим ИДУ (1), в котором ядро K подчинено условиям (10), $y \in Y$, а $x \in X$ – искомая обобщённая функция вида (6). С учётом соотношений (6), (12)–(14) преобразуем уравнение (1) к виду

$$(Az)(t) = y(t) - \sum_{j=1}^q \sum_{i=0}^{m_j-p-1} c_{ji} \psi_{ji}(t), \quad (15)$$

где $c_{ji} \equiv (-1)^i \gamma_{ji}$ ($i = \overline{0, m_j - p - 1}$, $j = \overline{1, q}$). Наша задача заключается в нахождении функции $z \in C_{-1}^{(p)}$ и произвольных постоянных c_{ji} .

Лемма 4. Если ядро K (по s) принадлежит классу C , а ядро K (по t) и функция y принадлежат пространству Y , то ИДУ (1) ($A : C_{-1}^{(p)} \rightarrow Y$) эквивалентно в $C_{-1}^{(p)}$ ИДУ

$$Bx \equiv (Dx)(t) + \int_{-1}^1 (T_t K)(t, s)x(s) ds = (Ty)(t)$$

и соотношениям

$$\int_{-1}^1 \varphi_{lk}(s)x(s) ds = y^{\{k\}}(t_l) \quad (k = \overline{p, m_l - 1}, \quad l = \overline{1, q}).$$

Доказательство. В силу равенства (3) очевидно, что для любой функции $g \in Y$ имеет место эквивалентность

$$g = 0 \Leftrightarrow Tg = 0, \quad g^{\{k\}}(t_l) = 0 \quad (k = \overline{p, m_l - 1}, \quad l = \overline{1, q}). \quad (16)$$

Тогда, взяв в (16) $g \equiv Ax - y \in Y$ ($x \in C_{-1}^{(p)}$, $y \in Y$), убеждаемся в справедливости утверждения леммы. Лемма доказана.

Из этой леммы следует, что уравнение (15) равносильно ИДУ

$$(Bz)(t) = (Ty)(t) - \sum_{j=1}^q \sum_{i=0}^{m_j-p-1} c_{ji}(T\psi_{ji})(t) \quad (17)$$

в пространстве $C_{-1}^{(p)}$ и соотношениям

$$y^{\{k\}}(t_l) - \int_{-1}^1 \varphi_{lk}(s)z(s)ds - \sum_{j=1}^q \sum_{i=0}^{m_j-p-1} c_{ji}\psi_{ji}^{\{k\}}(t_l) = 0 \quad (k = \overline{p, m_l-1}, \quad l = \overline{1, q}). \quad (18)$$

Очевидно, что подстановка $z^{(p)} \equiv w(t) \in C$ равносильным образом приводит ИДУ (17) к следующему уравнению Фредгольма второго рода в пространстве C :

$$w(t) + \int_{-1}^1 G(t, \rho)w(\rho)d\rho = (Ty)(t) - \sum_{j=1}^q \sum_{i=0}^{m_j-p-1} c_{ji}(T\psi_{ji})(t), \quad (19)$$

где

$$G(t, \rho) \equiv \frac{1}{(p-1)!} \int_{\rho}^1 (T_t K)(t, s)(s-\rho)^{p-1} ds,$$

причём $G \in C(I^2)$.

Пусть $\lambda = -1$ не является собственным значением уравнения (19) (или ядра G) и R – разрешающий оператор этого уравнения. Тогда функция

$$w^*(t) \equiv (RTy)(t) - \sum_{j=1}^q \sum_{i=0}^{m_j-p-1} c_{ji}(RT\psi_{ji})(t)$$

является единственным непрерывным решением уравнения (19). Следовательно,

$$z^*(t) \equiv (Jw^*)(t) = (JRTy)(t) - \sum_{j=1}^q \sum_{i=0}^{m_j-p-1} c_{ji}(JRT\psi_{ji})(t)$$

– единственное гладкое решение ИДУ (17), которое будет решением и исходного уравнения (15), если в силу (18) постоянные $\{c_{ji}\}$ удовлетворяют квадратной системе линейных алгебраических уравнений (СЛАУ)

$$\sum_{j=1}^q \sum_{i=0}^{m_j-p-1} c_{ji}(Q\psi_{ji})^{\{k\}}(t_l) = (Qy)^{\{k\}}(t_l) \quad (k = \overline{p, m_l-1}, \quad l = \overline{1, q}), \quad (20)$$

где оператор $Q \equiv E - KJRT$ отображает Y в Y , а E – единичный оператор в Y .

Таким образом, доказана

Теорема 2. Пусть выполнены следующие условия:

- а) ядро K удовлетворяет условиям (10), и $y \in Y$;
- б) число $\lambda = -1$ не является собственным значением ядра $G(t, \rho)$;
- с) определитель СЛАУ (20) отличен от нуля.

Тогда для любой правой части $y \in Y$ ИДУ(1) допускает единственное обобщённое решение $x^* \in X$. Это решение представляется формулой

$$x^*(t) = (JRTy)(t) - \sum_{j=1}^q \sum_{i=0}^{m_j-p-1} c_{ji}^*(JRT\psi_{ji})(t) + \sum_{j=1}^q \sum_{i=0}^{m_j-p-1} (-1)^i c_{ji}^* \delta^{\{i\}}(t - t_j), \quad (21)$$

где $\{c_{ji}^*\}$ – единственное решение СЛАУ (20).

Следствие 2. В условиях теоремы 2 интегро-дифференциальный оператор $A : X \rightarrow Y$, определённый равенством (1), непрерывно обратим.

Теперь построим пример, который, во-первых, хорошо иллюстрирует содержание теоремы 2 и, во-вторых, играет полезную роль при оптимизации прямых проекционных методов решения ИДУ вида (1).

Пример 1. Ради упрощения выкладок в ИДУ (1) будем считать $q = 1$, $t_1 = 0$, т.е. рассмотрим уравнение вида

$$Ax \equiv t^m x^{(p)}(t) + \int_{-1}^1 K(t, s)x(s) ds = y(t) \quad (t \in I), \quad (22)$$

в котором $p < m$, $y \in Y \equiv \{y \in C\{m; 0\} : y^{\{i\}}(0) = 0 \ (i = \overline{0, p-1})\}$,

$$K(t, s) = K^*(t, s) \equiv \sum_{j=p}^{m-1} (j!(j-p)!)^{-1} t^j s^{j-p},$$

а $x(t) \equiv z(t) + \sum_{i=0}^{m-p-1} \gamma_i \delta^{\{i\}}(t)$ – искомая обобщённая функция. В обозначениях теоремы 2 последовательно заключаем, что справедливы импликации:

$$T_t K = 0 \Rightarrow G(t, \rho) \equiv 0 \Rightarrow R = E : C \rightarrow C \Rightarrow Qy = y - KJT y \quad (y \in Y);$$

$$\psi_i(t) \equiv K_s^{\{i\}}(t, 0) = ((i+p)!)^{-1} t^{i+p} \quad (i = \overline{0, m-p-1}) \Rightarrow T\psi_i = 0 \Rightarrow Q\psi_i = \psi_i \quad (i = \overline{0, m-p-1}).$$

Следовательно, СЛАУ (20) принимает вид

$$\sum_{i=0}^{m-p-1} c_i \delta_{k, i+p} = (y - KJT y)^{\{k\}}(0) \quad (k = \overline{p, m-1}),$$

где $\delta_{k, i+p}$ – символ Кронекера, откуда получаем, что $c_i^* = (y - KJT y)^{\{i+p\}}(0) \ (i = \overline{0, m-p-1})$. Тогда в силу (21) имеем

$$x^*(t) = (JT y)(t) + \sum_{i=0}^{m-p-1} (-1)^i (y - KJT y)^{\{i+p\}}(0) \delta^{\{i\}}(t). \quad (23)$$

Непосредственная подстановка элемента x^* в уравнение (22) показывает, что ИДУ (22) при $K = K^*$ имеет единственное обобщённое решение (23) для любой правой части $y \in Y$.

4. Обобщённый метод коллокации (ОМК) приближённого решения ИДУ (1).

Пусть задано ИДУ (1). Ради сокращения громоздких выкладок и упрощения формулировок, не ограничивая при этом общности идей, методов и результатов, всюду в дальнейшем будем считать $q = 1$, $t_1 = 0$, т.е. рассмотрим ИДУ вида

$$(Ax)(t) \equiv (Vx)(t) + (Kx)(t) = y(t) \quad (t \in I), \quad (24)$$

$$V \equiv UD, \quad Df \equiv f^{(p)}(t), \quad Ug \equiv t^m g(t), \quad Kx \equiv \int_{-1}^1 K(t, s)x(s) ds,$$

где $p \in \mathbb{N} \cup \{0\}$, $m \in \mathbb{N}$, $p < m$; $y \in Y \equiv C\{m, p; 0\} \equiv \{y \in C\{m; 0\} : y^{\{i\}}(0) = 0 \ (i = \overline{0, p-1})\}$, ядро K обладает следующими свойствами:

$$K(\cdot, s) \in C, \quad K(t, \cdot) \in Y, \quad \psi_i(t) \equiv K_s^{\{i\}}(t, 0) \in Y \quad (i = \overline{0, m-p-1}), \quad (25)$$

а $x \in X$ – искомый элемент. В данном случае, согласно определению (6), $X \equiv D_{-1}^{(p)}\{m; 0\}$ – пространство обобщённых функций $x(t)$ вида

$$x(t) \equiv z(t) + \sum_{i=0}^{m-p-1} \gamma_i \delta^{\{i\}}(t) \quad (z \in C_{-1}^{(p)}, \quad \gamma_i \in \mathbb{R}).$$

Приближённое решение уравнения (24) будем искать в виде

$$x_n \equiv x_n(t; \{c_j\}) \equiv g_n(t) + \sum_{i=0}^{m-p-1} c_{i+n} \delta^{\{i\}}(t), \quad (26)$$

$$g_n(t) \equiv (Jz_n)(t), \quad z_n(t) \equiv \sum_{i=0}^{n-1} c_i t^i \quad (n = 2, 3, \dots). \quad (27)$$

Неизвестные коэффициенты $c_j = c_j^{(n)}$ ($j = \overline{0, n+m-p-1}$) найдём, согласно ОМК, из квадратной СЛАУ $(n+m-p)$ -го порядка:

$$(T\rho_n)(\nu_k) = 0 \quad (k = \overline{1, n}), \quad \rho_n^{\{i\}}(0) = 0 \quad (i = \overline{p, m-1}), \quad (28)$$

где $\rho_n(t) \equiv \rho_n^A(t) \equiv (Ax_n - y)(t)$ – невязка приближённого решения, а $\{\nu_k\} \subset I$ – система узлов Чебышёва первого (или второго) рода.

Для вычислительной схемы (24)–(28) справедлива

Теорема 3. Пусть однородное ИДУ $Ax = 0$ имеет в X лишь нулевое решение (например, в условиях теоремы 2), а функции $h \equiv T_t K$ (по t), $f_i \equiv T\psi_i$ ($i = \overline{0, m-p-1}$) и Ty принадлежат классу Дини–Липшица. Тогда при всех $n \in \mathbb{N}$ ($n \geq n_0$) СЛАУ (28) обладает единственным решением $\{c_j^*\}$ и последовательность приближённых решений $x_n^* \equiv x_n(t; \{c_j^*\})$ сходится к точному решению $x^* = A^{-1}y$ по норме пространства X со скоростью

$$\Delta x_n^* \equiv \|x_n^* - x^*\| = O\left\{\left[E_{n-1}^t(h) + \sum_{i=0}^{m-p-1} E_{n-1}(f_i) + E_{n-1}(Ty)\right] \ln n\right\}, \quad (29)$$

где $E_l(f)$ – наилучшее равномерное приближение функции $f \in C$ алгебраическими полиномами степени не выше l , а через $E_l^t(\cdot)$ обозначен функционал $E_l(\cdot)$, применённый по переменной t .

Доказательство. Очевидно, что ИДУ (24) представляется в виде линейного операторного уравнения

$$Ax \equiv Vx + Kx = y \quad (x \in X \equiv D_{-1}^{(p)}\{m; 0\}, \quad y \in Y \equiv C\{m, p; 0\}), \quad (30)$$

в котором оператор $A : X \rightarrow Y$ непрерывно обратим.

Систему (26)–(28) требуется записать также в операторной форме. С этой целью построим соответствующие конечномерные подпространства. Именно, через $X_n \subset X$ обозначим $(n+m-p)$ -мерное подпространство элементов вида (26), а за $Y_n \subset Y$ примем класс $\text{span}\{t^i\}_p^{n+m-1}$. Далее введём линейный оператор $\Gamma_n \equiv \Gamma_{n+m-p} : Y \rightarrow Y_n$ согласно правилу

$$\Gamma_n y \equiv \Gamma_{n+m-p}(y; t) \equiv (UL_n Ty)(t) + \sum_{i=p}^{m-1} y^{\{i\}}(0) \frac{t^i}{i!}, \quad (31)$$

где $L_n : C \rightarrow \Pi_{n-1} \equiv \text{span}\{t^i\}_0^{n-1}$ представляет собой интерполяционный оператор Лагранжа по системе узлов $\{\nu_k\}_1^n$.

Покажем теперь, что система (26)–(28) равносильна линейному уравнению

$$A_n x_n \equiv Vx_n + \Gamma_n Kx_n = \Gamma_n y \quad (x_n \in X_n, \quad \Gamma_n y \in Y_n). \quad (32)$$

Пусть $x_n^* \equiv x_n(t; \{c_j^*\})$ – решение уравнения (32), т.е. $Vx_n^* + \Gamma_n \eta_n^* = 0$ ($\eta_n^* \equiv Kx_n^* - y$). В силу равенств (26), (27) и (31) последнее означает, что

$$(U(z_n^* + L_n T \eta_n^*))(t) + \sum_{i=p}^{m-1} (\eta_n^*)^{\{i\}}(0) \frac{t^i}{i!} \equiv 0. \quad (33)$$

На основании равенства (3) при $q = 1$, $t_1 = 0$ с учётом того, что $L_n^2 = L_n$, очевидна эквивалентность тождества (33) системе

$$(L_n(z_n^* + T \eta_n^*))(t) \equiv 0, \quad (\eta_n^*)^{\{i\}}(0) = 0 \quad (i = \overline{p, m-1}). \quad (34)$$

Далее, согласно структуре уравнения (30) и равенствам (26), (27) имеем

$$(\rho_n^*)^{\{i\}}(0) = (\eta_n^*)^{\{i\}}(0) \quad (i = \overline{p, m-1}, \quad \rho_n^* \equiv Ax_n^* - y)$$

и

$$T \rho_n^* = T(Vx_n^* + \eta_n^*) = z_n^* + T \eta_n^*.$$

Поэтому в системе (34) тривиальность интерполяционного полинома означает, что

$$(T \rho_n^*)(\nu_k) = (z_n^* + T \eta_n^*)(\nu_k) = 0 \quad (k = \overline{1, n}).$$

Следовательно, система (34) принимает вид

$$(T \rho_n^*)(\nu_k) = 0 \quad (k = \overline{1, n}), \quad (\rho_n^*)^{\{i\}}(0) = 0 \quad (i = \overline{p, m-1}).$$

Итак, СЛАУ (28) имеет решение $\{c_j^*\}_0^{n+m-p-1}$, т.е. решение уравнения (32) является решением системы (26)–(28).

Для получения обратного утверждения достаточно провести те же рассуждения в обратном порядке.

Таким образом, для доказательства теоремы 3 достаточно установить существование, единственность и сходимости решений уравнений (32). В этих целях нам понадобится аппроксимативное свойство оператора Γ_n , которое содержит

Лемма 5. Для любой функции $y \in Y$ справедлива оценка

$$\|y - \Gamma_n y\|_Y \leq d_1 E_{n-1}(Ty) \ln n \quad (n = 2, 3, \dots) \quad (35)$$

(здесь и далее d_i ($i = \overline{1, 2}$) – некоторые константы, значения которых не зависят от числа n).

Справедливость леммы 5 легко следует из представления (3) (при $q = 1$, $t_1 = 0$), определений (31), (2) и оценки (см., например, [8, с. 107]) $\|f - L_n f\|_C \leq d_1 E_{n-1}(f) \ln n$ ($f \in C$).

Рассмотрим теперь вопрос о близости операторов A и A_n на подпространстве X_n . Используя уравнения (24), (32) и оценку (35), для произвольного элемента $x_n \in X_n$ находим, что

$$\|Ax_n - A_n x_n\|_Y = \|Kx_n - \Gamma_n Kx_n\|_Y \leq d_1 E_{n-1}(TKx_n) \ln n. \quad (36)$$

В силу равенств (14) и (26) имеем

$$(Kx_n)(t) = (Kg_n)(t) + \sum_{i=0}^{m-p-1} (-1)^i c_{i+n} \psi_i(t).$$

Следовательно,

$$TKx_n = \int_{-1}^1 h(t, s) g_n(s) ds + \sum_{i=0}^{m-p-1} (-1)^i c_{i+n} f_i(t). \quad (37)$$

Для полиномиального приближения функции $TKx_n \in C$ построим следующий элемент:

$$(P_{n-1}x_n)(t) \equiv \int_{-1}^1 h_{n-1}^t(t, s)g_n(s) ds + \sum_{i=0}^{m-p-1} (-1)^i c_{i+n} f_{n-1}^i(t), \quad (38)$$

где $h_{n-1}^t(t, s)$ и $f_{n-1}^i(t)$ – полиномы степени $n-1$ наилучшего равномерного приближения для $h(t, s)$ (по t) и $f_i(t)$ соответственно. Из вида функции (38) следует, что $P_{n-1}x_n \in \Pi_{n-1}$.

На основании выражений (37) и (38), леммы 3 и определения (8) последовательно выводим промежуточную оценку

$$\begin{aligned} E_{n-1}(TKx_n) &\leq \|TKx_n - P_{n-1}x_n\|_C \equiv \\ &\equiv \max_{t \in I} \left| \int_{-1}^1 (h - h_{n-1}^t)(t, s)g_n(s) ds + \sum_i (-1)^i c_{i+n} (f_i - f_{n-1}^i)(t) \right| \leq \\ &\leq 2\|g_n\|_C E_{n-1}^t(h) + \sum_i |c_{i+n}| E_{n-1}(f_i) \leq 2\|g_n\|_{(p)} E_{n-1}^t(h) + \|x_n\|_X \sum_i E_{n-1}(f_i) \leq \\ &\leq 2\|x_n\|_X E_{n-1}^t(h) + 2\|x_n\|_X \sum_i E_{n-1}(f_i) = 2 \left(E_{n-1}^t(h) + \sum_i E_{n-1}(f_i) \right) \|x_n\|. \end{aligned} \quad (39)$$

Из неравенств (36) и (39) следует искомая оценка близости операторов A и A_n :

$$\varepsilon_n \equiv \|A - A_n\|_{X_n \rightarrow Y} \leq d_2 \left(E_{n-1}^t(h) + \sum_i E_{n-1}(f_i) \right) \ln n. \quad (40)$$

Обозначим пространство

$$H_\alpha^r(S) \equiv \{f \in C^{(r)}(I) : \omega(f^{(r)}; \Delta) \leq S\Delta^\alpha, \quad S \equiv \text{const} > 0\},$$

где $\omega(f; \Delta)$ – модуль непрерывности функции $f \in C$ с шагом Δ ($0 < \Delta \leq 2$).

На основании оценок (40) и (35) из теоремы 7 [8, с. 19] вытекает утверждение теоремы 3 с оценкой погрешности (29).

Следствие 3. Если функции h (по t), f_i и Ty принадлежат пространству $H_\alpha^r(S)$, то в условиях теоремы 3 справедлива оценка $\Delta x_n^* = O(n^{-r-\alpha} \ln n)$ ($r+1 \in \mathbb{N}$, $\alpha \in (0, 1]$).

5. Заключительные замечания.

1. В силу определения нормы в пространстве $D_{-1}^{(p)}\{m; 0\}$ нетрудно заметить, что из сходимости последовательности (x_n^*) к $x^* = A^{-1}y$ в метрике $D_{-1}^{(p)}\{m; 0\}$ следует обычная сходимость в пространстве обобщённых функций, т.е. слабая сходимость.

2. При приближении решений операторных уравнений $Ax = y$ возникает естественный вопрос о скорости сходимости невязки $\rho_n^*(t) \equiv (Ax_n^* - y)(t)$ исследуемого метода. Один из результатов в этом направлении легко вытекает из основной теоремы 3, а именно: если исходные данные уравнения (24) принадлежат классу H_α^r ($0 < \alpha \leq 1$, $r = 0, 1, 2, \dots$), то в условиях теоремы 3 справедлива оценка $\|\rho_n^*\|_Y = O(n^{-r-\alpha} \ln n)$.

3. При $p = 0$ исследуемое ИДУ (24) является интегральным уравнением третьего рода с оператором $A : D\{m; 0\} \rightarrow C\{m; 0\}$, а прямой проекционный метод (26)–(28) – специальным для уравнения третьего рода вариантом ОМК. Следовательно, теорема 3 содержит в себе известные результаты [5, с. 105, 106] по обоснованию специального варианта ОМК для решения уравнений третьего рода в классе $D\{m; 0\}$ обобщённых функций.

4. Так как в условиях теоремы 3 аппроксимирующие операторы A_n обладают свойством вида $\|A_n^{-1}\| = O(1)$ ($A_n^{-1} : Y_n \rightarrow X_n$, $n \geq n_1$), то очевидно [8, с. 23, 24], что предложенный в данной работе прямой метод для ИДУ (24) устойчив относительно малых возмущений исходных данных. Последнее позволяет найти численное решение исследуемых уравнений на ЭВМ с любой наперёд заданной степенью точности. Более того, если ИДУ (24) хорошо обусловлено, то хорошо обусловленной является также СЛАУ (28).

СПИСОК ЛИТЕРАТУРЫ

1. *Bart G.R., Warnock R.L.* Linear integral equations of the third-kind // *SIAM J. Math. Anal.* 1973. V. 4. № 4. P. 609–622.
2. *Кейз К.М., Цвайфель П.Ф.* Линейная теория переноса. М., 1972.
3. *Бжизхатлов Х.Г.* Об одной краевой задаче со смещением // *Дифференц. уравнения.* 1973. Т. 9. № 1. С. 162–165.
4. *Расламбеков С.Н.* Сингулярное интегральное уравнение первого рода в исключительном случае в классах обобщенных функций // *Изв. вузов. Математика.* 1983. № 10. С. 51–56.
5. *Габбасов Н.С.* Методы решения интегральных уравнений Фредгольма в пространствах обобщённых функций. Казань, 2006.
6. *Замалиев Р.Р.* О прямых методах решения интегральных уравнений третьего рода с особенностями в ядре: дис. ... канд. физ.-мат. наук. Казань, 2012.
7. *Абдурахман.* Интегральное уравнение третьего рода с особым дифференциальным оператором в главной части: дис. ... канд. физ.-мат. наук. Ростов-на-Дону, 2003.
8. *Габдуллаев Б.Г.* Оптимальные аппроксимации решений линейных задач. Казань, 1980.
9. *Прессдорф З.* Сингулярное интегральное уравнение с символом, обращающимся в нуль в конечном числе точек // *Мат. исследования.* 1972. Т. 7. № 1. С. 116–132.
10. *Никольский С.М.* Квадратурные формулы. М., 1988.

Набережночелнинский институт
Казанского (Приволжского) федерального университета

Поступила в редакцию 15.01.2021 г.
После доработки 15.01.2021 г.
Принята к публикации 27.04.2021 г.

ЧИСЛЕННЫЕ МЕТОДЫ

УДК 519.642+517.956.3

ЧИСЛЕННЫЕ МЕТОДЫ РЕШЕНИЯ НЕЛИНЕЙНОГО ОПЕРАТОРНОГО УРАВНЕНИЯ, ВОЗНИКАЮЩЕГО В ОБРАТНОЙ КОЭФФИЦИЕНТНОЙ ЗАДАЧЕ

© 2021 г. С. В. Гаврилов, А. М. Денисов

Рассматривается обратная задача для линейной системы уравнений в частных производных. Обратная задача состоит в определении двух неизвестных коэффициентов системы по дополнительной информации об одной из компонент решения. Обратная задача сводится к нелинейному операторному уравнению для одного из неизвестных коэффициентов. Для численного решения этого операторного уравнения применяются метод последовательных приближений и метод Ньютона. Приводятся результаты расчётов, иллюстрирующие сходимость численных методов решения обратной задачи.

DOI: 10.31857/S0374064121070049

1. Введение. Рассмотрим задачу определения функций $u(x, t)$ и $a(x, t)$ таких, что

$$u_x + a_t = 0, \quad 0 \leq x \leq l, \quad 0 \leq t \leq T, \quad (1.1)$$

$$a_t = \gamma(t)(\varphi(t)u - a), \quad 0 \leq x \leq l, \quad 0 \leq t \leq T, \quad (1.2)$$

$$u(0, t) = \mu(t), \quad 0 \leq t \leq T, \quad (1.3)$$

$$a(x, 0) = \psi(x), \quad 0 \leq x \leq l. \quad (1.4)$$

Эта задача представляет собой математическую модель процесса динамики сорбции [1, с. 174; 2, с. 6] в предположении, что свойства поглощающего вещества меняются со временем.

В работе [3] изучены вопросы существования и единственности решения следующей обратной задачи.

Пусть функции $\mu(t)$ и $\psi(x)$ заданы, а функции $\gamma(t)$ и $\varphi(t)$ неизвестны. Требуется определить $\gamma(t)$, $\varphi(t)$, $u(x, t)$ и $a(x, t)$, если задана следующая дополнительная информация об одной из компонент решения задачи (1.1)–(1.4):

$$u(l, t) = g(t), \quad 0 \leq t \leq T, \quad (1.5)$$

$$u_x(l, t) = p(t), \quad 0 \leq t \leq T. \quad (1.6)$$

Будем предполагать, что для известных функций $\mu(t)$, $\psi(x)$, $g(t)$ и $p(t)$ выполняются

Условия А. Функции $\mu(t)$, $\psi(x)$, $g(t)$ и $p(t)$ таковы, что: $\mu, g, p \in C[0, T]$; $\psi \in C[0, l]$; $\mu(t) > 0$, $g(t) > 0$, $p(t) < 0$ для $0 \leq t \leq T$; $\psi(x) \geq 0$ для $0 \leq x \leq l$, $\psi(l) = 0$, $\psi(x)$ не равна нулю тождественно.

Дадим определение решения обратной задачи. Пусть $t_0 \in (0, T]$. Введём прямоугольник $Q_{t_0} = \{(x, t) : 0 \leq x \leq l, \quad 0 \leq t \leq t_0\}$.

Определение. Четвёрка функций $(\gamma(t), \varphi(t), u(x, t), a(x, t))$ называется *решением обратной задачи* при $t \in [0, t_0]$, если $\gamma, \varphi \in C[0, t_0]$, $u, u_x, a, a_t \in C(Q_{t_0})$, $\gamma(t) > 0$, $\varphi(t) > 0$ для $0 \leq t \leq t_0$, $u(x, t) > 0$, $a(x, t) \geq 0$ для $(x, t) \in Q_{t_0}$, $\gamma(t)$, $\varphi(t)$, $u(x, t)$, $a(x, t)$ удовлетворяют уравнениям (1.1), (1.2) и условиям (1.3)–(1.6) в Q_{t_0} .

Приведём некоторые результаты работы [3], которые будут использованы в дальнейшем.

Для функции $u(x, t)$ при заданной функции $\gamma(t)$ рассмотрим интегральное уравнение

$$u(x, t) = \mu(t) \exp\{-R(t; \gamma)x\} + \gamma(t) \exp\left\{-\int_0^t \gamma(\theta) d\theta\right\} \int_0^x \exp\{-R(t; \gamma)(x-s)\} \psi(s) ds + \\ + \gamma(t) \int_0^x \int_0^t \exp\left\{-R(t; \gamma)(x-s) - \int_\tau^t \gamma(\theta) d\theta\right\} R(\tau; \gamma) u(s, \tau) d\tau ds, \quad (x, t) \in Q_{t_0}, \quad (1.7)$$

где

$$R(t; \gamma) = -\left[p(t) + \gamma(t) \int_0^t p(\tau) d\tau\right] (g(t))^{-1}.$$

При выполнении условий А решение интегрального уравнения (1.7) существует и единственно. Чтобы подчеркнуть зависимость этого решения от функции $\gamma(t)$, будем обозначать его через $u(x, t; \gamma)$.

Определим оператор

$$(A\gamma)(t) = \left[g(t) - \mu(t) \exp\{-R(t; \gamma)l\} - \gamma(t) \int_0^l \int_0^t H(l, s, t, \tau; \gamma) R(\tau; \gamma) u(s, \tau; \gamma) d\tau ds\right] \times \\ \times \left(\int_0^l H(l, s, t, 0; \gamma) \psi(s) ds\right)^{-1}, \quad 0 \leq t \leq t_0, \quad (1.8)$$

где

$$H(x, s, t, \tau; \gamma) = \exp\left\{-R(t; \gamma)(x-s) - \int_\tau^t \gamma(\theta) d\theta\right\}.$$

В работе [3] показано, что решение обратной задачи сводится к решению нелинейного операторного уравнения

$$\gamma(t) = (A\gamma)(t), \quad 0 \leq t \leq t_0. \quad (1.9)$$

Данная статья посвящена численным методам решения уравнения (1.9) и сформулированной обратной задачи. Для решения нелинейного операторного уравнения (1.9) используются два итерационных метода: метод последовательных приближений и метод Ньютона. Итерационным методам решения операторных уравнений, к которым сводятся обратные задачи, посвящено достаточно много работ (см., например, [4–11]). Большое их число обусловлено особенностями и спецификой каждой конкретной обратной задачи.

Рассмотрим метод последовательных приближений для решения операторного уравнения (1.9).

Предположим, что выполнено неравенство

$$g(0) - \mu(0) \exp\{(p(0)/g(0))l\} > 0. \quad (1.10)$$

Определим положительную постоянную

$$\gamma_0 = (g(0) - \mu(0) \exp\{(p(0)/g(0))l\}) \left(\int_0^l \exp\{(p(0)/g(0))(l-s)\} \psi(s) ds\right)^{-1}.$$

Введём множество функций

$$\Gamma_0 = \{\gamma(t) : \gamma \in C[0, t_0], \quad \gamma_0/2 \leq \gamma(t) \leq 3\gamma_0/2, \quad 0 \leq t \leq t_0\}.$$

Рассмотрим последовательность функций $\gamma_n(t)$, $n = 0, 1, 2, \dots$, определяемую рекуррентно посредством метода последовательных приближений для решения уравнения (1.9):

$$\gamma_0(t) \in \Gamma_0, \quad \gamma_{n+1}(t) = (A\gamma_n)(t), \quad n = 0, 1, 2, \dots \quad (1.11)$$

Из результатов работы [3] следует утверждение о сходимости метода последовательных приближений.

Теорема 1. Пусть для функций $\mu(t)$, $\psi(x)$, $g(t)$ и $p(t)$ выполнены условия A и неравенство (1.10). Тогда найдётся $t_0 \in (0, T]$ такое, что для любой функции $\gamma_0(t) \in \Gamma_0$ последовательность функций $\gamma_n(t)$ принадлежит множеству Γ_0 и при $n \rightarrow \infty$ равномерно сходится к непрерывной функции $\bar{\gamma}(t)$, являющейся решением уравнения (1.9).

2. Метод Ньютона. Рассмотрим вопрос о применении метода Ньютона для численного решения исследуемой обратной задачи. Как уже отмечалось, для этого достаточно применить метод Ньютона для решения нелинейного операторного уравнения (1.9).

Для построения метода Ньютона нужно знать производную оператора, определяемого формулой (1.8). Вначале рассмотрим вопрос о дифференцируемости по параметру решения интегрального уравнения (1.7).

Пусть функции $\gamma(t)$, $\gamma_\Delta(t)$ и число ξ таковы, что функции $\gamma(t)$ и $\gamma(t) + \xi\gamma_\Delta(t)$ положительны и непрерывны на отрезке $[0, t_0]$ для всех $\xi \in (-\varepsilon, \varepsilon)$.

Лемма. Если выполнены условия A , то у решения $u(x, t; \gamma + \xi\gamma_\Delta)$ уравнения (1.7) существует частная производная $\left. \frac{\partial u}{\partial \xi}(x, t; \gamma + \xi\gamma_\Delta) \right|_{\xi=0}$.

Доказательство. Рассмотрим функцию

$$v(x, t; \gamma, \gamma_\Delta, \xi) = \frac{u(x, t; \gamma + \xi\gamma_\Delta) - u(x, t; \gamma)}{\xi}.$$

Так как функции $u(x, t; \gamma + \xi\gamma_\Delta)$ и $u(x, t; \gamma)$ являются решением уравнения (1.7) для $\gamma(t) + \xi\gamma_\Delta(t)$ и $\gamma(t)$ соответственно, то $v(x, t; \gamma, \gamma_\Delta, \xi)$ удовлетворяет уравнению

$$\begin{aligned} v(x, t; \gamma, \gamma_\Delta, \xi) &= \frac{F_1(x, t; \gamma, \gamma_\Delta, \xi) - F_1(x, t; \gamma, 0, 0)}{\xi} + \\ &+ \int_0^x \int_0^t \frac{F_2(x, t, s, \tau; \gamma, \gamma_\Delta, \xi) - F_2(x, t, s, \tau; \gamma, 0, 0)}{\xi} u(s, \tau; \gamma + \xi\gamma_\Delta) d\tau ds + \\ &+ \int_0^x \int_0^t F_2(x, t, s, \tau; \gamma, 0, 0) v(s, \tau; \gamma, \gamma_\Delta, \xi) d\tau ds, \quad (x, t) \in Q_{t_0}, \quad (x, t) \in Q_{t_0}, \end{aligned} \quad (2.1)$$

где

$$F_1(x, t; \gamma, \gamma_\Delta, \xi) = \mu(t) \exp\{-R(t; \gamma + \xi\gamma_\Delta)x\} + (\gamma(t) + \xi\gamma_\Delta(t)) \int_0^x H(x, s, t, 0; \gamma + \xi\gamma_\Delta) \psi(s) ds,$$

а $F_2(x, t, s, \tau; \gamma, \gamma_\Delta, \xi) = (\gamma(t) + \xi\gamma_\Delta(t))H(x, s, t, \tau; \gamma + \xi\gamma_\Delta)R(\tau; \gamma + \xi\gamma_\Delta)$. Переходя к пределу при $\xi \rightarrow 0$, получаем, что

$$\lim_{\xi \rightarrow 0} [F_1(x, t; \gamma, \gamma_\Delta, \xi) - F_1(x, t; \gamma, 0, 0)]/\xi = F_3(x, t; \gamma, \gamma_\Delta), \quad (2.2)$$

$$\lim_{\xi \rightarrow 0} [F_2(x, t, s, \tau; \gamma, \gamma_\Delta, \xi) - F_2(x, t, s, \tau; \gamma, 0, 0)]/\xi = F_4(x, t, s, \tau; \gamma, \gamma_\Delta), \quad (2.3)$$

где

$$F_3(x, t; \gamma, \gamma_\Delta) = \mu(t) \exp\{-R(t; \gamma)x\} \frac{1}{g(t)} \int_0^t p(\theta) d\theta x \gamma_\Delta(t) + \gamma_\Delta(t) \int_0^x H(x, s, t, 0; \gamma) \psi(s) ds + \\ + \gamma(t) \int_0^x H(x, s, t, 0; \gamma) \left(\frac{1}{g(t)} \int_0^t p(\theta) d\theta (x-s) \gamma_\Delta(t) - \int_0^t \gamma_\Delta(\theta) d\theta \right) \psi(s) ds,$$

а

$$F_4(x, t, s, \tau; \gamma, \gamma_\Delta) = \gamma_\Delta(t) H(x, s, t, \tau; \gamma) R(\tau; \gamma) + \\ + \gamma(t) H(x, s, t, \tau; \gamma) \left(\frac{1}{g(t)} \int_0^t p(\theta) d\theta (x-s) \gamma_\Delta(t) - \int_\tau^t \gamma_\Delta(\theta) d\theta \right) R(\tau; \gamma) - \\ - \gamma(t) H(x, s, t, \tau; \gamma) \frac{1}{g(\tau)} \int_0^\tau p(\theta) d\theta \gamma_\Delta(\tau).$$

Уравнение (2.1) представляет собой интегральное уравнение Вольтерры второго рода для функции $v(x, t; \gamma, \gamma_\Delta, \xi)$. Разрешив его относительно $v(x, t; \gamma, \gamma_\Delta, \xi)$, перейдя к пределу при $\xi \rightarrow 0$ и используя формулы (2.2), (2.3), получим, что производная $\partial u(x, t; \gamma + \xi \gamma_\Delta) / \partial \xi$ существует при $\xi = 0$. Обозначим её $w(x, t; \gamma, \gamma_\Delta)$. Из уравнения (2.1) следует, что $w(x, t; \gamma, \gamma_\Delta)$ является решением интегрального уравнения

$$w(x, t; \gamma, \gamma_\Delta) = F_3(x, t; \gamma, \gamma_\Delta) + \int_0^x \int_0^t F_4(x, t, s, \tau; \gamma, \gamma_\Delta) u(s, \tau; \gamma) d\tau ds + \\ + \int_0^x \int_0^t F_2(x, t, s, \tau; \gamma, 0, 0) w(s, \tau; \gamma, \gamma_\Delta) d\tau ds, \quad (x, t) \in Q_{t_0}. \quad (2.4)$$

Лемма доказана.

Рассмотрим вопрос о дифференцируемости по Гато оператора A , определяемого равенством (1.8). Из результатов [3] следует, что существует такое $t_0 \in (0, T]$, что оператор A отображает множество Γ_0 в себя. Далее будем предполагать что число t_0 удовлетворяет этому условию.

Введём множество

$$\Gamma_{00} = \{\gamma(t) : \gamma \in C[0, t_0], \quad \gamma_0/2 < \gamma(t) < 3\gamma_0/2, \quad 0 \leq t \leq t_0\}.$$

Теорема 2. Если выполнены условия A и неравенство (1.10), то для любой функции $\gamma \in \Gamma_{00}$ оператор A дифференцируем по Гато в Γ_{00} .

Доказательство. Пусть $\gamma(t)$ – произвольная функция из Γ_{00} , а функция $\gamma_\Delta(t)$ и число ξ таковы, что функция $\gamma(t) + \xi \gamma_\Delta(t)$ также принадлежит множеству Γ_{00} . Покажем, что при $\xi \rightarrow 0$ функции $((A(\gamma + \xi \gamma_\Delta))(t) - (A\gamma)(t))/\xi$ равномерно сходятся на отрезке $[0, t_0]$.

Рассмотрим функции

$$\frac{1}{\xi} \left[\left(\int_0^l H(l, s, t, 0; \gamma + \xi \gamma_\Delta) \psi(s) ds \right)^{-1} - \left(\int_0^l H(l, s, t, 0; \gamma) \psi(s) ds \right)^{-1} \right].$$

При $\xi \rightarrow 0$ они равномерно на отрезке $[0, t_0]$ сходятся к функции

$$F_5(t; \gamma, \gamma_\Delta) = - \left(\int_0^l H(l, s, t, 0; \gamma) \psi(s) ds \right)^{-2} \times \\ \times \int_0^l H(l, s, t, 0; \gamma) \left[\frac{1}{g(t)} \int_0^t p(\theta) d\theta (l-s) \gamma_\Delta(t) - \int_0^t \gamma_\Delta(\theta) d\theta \right] \psi(s) ds. \quad (2.5)$$

Из леммы следует, что при $\xi \rightarrow 0$ функции

$$\gamma(t) \int_0^l \int_0^t H(l, s, t, \tau; \gamma) R(\tau; \gamma) \frac{u(s, \tau; \gamma + \xi \gamma_\Delta) - u(s, \tau; \gamma)}{\xi} d\tau ds$$

равномерно на отрезке $[0, t_0]$ сходятся к функции

$$\gamma(t) \int_0^l \int_0^t H(l, s, t, \tau; \gamma) R(\tau; \gamma) w(s, \tau; \gamma, \gamma_\Delta) d\tau ds. \quad (2.6)$$

Из формул (2.3), (2.5) и (2.6) вытекает, что

$$\lim_{\xi \rightarrow 0} \frac{(A(\gamma + \xi \gamma_\Delta))(t) - (A\gamma)(t)}{\xi} = \left(\int_0^l H(l, s, t, 0; \gamma) \psi(s) ds \right)^{-1} \times \\ \times \left[- \frac{\mu(t)}{g(t)} \int_0^t p(\theta) d\theta \exp\{-R(t; \gamma)l\} l \gamma_\Delta(t) - \int_0^l \int_0^t F_4(l, t, s, \tau; \gamma, \gamma_\Delta) u(s, \tau; \gamma) d\tau ds \right] - \\ - \left(\gamma(t) \int_0^l \int_0^t H(l, s, t, \tau; \gamma) R(\tau; \gamma) w(s, \tau; \gamma, \gamma_\Delta) d\tau ds \right) \left(\int_0^l H(l, s, t, 0; \gamma) \psi(s) ds \right)^{-1} + \\ + \left[g(t) - \mu(t) \exp\{-R(t; \gamma)l\} - \gamma(t) \int_0^l \int_0^t H(l, s, t, \tau; \gamma) R(\tau; \gamma) u(s, \tau; \gamma) d\tau ds \right] F_5(t; \gamma, \gamma_\Delta), \quad (2.7)$$

где функция $w(x, t; \gamma, \gamma_\Delta)$ определяется из уравнения (2.4). Таким образом, оператор A дифференцируем по Гато на множестве Γ_{00} и его производная $A'[\gamma]\gamma_\Delta$ определяется правой частью равенства (2.7). Теорема доказана.

Итерационный процесс, соответствующий методу Ньютона [12, с. 669], определяется следующим образом. Задаётся функция $\gamma_0(t)$. Последующие функции $\gamma_{n+1}(t)$, $n = 0, 1, \dots$, определяются по формуле $\gamma_{n+1}(t) = \gamma_n(t) + \gamma_{\Delta n}(t)$, где $\gamma_{\Delta n}(t)$ – решение линейного интегрального уравнения

$$\gamma_{\Delta n} - A'[\gamma_n]\gamma_{\Delta n} = -\gamma_n + A\gamma_n. \quad (2.8)$$

3. Вычислительные эксперименты. Приведём результаты некоторых вычислительных экспериментов, в которых для решения исследуемой обратной задачи использовались метод последовательных приближений (1.11) и метод Ньютона (2.8).

Общая схема вычислительных экспериментов была следующей. На отрезке $[0, T]$ задавались функции $\mu(t)$, $\gamma(t)$ и $\varphi(t)$, а на отрезке $[0, l]$ – функция $\psi(x)$. С этими функциями

решалась задача (1.1)–(1.4) и определялись функции $g(t) = u(l, t)$ и $p(t) = u_x(l, t)$. Затем с функциями $\mu(t)$, $g(t)$, $p(t)$ и $\psi(x)$ итерационными методами (1.11) и (2.8) решалось операторное уравнение (1.9) и находилась приближённая функция $\tilde{\gamma}(t)$. Для определения приближённой функции $\tilde{\varphi}(t)$ использовалась полученная в [3] формула

$$\tilde{\varphi}(t) = - \left[p(t) + \tilde{\gamma}(t) \int_0^t p(\tau) d\tau \right] (g(t)\tilde{\gamma}(t))^{-1}. \quad (3.1)$$

При приближённом решении операторного уравнения (1.9) для обоих итерационных методов использовалось одинаковое начальное приближение $\gamma_0(t) = \gamma_0$ и один и тот же критерий останова

$$\|\gamma_{n+1}(t) - \gamma_n(t)\|_{C[0,T]} (\|\gamma_n(t)\|_{C[0,T]})^{-1} \leq \delta.$$

В первом вычислительном эксперименте $T = 0.5$, $l = 1$ и

$$\mu(t) = 1 + t, \quad \psi(x) = 2 - 2x, \quad \gamma(t) = 3 + \sin(2\pi t), \quad \varphi(t) = 0.5 + 0.1 \cos(2\pi t).$$

На рис. 1 приведены функция $\gamma(t) = 3 + \sin(2\pi t)$, первая $\gamma_1^I(t)$ и вторая $\gamma_2^I(t)$ итерации метода последовательных приближений, а также первая $\gamma_1^N(t)$ и вторая $\gamma_2^N(t)$ итерации метода Ньютона. При величине $\delta = 0.001$ метод последовательных приближений остановился на 9-м шаге итераций, а метод Ньютона – на 6-м шаге итераций. Полученные при этом приближённые решения $\tilde{\gamma}^I(t) = \gamma_9^I(t)$ и $\tilde{\gamma}^N(t) = \gamma_6^N(t)$, в рамках масштаба рис. 1, визуально совпадают с точным решением $\gamma(t) = 3 + \sin(2\pi t)$ и поэтому не показаны.

На рис. 2 изображены функция $\varphi(t) = 0.5 + 0.1 \cos(2\pi t)$ и функции $\varphi_1^I(t)$, $\varphi_2^I(t)$, $\varphi_1^N(t)$, $\varphi_2^N(t)$, полученные при подстановке функций $\gamma_1^I(t)$, $\gamma_2^I(t)$, $\gamma_1^N(t)$, $\gamma_2^N(t)$ в формулу (3.1) соответственно. Найденные аналогично приближённые решения $\tilde{\varphi}^I(t) = \varphi_9^I(t)$ и $\tilde{\varphi}^N(t) = \varphi_6^N(t)$ визуально совпадают с точным решением $\varphi(t) = 0.5 + 0.1 \cos(2\pi t)$ и поэтому не показаны.

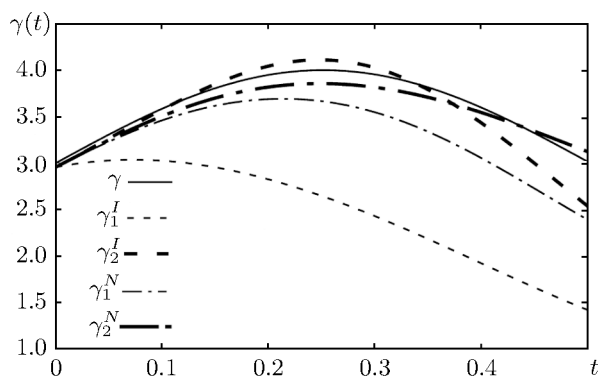


Рис. 1. Результаты первого вычислительного эксперимента: точная функция $\gamma(t)$ и функции, определённые на первых двух итерациях.

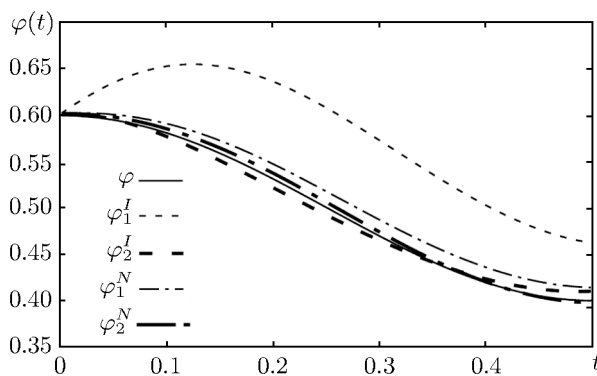


Рис. 2. Результаты первого вычислительного эксперимента: точная функция $\varphi(t)$ и функции, определённые на первых двух итерациях.

В втором вычислительном эксперименте $T = 0.5$, $l = 1$ и

$$\mu(t) = 1 + t, \quad \psi(x) = 2 - 2x, \quad \gamma(t) = 1.5 - 0.5 \sin(\pi t), \quad \varphi(t) = 2 + 0.5 \sin(2\pi t).$$

Величина δ выбиралась равной 0.001. Аналогично рис. 1 на рис. 3 показаны значения точной функции $\gamma(t)$, а также функции, полученные на первых двух итерациях для обоих методов. Критерий сходимости был удовлетворён на 7-м шаге метода последовательных приближений и на 9-м шаге метода Ньютона. В рамках масштаба рис. 3 приближённые решения, найденные на заключительной итерации обоих методов $\tilde{\gamma}^I(t) = \gamma_7^I(t)$ и $\tilde{\gamma}^N(t) = \gamma_9^N(t)$, визуально совпадают с точным решением.

На рис. 4 изображена функция $\varphi(t) = 2 + 0.5 \sin(2\pi t)$, а также соответствующие первым двум итерациям функции $\varphi_1^I(t)$, $\varphi_2^I(t)$, $\varphi_1^N(t)$, $\varphi_2^N(t)$. Функции $\tilde{\varphi}^I(t) = \varphi_7^I(t)$ и $\tilde{\varphi}^N(t) = \varphi_9^N(t)$, найденные по формуле (3.1), совпадают на рисунке с $\varphi(t) = 2 + 0.5 \sin(2\pi t)$.

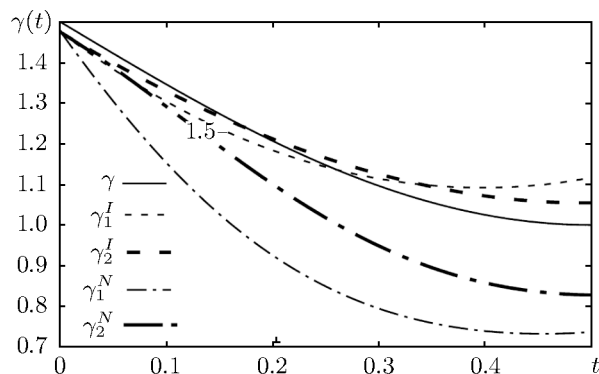


Рис. 3. Результаты второго вычислительного эксперимента: точная функция $\gamma(t)$ и функции, определённые на первых двух итерациях.

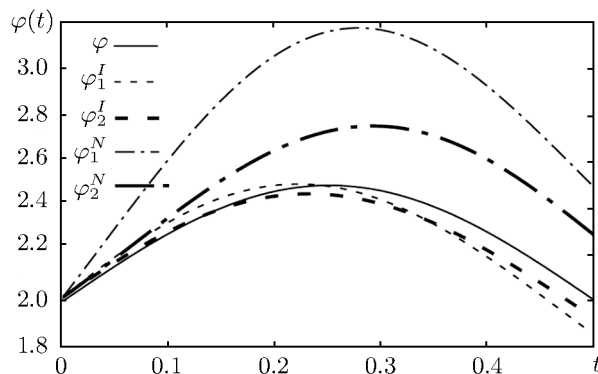


Рис. 4. Результаты второго вычислительного эксперимента: точная функция $\varphi(t)$ и функции, определённые на первых двух итерациях.

Приведённые примеры, а также ряд других численных расчётов позволяют сделать вывод о достаточно быстрой сходимости обоих методов и об отсутствии существенных преимуществ в скорости сходимости одного из методов в сравнении с другим.

Работа выполнена при финансовой поддержке Минобрауки РФ в рамках реализации программы Московского центра фундаментальной и прикладной математики по соглашению № 075-15-2019-1621.

СПИСОК ЛИТЕРАТУРЫ

1. Тихонов А.Н., Самарский А.А. Уравнения математической физики. М., 1999.
2. Денисов А.М., Лукшин А.В. Математические модели неравновесной динамики сорбции. М., 1989.
3. Денисов А.М. Существование и единственность решения системы нелинейных интегральных уравнений // Дифференц. уравнения. 2020. Т. 56. № 9. С. 1174–1181.
4. Бимуратов С.Ш., Кабанikhин С.И. Решение одномерной обратной задачи электродинамики методом Ньютона–Канторовича // Журн. вычислит. математики и мат. физики. 1992. Т. 32. № 12. С. 1900–1915.
5. Monch L. A Newton method for solving inverse scattering problem for a sound-hard obstacle // Inverse problems. 1996. V. 12. № 3. P. 309–324.
6. Kabanikhin S.I., Scherzer O., Shichlenin M.A. Iteration method for solving a two-dimensional inverse problem for hyperbolic equation // J. of Inverse and Ill-Posed Problems. 2003. V. 11. № 1. P. 1–23.
7. Самарский А.А., Вабшцевич П.Н. Численные методы решения обратных задач математической физики. М., 2004.
8. Yan-Bo Ma. Newton method for estimation of the Robin coefficient // J. Nonlin. Sci. Appl. 2015. V. 8. № 5. P. 660–669.
9. Денисов А.М. Итерационный метод решения обратной коэффициентной задачи для гиперболического уравнения // Дифференц. уравнения. 2017. Т. 53. № 7. С. 943–949.
10. Баев А.В., Гаврилов С.В. Итерационный метод решения обратной задачи рассеяния для системы уравнений акустики в слоисто-неоднородной среде с поглощением // Вестн. Моск. гос. ун-та. Сер. 15. Вычислит. математика и кибернетика. 2018. № 2. С. 7–14.
11. Денисов А.М. Итерационный метод решения обратной задачи для гиперболического уравнения с малым параметром при старшей производной // Дифференц. уравнения. 2019. Т. 55. № 7. С. 973–981.
12. Канторович Л.В., Акилов Г.П. Функциональный анализ. М., 1977.

Московский государственный университет
им. М.В. Ломоносова

Поступила в редакцию 03.02.2021 г.
После доработки 03.02.2021 г.
Принята к публикации 27.04.2021 г.

УДК 519.633+517.962.2+517.958

К ВОПРОСУ ОБ ЭКВИВАЛЕНТНОСТИ РАЗНОСТНЫХ МЕТОДОВ РЕШЕНИЯ ЗАДАЧИ СТЕФАНА НА ПОДВИЖНЫХ И ФИКСИРОВАННЫХ СЕТКАХ

© 2021 г. А. О. Гусев, О. В. Щерица, О. С. Мажорова

Для двумерной задачи Стефана представлены разностные схемы, полученные методами с явным выделением границы фазового перехода. Показано, что разностная схема, построенная на подвижной сетке, и разностная схема, построенная методом выпрямления фронта, эквивалентны.

DOI: 10.31857/S0374064121070050

Введение. Работа посвящена построению разностных схем для численного решения классической задачи Стефана. Задача состоит в определении поля температуры и положения границы раздела фаз в области, содержащей два агрегатных состояния чистого вещества. На движущейся границе раздела фаз задаётся фиксированная температура фазового перехода и условие баланса внутренней энергии. В работе рассматриваются методы решения задачи Стефана, основанные на процедуре явного выделения границы фазового перехода. При таком подходе положение границы раздела фаз определяется положением закреплённых на ней узлов сетки. Это обеспечивается за счёт либо использования подвижных сеток, согласованных в исходном пространстве с формой границы фазового перехода [1–8], либо с помощью динамической замены переменных [9–17]; замена переменных выбирается так, чтобы в новых координатах расчётная область была регулярной, с фиксированными границами, совпадающими с координатными линиями – метод выпрямления фронта [18]. В первом случае осуществляется аппроксимация исходных дифференциальных уравнений, во втором – уравнений, полученных в результате замены переменных. На уровне постановки дифференциальной задачи оба эти подхода эквивалентны. В работе [19] для одномерной термодиффузионной задачи Стефана построено семейство разностных схем, для которых подходы, основанные на использовании подвижной и фиксированной сеток, эквивалентны.

В данной работе консервативные разностные схемы на подвижной и фиксированной сетках построены для двумерной задачи Стефана. Доказана алгебраическая эквивалентность построенных схем. Преобразование одной схемы в другую осуществляется с помощью замены переменных, аналогично тому, как это делается в дифференциальной задаче.

1. Постановка задачи. Задачу о фазовом переходе в чистом веществе рассмотрим в декартовой системе координат Oxy в прямоугольнике $\Omega = \Omega(t, x, y) = [0, H_1] \times [0, H_2]$. В подобласти $\Omega^1(t, x, y) = \{(x, y) : x \in [0, H_1], y \in [0, \zeta(t, x)]\}$ находится твёрдая фаза, а в оставшейся части прямоугольника Ω , т.е. в подобласти $\Omega^2(t, x, y) = \{(x, y) : x \in [0, H_1], y \in [\zeta(t, x), H_2]\}$, располагается жидкая фаза. Здесь и далее $y = \zeta(t, x)$ – дифференцируемая функция, представляющая собой в каждый момент времени t границу раздела фаз, положение которой меняется в ходе процесса (см. левую часть рис. 1).

Твёрдая и жидкая фазы имеют одинаковую плотность $\rho^{cr} = \rho^{lq} = \rho$ и удельную теплоёмкость $c_p^{cr} = c_p^l = c_p$, но раз-

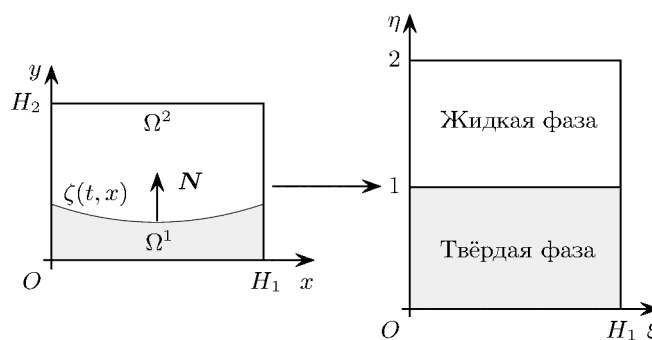


Рис. 1. Преобразование системы координат.

личные коэффициенты теплопроводности – k^{cr} и k^{l} соответственно. Распределение температуры в области описывается уравнением теплопроводности

$$\partial_t T = \nabla \cdot (\kappa \nabla T). \quad (1)$$

Здесь $\nabla = (\partial_x, \partial_y)$, $\partial_x = \partial/\partial x$, $\partial_y = \partial/\partial y$, $\partial_t = \partial/\partial t$. Коэффициент температуропроводности является кусочно-постоянной функцией: $\kappa(x, y) = \kappa^{\text{cr}}$ при $(x, y) \in \Omega^{\text{l}}$, $\kappa(x, y) = \kappa^{\text{l}}$ при $(x, y) \in \Omega^{\text{g}}$, где $\kappa^{\text{cr}} = k^{\text{cr}}/(c_p \rho)$ и $\kappa^{\text{l}} = k^{\text{l}}/(c_p \rho)$.

На межфазной границе температура равна температуре плавления чистого вещества

$$T|_{y=\zeta(t,x)} = T_{\text{melt}} \quad (2)$$

и выполняется закон сохранения энергии (условие Стефана):

$$(k^{\text{cr}} \nabla T \cdot \mathbf{N})|_{y=\zeta(t,x)-0} - (k^{\text{l}} \nabla T \cdot \mathbf{N})|_{y=\zeta(t,x)+0} = \lambda \rho v^{\text{ph}}(\mathbf{e}_y \cdot \mathbf{N}). \quad (3)$$

Здесь λ – скрытая теплота плавления, $v^{\text{ph}} = v^{\text{ph}}(t, x) = \partial_t \zeta$ – скорость движения границы раздела фаз, $\mathbf{N}$ – единичная нормаль к межфазной границе, направленная в жидкую фазу, $\mathbf{e}_y$ – единичный вектор оси y .

На границе области Ω задана температура

$$T|_{\partial\Omega} = T_{\text{b}}. \quad (4)$$

2. Метод выпрямления фронта.

2.1. Замена переменных. Для решения задачи с внутренней подвижной границей, положение которой необходимо определять в ходе решения задачи, применим метод выпрямления фронта [18]. Основу этого метода составляет динамическая замена переменных специального вида, при которой физическая область $\Omega(t, x, y)$ отображается в расчётную область $\Omega(\tilde{t}, \xi, \eta)$ так, что в новой системе координат положение границы раздела фаз фиксировано и совпадает с координатной линией $\eta = \text{const}$.

В данной работе используется замена переменных, при которой области Ω^m , $m = 1, 2$, в новой системе координат становятся прямоугольниками (рис. 1), а границы областей $y = Y_0 = 0$, $y = Y_1 = \zeta(t, x)$ и $y = Y_2 = H_2$ переходят в прямые $\eta = 0$, $\eta = 1$ и $\eta = 2$ соответственно. Связь между системами координат имеет вид

$$\tilde{t} = t, \quad \xi = x, \quad \eta = (y - Y_{m-1})/l^m + m - 1, \quad m = 1, 2, \quad (5)$$

где $l^m = l^m(t, \xi) = Y_m - Y_{m-1}$ – толщина зоны Ω^m .

Запишем задачу (1)–(4) в неподвижной системе координат [20]. Частные производные от температуры преобразуются следующим образом:

$$\frac{\partial T}{\partial t} = \frac{\partial T}{\partial \tilde{t}} + \frac{\partial \eta}{\partial t} \frac{\partial T}{\partial \eta}, \quad \frac{\partial T}{\partial x} = \frac{\partial T}{\partial \xi} + \frac{\partial \eta}{\partial x} \frac{\partial T}{\partial \eta}, \quad \frac{\partial T}{\partial y} = \frac{\partial \eta}{\partial y} \frac{\partial T}{\partial \eta}. \quad (6)$$

Для вычисления метрических коэффициентов $\partial \eta / \partial t$, $\partial \eta / \partial x$, $\partial \eta / \partial y$ воспользуемся обратным преобразованием координат

$$t = \tilde{t}, \quad x = \xi, \quad y = \varphi^m(\tilde{t}, \xi, \eta) = Y_{m-1} + l^m(\eta - m + 1), \quad m = 1, 2. \quad (7)$$

Якобиан преобразования (7): $J^m = \partial(t, x, y) / \partial(\tilde{t}, \xi, \eta) = l^m$, $m = 1, 2$. Нетрудно показать [21, 22], что справедливы равенства

$$\frac{\partial \eta}{\partial t} = -\frac{1}{J^m} \frac{\partial y}{\partial \tilde{t}} = -\frac{1}{l^m} \frac{\partial \varphi^m}{\partial \tilde{t}}, \quad \frac{\partial \eta}{\partial x} = -\frac{1}{J^m} \frac{\partial y}{\partial \xi} = -\frac{1}{l^m} \frac{\partial \varphi^m}{\partial \xi}, \quad \frac{\partial \eta}{\partial y} = \frac{1}{J^m} \frac{\partial x}{\partial \xi} = \frac{1}{l^m}. \quad (8)$$

Дифференциальные операторы в уравнениях (1)–(4) преобразуем с помощью соотношений (6), (8) и умножим затем их на J^m . В дальнейшем для краткости будем опускать волну над

$\tilde{t}$ и верхний индекс m у функций l^m , φ^m , $\varphi_t^m = \partial\varphi^m/\partial t$, $\varphi_\xi^m = \partial\varphi^m/\partial\xi$. Для производной по времени получим

$$\mathcal{T}^{(\xi,\eta)}(T) = \partial_t(lT) - \partial_\eta(\varphi_t T). \quad (9)$$

Для диссипативных членов имеем

$$\mathcal{L}^{(\xi,\eta)}(T) = \partial_\xi \mathcal{W}^\xi + \partial_\eta \mathcal{W}^\eta. \quad (10)$$

Поток $\mathcal{W}^{(\xi,\eta)} = (\mathcal{W}^\xi, \mathcal{W}^\eta)$ вычисляется по формуле

$$\mathcal{W}^\xi = L^{\xi\xi} \partial_\xi T + L^{\xi\eta} \partial_\eta T, \quad \mathcal{W}^\eta = L^{\eta\xi} \partial_\xi T + L^{\eta\eta} \partial_\eta T,$$

где метрические коэффициенты $L^{\xi\xi}$, $L^{\xi\eta}$, $L^{\eta\xi}$, $L^{\eta\eta}$ вычисляются следующим образом:

$$L^{\xi\xi} = \kappa l, \quad L^{\xi\eta} = L^{\eta\xi} = \kappa(-\varphi_\xi), \quad L^{\eta\eta} = \kappa(1 + \varphi_\xi^2)/l.$$

Важно отметить, что метрические коэффициенты зависят от неизвестного положения границы раздела фаз.

Таким образом, в расчётной области $\Omega(t, \xi, \eta)$ уравнение теплопроводности (1) имеет вид

$$\mathcal{T}^{(\xi,\eta)}(T) = \mathcal{L}^{(\xi,\eta)}(T). \quad (11)$$

2.2. Сетка и сеточные функции. В расчётной области $\Omega(t, \xi, \eta)$ введём прямоугольную сетку $\omega_h^{(\xi,\eta)} = \omega_h^\xi \times \omega_h^\eta$, где $\omega_h^\xi = \{\xi_i, i = \overline{0, M}, \xi_0 = 0, \xi_M = H_1\}$, $\omega_h^\eta = \{\eta_j, j = \overline{0, N}, \eta_0 = 0, \eta_{j^*} = 1, \eta_N = 2\}$, шаги сетки: $h_{i+1/2}^\xi = \xi_{i+1} - \xi_i$, $h_{j+1/2}^\eta = \eta_{j+1} - \eta_j$. Введём также полуцелые узлы: $\xi_{i+1/2} = (\xi_{i+1} + \xi_i)/2$, $\eta_{j+1/2} = (\eta_{j+1} + \eta_j)/2$. Расчётную область $\Omega(t, \xi, \eta)$ разобьём на прямоугольные ячейки $S_{ij}^{(\xi,\eta)} = [\xi_{i-1/2}, \xi_{i+1/2}] \times [\eta_{j-1/2}, \eta_{j+1/2}]$; длины граней ячейки $S_{ij}^{(\xi,\eta)}$ равны $\tilde{h}_i^\xi = 0.5(h_{i+1/2}^\xi + h_{i-1/2}^\xi)$ и $\tilde{h}_j^\eta = 0.5(h_{j+1/2}^\eta + h_{j-1/2}^\eta)$, её площадь $dS_{ij}^{(\xi,\eta)} = \tilde{h}_i^\xi \tilde{h}_j^\eta$. Также рассмотрим ячейки $S_{i+1/2, j+1/2}^{(\xi,\eta)}$ с центрами в точках $(\xi_{i+1/2}, \eta_{j+1/2})$ (рис. 2). Сетка по времени $\omega_\tau = \{t_0 = 0, t_{k+1} = t_k + \tau, k = 0, 1, \dots\}$, где τ – шаг по времени.

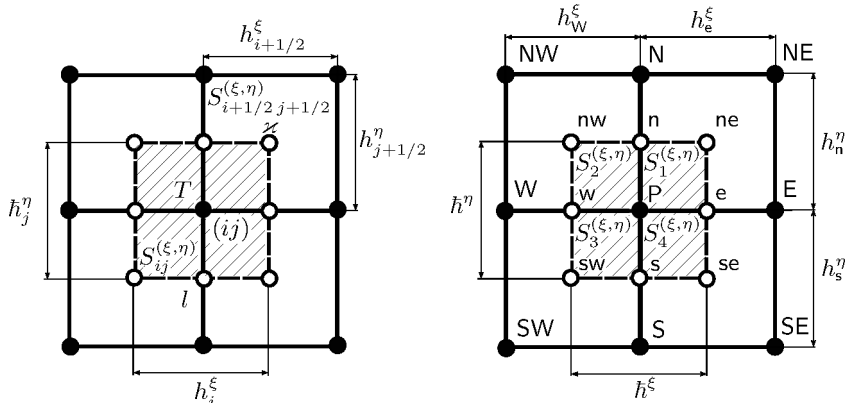


Рис. 2. Сетка и сеточные функции.

Сеточную функцию $T_{ij}^k = T(t_k, \xi_i, \eta_j)$ будем относить к узлам сетки $\omega_h^{(\xi,\eta)}$. Доопределим эту функцию внутри расчётных ячеек: $T(t_k, \xi, \eta) = T_{ij}^k$ при $(\xi, \eta) \in S_{ij}^{(\xi,\eta)}$. Длины фаз l и скорость движения границы раздела v^{ph} внутри ячеек доопределим линейным образом:

$$l(t_k, \xi, \eta_{j+1/2}) = l_{ij+1/2}^k + [l_{i+1, j+1/2}^k - l_{ij+1/2}^k](\xi - \xi_i)/h_{i+1/2}^\xi, \\ v^{\text{ph}}(t_k, \xi) = (v^{\text{ph}})_i^k + [(v^{\text{ph}})_{i+1}^k - (v^{\text{ph}})_i^k](\xi - \xi_i)/h_{i+1/2}^\xi \quad \text{при} \quad \xi \in [\xi_i, \xi_{i+1}]. \quad (12)$$

Температуропроводность $\varkappa$ и метрические коэффициенты $L^{\xi\xi}$, $L^{\xi\eta}$, $L^{\eta\eta}$ относятся к центрам ячеек $S_{i+1/2, j+1/2}^{(\xi, \eta)}$ (см. рис. 2).

Для обозначения узлов сетки также будем использовать локальную географическую нотацию [23]. Доопределим температуру в узлах $\mathbf{n}$ и $\mathbf{s}$ следующим образом: $T_{\mathbf{n}} = 0.5[T_{\mathbf{P}} + T_{\mathbf{N}}]$, $T_{\mathbf{s}} = 0.5[T_{\mathbf{P}} + T_{\mathbf{S}}]$. Введём сеточные операторы разностного дифференцирования. Для пространственных производных в направлении оси ξ имеем $T_{\mathbf{P}, \xi} = (T_{\mathbf{E}} - T_{\mathbf{P}})/h_{\mathbf{e}}^{\xi}$, $T_{\mathbf{P}, \bar{\xi}} = T_{\mathbf{W}, \xi}$; разностные аппроксимации пространственных производных в направлении оси η : $T_{\mathbf{P}, \eta} = (T_{\mathbf{N}} - T_{\mathbf{P}})/h_{\mathbf{n}}^{\eta}$, $T_{\mathbf{P}, \bar{\eta}} = T_{\mathbf{S}, \eta}$, $T_{\mathbf{P}, \tilde{\eta}} = (T_{\mathbf{n}} - T_{\mathbf{P}})/(0.5h_{\mathbf{n}}^{\eta})$, $T_{\mathbf{P}, \bar{\tilde{\eta}}} = T_{\mathbf{S}, \tilde{\eta}}$. Разностную производную по времени обозначим через $T_t = (\hat{T}_{\mathbf{P}} - T_{\mathbf{P}})/\tau$, где $\hat{T}_{\mathbf{P}} = T(t_k + \tau, \xi_{\mathbf{P}}, \eta_{\mathbf{P}})$. Контрольный объём $S_{\mathbf{P}}^{(\xi, \eta)} = \bigcup_{q=1}^4 S_q^{(\xi, \eta)}$.

2.3. Разностная схема. Построим консервативную разностную схему с помощью интегро-интерполяционного метода. Проинтегрируем уравнение (11) по ячейке $S_{\mathbf{P}}^{(\xi, \eta)}$:

$$\int_{t_k}^{t_{k+1}} \int_{S_{\mathbf{P}}^{(\xi, \eta)}} \mathcal{T}^{(\xi, \eta)}(T) d\xi d\eta dt = \int_{t_k}^{t_{k+1}} \int_{S_{\mathbf{P}}^{(\xi, \eta)}} \mathcal{L}^{(\xi, \eta)}(T) d\xi d\eta dt.$$

Подробное описание построения разностной схемы приведено в работе [20].

Аппроксимация диссипативных членов. Для диссипативных членов получаем

$$\mathcal{L}_{\mathbf{P}}^{(\xi, \eta)}(T) dS_{\mathbf{P}}^{(\xi, \eta)} = (\mathcal{W}_{\mathbf{n}}^{\eta} - \mathcal{W}_{\mathbf{s}}^{\eta}) \hbar^{\xi} + (\mathcal{W}_{\mathbf{e}}^{\xi} - \mathcal{W}_{\mathbf{w}}^{\xi}) \hbar^{\eta}.$$

Аппроксимация потоков $\mathcal{W}_{\mathbf{n}}^{\eta}$, $\mathcal{W}_{\mathbf{e}}^{\xi}$ имеет вид

$$\mathcal{W}_{\mathbf{n}}^{\eta} \hbar^{\xi} = 0.5[h_{\mathbf{e}}^{\xi} L_{\mathbf{ne}}^{\eta\eta} + h_{\mathbf{w}}^{\xi} L_{\mathbf{nw}}^{\eta\eta}] T_{\mathbf{P}, \eta} + 0.5[h_{\mathbf{e}}^{\xi} L_{\mathbf{ne}}^{\eta\xi} T_{\mathbf{n}, \xi} + h_{\mathbf{w}}^{\xi} L_{\mathbf{nw}}^{\eta\xi} T_{\mathbf{n}, \bar{\xi}}], \quad (13)$$

$$\mathcal{W}_{\mathbf{e}}^{\xi} \hbar^{\eta} = 0.5[h_{\mathbf{n}}^{\eta} L_{\mathbf{ne}}^{\xi\xi} + h_{\mathbf{s}}^{\eta} L_{\mathbf{se}}^{\xi\xi}] T_{\mathbf{P}, \xi} + 0.5[h_{\mathbf{n}}^{\eta} L_{\mathbf{ne}}^{\xi\eta} T_{\mathbf{e}, \eta} + h_{\mathbf{s}}^{\eta} L_{\mathbf{se}}^{\xi\eta} T_{\mathbf{e}, \bar{\eta}}], \quad (14)$$

здесь $T_{\mathbf{n}, \xi} = 0.5(T_{\mathbf{P}, \xi} + T_{\mathbf{N}, \xi})$, $T_{\mathbf{n}, \bar{\xi}} = 0.5(T_{\mathbf{P}, \bar{\xi}} + T_{\mathbf{N}, \bar{\xi}})$, $T_{\mathbf{e}, \eta} = 0.5(T_{\mathbf{P}, \eta} + T_{\mathbf{E}, \eta})$, $T_{\mathbf{e}, \bar{\eta}} = 0.5(T_{\mathbf{P}, \bar{\eta}} + T_{\mathbf{E}, \bar{\eta}})$. Метрические коэффициенты вычисляются следующим образом:

$$L_{\mathbf{ne}}^{\xi\xi} = \varkappa_{\mathbf{ne}} l_{\mathbf{ne}}, \quad L_{\mathbf{se}}^{\xi\xi} = \varkappa_{\mathbf{se}} l_{\mathbf{se}}, \quad L_{\mathbf{ne}}^{\eta\eta} = \varkappa_{\mathbf{ne}} [1 + (\varphi_{\mathbf{n}, \xi})^2] / l_{\mathbf{n}}, \quad L_{\mathbf{nw}}^{\eta\eta} = \varkappa_{\mathbf{nw}} [1 + (\varphi_{\mathbf{n}, \xi})^2] / l_{\mathbf{n}},$$

$$L_{\mathbf{ne}}^{\xi\eta} = L_{\mathbf{ne}}^{\eta\xi} = \varkappa_{\mathbf{ne}} (-\varphi_{\mathbf{n}, \xi}), \quad L_{\mathbf{se}}^{\xi\eta} = \varkappa_{\mathbf{se}} (-\varphi_{\mathbf{s}, \xi}), \quad L_{\mathbf{nw}}^{\eta\xi} = \varkappa_{\mathbf{nw}} (-\varphi_{\mathbf{n}, \bar{\xi}}).$$

Потоки $\mathcal{W}_{\mathbf{s}}^{\eta}$ и $\mathcal{W}_{\mathbf{w}}^{\xi}$ вычисляются аналогично.

На межфазной границе в разностной аппроксимации диссипативных членов возникает слабое, обеспечивающее выполнение условия Стефана (3). Это слабое имеет вид

$$\mathcal{E}_{\mathbf{P}^*}^{(\xi, \eta)}(T) dS_{\mathbf{P}^*}^{(\xi, \eta)} = \hbar^{\xi} \tilde{\lambda} (\bar{\mathbf{v}}^{\text{ph}})^{\eta}.$$

Здесь $\mathbf{P}^*$ – индекс узлов сетки, лежащих на границе раздела фаз, $\tilde{\lambda} = \lambda \rho / c_{\mathbf{p}}$,

$$(\bar{\mathbf{v}}^{\text{ph}})^{\eta} = (0.5[v_{\mathbf{P}}^{\text{ph}} + v_{\mathbf{w}}^{\text{ph}}] h_{\mathbf{w}}^{\xi} + 0.5[v_{\mathbf{P}}^{\text{ph}} + v_{\mathbf{e}}^{\text{ph}}] h_{\mathbf{e}}^{\xi}) / (2\hbar^{\xi}). \quad (15)$$

Значения $v_{\mathbf{w}}^{\text{ph}}$ и $v_{\mathbf{e}}^{\text{ph}}$ вычисляются по формуле (12).

С помощью несложных, но достаточно громоздких преобразований показывается, что предложенная аппроксимация диссипативных членов задаёт самосопряжённый отрицательно определённый разностный оператор.

Аппроксимация производной по времени. Построим пространственную аппроксимацию производной по времени. Для этого проинтегрируем выражение (9) по ячейке $S_{\mathbf{P}}^{(\xi, \eta)}$. Для

удобства интеграл по ячейке $S^{(\xi,\eta)}$ представим в виде суммы интегралов по ячейкам $S_q^{(\xi,\eta)}$, $q = \overline{1,4}$:

$$\int_{S_P^{(\xi,\eta)}} \mathcal{T}^{(\xi,\eta)}(T) d\xi d\eta = \sum_{q=1}^4 \left[\frac{\partial}{\partial t} \int_{S_q^{(\xi,\eta)}} lT d\xi d\eta - \int_{S_q^{(\xi,\eta)}} \frac{\partial(\varphi_t T)}{\partial \eta} d\xi d\eta \right]. \quad (16)$$

Интегрируя первое слагаемое в (16) по ячейке $S_1^{(\xi,\eta)}$, с учётом формулы (12) получаем

$$\int_{S_1^{(\xi,\eta)}} lT d\xi d\eta = [l_n + l_{ne}]T_P(h_e^\xi h_n^\eta)/8.$$

Интегралы по ячейкам $S_2^{(\xi,\eta)}$, $S_3^{(\xi,\eta)}$ и $S_4^{(\xi,\eta)}$ вычисляются аналогичным образом. Для второго слагаемого в (16) имеем

$$\begin{aligned} \int_{S_1^{(\xi,\eta)}} \frac{\partial(\varphi_t T)}{\partial \eta} d\xi d\eta + \int_{S_4^{(\xi,\eta)}} \frac{\partial(\varphi_t T)}{\partial \eta} d\xi d\eta &= 0.5[(\varphi_t)_n + (\varphi_t)_{ne}]T_n h_e^\xi/2 - 0.5[(\varphi_t)_s + (\varphi_t)_{se}]T_s h_e^\xi/2, \\ \int_{S_2^{(\xi,\eta)}} \frac{\partial(\varphi_t T)}{\partial \eta} d\xi d\eta + \int_{S_3^{(\xi,\eta)}} \frac{\partial(\varphi_t T)}{\partial \eta} d\xi d\eta &= 0.5[(\varphi_t)_n + (\varphi_t)_{nw}]T_n h_w^\xi/2 - 0.5[(\varphi_t)_s + (\varphi_t)_{sw}]T_s h_w^\xi/2. \end{aligned}$$

Таким образом, пространственная аппроксимация производной по времени принимает вид

$$\int_{S_P^{(\xi,\eta)}} \mathcal{T}^{(\xi,\eta)}(T) d\xi d\eta = \frac{\partial}{\partial t}(\bar{l}_P T_P) dS_P^{(\xi,\eta)} - [(\overline{\varphi_t})_n T_n - (\overline{\varphi_t})_s T_s] h^\xi, \quad (17)$$

где

$$\begin{aligned} \bar{l}_P &= [(0.5[l_n + l_{ne}]h_e^\xi + 0.5[l_n + l_{nw}]h_w^\xi)h_n^\eta + (0.5[l_s + l_{se}]h_e^\xi + 0.5[l_s + l_{sw}]h_w^\xi)h_s^\eta]/(4h^\xi h^\eta), \\ (\overline{\varphi_t})_n &= (0.5[(\varphi_t)_n + (\varphi_t)_{ne}]h_e^\xi + 0.5[(\varphi_t)_n + (\varphi_t)_{nw}]h_w^\xi)/(2h^\xi). \end{aligned}$$

Пространственные аппроксимации операторов (9) и (10) проинтегрируем по отрезку $[t_k, t_{k+1}]$, получившиеся выражения разделим на шаг по времени τ . В результате получим следующую аппроксимацию по времени оператора (17):

$$\mathcal{T}_P^{(\xi,\eta)}(T) dS_P^{(\xi,\eta)} = (\bar{l}_P T_P)_t dS_P^{(\xi,\eta)} - [(\overline{\varphi_t})_n \hat{T}_n - (\overline{\varphi_t})_s \hat{T}_s] h^\xi. \quad (18)$$

Во внутренних точках сетки разностная схема для задачи (1)–(4), полученная с помощью метода выпрямления фронта, имеет вид

$$\mathcal{T}_P^{(\xi,\eta)}(T) dS_P^{(\xi,\eta)} = [\mathcal{L}_P^{(\xi,\eta)}(\hat{T}) + \mathcal{E}_{P^*}^{(\xi,\eta)}(\hat{T})\delta_{PP^*}] dS_P^{(\xi,\eta)}. \quad (19)$$

Здесь $\delta_{PP^*} = 1$, если $P = P^*$, и $\delta_{PP^*} = 0$ в противном случае. Разностная схема (19) является консервативной, метрические коэффициенты $L^{\xi\xi}$, $L^{\xi\eta}$, $L^{\eta\eta}$ в ней взяты с верхнего временного слоя, соответствующая система нелинейных сеточных уравнений решается относительно вектора неизвестных, компонентами которого являются температуры в твёрдой и жидкой фазах и скорость межфазной границы.

Несложно показать (см. [19]), что в регулярной точке выражение (18) преобразуется к виду

$$\mathcal{T}_P^{(\xi,\eta)}(T) dS_P^{(\xi,\eta)} = \bar{l}_P T_{P,t} dS_P^{(\xi,\eta)} - 0.5[h_n^\eta (\overline{\varphi_t})_n \hat{T}_{\bar{\eta}} + h_s^\eta (\overline{\varphi_t})_s \hat{T}_{\bar{\eta}}] h^\xi. \quad (20)$$

Заменив в системе уравнений (19) её левую часть представлением (20), получим разностную схему, аппроксимирующую недивергентную форму записи задачи (11) в расчётной области.

3. Метод расчёта на подвижной сетке.

3.1. Сетка и сеточные функции. В физической области $\Omega(t, x, y)$ введём подвижную сетку $\omega_h^{(x,y)}(t) = \{(x_i, y_j(t, x_i)), i = \overline{0, M}, j = \overline{0, N}\}$ таким образом, чтобы $y_0(t, x_i) = 0$, $y_N(t, x_i) = H_2$, $y_j^*(t, x_i) = \zeta(t, x_i)$ для любого x_i . Сетка по времени $\omega_\tau = \{t_0 = 0, t_{k+1} = t_k + \tau, k = 0, 1, \dots\}$, где τ – шаг по времени.

Утверждение 1. Если в начальный момент времени координаты узлов сеток $\omega_h^{(x,y)}(0)$ и $\omega_h^{(\xi,\eta)}$ связаны соотношениями

$$x_i = \xi_i, \quad y_j(0, x_i) = \varphi(0, \xi_i, \eta_j), \quad i = \overline{0, M}, \quad j = \overline{0, N}, \quad (21)$$

и движение узлов сетки $\omega_h^{(x,y)}(t)$ определяется выражением

$$y_j(t_{k+1}, x_i) = y_j(t_k, x_i) + \tau v_j^y(x_i), \quad v_j^y(x_i) = \begin{cases} y_j v_i^{\text{ph}} / \zeta_i, & 0 < j \leq j^*, \\ (H_2 - y_j) v_i^{\text{ph}} / (H_2 - \zeta_i), & j^* < j < N, \end{cases} \quad (22)$$

то координаты узлов сеток $\omega_h^{(x,y)}(t)$ и $\omega_h^{(\xi,\eta)}$ связаны соотношениями

$$x_i = \xi_i, \quad y_j(t_k, x_i) = \varphi(t_k, \xi_i, \eta_j), \quad i = \overline{0, M}, \quad j = \overline{0, N}, \quad (23)$$

для любого $k > 0$.

Доказательство. Для определённости рассмотрим узлы сетки, расположенные в твёрдой фазе. Доказательство утверждения проведём методом математической индукции. Из формулировки утверждения следует, что в начальный момент времени (т.е. при $k = 0$) равенство выполнено. Пусть это равенство выполнено на k -м временном слое, т.е. $y_j(t_k, x_i) = \varphi(t_k, \xi_i, \eta_j) = \eta_j l(t_k, \xi_i)$. Покажем, что оно выполняется и в момент времени $t_{k+1} = t_k + \tau$. Из равенств (5) и (22) следует, что $v_j^y(x_i) = \eta_j v_i^{\text{ph}}$, поэтому в момент времени t_{k+1} имеем

$$y_j(t_{k+1}, x_i) = y_j(t_k, x_i) + \tau v_j^y(x_i) = [l(t_k, \xi_i) + \tau v_i^{\text{ph}}] \eta_j = l(t_{k+1}, \xi_i) \eta_j = \varphi(t_{k+1}, \xi_i, \eta_j).$$

Для узлов сетки, лежащих в жидкой фазе, утверждение доказывается аналогично. Утверждение доказано.

В дальнейшем в физической области будем использовать сетку $\omega_h^{(x,y)}(t)$, удовлетворяющую условиям (21), (22). Полуцелые узлы подвижной сетки определим равенствами

$$x_{i+1/2} = \xi_{i+1/2}, \quad y_j(t, x_{i+1/2}) = \varphi(t, \xi_{i+1/2}, \eta_j), \quad y_{j+1/2}(t, x_{i+1/2}) = \varphi(t, \xi_{i+1/2}, \eta_{j+1/2}).$$

Шаги подвижной сетки

$$h_{i+1/2}^x = x_{i+1} - x_i, \quad h_{ij+1/2}^y(t_k) = y_{j+1}(t_k, x_i) - y_j(t_k, x_i),$$

$$h_{i+1/2, j+1/2}^y(t_k) = y_{j+1}(t_k, x_{i+1/2}) - y_j(t_k, x_{i+1/2}), \quad \bar{h}_i^x = x_{i+1/2} - x_{i-1/2},$$

$$\bar{h}_{i+1/2, j}^y(t_k) = y_{j+1/2}(t_k, x_{i+1/2}) - y_{j-1/2}(t_k, x_{i+1/2}).$$

В общем случае $h_{ij+1/2}^y \neq h_{i+1/2, j+1/2}^y \neq h_{i+1, j+1/2}^y$.

Сетка $\omega_h^{(x,y)}(t)$ разбивает физическую область на четырёхугольные ячейки $S_{i+1/2, j+1/2}^{(x,y)}(t)$ с вершинами в точках $(x_{i+\alpha}, y_{j+\beta}(t, x_{i+\alpha}))$, $\alpha, \beta = 0, 1$. Соединим середины противоположных границ данных ячеек отрезками прямых. Полученные отрезки разбивают область на шестиугольные ячейки $S_{ij}^{(x,y)}(t)$, $i = \overline{0, M}$, $j = \overline{0, N}$ (заштрихованная область на рис. 3). Из соотношения (23) следует, что $S_{ij}^{(x,y)}(t_k) = \{(x = \xi, y = \varphi(t_k, \xi, \eta)) : (\xi, \eta) \in S_{ij}^{(\xi,\eta)}\}$. Поэтому температуру внутри расчётной ячейки $S_{ij}^{(x,y)}(t)$ доопределим так же, как на фиксированной сетке $\omega_h^{(\xi,\eta)}$.

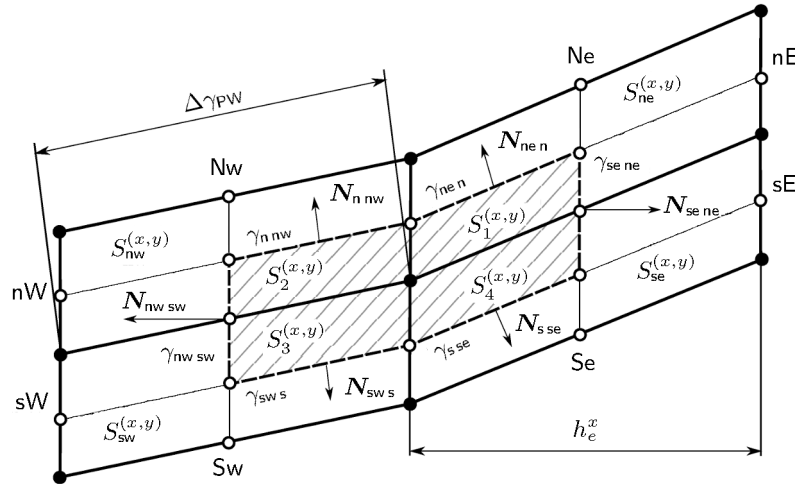


Рис. 3. Подвижная сетка.

Далее для обозначения узлов сетки будем использовать локальную географическую нотацию. Обозначим $S_p^{(x,y)} = \bigcup_{q=1}^4 S_q^{(x,y)}$ (рис. 3) и $dS_q^{(x,y)}$ – площадь ячейки $S_q^{(x,y)}$. Границу ячейки $S_p^{(x,y)}$ будем считать положительно ориентированной и введём упорядоченное в соответствии с ориентацией границы множество вершин $V = \{ne, n, nw, sw, s, se\}$. Пусть $\gamma_{\alpha\beta}$ – участок границы с началом в точке α и концом в точке β ($\alpha, \beta \in V$); $\Delta\gamma_{\alpha\beta}$ – длина отрезка $\gamma_{\alpha\beta}$, $\Delta\gamma_{\alpha\beta} = \sqrt{\delta x_{\alpha\beta}^2 + \delta y_{\alpha\beta}^2}$, где $\delta x_{\alpha\beta} = x_\beta - x_\alpha$, $\delta y_{\alpha\beta} = y_\beta - y_\alpha$; $N_{\alpha\beta}$ – вектор внешней единичной нормали к границе $\gamma_{\alpha\beta}$ с координатами $(N_{\alpha\beta}^x, N_{\alpha\beta}^y)$; $\mu_{\alpha\beta}$ – единичный направляющий вектор границы, $\mu_{\alpha\beta} = (\mu_{\alpha\beta}^x, \mu_{\alpha\beta}^y)$. В нашем случае

$$\mu_{\alpha\beta}^x = -N_{\alpha\beta}^y, \quad \mu_{\alpha\beta}^y = N_{\alpha\beta}^x, \quad (24)$$

где $N_{\alpha\beta}^x = \delta y_{\alpha\beta} / \Delta\gamma_{\alpha\beta}$, $N_{\alpha\beta}^y = -\delta x_{\alpha\beta} / \Delta\gamma_{\alpha\beta}$. Производная вдоль участка границы $\gamma_{\alpha\beta}$ вычисляется по формуле $T_{\mu_{\alpha\beta}} = (T_\beta - T_\alpha) / \Delta\gamma_{\alpha\beta}$.

Наряду с множеством V введём упорядоченное множество $\tilde{V} = \{nE, n, nW, sW, s, sE\}$, образованное серединами вертикальных границ ячеек основной сетки $S_\nu^{(x,y)}$, $\nu \in \{ne, nw, sw, se\}$. Доопределим функцию T в серединах границ ячейки основной сетки $S_\nu^{(x,y)}$ как полусумму её значений на концах соответствующих границ, т.е. $T_n = 0.5[T_P + T_N]$, $T_{nE} = 0.5[T_E + T_{NE}]$, $T_e = 0.5[T_P + T_E]$, $T_{Ne} = 0.5[T_N + T_{NE}]$ и т.д.; значение температуры в центре ячейки основной сетки $S_\nu^{(x,y)}$ равно среднему значению, вычисленному по значениям T в вершинах данной ячейки, т.е. $T_{ne} = 0.25[T_P + T_E + T_N + T_{NE}]$ и т.д. При такой переинтерполяции на “северных” и “южных” границах ячейки $S_p^{(x,y)}$ для сонаправленных векторов $\mu_{\alpha\beta}$ и $\mu_{\tilde{\alpha}\tilde{\beta}}$, $\alpha, \beta \in V$, $\tilde{\alpha}, \tilde{\beta} \in \tilde{V}$, имеем

$$T_{\mu_{\alpha\beta}} = T_{\mu_{\tilde{\alpha}\tilde{\beta}}} = (T_{\tilde{\beta}} - T_{\tilde{\alpha}}) / (2\Delta\gamma_{\alpha\beta}), \quad (25)$$

например, $T_{\mu_{ne n}} = (T_n - T_{nE}) / (2\Delta\gamma_{ne n}) = ([T_P + T_N] - [T_{NE} + T_E]) / (4\Delta\gamma_{ne n})$. Производную $\partial_y T$ в центрах ячеек основной сетки $S_\nu^{(x,y)}$ аппроксимируем следующим образом:

$$[\partial_y T]_{ne}^h = T_{e,y} = \frac{T_{Ne} - T_e}{h_{ne}^y}, \quad [\partial_y T]_{se}^h = T_{e,\bar{y}} = \frac{T_e - T_{Se}}{h_{se}^y}, \quad [\partial_y T]_{nw}^h = T_{w,y}, \quad [\partial_y T]_{sw}^h = T_{w,\bar{y}}. \quad (26)$$

Остальные разностные аппроксимации пространственных производных на подвижной сетке вычисляются стандартным образом.

3.2. Разностная схема. Разностную схему будем строить с помощью интегро-интерполяционного метода. Проинтегрируем уравнения (1)–(4) по контрольному объёму $S_P^{(x,y)}(t)$:

$$\int_{t_k}^{t_{k+1}} \int_{S_P^{(x,y)}(t)} \frac{\partial T}{\partial t} dx dy dt = \int_{t_k}^{t_{k+1}} \int_{S_P^{(x,y)}(t)} \nabla \cdot (\varkappa \nabla T) dx dy dt.$$

Аппроксимация диссипативных членов. Построим пространственную аппроксимацию диссипативных членов

$$\mathcal{L}_P^{(x,y)}(T) dS_P^{(x,y)} = \int_{S_P^{(x,y)}(t)} \left[\frac{\partial}{\partial x} \mathcal{W}^x + \frac{\partial}{\partial y} \mathcal{W}^y \right] dx dy.$$

Здесь $\mathcal{W}^x = \varkappa \partial_x T$, $\mathcal{W}^y = \varkappa \partial_y T$. В силу формулы Остроградского–Гаусса имеем

$$\int_{S_P^{(x,y)}(t)} \left[\frac{\partial}{\partial x} \mathcal{W}^x + \frac{\partial}{\partial y} \mathcal{W}^y \right] dx dy = \sum_{\alpha, \beta \in V_{\gamma_{\alpha\beta}}} \int [\mathcal{W}^x N_{\alpha\beta}^x + \mathcal{W}^y N_{\alpha\beta}^y] d\gamma.$$

Вычислим интегралы вдоль “горизонтальных” границ $\gamma_{ne n}$, $\gamma_{n nw}$, $\gamma_{sw s}$, $\gamma_{s se}$. Воспользовавшись тем, что

$$\frac{\partial T}{\partial x} = \frac{1}{\mu_{\alpha\beta}^x} \left(\frac{\partial T}{\partial \mu_{\alpha\beta}} - \mu_{\alpha\beta}^y \frac{\partial T}{\partial y} \right),$$

запишем интеграл вдоль границы следующим образом:

$$\int_{\gamma_{\alpha\beta}} \left[\varkappa \frac{\partial T}{\partial x} N_{\alpha\beta}^x + \varkappa \frac{\partial T}{\partial y} N_{\alpha\beta}^y \right] d\gamma = \int_{\gamma_{\alpha\beta}} \frac{N_{\alpha\beta}^x}{\mu_{\alpha\beta}^x} \varkappa \frac{\partial T}{\partial \mu_{\alpha\beta}} d\gamma + \int_{\gamma_{\alpha\beta}} \left[N_{\alpha\beta}^y - \frac{\mu_{\alpha\beta}^y}{\mu_{\alpha\beta}^x} N_{\alpha\beta}^x \right] \varkappa \frac{\partial T}{\partial y} d\gamma. \quad (27)$$

Для интегралов в правой части равенства (27) с учётом соотношений (24) получаем

$$\int_{\gamma_{\alpha\beta}} \frac{N_{\alpha\beta}^x}{\mu_{\alpha\beta}^x} \varkappa \frac{\partial T}{\partial \mu_{\alpha\beta}} d\gamma = \int_{\gamma_{\alpha\beta}} -\frac{N_{\alpha\beta}^x}{N_{\alpha\beta}^y} \varkappa \frac{\partial T}{\partial \mu_{\alpha\beta}} d\gamma \approx \Delta \gamma_{\alpha\beta} \left(\frac{\delta y_{\alpha\beta}}{\delta x_{\alpha\beta}} [\varkappa]^h T_{\mu_{\alpha\beta}} \right), \quad (28)$$

$$\begin{aligned} \int_{\gamma_{\alpha\beta}} \left[N_{\alpha\beta}^y - \frac{\mu_{\alpha\beta}^y}{\mu_{\alpha\beta}^x} N_{\alpha\beta}^x \right] \varkappa \frac{\partial T}{\partial y} d\gamma &= \int_{\gamma_{\alpha\beta}} \frac{(N_{\alpha\beta}^x)^2 + (N_{\alpha\beta}^y)^2}{N_{\alpha\beta}^y} \varkappa \frac{\partial T}{\partial y} d\gamma = \int_{\gamma_{\alpha\beta}} \frac{1}{N_{\alpha\beta}^y} \varkappa \frac{\partial T}{\partial y} d\gamma \approx \\ &\approx \Delta \gamma_{\alpha\beta} \left(-\frac{\Delta \gamma_{\alpha\beta}}{\delta x_{\alpha\beta}} [\varkappa]^h [\partial_y T]_P^h \right) = -\delta x_{\alpha\beta} [1 + (\delta y_{\alpha\beta} / \delta x_{\alpha\beta})^2] [\varkappa]^h [\partial_y T]_P^h. \end{aligned} \quad (29)$$

Здесь $[\varkappa]^h = \varkappa_{ne}$ для $\gamma_{ne n}$, $[\varkappa]^h = \varkappa_{nw}$ для $\gamma_{n nw}$, $[\varkappa]^h = \varkappa_{se}$ для γ_{ses} и $[\varkappa]^h = \varkappa_{sw}$ для $\gamma_{sw s}$; $[\partial_y T]_P^h$ – разностная аппроксимация оператора $\partial_y T$, $[\partial_y T]_P^h = T_{P,y}$ для “северных” границ $\gamma_{ne n}$ и $\gamma_{n nw}$, $[\partial_y T]_P^h = T_{P,\bar{y}}$ для “южных” границ γ_{ses} и $\gamma_{sw s}$.

Преобразуем теперь правые части в (28) и (29) к виду, удобному для сравнения с аппроксимацией теплового потока (13) через образ отрезка $\gamma_{\alpha\beta}$ в расчётной системе координат (t, ξ, η) . На границе $\gamma_{ne n}$ для ориентированной разности имеем $\delta x_{ne n} = -h_e/2$; тогда $\delta y_{ne n} / \delta x_{ne n} = y_{n,x} = (y_{ne} - y_n) / h_e^x$ и, учитывая равенство (25), получаем $T_{\mu_{ne n}} = -T_{n,\mu}$, где

$$T_{n,\mu} = (T_{nE} - T_n) / (2\Delta \gamma_{ne n}) = ([T_E + T_{NE}] - [T_P + T_N]) / (4\Delta \gamma_{ne n}).$$

На $\gamma_{n\text{nw}}$ справедливы равенства

$$\delta x_{n\text{nw}} = -h_w/2, \quad \delta y_{n\text{nw}}/\delta x_{n\text{nw}} = y_{n,\bar{x}} = (y_n - y_{n\text{w}})/h_w^x \quad \text{и} \quad T_{\mu_{n\text{nw}}} = -T_{n,\bar{\mu}},$$

где $T_{n,\bar{\mu}} = [T_n - T_{n\text{w}}]/(2\Delta\gamma_{ne})$. Таким образом, поток $\mathcal{I}_n$ через “северную” границу представляется в виде

$$\mathcal{I}_n = 0.5[h_e^x M_{ne}^{yy} + h_w^x M_{nw}^{yy}]T_{P,y} + \Delta\gamma_{ne} M_{ne}^{y\mu} T_{n,\mu} + \Delta\gamma_{n\text{nw}} M_{nw}^{y\mu} T_{n,\bar{\mu}}, \quad (30)$$

где коэффициенты M^{yy} и $M^{y\mu}$ вычисляются по формулам

$$M_{ne}^{yy} = \varkappa_{ne}[1 + (y_{n,x})^2], \quad M_{nw}^{yy} = \varkappa_{nw}[1 + (y_{n,\bar{x}})^2], \quad M_{ne}^{y\mu} = \varkappa_{ne}(-y_{n,x}), \quad M_{nw}^{y\mu} = \varkappa_{nw}(-y_{n,\bar{x}}).$$

На южных границах γ_{sws} и γ_{sse} выполняются равенства:

$$\delta x_{sws} = h_w/2, \quad \delta y_{sws}/\delta x_{sws} = y_{s,\bar{x}}, \quad T_{\mu_{sws}} = T_{s,\bar{\mu}} = [T_s - T_{sw}]/(2\Delta\gamma_{sws});$$

$$\delta x_{sse} = h_e/2, \quad \delta y_{sse}/\delta x_{sse} = y_{s,x}, \quad T_{\mu_{sse}} = T_{s,\mu} = [T_{se} - T_s]/(2\Delta\gamma_{sse}).$$

Поэтому поток $\mathcal{I}_s$ через “южную” границу представляется в виде

$$\mathcal{I}_s = -0.5[h_e^x M_{se}^{yy} + h_w^x M_{sw}^{yy}]T_{P,\bar{y}} - \Delta\gamma_{sse} M_{se}^{y\mu} T_{s,\mu} - \Delta\gamma_{sws} M_{sw}^{y\mu} T_{s,\bar{\mu}}, \quad (31)$$

где коэффициенты M^{yy} и $M^{y\mu}$ вычисляются по формулам

$$M_{se}^{yy} = \varkappa_{se}[1 + (y_{s,x})^2], \quad M_{sw}^{yy} = \varkappa_{sw}[1 + (y_{s,\bar{x}})^2], \quad M_{se}^{y\mu} = \varkappa_{se}(-y_{s,x}), \quad M_{sw}^{y\mu} = \varkappa_{sw}(-y_{s,\bar{x}}).$$

Запишем интегралы вдоль вертикальных границ γ_{nwsw} и γ_{sene} . В этом случае $N_{\alpha\beta}^x = \delta y_{\alpha\beta}/|\delta y_{\alpha\beta}|$, $N_{\alpha\beta}^y = 0$ и

$$\int_{\gamma_{\alpha\beta}} \left[\varkappa \frac{\partial T}{\partial x} N_{\alpha\beta}^x + \varkappa \frac{\partial T}{\partial y} N_{\alpha\beta}^y \right] d\gamma = \int_{\gamma_{\alpha\beta}} \frac{\delta y_{\alpha\beta}}{|\delta y_{\alpha\beta}|} \varkappa \frac{\partial T}{\partial x} dy.$$

Как и выше, производную $\partial_x T$ выразим через производную по некоторому направлению. В качестве такого направления выберем $\mu = \mu_{PE}$ на “восточной” границе γ_{sene} и $\mu = \mu_{PW}$ на “западной” границе γ_{nwsw} . Для “восточной” границы γ_{sene} получим

$$\int_{\gamma_{sene}} \frac{\delta y_{sene}}{|\delta y_{sene}|} \varkappa \frac{\partial T}{\partial x} dy = \int_{\gamma_{sene}} \frac{1}{\mu_{PE}^x} \varkappa \frac{\partial T}{\partial \mu_{PE}} dy + \int_{\gamma_{sene}} \left(-\frac{\mu_{PE}^y}{\mu_{PE}^x} \right) \varkappa \frac{\partial T}{\partial y} dy. \quad (32)$$

Первое слагаемое в правой части равенства (32) аппроксимируем следующим образом:

$$\begin{aligned} \int_{\gamma_{sene}} \frac{1}{\mu_{PE}^x} \varkappa \frac{\partial T}{\partial \mu_{PE}} dy &= \int_{\gamma_{sene}} \frac{\Delta\gamma_{PE}}{\delta x_{PE}} \varkappa \frac{\partial T}{\partial \mu_{PE}} dy \approx \frac{1}{2} [h_{ne}^y \varkappa_{ne} + h_{se}^y \varkappa_{se}] \frac{\Delta\gamma_{PE}}{\delta x_{PE}} T_{\mu_{PE}} = \\ &= 0.5 [h_{ne}^y \varkappa_{ne} + h_{se}^y \varkappa_{se}] (|\delta x_{PE}|/\delta x_{PE}) \sqrt{1 + (\delta y_{PE}/\delta x_{PE})^2} T_{\mu_{PE}}. \end{aligned} \quad (33)$$

Для аппроксимации второго слагаемого в правой части равенства (32) воспользуемся следующей квадратурной формулой:

$$\begin{aligned} \int_{\gamma_{sene}} \left(-\frac{\mu_{PE}^y}{\mu_{PE}^x} \right) \varkappa \frac{\partial T}{\partial y} dy &\approx \frac{1}{2} \left[h_{ne}^y \left(-\frac{\mu_{ne}^y}{\mu_{ne}^x} \right) \varkappa_{ne} \frac{\partial T}{\partial y} \Big|_{ne} + h_{se}^y \left(-\frac{\mu_{se}^y}{\mu_{se}^x} \right) \varkappa_{se} \frac{\partial T}{\partial y} \Big|_{se} \right] = \\ &= 0.5 [h_{ne}^y (-\delta y_{ne}/\delta x_{ne}) \varkappa_{ne} [\partial_y T]_{ne}^h + h_{se}^y (-\delta y_{se}/\delta x_{se}) \varkappa_{se} [\partial_y T]_{se}^h]. \end{aligned} \quad (34)$$

Интеграл вдоль “западной” границы $\gamma_{nw\ sw}$ вычисляется аналогично:

$$\int_{\gamma_{nw\ sw}} \frac{\delta y_{nw\ sw}}{|\delta y_{nw\ sw}|} \kappa \frac{\partial T}{\partial x} dy = -0.5[h_{nw}^y \kappa_{nw} + h_{sw}^y \kappa_{sw}] (|\delta x_{pw}|/\delta x_{pw}) \sqrt{1 + (\delta y_{pw}/\delta x_{pw})^2} T_{\mu_{pw}} - \\ - 0.5[h_{nw}^y (-\delta y_{n\ nw}/\delta x_{n\ nw}) \kappa_{nw} [\partial_y T]_{nw}^h + h_{sw}^y (-\delta y_{s\ sw}/\delta x_{s\ sw}) \kappa_{sw} [\partial_y T]_{sw}^h]. \quad (35)$$

Преобразуем правые части равенств (33), (34) и (35) к виду, удобному для сравнения с аппроксимацией теплового потока (14) в расчётной области. На границе $\gamma_{se\ ne}$ имеем: $\delta x_{pe} = h_e^x$, $\delta y_{pe}/\delta x_{pe} = y_{p,x} = [y_e - y_p]/h_e^x$, $T_{\mu_{pe}} = T_{p,\mu} = [T_e - T_p]/\Delta\gamma_{pe}$ и, учитывая (26), $[\partial_y T]_{ne}^h = T_{e,y}$, $[\partial_y T]_{se}^h = T_{e,\bar{y}}$. Поэтому поток $\mathcal{I}_e$ через “восточную” границу ячейки можно представить в виде

$$\mathcal{I}_e = 0.5[h_{ne}^y M_{ne}^{\mu\mu} + h_{se}^y M_{se}^{\mu\mu}] T_{p,\mu} + 0.5[h_{ne}^y M_{ne}^{\mu y} T_{e,y} + h_{se}^y M_{se}^{\mu y} T_{e,\bar{y}}], \quad (36)$$

где

$$M_{ne}^{\mu\mu} = \kappa_{ne} \sqrt{1 + (y_{p,x})^2}, \quad M_{se}^{\mu\mu} = \kappa_{se} \sqrt{1 + (y_{p,x})^2}, \quad \mathcal{M}_{ne}^{\mu y} = \mathcal{M}_{ne}^{y\mu}, \quad \mathcal{M}_{se}^{\mu y} = \mathcal{M}_{se}^{y\mu}.$$

На границе $\gamma_{nw\ sw}$ имеем: $\delta x_{pw} = -h_w^x$, $\delta y_{pw}/\delta x_{pw} = y_{p,\bar{x}} = [y_p - y_w]/h_w^x$, $T_{\mu_{pw}} = -T_{p,\bar{\mu}}$, где $T_{p,\bar{\mu}} = [T_p - T_w]/\Delta\gamma_{pw}$ и $[\partial_y T]_{nw}^h = T_{w,y}$, $[\partial_y T]_{sw}^h = T_{w,\bar{y}}$. Таким образом, поток $\mathcal{I}_w$ через “западную” границу можно представить в виде

$$\mathcal{I}_w = -0.5[h_{nw}^y M_{nw}^{\mu\mu} + h_{sw}^y M_{sw}^{\mu\mu}] T_{p,\bar{\mu}} - 0.5[h_{nw}^y M_{nw}^{\mu y} T_{w,y} + h_{sw}^y M_{sw}^{\mu y} T_{w,\bar{y}}], \quad (37)$$

где

$$M_{nw}^{\mu\mu} = \kappa_{nw} \sqrt{1 + (y_{p,\bar{x}})^2}, \quad M_{sw}^{\mu\mu} = \kappa_{sw} \sqrt{1 + (y_{p,\bar{x}})^2}, \quad \mathcal{M}_{nw}^{\mu y} = \mathcal{M}_{nw}^{y\mu}, \quad \mathcal{M}_{sw}^{\mu y} = \mathcal{M}_{sw}^{y\mu}.$$

Аппроксимация диссипативных членов на подвижной сетке запишется следующим образом:

$$\mathcal{L}_p^{(x,y)}(T) dS_p^{(x,y)} = \mathcal{I}_n + \mathcal{I}_s + \mathcal{I}_e + \mathcal{I}_w. \quad (38)$$

На фронте кристаллизации в аппроксимации диссипативных членов возникает дополнительное слагаемое. Из условия Стефана (3) следует, что для скачка теплового потока справедливо равенство

$$\mathcal{E}_{p*}^{(x,y)}(T) dS_{p*}^{(x,y)} = \int_{\gamma_{wp}} \tilde{\lambda} v^{\text{ph}} N_{wp}^y d\gamma + \int_{\gamma_{pe}} \tilde{\lambda} v^{\text{ph}} N_{pe}^y d\gamma \approx \hbar^x \tilde{\lambda} (\bar{v}^{\text{ph}})^y, \quad (39)$$

здесь

$$(\bar{v}^{\text{ph}})^y = (0.5[v_p^{\text{ph}} + v_w^{\text{ph}}]h_w^x + 0.5[v_p^{\text{ph}} + v_e^{\text{ph}}]h_e^x)/(2\hbar_p^x). \quad (40)$$

Аппроксимация производной по времени. Для построения аппроксимации производной по времени воспользуемся транспортной теоремой Рейнольдса:

$$\frac{d}{dt} \int_{S_p^{(x,y)}(t)} T dx dy = \int_{S_p^{(x,y)}(t)} \frac{\partial T}{\partial t} dx dy + \int_{\partial S_p^{(x,y)}(t)} (Tv \cdot N) d\gamma, \quad (41)$$

где v – скорость движения границы ячейки $S_p^{(x,y)}(t)$. Проинтегрировав соотношение (41) по отрезку $[t_k, t_{k+1}]$, получим

$$\int_{t_k}^{t_{k+1}} \int_{S_p^{(x,y)}(t)} \frac{\partial T}{\partial t} dx dy dt = \int_{t_k}^{t_{k+1}} \frac{d}{dt} \int_{S_p^{(x,y)}(t)} T dx dy dt - \int_{t_k}^{t_{k+1}} \int_{\partial S_p^{(x,y)}(t)} (Tv \cdot N) d\gamma dt. \quad (42)$$

Рассмотрим слагаемые, стоящие в правой части равенства (42). Первое слагаемое аппроксимируем следующим образом:

$$\int_{t_k}^{t_{k+1}} \frac{d}{dt} \int_{S_P^{(x,y)}(t)} T dx dy dt = \widehat{T}_P \sum_{q=1}^4 \widehat{dS}_q^{(x,y)} - T_P \sum_q dS_q^{(x,y)}. \quad (43)$$

Площади ячеек $dS_q^{(x,y)}$ и $\widehat{dS}_q^{(x,y)}$ будем вычислять по формуле трапеций, используя значения координат вершин этих ячеек на слое t_k и t_{k+1} соответственно. Для $dS_q^{(x,y)}$, $q = \overline{1,4}$, имеем

$$dS_1^{(x,y)} = \frac{(y_n - y_P) + (y_{ne} - y_e)}{2} \frac{h_e^x}{2} = \frac{[h_n^y + h_{ne}^y] h_e^x}{8},$$

$$dS_2^{(x,y)} = \frac{[h_n^y + h_{nw}^y] h_w^x}{8}, \quad dS_3^{(x,y)} = \frac{[h_s^y + h_{sw}^y] h_w^x}{8}, \quad dS_4^{(x,y)} = \frac{[h_s^y + h_{se}^y] h_e^x}{8}.$$

Расчётная ячейка $S_P^{(x,y)}(t)$ перемещается только в направлении оси y , т.е. $v = (0, v^y)$. Поэтому для второго слагаемого в правой части соотношения (42) справедливо равенство

$$\int_{t_k}^{t_{k+1}} \int_{\partial S_P^{(x,y)}(t)} (Tv \cdot N) d\gamma dt = \int_{t_k}^{t_{k+1}} \sum_{\alpha, \beta \in V_{\gamma_{\alpha\beta}}} \int T v^y N_{\alpha\beta}^y d\gamma dt. \quad (44)$$

Для интегралов вдоль “горизонтальных” границ получаем

$$\int_{t_k}^{t_{k+1}} \int_{\gamma_{\alpha\beta}} T v^y N_{\alpha\beta}^y d\gamma dt \approx \tau \left(\Delta \gamma_{\alpha\beta} [T]^h \frac{v_\alpha^y + v_\beta^y}{2} N_{\alpha\beta}^y \right) \Big|_{t=t_{k+1}} = \tau \left(-\delta x_{\alpha\beta} [\widehat{T}]^h \frac{v_\alpha^y + v_\beta^y}{2} \right). \quad (45)$$

Здесь $[\widehat{T}]^h = \widehat{T}_n$ для “северных” границ γ_{nen} , γ_{nnw} и $[\widehat{T}]^h = \widehat{T}_s$ для “южных” границ γ_{sws} , γ_{sse} . Скорости движения v_α^y , $\alpha \in V$, вершин контрольного объёма равны средним значениям, вычисленным по соответствующим значениям скоростей узлов сетки, например,

$$v_n^y = 0.5[v_P^y + v_N^y], \quad v_{ne} = 0.25[v_P^y + v_N^y + v_E^y + v_{NE}^y]$$

и т.д. Суммируя равенства (45) по “северным” и “южным” границам ячейки, приходим к следующей формуле для приближённого вычисления интеграла (44):

$$\int_{t_k}^{t_{k+1}} \int_{\partial S_P^{(x,y)}(t)} (Tv \cdot N) d\gamma dt \approx -\tau (\bar{v}_n^y \widehat{T}_n - \bar{v}_s^y \widehat{T}_s) \bar{h}_P^x. \quad (46)$$

Здесь скорость $\bar{v}^y$ вычисляется по формулам

$$\bar{v}_n^y = (0.5[v_n^y + v_{ne}^y] h_e^x + 0.5[v_n^y + v_{nw}^y] h_w^x) / (2\bar{h}_P^x), \quad \bar{v}_s^y = (0.5[v_s^y + v_{se}^y] h_e^x + 0.5[v_s^y + v_{sw}^y] h_w^x) / (2\bar{h}_P^x).$$

Разделив равенства (43) и (46) на шаг τ , найдём аппроксимацию производной по времени на подвижной сетке $\omega^{(x,y)}(t)$:

$$\mathcal{T}_P^{(x,y)}(T) dS_P^{(x,y)} = (T_P dS_P^{(x,y)})_t - (\bar{v}_n^y \widehat{T}_n - \bar{v}_s^y \widehat{T}_s) \bar{h}_P^x. \quad (47)$$

Разностный оператор (47) можно преобразовать к виду

$$\mathcal{T}_P^{(x,y)}(T) dS_P^{(x,y)} = (T_P dS_P^{(x,y)})_t - 0.5[\widehat{h}_n^y (\widehat{T}_P \bar{v}_P^y)_{\bar{y}} + \widehat{h}_s^y (\widehat{T}_P \bar{v}_P^y)_{\bar{y}}] \bar{h}_P^x, \quad (48)$$

где $\bar{v}_P^y = (0.5[v_P^y + v_e^y] h_e^x + 0.5[v_P^y + v_w^y] h_w^x) / (2\bar{h}_P^x)$.

По аналогии с [19] преобразуем аппроксимацию производной по времени (48) к недивергентному виду. Рассмотрим верхнюю часть ячейки $S_P^{(x,y)}$, т.е. $S_1^{(x,y)} \cup S_2^{(x,y)}$. Воспользуемся формулами разностного дифференцирования произведения [24]. Для разностной производной по времени получаем

$$([dS_1^{(x,y)} + dS_2^{(x,y)}]T_P)_t = (dS_1^{(x,y)} + dS_2^{(x,y)})T_{P,t} + \left(\frac{h_e^x[h_{n,t}^y + h_{ne,t}^y]}{8} + \frac{h_w^x[h_{n,t}^y + h_{nw,t}^y]}{8} \right) \hat{T}_P.$$

Для разностной производной по переменной y имеем

$$(\hat{T} \bar{v}_P^y)_{\tilde{y}} = \bar{v}_{P,\tilde{y}}^y \hat{T}_P + \bar{v}_n^y \hat{T}_{P,\tilde{y}}.$$

В силу соотношения (23) производная по времени от шага по пространству выражается равенствами

$$0.5(h_{n,t}^y + h_{ne,t}^y) = [y_{n,t} - y_{P,t}] + [y_{ne,t} - y_{e,t}] = [v_n^y - v_P^y] + [v_{ne}^y - v_e^y],$$

$$0.5(h_{n,t}^y + h_{nw,t}^y) = [y_{n,t} - y_{P,t}] + [y_{nw,t} - y_{w,t}] = [v_n^y - v_P^y] + [v_{nw}^y - v_w^y],$$

из которых вытекает, что

$$\left(\frac{h_e^x[h_{n,t}^y + h_{ne,t}^y]}{8} + \frac{h_w^x[h_{n,t}^y + h_{nw,t}^y]}{8} \right) \hat{T}_P = \frac{\hat{h}_P^x \hat{h}_n^y}{2} \bar{v}_{P,\tilde{y}}^y \hat{T}_P.$$

Откуда, приводя подобные члены, в регулярной точке сетки получаем равенство

$$([dS_1^{(x,y)} + dS_2^{(x,y)}]T_P)_t - \frac{\hat{h}_P^x \hat{h}_n^y}{2} (\hat{T}_P \bar{v}_P^y)_{\tilde{y}} = (dS_1^{(x,y)} + dS_2^{(x,y)})T_{P,t} - \frac{\hat{h}_P^x \hat{h}_n^y}{2} \bar{v}_n^y \hat{T}_{P,\tilde{y}}. \quad (49)$$

Аналогично, для нижней части ячейки $S_P^{x,y}$ имеем

$$([dS_3^{(x,y)} + dS_4^{(x,y)}]T_P)_t - \frac{\hat{h}_P^x \hat{h}_s^y}{2} (\hat{T}_P \bar{v}_P^y)_{\tilde{y}} = (dS_3^{(x,y)} + dS_4^{(x,y)})T_{P,t} - \frac{\hat{h}_P^x \hat{h}_s^y}{2} \bar{v}_s^y \hat{T}_{P,\tilde{y}}. \quad (50)$$

Используя выражения (48)–(50), получим недивергентную форму записи аппроксимации производной по времени:

$$\mathcal{T}_P^{(x,y)}(T) dS_P^{(x,y)} = T_{P,t} dS_P^{(x,y)} - 0.5[\hat{h}_n^y \bar{v}_n^y \hat{T}_{P,\tilde{y}} + \hat{h}_s^y \bar{v}_s^y \hat{T}_{P,\tilde{y}}] \hat{h}_P^x. \quad (51)$$

Во внутренних точках сетки $\omega^{(x,y)}(t)$ разностная схема для задачи (1)–(4) имеет вид

$$\mathcal{T}_P^{(x,y)}(T) \widehat{dS}_P^{(x,y)} = [\mathcal{L}_P^{(x,y)}(\hat{T}) + \mathcal{E}_{P*}^{(x,y)}(\hat{T}) \delta_{PP*}] \widehat{dS}_P^{(x,y)}. \quad (52)$$

Разностная схема (52) является консервативной, коэффициенты $M^{\tau\tau}$, $M^{\tau y}$, M^{yy} в ней взяты с верхнего временного слоя, а соответствующая система нелинейных сеточных уравнений решается относительно вектора неизвестных, компонентами которого являются температуры в твёрдой и жидкой фазах и скорость межфазной границы.

4. Эквивалентность разностных схем.

Утверждение 2. Разностная схема (19), построенная с помощью метода выпрямления фронта, и разностная схема (52), построенная на подвижной сетке, удовлетворяющей соотношению (23), алгебраически эквивалентны.

Доказательство. Выполним в разностной схеме (52) дискретный аналог замены переменных (5). Нетрудно проверить, что шаги сеток $\omega_h^{(x,y)}$ и $\omega_h^{(\xi,\eta)}$ связаны соотношениями

$$h_e^x = h_e^\xi, \quad h_w^x = h_w^\xi, \quad (53)$$

$$\begin{aligned} h_n^y &= y_N - y_P = \varphi(t_k, \xi_P, \eta_N) - \varphi(t_k, \xi_P, \eta_P) = l_n(\eta_N - \eta_P) = l_n h_n^\eta, \\ h_\alpha^y &= l_\alpha h_\alpha^\eta, \quad \hat{h}_\alpha^y = \hat{l}_\alpha h_\alpha^\eta, \quad \alpha \in V. \end{aligned} \quad (54)$$

В силу равенств (7), (22) скорости движения вершин контрольного объема v_α^y , $\alpha \in V$, можно записать следующим образом:

$$v_\alpha^y = \varphi_{\alpha,t}, \quad \alpha \in V. \quad (55)$$

Выполним замену переменных в разностном операторе $\mathcal{L}_P^{(x,y)}(\hat{T}) \widehat{dS}_P^{(x,y)}$. Последовательно рассмотрим потоки через все границы ячейки $S_P^{(x,y)}$. Начнём с “северной” границы, поток через которую определяется выражением (30). Замена переменных в разностных аналогах смешанных производных, входящих в (30), приводит к следующему равенству:

$$\widehat{\Delta\gamma}_{ne n} M_{ne}^{y\mu} \hat{T}_{n,\mu} + \widehat{\Delta\gamma}_{n nw} M_{nw}^{y\mu} \hat{T}_{n,\bar{\mu}} = 0.5[h_e^\xi L_{ne}^{\eta\xi} \hat{T}_{n,\xi} + h_w^\xi L_{nw}^{\eta\xi} \hat{T}_{n,\bar{\xi}}], \quad (56)$$

потому что $\widehat{\Delta\gamma}_{ne n} \hat{T}_{n,\mu} = 0.5 h_e^\xi \hat{T}_{n,\xi}$, $\widehat{\Delta\gamma}_{n nw} \hat{T}_{n,\bar{\mu}} = 0.5 h_w^\xi \hat{T}_{n,\bar{\xi}}$ и

$$\begin{aligned} M_{ne}^{y\mu} &= \varkappa_{ne}(-\hat{y}_{n,x}) = \varkappa_{ne}(\hat{y}_{nE} - \hat{y}_n)/h_e^x = -\varkappa_{ne}[\varphi(t_{k+1}, \xi_E, \eta_n) - \varphi(t_{k+1}, \xi_P, \eta_n)]/h_e^\xi = \\ &= \varkappa_{ne}(-\hat{\varphi}_{n,\xi}) = L_{ne}^{\eta\xi}, \quad M_{nw}^{y\mu} = L_{nw}^{\eta\xi}. \end{aligned}$$

С помощью соотношений (54) преобразуем члены, связанные с $\hat{T}_{P,y}$, к виду

$$\begin{aligned} M_{ne}^{yy} \hat{T}_{P,y} &= \varkappa_{ne}[1 + (\hat{y}_{n,x})^2] \hat{T}_{P,y} = \varkappa_{ne}[1 + (\hat{\varphi}_{n,\xi})^2] (\hat{T}_N - \hat{T}_P) / \hat{h}_n^y = \varkappa_{ne}[1 + (\hat{\varphi}_{n,\xi})^2] / (\hat{l}_n) \hat{T}_{P,\eta} = L_{ne}^{yy} \hat{T}_{P,\eta}, \\ M_{nw}^{yy} \hat{T}_{P,y} &= L_{nw}^{yy} \hat{T}_{P,\eta}, \end{aligned}$$

откуда

$$0.5[h_e^x M_{ne}^{yy} + h_w^x M_{nw}^{yy}] \hat{T}_{P,y} = 0.5[h_e^\xi L_{ne}^{\eta\eta} + h_w^\xi L_{nw}^{\eta\eta}] \hat{T}_{P,\eta}. \quad (57)$$

Аналогично выполним замену переменных в выражении (31) для потока через “южную” границу:

$$\widehat{\Delta\gamma}_{sse} M_{se}^{y\mu} \hat{T}_{s,\mu} + \widehat{\Delta\gamma}_{sw s} M_{sw}^{y\mu} \hat{T}_{s,\bar{\mu}} = 0.5[h_e^\xi L_{se}^{\eta\xi} \hat{T}_{s,\xi} + h_w^\xi L_{sw}^{\eta\xi} \hat{T}_{s,\bar{\xi}}], \quad (58)$$

$$0.5[h_e^x M_{se}^{yy} + h_w^x M_{sw}^{yy}] \hat{T}_{P,\bar{y}} = 0.5[h_e^\xi L_{se}^{\eta\eta} + h_w^\xi L_{sw}^{\eta\eta}] \hat{T}_{P,\bar{\eta}}. \quad (59)$$

Рассмотрим поток (36) через “восточную” границу. Так как $\hat{h}_{ne}^y \hat{T}_{e,y} = h_n^\eta \hat{T}_{e,\eta}$, $\hat{h}_{se}^y \hat{T}_{e,\bar{y}} = h_s^\eta \hat{T}_{e,\bar{\eta}}$ и $M_{ne}^{\mu y} = L_{ne}^{\xi\eta}$, $M_{se}^{\mu y} = L_{se}^{\xi\eta}$, то для смешанных производных в (36) будем иметь

$$0.5[\hat{h}_{ne}^y M_{ne}^{\mu y} \hat{T}_{e,y} + \hat{h}_{se}^y M_{se}^{\mu y} \hat{T}_{e,\bar{y}}] = 0.5[h_n^\eta L_{ne}^{\xi\eta} \hat{T}_{e,\eta} + h_s^\eta L_{se}^{\xi\eta} \hat{T}_{e,\bar{\eta}}]. \quad (60)$$

Для производной в направлении μ получаем

$$0.5[\hat{h}_{ne}^y M_{ne}^{\mu\mu} + \hat{h}_{se}^y M_{se}^{\mu\mu}] \hat{T}_{P,\mu} = 0.5[h_n^\eta L_{ne}^{\xi\xi} + h_s^\eta L_{se}^{\xi\xi}] \hat{T}_{P,\xi}. \quad (61)$$

Здесь воспользовались тем, что в силу соотношений (53) и (54) справедливы равенства

$$\begin{aligned} \hat{h}_{ne}^y M_{ne}^{\mu\mu} \hat{T}_{P,\mu} &= \hat{h}_{ne}^y \varkappa_{ne} \sqrt{1 + (\hat{y}_{P,x})^2} \hat{T}_{P,\mu} = \hat{h}_{ne}^y \varkappa_{ne} (\hat{T}_E - \hat{T}_P) / h_e^x = \hat{l}_{ne} h_n^\eta \varkappa_{ne} \hat{T}_{P,\xi} = h_n^\eta L_{ne}^{\xi\xi} \hat{T}_{P,\xi}, \\ \hat{h}_{se}^y M_{se}^{\mu\mu} \hat{T}_{P,\mu} &= h_s^\eta L_{se}^{\xi\xi} \hat{T}_{P,\xi}. \end{aligned}$$

Для потока (37) на “западной” границе имеем

$$0.5[\hat{h}_{nw}^y M_{nw}^{\mu y} \hat{T}_{w,y} + \hat{h}_{sw}^y M_{sw}^{\mu y} \hat{T}_{w,\bar{y}}] = 0.5[h_n^\eta L_{nw}^{\xi\eta} \hat{T}_{w,\eta} + h_s^\eta L_{sw}^{\xi\eta} \hat{T}_{w,\bar{\eta}}], \quad (62)$$

$$0.5[\hat{h}_{nw}^y M_{nw}^{\mu\mu} + \hat{h}_{sw}^y M_{sw}^{\mu\mu}] \hat{T}_{P,\bar{\mu}} = 0.5[h_n^\eta L_{nw}^{\xi\xi} + h_s^\eta L_{sw}^{\xi\xi}] \hat{T}_{P,\bar{\xi}}. \quad (63)$$

Учитывая соотношения (56)–(63) в формуле (38), получаем

$$\mathcal{L}_P^{(x,y)}(\hat{T}) \widehat{dS}_P^{(x,y)} = \mathcal{L}_P^{(\xi,\eta)}(\hat{T}) dS_P^{(\xi,\eta)}.$$

На фронте кристаллизации в уравнение (52) входит слагаемое, отвечающее за выполнение условия Стефана и описывающее скачок теплового потока при изменении агрегатного состояния вещества. Из формул (15) и (40) следует, что $(\bar{v}^{\text{ph}})^y = (\bar{v}^{\text{ph}})^\eta$, откуда вытекает, что на границе раздела фаз выполняется равенство

$$\mathcal{E}^{(x,y)}(\hat{T}) \widehat{dS}_P^{(x,y)} = \mathcal{E}^{(\xi,\eta)}(\hat{T}) dS_P^{(\xi,\eta)}. \quad (64)$$

Выполним замену переменных в $\mathcal{T}_P^{(x,y)}(T) \widehat{dS}_P^{(x,y)}$. В силу соотношений (53) и (54) имеем

$$dS_1^{(x,y)} = (h_n^y + h_{ne}^y) h_e^x / 8 = (l_n + l_{ne}) h_e^\xi h_n^\eta / 8.$$

Площади остальных четвертей ячейки преобразуются аналогичным образом. Поэтому первое слагаемое в аппроксимации (51) запишется в виде

$$T_{P,t} dS_P^{(x,y)} = T_{P,t} \bar{l}_P dS_P^{(\xi,\eta)}. \quad (65)$$

Воспользовавшись формулой (55), для второго слагаемого в (51) получим

$$0.5[\hat{h}_n^y \bar{v}_n^y \hat{T}_{P,\tilde{y}} + \hat{h}_s^y \bar{v}_s^y \hat{T}_{P,\bar{y}}] \hat{h}_P^x = 0.5[h_n^\eta (\bar{\varphi}_t)_n \hat{T}_{P,\tilde{\eta}} + h_s^\eta (\bar{\varphi}_t)_s \hat{T}_{P,\bar{\eta}}] \hat{h}_P^\xi. \quad (66)$$

Замена слагаемых в аппроксимации (51) их выражениями (65) и (66) приводит к равенству

$$\mathcal{T}_P^{(x,y)}(T) \widehat{dS}_P^{(x,y)} = \mathcal{T}_P^{(\xi,\eta)}(T) dS_P^{(\xi,\eta)}.$$

В результате показано, что

$$[\mathcal{T}_P^{(x,y)}(T) - \mathcal{L}_P^{(x,y)}(\hat{T}) - \mathcal{E}_P^{(x,y)}(\hat{T}) \delta_{PP*}] \widehat{dS}_P^{(x,y)} = [\mathcal{T}_P^{(\xi,\eta)}(T) - \mathcal{L}_P^{(\xi,\eta)}(\hat{T}) - \mathcal{E}_P^{(\xi,\eta)}(\hat{T}) \delta_{PP*}] dS_P^{(\xi,\eta)}.$$

Таким образом, доказано, что схемы (52) и (19) алгебраически эквивалентны.

Заключение. Для классической двухфазной задачи Стефана выделен класс консервативных разностных схем, для которых метод выпрямления фронта и метод, основанный на использовании подвижных сеток, алгебраически эквивалентны. Получены дивергентная и недивергентная формы записи, аппроксимирующие соответствующие формы записи исходной дифференциальной задачи. Для разностного оператора, аппроксимирующего на неподвижной сетке диссипативные члены в уравнении теплопроводности, можно доказать свойства самосопряжённости и отрицательной определённости. Доказательство указанных свойств является несложным, но громоздким и поэтому в работе не приводится. В силу алгебраической эквивалентности предложенных разностных схем аналогичными свойствами обладает и разностный оператор на подвижной сетке. Таким образом, построенный класс схем наследует основные свойства исходной дифференциальной задачи.

СПИСОК ЛИТЕРАТУРЫ

1. Бакирова О.И. Численное моделирование процесса зонной плавки на основе решения задачи о фазовом переходе в бинарной системе // Математическое моделирование. Получение металлов и полупроводниковых структур. М., 1986. С. 142–158.
2. Дегтярев Л.М., Дроздов В.В., Иванов Т.С. Метод адаптивных к решению сеток в сингулярно-возмущенных одномерных краевых задачах // Дифференц. уравнения. 1987. Т. 23. № 7. С. 1160–1168.

3. Фрязинов И.В., Забелина М.П. Сеточный метод решения задачи Стефана для бинарной системы // Дифференц. уравнения. 1987. Т. 23. № 7. С. 1188–1197.
4. Vermolen F.J., Vuk C. A mathematical model for the dissolution of particles in multi-component alloys // J. of Comput. and Appl. Math. 2000. V. 126. P. 233–254.
5. Muray W.D., Landis F. Numerical and machine solutions of the transient heat conduction problems involving melting or freezing // J. Heat Transfer. 1959. V. 81. P. 106–112.
6. Kutluay S. Numerical schemes for one-dimensional Stefan-like problems with a forcing term // Appl. Math. and Comput. 2005. V. 168. P. 1159–1168.
7. Мажукин А.В., Мажукин В.И. Динамическая адаптация в параболических уравнениях // Журн. вычислит. математики и мат. физики. 2007. Т. 47. № 11. С. 1913–1936.
8. Sekhon M., Lent B., Dost S. Numerical study of liquid phase diffusion growth of SiGe subjected to accelerated crucible rotation // J. of Crystal Growth. 2016. V. 438. P. 90–98.
9. Friaazinov I.V., Marchenko M.P., Mazhorova O.S. Transient simulation of crystal growth by Bridgman technique under normal and low gravity // Proc. of Microgravity Science and Low Gravity / Aerospace congress. Moscow, 1991. P. 179–187.
10. Lan C.W., Chen F.C. A finite volume method for solute segregation in directional solidification and comparison with a finite element method // Comput. Methods Appl. Mech. Eng. 1996. V. 131. P. 191–207.
11. Zhou Yu., North T.H. Kinetic modelling of diffusion-controlled, two-phase moving interface problems // Model. and Simul. in Material Science and Engineering. 1993. V. 1. № 4. P. 505–516.
12. Illingworth T.C., Golosnoy I.O. Numerical solutions of diffusion-controlled moving boundary problems which conserve solute // J. of Comput. Physics. 2005. V. 209. P. 207–225.
13. Pandelaers L., Verhaeghe F., Wollants P., Blanpain B. An implicit conservative scheme for coupled heat and mass transfer problems with multiple moving interfaces // Int. J. of Heat and Mass Transfer. 2011. V. 54. № 5–6. P. 1039–1045.
14. Самарский А.А., Вабищевич П.Н. Вычислительная теплопередача. М., 2003.
15. Мажорова О.С., Попов Ю.П., Щерица О.В. Алгоритм расчета задачи о фазовом переходе в многокомпонентной системе // Дифференц. уравнения. 2004. Т. 40. № 7. С. 1051–1060.
16. Мажорова О.С., Попов Ю.П., Щерица О.В. Метод численного решения задач кристаллизации многокомпонентных растворов / Препринт Ин-та прикладной математики им. М.В. Келдыша РАН. М., 2002. № 18.
17. Nobeoka M., Takagi Y., Okano Y et al. Numerical simulation of InGaSb crystal growth by temperature gradient method under normal- and micro-gravity fields // J. of Crystal Growth. 2014. V. 385. P. 66–71.
18. Landau H.G. Heat conduction in a melting solid // J. of Appl. Math. 1950. V. 8. P. 81–94.
19. Мажорова О.С., Попов Ю.П., Щерица О.В. Консервативные разностные схемы для термо-диффузионной задачи Стефана // Дифференц. уравнения. 2013. Т. 49. № 7. С. 897–905.
20. Gusev A.O., Shcheritsa O.V., Mazhorova O.S. Conservative finite volume strategy for investigation of solution crystal growth techniques // Computers & Fluids. 2020. V. 202. P. 104501.
21. Vinokur M. Conservation equations of gasdynamics in curvilinear coordinate system // J. of Comput. Phys. 1974. V. 14. P. 105–125.
22. Steger J. Implicit finite-difference simulation of flow about arbitrary two-dimensional geometries // American Institute of Aeronautics and Astronautics Journal. 1978. V. 16. № 7. P. 679–686.
23. Patankar S.V. Numerical heat transfer and fluid flow. New York, 1981.
24. Самарский А.А. Введение в теорию разностных схем. М., 1971.

Институт прикладной математики
им. М.В. Келдыша РАН, г. Москва

Поступила в редакцию 10.03.2021 г.
После доработки 10.03.2021 г.
Принята к публикации 27.04.2021 г.

ЧИСЛЕННЫЕ МЕТОДЫ

УДК 519.633.2+517.958:536.71

СПЕКТРАЛЬНЫЕ УСЛОВИЯ УСТОЙЧИВОСТИ
ЯВНОЙ ТРЁХСЛОЙНОЙ РАЗНОСТНОЙ СХЕМЫ
ДЛЯ МНОГОМЕРНОГО УРАВНЕНИЯ ПЕРЕНОСА
С ВОЗМУЩЕНИЯМИ

© 2021 г. А. А. Злотник, Б. Н. Четверушкин

Изучаются разностные схемы, связанные с упрощённой линеаризованной многомерной гиперболической квазигазодинамической системой дифференциальных уравнений. Показывается, что явную двухслойную векторную разностную схему с релаксацией потоков для гиперболического уравнения второго порядка с переменными коэффициентами, являющегося возмущением уравнения переноса с параметром-множителем при старших производных, можно свести к явной трёхслойной разностной схеме. Анализируется спектральное условие равномерной по времени устойчивости такой явной трёхслойной разностной схемы в случае постоянных коэффициентов и выводятся как достаточные, так и необходимые условия его справедливости, в том числе в форме условий типа Куранта на отношение шагов по времени и по пространству.

DOI: 10.31857/S0374064121070062

Введение. Гиперболическая квазигазодинамическая (ГКГД) система уравнений представляет собой специальным образом возмущённую систему уравнений газовой динамики, содержащую слагаемые с производными второго порядка по пространственным и временной переменным с малым параметром-множителем $\tau > 0$. Она применяется для построения семейства трёхслойных и двухслойных векторных разностных методов численного решения различных задач газовой динамики [1]. Отметим, что параболические КГД системы представлены в [2, 3]. В [4, 5] установлен набор математических свойств ГКГД системы, в том числе выведены равномерные по времени оценки для соответствующей линеаризованной на постоянном решении системы. Изучались также гиперболическое второго порядка возмущение с параметром τ параболической начально-краевой задачи без конвективных членов [6, 7] и с ними [8] и устойчивость соответствующих трёхслойного с весом и двухслойного векторного численных методов [9]. Равномерная по времени устойчивость неявных трёхслойной с весом и двухслойной векторной разностных схем для линеаризованной ГКГД системы доказана в [10].

На практике наибольший интерес представляют явные разностные схемы для ГКГД системы. Однако обоснование равномерной по τ устойчивости их линеаризаций в настоящее время отсутствует. Применить для этого энергетический метод подобно [4, 10] пока не удалось ввиду достаточно сложной структуры линеаризованных схем, в том числе и по причине наличия в них аппроксимаций слагаемых как с первыми, так и со вторыми производными по пространственным и временной переменным, при этом перед вторыми производными присутствует параметр-множитель τ .

Чтобы продвинуться в изучении этого вопроса, в настоящей работе вместо ГКГД системы уравнений рассматривается упрощённый случай одного гиперболического уравнения второго порядка с переменными коэффициентами, являющегося возмущением уравнения переноса с параметром τ при вторых производных. Во-первых, показывается, что явную двухслойную векторную разностную схему с релаксацией потоков типа [11] для такого уравнения можно свести к явной трёхслойной разностной схеме. В этой трёхслойной схеме параметр τ перед аппроксимацией второй производной по t умножен на некоторый множитель больше единицы. Во-вторых, и это главное, анализируется спектральное условие равномерной по времени устойчивости такой явной трёхслойной схемы в случае постоянных коэффициентов. Выводятся как

достаточные, так и близкие к ним необходимые условия справедливости спектрального условия, в том числе в форме условий типа Куранта на отношение шагов по времени и пространству. Эти условия не зависят от τ . В число необходимых условий входит условие преобладания коэффициентов вязких слагаемых над коэффициентами переноса. Вывод основан на анализе расположения на комплексной плоскости корней соответствующего характеристического многочлена с комплексными коэффициентами с помощью обобщённого критерия Рауса–Гурвица [12, гл. XVI, § 19]. Предварительно анализируются также критерии аналогичного спектрального условия устойчивости для самого дифференциального уравнения и разностной схемы с обнулением параметра τ . Для уравнения таким критерием является условие своего рода преобладания матрицы вязких слагаемых над вектором коэффициентов переноса; условия подобного рода играли существенную роль в [4, 10]. Для указанной схемы критерием служит условие типа Куранта. Полученные результаты являются достаточно обнадеживающими для последующего анализа соответствующих свойств разностных схем для ГКГД системы.

1. Гиперболическое дифференциальное уравнение второго порядка и сведение явной двухслойной векторной разностной схемы с релаксацией потоков для него к трёхслойной схеме. Рассмотрим гиперболическое дифференциальное уравнение второго порядка с параметром-множителем $\tau > 0$ при старших производных

$$\tau \partial_t^2 u + \partial_t u + \partial_i(b_i(x)u) - \tau \operatorname{div}(A(x)\nabla u) = f. \quad (1)$$

Искомая функция $u = u(x, t)$ определена и уравнение рассматривается при $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, $t \geq 0$, а $n \geq 1$. Здесь и ниже $A(x) = \operatorname{diag}\{a_1(x), \dots, a_n(x)\}$ – диагональная матрица, операторы div и ∇ берутся по x , а по повторяющимся индексам i, j предполагается суммирование от 1 до n .

Введём равномерную сетку ω_{kh} с узлами $x_{kl} = lh_k$, $l \in \mathbb{Z}$, и шагом $h_k > 0$, сетку, сдвинутую на $h_k/2$, с узлами $x_{k(l-1/2)} = (l - (1/2))h_k$, $l \in \mathbb{Z}$, и сеточные операторы

$$s_k v_{l-1/2} = \frac{v_{l-1} + v_l}{2}, \quad \delta_k v_{l-1/2} = \frac{v_l - v_{l-1}}{h_k}, \quad \delta_k^* w_l = \frac{w_{l+1/2} - w_{l-1/2}}{h_k}, \quad \Lambda_k v_l = \frac{v_{l+1} - 2v_l + v_{l-1}}{h_k^2},$$

где $v_l = v(x_{kl})$, $w_{l-1/2} = w(x_{k(l-1/2)})$, $1 \leq k \leq n$. Положим также $h = (h_1, \dots, h_n)$ и введём сетку $\omega_h := \omega_{1h} \times \dots \times \omega_{nh}$ и операторы $s = (s_1, \dots, s_n)$, $\nabla_h = (\delta_1, \dots, \delta_n)$.

Введём равномерную сетку $\bar{\omega}^{h_t}$ с узлами $t_m = mh_t$, $m \geq 0$, и шагом $h_t > 0$ и разностные операторы

$$\bar{\delta}_t y = \frac{y - \check{y}}{h_t}, \quad \delta_t y = \frac{\hat{y} - y}{h_t}, \quad \delta_t^* y = \frac{\hat{y} - \check{y}}{2h_t}, \quad \Lambda_t y = \delta_t \bar{\delta}_t y = \frac{\hat{y} - 2y + \check{y}}{h_t^2},$$

где $y^m = y(t_m)$, $\check{y}^m = y^{m-1}$ и $\hat{y}^m = y^{m+1}$. Пусть $\omega^{h_t} = \bar{\omega}^{h_t} \setminus \{0\}$.

Рассмотрим явную двухслойную по t векторную разностную схему с релаксацией потоков типа [11] для уравнения (1), симметричную и трёхточечную по $x_1, \dots, x_n$:

$$\hat{\varphi} = e^{-h_t/\tau} \varphi - (1 - e^{-h_t/\tau})(Bsv - \tau A \nabla_h v), \quad (2)$$

$$\delta_t v = \operatorname{div}_h \hat{\varphi} + f. \quad (3)$$

Искомыми являются функция $v \approx u$, определённая на сетке $\omega_h \times \bar{\omega}^{h_t}$, и вспомогательная вектор-функция $\varphi = (\varphi_1, \dots, \varphi_n)^T$ с компонентами φ_k , определёнными на сетке, сдвинутой на $h_k/2$ по x_k относительно $\omega_h \times \bar{\omega}^{h_t}$. Здесь $B(x) = \operatorname{diag}\{b_1(x), \dots, b_n(x)\}$ – диагональная матрица, $\operatorname{div}_h \varphi = \delta_1^* \varphi_1 + \dots + \delta_n^* \varphi_n$ и уравнения схемы записываются на тех же сетках, на которых определены сами искомые функции.

Сведём двухслойную векторную схему к трёхслойной схеме стандартного типа для v .

Теорема 1. Функция v удовлетворяет явной по времени симметричной и трёхточечной по всем переменным $x_1, \dots, x_n$, t (трёхслойной по t) разностной схеме

$$\tau \alpha_0 \left(\frac{h_t}{2\tau} \right) \Lambda_t v + \delta_t^* v + \operatorname{div}_h(Bsv) - \tau \operatorname{div}_h(A \nabla_h v) = \tilde{f} \quad \text{на } \omega_h \times \omega^{h_t}, \quad (4)$$

где $\alpha_0(\zeta) := \zeta \operatorname{cth} \zeta$ и свободный член имеет вид

$$\tilde{f} := \frac{f - e^{-h_t/\tau} \check{f}}{1 - e^{-h_t/\tau}}. \quad (5)$$

Доказательство. Применим оператор div_h к уравнению (2) и получим уравнение

$$\operatorname{div}_h \widehat{\varphi} = e^{-h_t/\tau} \operatorname{div}_h \varphi - (1 - e^{-h_t/\tau}) \operatorname{div}_h (Bsv - \tau A \nabla_h v).$$

Исключим из него функцию φ , выразив $\operatorname{div}_h \widehat{\varphi}$ и $\operatorname{div}_h \varphi = \bar{\delta}_t v - \check{f}$ в силу уравнения (3):

$$\delta_t v - f = e^{-h_t/\tau} (\bar{\delta}_t v - \check{f}) - (1 - e^{-h_t/\tau}) \operatorname{div}_h (Bsv - \tau A \nabla_h v).$$

Симметризуем запись последнего уравнения умножением его на $e^{h_t/(2\tau)}$:

$$e^{h_t/(2\tau)} \delta_t v - e^{-h_t/(2\tau)} \bar{\delta}_t v + 2 \operatorname{sh} \frac{h_t}{2\tau} \operatorname{div}_h (Bsv - \tau A \nabla_h v) = e^{h_t/(2\tau)} f - e^{-h_t/(2\tau)} \check{f}.$$

Далее воспользуемся простыми формулами

$$\delta_t v = \mathring{\delta}_t v + (h_t/2) \Lambda_t v, \quad \bar{\delta}_t v = \mathring{\delta}_t v - (h_t/2) \Lambda_t v$$

и перейдём к уравнению

$$h_t \operatorname{ch} \frac{h_t}{2\tau} \Lambda_t v + 2 \operatorname{sh} \frac{h_t}{2\tau} \mathring{\delta}_t v + 2 \operatorname{sh} \frac{h_t}{2\tau} \operatorname{div}_h (Bsv - \tau A \nabla_h v) = e^{h_t/(2\tau)} f - e^{-h_t/(2\tau)} \check{f}.$$

Поделив его на $2 \operatorname{sh}(h_t/(2\tau))$, окончательно получим для v явную трёхслойную по времени разностную схему (4), (5). Теорема доказана.

Отметим, что $\alpha_0(\zeta) > \max\{1, \zeta\}$ при $\zeta > 0$ и $\alpha_0(\zeta) = \zeta(1 + O(e^{-2\zeta}))$ при $\zeta \rightarrow +\infty$; при этом $|\alpha_0(\zeta) - \zeta| < 0.075, 0.015, 0.0027$ уже при $\zeta \geq 2, 3, 4$ соответственно. Обратим также внимание на то, что в пределе при $h_t/(2\tau) \rightarrow +\infty$ происходит замена $\tau \alpha_0(h_t/(2\tau))$ на $h_t/2$ и трёхслойная разностная схема (4) переходит в явную двухслойную схему

$$\delta_t v + \operatorname{div}_h (Bsv) - \tau \operatorname{div}_h (A \nabla_h v) = f,$$

поскольку $(h_t/2) \Lambda_t v + \mathring{\delta}_t v = \delta_t v$.

Обратим внимание на то, что в случае постоянных коэффициентов умножение разностной схемы (4) на τ и замена шагов $(h, h_t) \mapsto (h/\tau, h_t/\tau)$ приводит к частному случаю этой схемы с $\tau = 1$ (включая аргумент коэффициента $\alpha_0(h_t/(2\tau))$). Поэтому некоторые свойства этой схемы не зависят от значения τ . В том числе это относится к изучаемому ниже спектральному условию равномерной по времени устойчивости.

Существенно, что применённый выше способ сведения двухслойной схемы к трёхслойной можно использовать не только в рассмотренном скалярном случае, но и для самой ГКГД системы уравнений. Сведение других двухслойных векторных методов для гиперболических уравнений второго порядка к трёхслойным методам выполнялось, в том числе, в [10, 13, 14]. Отметим также, что большое количество двухслойных разностных схем для уравнения переноса приведено, например, в [15].

2. Многомерное уравнение переноса с возмущениями и явная трёхслойная разностная схема для него. Рассмотрим многомерное однородное уравнение переноса с возмущениями – слагаемыми, представляющими собой производные второго порядка от искомой функции по t и по x_i с параметрами-множителями:

$$\alpha_0 \tau \partial_t^2 u + \partial_t u + b_i \partial_i u - \tau a_{ij} \partial_i \partial_j u = 0, \quad (6)$$

где b_i, a_{ij} – постоянные коэффициенты, $\alpha_0 > 0, \tau > 0$ – параметры. Напомним, что по повторяющимся индексам i, j предполагается суммирование от 1 до n . Введение параметра

α_0 обусловлено видом схемы (4) (от него легко освободиться заменами $\alpha_0 \tau \mapsto \tau$, $a_{ij} \mapsto a_{ij}/\alpha_0$, но для наглядности этого делать не будем). Искомая функция $u(x, t)$ определена и уравнение (6) рассматривается на $\mathbb{R}^n \times [0, +\infty)$. Функции $u|_{t=0}$ и $\partial_t u|_{t=0}$ считаем заданными. Не ограничивая общности будем считать, что матрица $A = (a_{ij})_{i,j=1}^n$ симметричная. Пусть $b = (b_1, \dots, b_n)^T$.

Предварительно, во-первых, обратим внимание на то, что дифференцирование уравнения переноса $\partial_t u + b_i \partial_i u = f$ приводит к равенствам

$$\partial_t^2 u + b_j \partial_t \partial_j u = \partial_t f, \quad b_j \partial_j \partial_t u + b_j b_i \partial_j \partial_i u = b_j \partial_j f$$

и поэтому

$$\partial_t^2 u - b_i b_j \partial_j \partial_i u = \partial_t f - b_j \partial_j f.$$

Тем самым решение уравнения переноса удовлетворяет и уравнению второго порядка

$$\tau \partial_t^2 u + \partial_t u + b_i \partial_i u - \tau b_i b_j \partial_i \partial_j u = f + \tau(\partial_t f - b_j \partial_j f)$$

типа (6) с $\alpha_0 = 1$ при любом $\tau \neq 0$ (при $\tau = 0$ это само уравнение переноса). Возникшая матрица $b \otimes b$ с элементами $b_i b_j$ такова, что $b \otimes b \geq 0$, а $\text{Ker } b \otimes b$ совпадает с ортогональным дополнением к вектору b . Поэтому полученное уравнение является вырождающимся при $n \geq 2$ или при $n = 1$ и $b_1 = 0$.

Во-вторых, умножение уравнения (6) на τ и замена переменных $(x, t) \mapsto (x/\tau, t/\tau)$ приводит к частному случаю уравнения (6) с $\tau = 1$. Поэтому ряд свойств решений этого уравнения, в том числе анализируемое ниже свойство, не зависят от значения τ .

Рассмотрим комплексные частные решения уравнения (6), имеющие вид

$$u(x, t) = e^{ix \cdot \xi} w(\xi, t), \quad x, \xi = (\xi_1, \dots, \xi_n)^T \in \mathbb{R}^n, \quad t \geq 0,$$

где i – мнимая единица, а символ “ $\cdot$ ” обозначает скалярное произведение векторов. Их подстановка в уравнение (6) приводит к следующему обыкновенному дифференциальному уравнению (ОДУ) относительно w :

$$\alpha_0 \tau \partial_t^2 w + \partial_t w + \theta w = 0, \quad \theta := ib_i \xi_i + \tau a_{ij} \xi_j \xi_i = ib \cdot \xi + \tau A \xi \cdot \xi \in \mathbb{C}. \quad (7)$$

Изучим свойство равномерной ограниченности по t его решения. Пусть $\text{Ker } A$ и $\text{Im } A$ – ядро и образ матрицы A (т.е. соответствующего линейного оператора, действующего в $\mathbb{R}^n$). Напомним, что так как $A = A^T$, то $\mathbb{R}^n$ представляется в виде ортогональной прямой суммы подпространств $\text{Ker } A$ и $\text{Im } A$.

Теорема 2. Для справедливости оценки

$$\sup_{t \geq 0} |w(\xi, t)| < \infty \quad \text{для каждого } \xi \in \mathbb{R}^n \quad (8)$$

при любых $w(\xi, 0)$ и $\partial_t w(\xi, 0)$ необходимо, чтобы выполнялось свойство $A \geq 0$.

Критериями (т.е. необходимыми и достаточными условиями) справедливости указанной оценки, кроме $A \geq 0$, служат:

1) если $A > 0$, то $\alpha_0 |A^{-1/2} b|^2 = \alpha_0 A^{-1} b \cdot b \leq 1$;

2) если же $\text{Ker } A \neq \{0\}$, то $b \in \text{Im } A$ и при $\text{Im } A \neq \{0\}$ имеем $\alpha_0 \tilde{A}^{-1} b \cdot b \leq 1$, где $\tilde{A}$ – сужение A на $\text{Im } A$, а иначе $b = 0$ при $\text{Im } A = \{0\}$ (т.е. $A = 0$).

В частности, при $A = \text{diag } \{a_1, \dots, a_n\}$ критерием служит совокупность условий:

$$a_k \geq 0, \quad \text{причём если } a_k = 0, \text{ то } b_k = 0 \text{ при } 1 \leq k \leq n; \quad \alpha_0 \sum_{k: a_k > 0} \frac{b_k^2}{a_k} \leq 1. \quad (9)$$

Доказательство. Характеристическое уравнение для ОДУ (7) имеет вид

$$r_2(z) \equiv \alpha_0 \tau z^2 + z + \theta = 0, \quad z \in \mathbb{C}.$$

Свойство (8) эквивалентно тому, что оба корня многочлена $r_2(z)$ лежат в замкнутой левой полуплоскости $\bar{P}_l := \{z \in \mathbb{C}: z_R \leq 0\}$ и на мнимой оси нет кратных корней (т.е. $z = 0$ не является кратным корнем, что и имеет место). Здесь и ниже используется стандартная запись комплексных чисел в виде $z = z_R + iz_I$.

Для анализа последнего свойства воспользуемся обобщённым критерием Рауса–Гурвица для многочленов с комплексными коэффициентами [12, гл. XVI, § 19]. Для этого запишем вспомогательный многочлен $-ir_2(iz) = z + \theta_I + i(\alpha_0\tau z^2 - \theta_R)$ и определитель четвёртого порядка – результат (с точностью до знака) возникших двух многочленов с вещественными коэффициентами

$$V_4 = \begin{vmatrix} \alpha_0\tau & 0 & -\theta_R & 0 \\ 0 & 1 & \theta_I & 0 \\ 0 & \alpha_0\tau & 0 & -\theta_R \\ 0 & 0 & 1 & \theta_I \end{vmatrix}.$$

Его главный угловой минор второго порядка $V_2 = \alpha_0\tau$ положителен, а сам определитель равен

$$V_4 = \alpha_0\tau(\theta_R - \alpha_0\tau\theta_I^2) = \alpha_0\tau^2[A\xi \cdot \xi - \alpha_0(b \cdot \xi)^2].$$

Пусть сначала $V_4 \neq 0$. Тогда критерием того, что корни $r_2(z)$ лежат в открытой левой полуплоскости $P_l := \{z \in \mathbb{C}: z_R < 0\}$, является условие $V_4 > 0$. При $V_4 = 0$, как легко проверить, имеется чисто мнимый корень $-i\theta_I$, а поэтому второй корень равен $i\theta_I - 1/(\alpha_0\tau) \in P_l$. Поэтому критерием того, что оба корня лежат в $\bar{P}_l$, является условие $V_4 \geq 0$, т.е.

$$\alpha_0(b \cdot \xi)^2 \leq A\xi \cdot \xi \quad \text{для всех } \xi \in \mathbb{R}^n. \quad (10)$$

Его можно записать в виде $\alpha_0 b \otimes b \leq A$, и это условие своего рода преобладания A над b . В него не вошёл параметр τ , что естественно в соответствии со сказанным выше. Условия подобного рода играли существенную роль в работах [4, 10].

Из критерия (10) следует, что $A \geq 0$. Если $A > 0$, то его можно записать в виде

$$\alpha_0(A^{-1/2}b \cdot A^{1/2}\xi)^2 \leq |A^{1/2}\xi|^2 \quad \text{для всех } \xi \in \mathbb{R}^n.$$

Введя вектор $\eta := A^{1/2}\xi$, получим эквивалентный явный критерий из п. 1) теоремы.

Если же $\text{Ker } A \neq \{0\}$, то $b \perp \text{Ker } A$ в силу критерия (10), и поэтому $b \in \text{Im } A$, а сам он принимает вид

$$\alpha_0(b \cdot \xi)^2 \leq A\xi \cdot \xi = \tilde{A}\xi \cdot \xi \quad \text{для всех } \xi \in \text{Im } A,$$

что сводит п. 2) теоремы к п. 1), в котором в роли $\mathbb{R}^n$ выступает $\text{Im } A$ при $\text{Im } A \neq \{0\}$. Случай $\text{Im } A = \{0\}$ очевиден. Теорема доказана.

Рассмотрим теперь явную симметричную и трёхточечную по всем переменным $x_1, \dots, x_n$, t (трёхслойную по t) разностную схему для уравнения (6):

$$\alpha_0\tau\Lambda_tv + \delta_t^{\circ}v + b_i\delta_iv - \tau a_i\Lambda_iv = 0 \quad \text{на } \omega_h \times \omega^{h_t}, \quad (11)$$

в которой искомая функция v определена на $\omega_h \times \bar{\omega}^{h_t}$. Значения на исходном и первом слоях по времени $v(x_k, 0)$ и $v(x_k, h_t)$, где $x_k = (k_1h_1, \dots, k_nh_n)$, $k = (k_1, \dots, k_n) \in \mathbb{Z}^n$, считаем заданными. Здесь для большей наглядности взят случай $A = \text{diag}\{a_1, \dots, a_n\}$, причём $a_i > 0$ при $1 \leq i \leq n$. Так как $\text{div}_h(Bs) = b_i\delta_i$ и $\text{div}_h(A\nabla_h) = a_i\Lambda_i$ при постоянных коэффициентах b_i и a_i , то эта схема получается из схемы (4).

Рассмотрим комплексные частные решения разностной схемы (11), имеющие вид

$$v(x_k, t_m) = e^{ik \cdot \xi} w(\xi, t_m), \quad k \in \mathbb{Z}^n, \quad m \geq 0, \quad \xi = (\xi_1, \dots, \xi_n)^T \in [-\pi, \pi]^n.$$

Назовём *спектральным условием равномерной по времени устойчивости* свойство:

$$\sup_{m \geq 0} |w(\xi, t_m)| < \infty \quad \text{для каждого } \xi \in [-\pi, \pi]^n \quad (12)$$

при любых $w(\xi, 0)$ и $w(\xi, h_t)$ (ср. с (8) и [16, гл. 8, § 25]).

Функция $e^{ik \cdot \xi}$ является собственной функцией операторов δ_l и Λ_l , отвечающей собственным значениям $ih_l^{-1} \sin \xi_l$ и $-4h_l^{-2} \sin^2(\xi_l/2)$ соответственно. Поэтому функция w удовлетворяет трёхслойному разностному уравнению (фактически разностной схеме для ОДУ (7))

$$\alpha_0 \tau \Lambda_t w + \delta_t w + \left(i \frac{b_i}{h_i} \sin \xi_i + \tau \frac{4a_i}{h_i^2} \sin^2 \frac{\xi_i}{2} \right) w = 0 \quad \text{на } \omega^{h_t}. \quad (13)$$

В силу определений операторов Λ_t и δ_t после умножения на $h_t^2/(\alpha_0 \tau)$ последнее уравнение можно записать в рекуррентной форме

$$(1+d)\hat{w} - 2(1-\theta)w + (1-d)\check{w} = 0, \quad (14)$$

где

$$d := \frac{h_t}{2\alpha_0 \tau} > 0, \quad \theta \equiv \theta(\xi) = 2 \frac{h_t^2 a_i}{\alpha_0 h_i^2} \sin^2 \frac{\xi_i}{2} + i \frac{h_t^2 b_i}{2\alpha_0 \tau h_i} \sin \xi_i \in \mathbb{C}.$$

Характеристическое уравнение для разностного уравнения (14) имеет вид

$$p_2(q) \equiv (1+d)q^2 - 2(1-\theta)q + 1-d = 0. \quad (15)$$

Выполнение спектрального условия устойчивости (12) равносильно тому, что оба корня этого уравнения лежат в замкнутом единичном круге $|q| \leq 1$ на $\mathbb{C}$ и на его границе нет кратных корней.

Для сравнения изучим сначала трёхслойную симметричную схему

$$\delta_t v + b_i \delta_i v = 0 \quad \text{на } \omega_h \times \omega^{h_t}, \quad (16)$$

получающуюся из схемы (11) при $\tau = 0$.

Теорема 3. Для разностной схемы (16) критерий справедливости спектрального условия равномерной по времени устойчивости имеет вид условия типа Куранта

$$h_t \sum_{k=1}^n \frac{|b_k|}{h_k} < 1. \quad (17)$$

Доказательство. При $\tau = 0$ из схемы (13) следует иное, чем (14), рекуррентное уравнение

$$\hat{w} + 2ih_t \gamma w - \check{w} = 0, \quad \gamma \equiv \gamma(\xi) = \frac{b_i}{h_i} \sin \xi_i.$$

Соответствующее характеристическое уравнение – это $q^2 + 2ih_t \gamma q - 1 = 0$. Оно имеет корни

$$q_{1,2} = \begin{cases} -ih_t \gamma \pm \sqrt{1 - h_t^2 \gamma^2}, & \text{если } h_t |\gamma| \leq 1, \\ -i(h_t \gamma \pm \sqrt{h_t^2 \gamma^2 - 1}), & \text{если } h_t |\gamma| > 1. \end{cases}$$

В случае $h_t |\gamma| > 1$ максимальный из $|q_{1,2}|$ равен $h_t |\gamma| + \sqrt{h_t^2 \gamma^2 - 1} > 1$. В случае $h_t |\gamma| \leq 1$ имеем $|q_{1,2}| = 1$, но при $h_t |\gamma| = 1$ корень становится кратным. Это приводит к условию $h_t |\gamma| < 1$ при всех $\xi \in [-\pi, \pi]^n$ и в конечном итоге (при $\xi = ((\pi/2) \operatorname{sgn} b_1, \dots, (\pi/2) \operatorname{sgn} b_n)$) – к условию (17). Теорема доказана.

Перейдём теперь к анализу разностной схемы (11).

Теорема 4. Пусть $a_i > 0$ при $1 \leq i \leq n$. Для разностной схемы (11) справедливость спектрального условия её равномерной по времени устойчивости не зависит от значения параметра $\tau > 0$.

Для его справедливости необходимо выполнение двух условий типа Куранта на h_t :

$$\frac{h_t}{\sqrt{\alpha_0}} \left(\sum_{k=1}^n \frac{a_k}{h_k^2} \right)^{1/2} \leq 1, \quad h_t \sum_{k=1}^n \frac{|b_k|}{h_k} \leq 1, \quad (18)$$

второе из которых почти совпадает с условием (17). Необходимым является также условие

$$\alpha_0 \sum_{k=1}^n \frac{b_k^2}{a_k} \leq 1 \quad (19)$$

(ср. с условиями (9)).

Для справедливости указанного спектрального условия достаточно выполнение условия

$$\sum_{k=1}^n \max \left\{ \frac{h_t^2}{\alpha_0} \frac{a_k}{h_k^2}, \alpha_0 \frac{b_k^2}{a_k} \right\} \leq 1. \quad (20)$$

При $n = 1$ оно принимает вид

$$\frac{h_t}{\sqrt{\alpha_0}} \frac{\sqrt{a_1}}{h_1} \leq 1, \quad \alpha_0 b_1^2 \leq a_1 \quad (21)$$

и служит критерием.

Замечание 1. Прокомментируем условия из этой теоремы. Второе из условий (18) следует из первого условия и условия (19) в силу неравенства Коши–Шварца:

$$\left(h_t \sum_{k=1}^n \frac{|b_k|}{h_k} \right)^2 \leq \left(\alpha_0 \sum_{k=1}^n \frac{b_k^2}{a_k} \right) \frac{h_t^2}{\alpha_0} \sum_{k=1}^n \frac{a_k}{h_k^2}.$$

Как будет доказано, необходимым является и несколько более точное условие

$$h_t \sum_{k=1}^n \frac{|b_k|}{h_k} \leq \left[\left(2 - \frac{h_t^2}{\alpha_0} \sum_{k=1}^n \frac{a_k}{h_k^2} \right) \frac{h_t^2}{\alpha_0} \sum_{k=1}^n \frac{a_k}{h_k^2} \right]^{1/2}, \quad (22)$$

правая часть которого в силу первого из условий (18) не превосходит 1.

Более грубым достаточным условием, чем (20), очевидно, служит неравенство

$$\frac{h_t^2}{\alpha_0} \sum_{k=1}^n \frac{a_k}{h_k^2} + \alpha_0 \sum_{k=1}^n \frac{b_k^2}{a_k} \leq 1$$

(ср. с первым условием (18) и условием (19)).

Замечание 2. Можно использовать 1-нормы векторов коэффициентов $a := (a_1, \dots, a_n)$ и b и ввести при $a \neq 0$ и $b \neq 0$ специальные средние шаги по пространству $h_a > 0$ и $h_b > 0$ формулами

$$|a|_1 := \sum_{k=1}^n a_k, \quad |b|_1 := \sum_{k=1}^n |b_k|, \quad h_a := |a|_1^{1/2} \left(\sum_{k=1}^n \frac{a_k}{h_k^2} \right)^{-1/2}, \quad h_b := |b|_1 \left(\sum_{k=1}^n \frac{|b_k|}{h_k} \right)^{-1}.$$

При этом $h_a, h_b \in [h_{\min}, h_{\max}]$, где $h_{\min} = \min_{1 \leq k \leq n} h_k$, $h_{\max} = \max_{1 \leq k \leq n} h_k$. Тогда условия типа Куранта (18) примут следующий вид:

$$\sqrt{\frac{|a|_1}{\alpha_0}} \frac{h_t}{h_a} \leq 1, \quad |b|_1 \frac{h_t}{h_b} \leq 1.$$

Доказательство. Если корни уравнения (15) кратные, то их модуль равен $\sqrt{|1-d|/(1+d)}$ и меньше 1. Если $\theta = 0$, то корнями многочлена $p_2(q)$ являются 1 и $(1-d)/(1+d)$ (последний по модулю, очевидно, меньше 1). Следовательно, ниже можно считать, что $\theta \neq 0$ и поэтому $p_2(1) = 2\theta \neq 0$, и что эти корни простые.

Выполним дробно-рациональное преобразование $q(z) = (z+1)/(z-1)$, устанавливающее взаимно однозначное соответствие между проколотым замкнутым единичным кругом $\{q \in \mathbb{C}: |q| \leq 1, q \neq 1\}$ и замкнутой левой полуплоскостью $\bar{P}_l$. Соответствие между проколотой единичной окружностью $\{q \in \mathbb{C}: |q| = 1, q \neq 1\}$ и мнимой осью $z_R = 0$ также взаимно однозначное; отметим, что обратная функция $z^{(-1)}(q) = (q+1)/(q-1)$ совпадает с исходной. При этом

$$p_2(q(z))(z-1)^2 = 2r_2(z) := 2(\theta z^2 + 2dz + 2 - \theta).$$

Здесь $r_2(1) = 2(1+d) > 0$. Обозначим через z_1, z_2 корни многочлена $r_2(z)$ и получим условия их принадлежности полуплоскости $\bar{P}_l$.

Пусть сначала $\theta_I = 0$. При $\theta_R \neq 0$, как хорошо известно, $z_1, z_2 \in \bar{P}_l$ при следующих условиях на коэффициенты многочлена: $d/\theta_R \geq 0$ и $(2-\theta_R)/\theta_R \geq 0$. Так как $d > 0$, то это приводит к условию $0 < \theta_R \leq 2$ (причём $z_1, z_2 \in P_l$ при $\theta_R < 2$, либо $z_1 = -d < 0$ и $z_2 = 0$ при $\theta_R = 2$).

Пусть теперь $\theta_I \neq 0$ и поэтому $\xi \neq 0$. Снова запишем вспомогательный многочлен

$$-r_2(iz) = \theta_R z^2 + \theta_R - 2 + i(\theta_I z^2 - 2dz + \theta_I) \quad (23)$$

и определитель 4-го порядка – результат (с точностью до знака) двух многочленов с вещественными коэффициентами в правой части

$$V_4 = \begin{vmatrix} \theta_I & -2d & \theta_I & 0 \\ \theta_R & 0 & \theta_R - 2 & 0 \\ 0 & \theta_I & -2d & \theta_I \\ 0 & \theta_R & 0 & \theta_R - 2 \end{vmatrix}.$$

Его главный угловой минор второго порядка $V_2 = 2d\theta_R$ положителен при $\xi \neq 0$. Умножим 2-ю и 4-ю строки V_4 на θ_I/θ_R , вычтем их из 1-й и 3-й строк соответственно и разложим полученный определитель по первому столбцу:

$$V_4 = -\theta_R \begin{vmatrix} -2d & -2\theta_I/\theta_R & 0 \\ \theta_I & -2d & -2\theta_I/\theta_R \\ \theta_R & 0 & \theta_R - 2 \end{vmatrix} = -4\theta_R \left[d^2(\theta_R - 2) + \frac{\theta_I^2}{\theta_R} \right] = -4[d^2\theta_R(\theta_R - 2) + \theta_I^2].$$

(Несложно видеть, что итоговая формула верна и при $\theta_R = 0$). Напомним, что при $V_4 \neq 0$ критерием того, что корни $r_2(z)$ лежат в P_l , является условие $V_4 > 0$. При $V_4 = 0$ имеется чисто мнимый корень $z_1 = iy_1$. Он является общим корнем двух вещественных многочленов в правой части (23), откуда легко получается, что $y_1 = \theta_I/(d\theta_R)$. Тогда $z_2 = -iy_1^{-1}((2/\theta) - 1)$ и поэтому $z_{2,R} = -d/(\theta_R|\theta|^2) < 0$, т.е. $z_2 \in P_l$. Следовательно, критерием того, что оба корня лежат в $\bar{P}_l$ и на мнимой оси нет кратных корней, является условие $V_4 \geq 0$, т.е.

$$\theta_R^2 + (\theta_I/d)^2 \leq 2\theta_R \quad \text{для всех } \xi \in [-\pi, \pi]^n. \quad (24)$$

В соответствии с предыдущим анализом это условие справедливо и при $\theta_I = 0$. Полученный критерий геометрически означает принадлежность θ эллипсу на $\mathbb{C}$:

$$(\theta_R - 1)^2 + (\theta_I/d)^2 \leq 1 \quad \text{для всех } \xi \in [-\pi, \pi]^n. \quad (25)$$

Поскольку $\theta_I/d = b_i(h_t/h_i) \sin \xi_i$, то параметр τ не участвует в выведенном критерии.

Очевидно, что необходимым условием справедливости критерия (25) служат неравенства

$$0 \leq \theta_R \leq 2, \quad |\theta_I|/d \leq 1 \quad \text{для всех } \xi \in [-\pi, \pi]^n$$

(условия принадлежности θ прямоугольнику, объемлющему указанный эллипс). Максимум величины $\theta_R \geq 0$, очевидно, достигается при $\xi = (\pi, \dots, \pi)$, а максимум величины $|\theta_I|/d$ – при $\xi = ((\pi/2) \operatorname{sgn} b_1, \dots, (\pi/2) \operatorname{sgn} b_n)$. Это и приводит к условиям (18).

Положив $\xi = ((\pi/2) \operatorname{sgn} b_1, \dots, (\pi/2) \operatorname{sgn} b_n)$ в (24), приходим к неравенству

$$\frac{|\tilde{\theta}_I|}{d} = h_t \sum_{k=1}^n \frac{|b_k|}{h_k} \leq \sqrt{(2 - \tilde{\theta}_R) \tilde{\theta}_R} \quad \text{с} \quad \tilde{\theta}_R = \frac{h_t^2}{\alpha_0} \sum_{k=1}^n \frac{a_k}{h_k^2},$$

т.е. к необходимому условию (22).

Возьмём также $\xi_k = \varepsilon b_k h_k / (a_k h_t)$, $k = \overline{1, n}$, в критерии (24) и, разложив обе его части по степеням $\varepsilon \rightarrow 0$, получим

$$O(\varepsilon^4) + \varepsilon^2 \left(\sum_{k=1}^n \frac{b_k^2}{a_k} \right)^2 \leq \frac{\varepsilon^2}{\alpha_0} \sum_{k=1}^n \frac{b_k^2}{a_k} + O(\varepsilon^4).$$

Разделив обе части на ε^2 и перейдя к пределу при $\varepsilon \rightarrow 0$, выведем необходимое условие (19).

С другой стороны, имеем $|\sin \xi_k| = 2\sqrt{\sigma_k(1 - \sigma_k)}$, где $\sigma_k := \sin^2(\xi_k/2)$ пробегает отрезок $[0, 1]$, и по неравенству Коши–Шварца верна оценка

$$\left(\frac{\theta_I}{d} \right)^2 \leq 2\theta_R \alpha_0 \frac{b_i^2}{a_i} (1 - \sigma_i).$$

Поэтому после сокращения на $2\theta_R \neq 0$ (что эквивалентно $\sigma := (\sigma_1, \dots, \sigma_n) \neq 0$) заключаем, что критерий (24) следует из неравенства

$$\frac{a_i}{\alpha_0} \left(\frac{h_t}{h_i} \right)^2 \sigma_i + \alpha_0 \frac{b_i^2}{a_i} (1 - \sigma_i) \leq 1 \quad \text{для всех} \quad \sigma \in [0, 1]^n, \quad \sigma \neq 0. \quad (26)$$

Его левая часть является аффинной функцией σ простого вида, и поэтому, как нетрудно видеть, это неравенство эквивалентно неравенству (20), тем самым являющемуся достаточным условием справедливости критерия.

При $n = 1$ критерий (24) после деления на $2\theta_R$ принимает вид (26), и поэтому условия (21) становятся критерием.

Замечание 3. Отметим дополнительно, что в случае коэффициентов a_i произвольных знаков ситуация в теореме 4 становится аналогичной указанной в критерии (9). Во-первых, в доказательстве теоремы 4 при $\theta_I = 0$ показано, что условие $\theta_R \geq 0$ необходимо. При $\theta_I \neq 0$ и $\theta_R < 0$ имеем $V_2 < 0$ и $V_4 < 0$, значит, один из корней $r_2(z)$ лежит в правой полуплоскости $z_R > 0$. Поэтому и при $\theta_I \neq 0$ необходимо, чтобы $\theta_R \geq 0$ при всех $\xi \in [-\pi, \pi]^n$. Последнее влечёт за собой необходимость условия $a_i \geq 0$, $1 \leq i \leq n$. Во-вторых, если $a_k = 0$ при некотором k , то возьмём $\xi_i = 0$ при $i \neq k$, а $\sin \xi_k \neq 0$. Тогда $\theta_R = 0$ и $\theta_I/d = b_k(h_t/h_k) \sin \xi_k$. При $b_k \neq 0$ корни $r_2(z) = i(\theta_I z^2 - 2idz - \theta_I - 2i)$ таковы, что $z_1 + z_2 = 2id/\theta_I$, откуда $z_{1R} + z_{2R} = 0$, и при $z_1, z_2 \in \bar{P}_l$ имеем $z_{1R} = z_{2R} = 0$. Но тогда должно быть вещественным произведение $z_1 z_2 = -1 - (2i/\theta_I)$, что не так. Значит, при $a_k = 0$ условие $b_k = 0$ необходимо.

Поэтому теорема 4 сохраняет силу и при $a_i \geq 0$, $1 \leq i \leq n$, с той лишь разницей, что суммы в условиях (19) и (20) надо брать по всем $1 \leq k \leq n$ таким, что $a_k > 0$, и иметь в виду свойство: если $a_k = 0$, то $b_k = 0$.

В частном случае $b = 0$ имеем $\theta_I = 0$, и в теореме 4 первое из условий (18) становится критерием. Оно аналогично условию устойчивости трёхслойной разностной схемы с параметром τ из [9] в случае веса $\sigma = 0$. В этой схеме в данном случае надо взять $B_h = \alpha_0 I$, $B_{1h} = I$ (здесь I – единичный оператор), $A_h = -\tau a_i \Lambda_i$, а условие её устойчивости имеет вид $h_t(\alpha_h/2) \leq \varepsilon_1$ с произвольным $0 < \varepsilon_1 < 1$, где

$$\alpha_h := \frac{1}{\sqrt{\alpha_0}} \left(\sum_{k=1}^n \frac{4a_k}{h_k^2} \sin^2 \frac{\zeta_k}{2} \right)^{1/2} < \frac{2}{\sqrt{\alpha_0}} \left(\sum_{k=1}^n \frac{a_k}{h_k^2} \right)^{1/2}, \quad \zeta_k := \frac{\pi(N_k - 1)}{N_k},$$

причём N_k – число отрезков разбиения по x_k (в отличие от данной работы в [9] рассматривается начально-краевая задача), а $\sin^2(\zeta_k/2) \rightarrow 1$ и последнее неравенство переходит в равенство при $N_k \rightarrow +\infty$, $1 \leq k \leq n$. Случай $\varepsilon_1 = 1$ также возможен (с некоторым ослаблением используемых норм), см. в этой связи [17, замечание 1].

Работа выполнена при финансовой поддержке Российского научного фонда (проект 19-11-00104).

СПИСОК ЛИТЕРАТУРЫ

1. Четверушкин Б.Н. Гиперболическая квазигазодинамическая система // Мат. моделирование. 2018. Т. 30. № 2. С. 81–98.
2. Четверушкин Б.Н. Кинетические схемы и квазигазодинамическая система уравнений. М., 2004.
3. Елизарова Т.Г. Квазигазодинамические уравнения и методы расчета вязких течений. М., 2007.
4. Злотник А.А., Четверушкин Б.Н. О параболичности квазигазодинамической системы уравнений, ее гиперболической 2-го порядка модификации и устойчивости малых возмущений для них // Журн. вычислит. математики и мат. физики. 2008. Т. 48. № 3. С. 445–472.
5. Chetverushkin B.N., Zlotnik A.A. On some properties of multidimensional hyperbolic quasi-gasdynamics systems of equations // Russ. J. Math. Phys. 2017. V. 24. № 3. P. 299–309.
6. Сурначев М.Д., Тишкин В.Ф., Четверушкин Б.Н. О законах сохранения для гиперболизированных уравнений // Дифференц. уравнения. 2016. Т. 52. № 7. С. 859–865.
7. Chetverushkin B.N., Zlotnik A.A. On a hyperbolic perturbation of a parabolic initial-boundary value problem // Appl. Math. Lett. 2018. V. 83. P. 116–122.
8. Ильин А.А., Рыков Ю.Г. О близости траекторий для модельных квазигазодинамических уравнений. Линейный случай // Препринты ИПМ им. М.В. Келдыша. 2016. № 90. С. 1–14.
9. Злотник А.А., Четверушкин Б.Н. Устойчивость численных методов решения гиперболических уравнений второго порядка с малым параметром // Докл. РАН. Математика. Информатика. Процессы управления. 2020. Т. 490. № 1. С. 35–41.
10. Злотник А.А., Четверушкин Б.Н. Устойчивость неявных разностных схем для линеаризованной гиперболической квазигазодинамической системы уравнений // Дифференц. уравнения. 2020. Т. 56. № 7. С. 936–947.
11. Давыдов А.А., Четверушкин Б.Н., Шильников Е.В. Моделирование течений несжимаемой жидкости и слабосжимаемого газа на многоядерных гибридных вычислительных системах // Журн. вычислит. математики и мат. физики. 2010. Т. 50. № 12. С. 2275–2284.
12. Гантмахер Ф.Р. Теория матриц. М., 1988.
13. Злотник А.А. Оценки скорости сходимости проекционно-сеточных методов для гиперболических уравнений второго порядка // Вычислит. процессы и системы / Под ред. Г.И. Марчука. Вып. 8. М., 1991. С. 116–167.
14. Zlotnik A., Čiegis R. On properties of compact 4th order finite-difference schemes for the variable coefficient wave equation // Preprint. 2021. arXiv.org:2101.10575v2. P. 1–25.
15. Галанин М.П. Численное решение уравнения переноса // Будущее прикладной математики. Лекции для молодых исследователей / Под ред. Г.Г. Малинецкого. М., 2004. С. 78–116.
16. Годунов С.К., Рябенький В.С. Разностные схемы. М., 1977.
17. Zlotnik A., Kireeva O. On compact 4th order finite-difference schemes for the wave equation // Preprint. arXiv.org:2011.14104v2. P. 1–22.

Национальный исследовательский университет
“Высшая школа экономики”, г. Москва,
Институт прикладной математики
им. М.В. Келдыша РАН, г. Москва

Поступила в редакцию 08.02.2021 г.
После доработки 08.02.2021 г.
Принята к публикации 27.04.2021 г.

ЧИСЛЕННЫЕ МЕТОДЫ

УДК 519.644.5+517.956.224

КВАДРАТУРНАЯ ФОРМУЛА ДЛЯ ГАРМОНИЧЕСКОГО ПОТЕНЦИАЛА ДВОЙНОГО СЛОЯ

© 2021 г. П. А. Крутицкий, И. О. Резниченко

Выводится квадратурная формула для гармонического потенциала двойного слоя с непрерывной плотностью, заданной на гладкой поверхности (замкнутой или разомкнутой). Полученная квадратурная формула даёт более высокую точность вычислений, чем стандартная квадратурная формула, что подтверждается численными тестами. Преимущество новой квадратурной формулы особенно заметно вблизи поверхности, где значения, подсчитанные по стандартной квадратурной формуле, стремятся к бесконечности, тогда как новая формула обеспечивает приемлемую точность вычислений.

DOI: 10.31857/S0374064121070074

Введение. Гармонический потенциал двойного слоя используется при решении краевых задач для уравнения Лапласа методом интегральных уравнений. Такие задачи возникают в различных областях математической физики, например, в теории обтекания препятствий потоком идеальной жидкости, в электростатике, в стационарной теплопроводности, в теории фильтрации и т.д. Численное решение краевых задач с помощью потенциала двойного слоя обсуждалось в монографиях [1–3] и состоит из двух этапов. На первом этапе, численно решая граничное интегральное уравнение, находят плотность потенциала. На втором этапе, подставляя численное значение плотности в квадратурную формулу, находят решение краевой задачи в любой точке области. Однако квадратурные формулы для потенциалов, используемые в инженерных расчётах [4, гл. 2], не дают равномерной аппроксимации потенциала в области и не сохраняют свойство непрерывности потенциалов вплоть до границы области. Более того, вблизи некоторых точек на границе области потенциалы, подсчитанные по этим квадратурным формулам, стремятся к бесконечности, хотя сами потенциалы ограничены вблизи границы. При использовании стандартных квадратурных формул для повышения точности приходится либо уменьшать шаг, либо проводить дополнительные построения вблизи границы, что увеличивает стоимость вычислений. Актуальной является задача по получению улучшенных квадратурных формул, обеспечивающих повышенную точность вблизи границы.

В двумерном случае улучшенная квадратурная формула для потенциала простого слоя с плотностью, заданной на разомкнутых кривых и имеющей степенные особенности на их концах, построена в [5, 6]. Эта формула может применяться при нахождении численных решений краевых задач для уравнения Лапласа вне разрезов и разомкнутых кривых на плоскости с использованием метода потенциалов и граничных интегральных уравнений. Такие задачи изучались указанным методом в работах [7–9]. В [10] предложена улучшенная квадратурная формула для потенциала простого слоя в трёхмерном случае, обеспечивающая равномерную аппроксимацию и равномерную сходимость в области. Кроме того, эта формула сохраняет свойство непрерывности потенциала простого слоя при переходе через границу области.

В настоящей работе построена улучшенная квадратурная формула для гармонического потенциала двойного слоя. Формула обеспечивает более высокую точность вычислений, чем стандартная квадратурная формула. Преимущество улучшенной формулы по сравнению со стандартной особенно заметно на небольших расстояниях от границы.

1. Постановка задачи. Введём в пространстве трёхмерную декартову систему координат $x = (x_1, x_2, x_3) \in \mathbb{R}^3$. Пусть Γ – либо простая гладкая замкнутая поверхность класса C^2 , либо простая гладкая ограниченная разомкнутая ориентированная поверхность класса C^2 , содержащая свои предельные точки [11, гл. 14, § 1]. Если поверхность Γ замкнутая, то она

должна ограничивать объёмно-односвязную внутреннюю область [12, с. 201]. Предположим, что поверхность Γ параметризована так, что на неё отображается прямоугольник:

$$y = (y_1, y_2, y_3) \in \Gamma, \quad y_1 = y_1(u, v), \quad y_2 = y_2(u, v), \quad y_3 = y_3(u, v); \quad u \in [0, A], \quad v \in [0, B];$$

$$y_j(u, v) \in C^2([0, A] \times [0, B]), \quad j = 1, 2, 3. \quad (1)$$

Сферу, поверхность эллипсоида, гладкие поверхности фигур вращения, поверхность тора и многие другие более сложные поверхности можно параметризовать таким образом.

Введём N точек u_n с шагом h на отрезке $[0, A]$ и M точек v_m с шагом H на отрезке $[0, B]$ равенствами

$$u_n = (n + 1/2)h, \quad n = \overline{0, N-1} \quad \text{и} \quad v_m = (m + 1/2)H, \quad m = \overline{0, M-1}; \quad A = Nh, \quad B = MH.$$

Прямоугольник $[0, A] \times [0, B]$, который отображается на поверхность Γ , разобьём на NM маленьких прямоугольничков размера $h \times H$, тогда точки (u_n, v_m) являются серединами этих прямоугольничков.

Известно [11, гл. 14, § 1], что компоненты вектора нормали (не обязательно единичного) $\eta(y) = (\eta_1(y), \eta_2(y), \eta_3(y))$ в точке поверхности $y = (y_1, y_2, y_3) \in \Gamma$ выражаются через определители второго порядка по формулам

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

Положим $|\eta(y)| = \sqrt{(\eta_1(y))^2 + (\eta_2(y))^2 + (\eta_3(y))^2}$. Известно [11, гл. 14, §§ 1, 2], что

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

Потребуем, чтобы выполнялось неравенство

$$|\eta(y(u, v))| > 0 \quad \text{для всех} \quad (u, v) \in (0, A) \times (0, B). \quad (3)$$

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

Через $\mathbf{n}_y$ обозначим единичную нормаль в точке $y \in \Gamma$, т.е. $\mathbf{n}_y = \eta(y)/|\eta(y)|$. Производная вдоль нормали $\mathbf{n}_y$ имеет вид

$$\frac{\partial}{\partial \mathbf{n}_y} = |\eta(y)|^{-1} (\eta(y), \nabla_y).$$

Обозначим $|x - y(u, v)| = \sqrt{(x_1 - y_1(u, v))^2 + (x_2 - y_2(u, v))^2 + (x_3 - y_3(u, v))^2}$ и заметим, что

$$\frac{\partial}{\partial \mathbf{n}_y} |x - y| = \frac{1}{|\eta(y)|} \sum_{j=1}^3 \eta_j(y) \frac{y_j - x_j}{|x - y|} = \frac{(\eta(y), y - x)}{|\eta(y)| |x - y|},$$

$$\frac{\partial}{\partial \mathbf{n}_y} \frac{1}{|x - y|} = - \frac{(\eta(y), y - x)}{|\eta(y)| |x - y|^3} = \frac{-1}{|\eta(y)| |x - y|^3} \sum_{j=1}^3 \eta_j(y) (y_j - x_j).$$

Гармонический потенциал двойного слоя используется при решении краевых задач для уравнения Лапласа методом интегральных уравнений. Пусть $x \notin \Gamma$. Рассмотрим гармонический потенциал двойного слоя с заданной на поверхности Γ плотностью $\mu(y) \in C^0(\Gamma)$:

$$\mathcal{W}_0[\mu](x) = \frac{1}{4\pi} \int_{\Gamma} \mu(y) \frac{\partial}{\partial \mathbf{n}_y} \frac{1}{|x - y|} ds_y = \frac{1}{4\pi} \int_0^A du \int_0^B dv \mu(y(u, v)) |\eta(y(u, v))| \frac{\partial}{\partial \mathbf{n}_y} \frac{1}{|x - y(u, v)|} =$$

$$= -\frac{1}{4\pi} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \int_{u_n-h/2}^{u_n+h/2} du \int_{v_m-H/2}^{v_m+H/2} dv \frac{\mu(y(u,v))}{|x-y(u,v)|^3} \sum_{j=1}^3 \eta_j(y(u,v))(y_j(u,v) - x_j). \quad (4)$$

Пусть $\mu_{nm} = \mu(y(u_n, v_m))$, тогда

$$\mu(y(u, v)) = \mu_{nm} + o(1) \quad (5)$$

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

$$\mathcal{W}_0[\mu](x) \approx -\frac{1}{4\pi} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \mu_{nm} \int_{u_n-h/2}^{u_n+h/2} du \int_{v_m-H/2}^{v_m+H/2} dv \frac{1}{|x-y(u,v)|^3} \sum_{j=1}^3 \eta_j(y(u,v))(y_j(u,v) - x_j). \quad (6)$$

Таким образом, чтобы получить квадратурную формулу для гармонического потенциала двойного слоя при $x \notin \Gamma$, необходимо вычислить интеграл

$$\int_{u_n-h/2}^{u_n+h/2} du \int_{v_m-H/2}^{v_m+H/2} dv \frac{1}{|x-y(u,v)|^3} \sum_{j=1}^3 \eta_j(y(u,v))(y_j(u,v) - x_j), \quad (7)$$

который будем называть *каноническим интегралом*.

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

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

где $D_j = (y_j)'_u(u - u_n) + (y_j)'_v(v - v_m)$. Здесь и далее все производные по u и v берутся в точке (u_n, v_m) . Положим

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

тогда

$$y_j(u, v) - x_j = r_j + D_j + O(H^2 + h^2), \quad j = 1, 2, 3.$$

Следовательно,

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

где $U = u - u_n$, $V = v - v_m$,

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

Производные по u и v берутся в точке $u = u_n$, $v = v_m$. Несложно показывается [11, гл. 14, § 1], что справедливо равенство

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

Так как, согласно условию (3), $|\eta(y(u_n, v_m))| > 0$ для всех возможных n и m , то

$$\alpha^2 \beta^2 - \delta^2 > 0. \quad (8)$$

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

Применяя формулу Тейлора с центром разложения в точке (u_n, v_m) с остаточным членом в форме Пеано [11, гл. 10, § 5.3], находим

$$\eta_j(y(u, v)) = \eta_j(y(u_n, v_m)) + (\eta_j)'_u(u - u_n) + (\eta_j)'_v(v - v_m) + o(\sqrt{(u - u_n)^2 + (v - v_m)^2}).$$

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

Для вычисления выражения

$$\sum_{j=1}^3 \eta_j(y(u, v))(y_j(u, v) - x_j)$$

с учётом формул

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

означающих ортогональность вектора нормали и касательных векторов к поверхности (см. [11, гл. 14, § 1.2]), воспользуемся разложением по формуле Тейлора с центром разложения в точке (u_n, v_m) с остаточным членом в форме Пеано

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

тогда

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

где $U = u - u_n$, $V = v - v_m$,

$$\xi_1 = \sum_{j=1}^3 \left(\frac{1}{2} \eta_j(y(u_n, v_m))(y_j)''_{uu} + (\eta_j)'_u(y_j)'_u \right), \quad \xi_2 = \sum_{j=1}^3 \left(\frac{1}{2} \eta_j(y(u_n, v_m))(y_j)''_{vv} + (\eta_j)'_v(y_j)'_v \right),$$

$$\xi_3 = \sum_{j=1}^3 \left(\eta_j(y(u_n, v_m))(y_j)''_{uv} + (\eta_j)'_u(y_j)'_v + (\eta_j)'_v(y_j)'_u \right),$$

$$\xi_4 = \sum_{j=1}^3 (\eta_j)'_u r_j, \quad \xi_5 = \sum_{j=1}^3 (\eta_j)'_v r_j, \quad R = \sum_{j=1}^3 \eta_j(y(u_n, v_m)) r_j.$$

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

Из приведённых соотношений вытекает, что канонический интеграл (7) приближённо равен следующему интегралу, который обозначим через $K_{nm}(x)$:

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

Следовательно, чтобы вывести квадратурную формулу для потенциала двойного слоя, необходимо вычислить интеграл $K_{nm}(x)$ в явном виде.

2.1. Вычисление интегралов по dV . Введём обозначения

$$z = V + \delta U/\beta^2 + Q/\beta^2 = V + c, \quad c = \delta U/\beta^2 + Q/\beta^2, \quad (9)$$

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

Как показано выше, из неравенства (8) вытекает, что $\alpha^2 > 0$ и $\beta^2 > 0$. Кроме того, из неравенства (8) следует, что $a \neq 0$, поскольку a является квадратным трёхчленом по U , в котором коэффициент при U^2 положителен: $(\alpha^2 \beta^2 - \delta^2)/\beta^4 > 0$.

Покажем, что $a \geq 0$. Положим $\tilde{D}_j = (y_j)'_u U - (y_j)'_v c$, где c определено в (9), и проведём следующие преобразования с учётом введённых обозначений:

$$\begin{aligned} \sum_{j=1}^3 (r_j + \tilde{D}_j)^2 &= \sum_{j=1}^3 (r_j^2 + 2r_j \tilde{D}_j + \tilde{D}_j^2) = r^2 + 2 \sum_{j=1}^3 r_j \tilde{D}_j + \sum_{j=1}^3 \tilde{D}_j^2 = \\ &= r^2 + 2PU - 2Qc + \alpha^2 U^2 + \beta^2 c^2 - 2\delta U c = \\ &= \beta^2 \left(c^2 - 2 \frac{\delta U + Q}{\beta^2} c + \left(\frac{\delta U + Q}{\beta^2} \right)^2 \right) - \frac{(\delta U + Q)^2}{\beta^2} + \alpha^2 U^2 + 2PU + r^2 = \\ &= \beta^2 \left(-c + \frac{\delta U + Q}{\beta^2} \right)^2 - \frac{(\delta U + Q)^2}{\beta^2} + \alpha^2 U^2 + 2PU + r^2 = \\ &= -\frac{(\delta U + Q)^2}{\beta^2} + \alpha^2 U^2 + 2PU + r^2 = a\beta^2 \geq 0. \end{aligned} \quad (10)$$

Так как $\beta^2 > 0$, то, поделив полученное неравенство на β^2 , заключаем, что $a \geq 0$. Следовательно, квадратный трёхчлен a неотрицательный.

Обозначим

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

Применяя введённые обозначения, вычислим в $K_{nm}(x)$ интеграл по dV , переходя к переменной z :

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

где использованы интегралы 1.2.43.17–1.2.43.19 из справочника [13]. Заметим, что

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

где функция $\theta_{nm}(x)$ найдена в явном виде в работе [10]. Тогда интеграл $K_{nm}(x)$ можно представить в виде

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

Так как функция $\theta_{nm}(x)$ найдена в [10], то задача сводится к вычислению интеграла

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

Достаточно вычислить интеграл $J(H)$. Интеграл $J(-H)$ вычисляется по тем же формулам, что и интеграл $J(H)$, с заменой в них H на $-H$.

Вычислим интеграл $J(H)$. Распишем величины, входящие в подынтегральную функцию, в виде многочленов по U :

$$\begin{aligned} a &= C_2 U^2 + C_1 U + C_0, \quad C_2 = (\alpha^2 - \delta^2/\beta^2)/\beta^2, \quad C_1 = (2P - 2\delta Q/\beta^2)/\beta^2, \quad C_0 = (r^2 - Q^2/\beta^2)/\beta^2; \\ z_+^2 + a &= B_2 U^2 + B_1 U + B_0, \quad B_2 = \alpha^2/\beta^2, \quad B_1 = (H\delta + 2P)/\beta^2, \quad B_0 = H^2/4 + (HQ + r^2)/\beta^2; \\ b &= A_2 U^2 + A_1 U + A_0, \quad A_2 = \xi_1 - \xi_3 \delta/\beta^2 + \xi_2 \delta^2/\beta^4, \quad A_1 = \xi_4 - \xi_5 \delta/\beta^2 - \xi_3 Q/\beta^2 + 2\xi_2 \delta Q/\beta^4, \\ & \quad A_0 = R - \xi_5 Q/\beta^2 + \xi_2 Q^2/\beta^4; \\ bz_+ &= (A_2 U^2 + A_1 U + A_0)(\delta U/\beta^2 + H/2 + Q/\beta^2) = E_3 U^3 + E_2 U^2 + E_1 U + E_0, \quad E_3 = A_2 \delta/\beta^2, \\ E_2 &= A_2(H/2 + Q/\beta^2) + A_1 \delta/\beta^2, \quad E_1 = A_1(H/2 + Q/\beta^2) + A_0 \delta/\beta^2, \quad E_0 = A_0(H/2 + Q/\beta^2); \\ \xi_2 z_+ + \xi_3 U + \xi_5 - 2\xi_2 c &= F_1 U + F_0, \quad F_1 = \xi_3 - \xi_2 \delta/\beta^2, \quad F_0 = \xi_5 - \xi_2 Q/\beta^2 + \xi_2 H/2. \end{aligned}$$

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

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

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

$$J_2 = \int_{-h/2}^{h/2} dU \frac{F_1 U + F_0}{\sqrt{B_2 U^2 + B_1 U + B_0}}. \quad (11)$$

2.2. Вычисление интегралов (11). Используя деление многочленов и учитывая, что $C_2 > 0$ в силу (8), приведём интеграл J_1 к виду

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

где

$$J_{11} = \int_{-h/2}^{h/2} dU \frac{L_1 U + L_0}{\sqrt{B_2 U^2 + B_1 U + B_0}},$$

$$J_{12} = \int_{-h/2}^{h/2} dU \frac{S_1 U + S_0}{(C_2 U^2 + C_1 U + C_0) \sqrt{B_2 U^2 + B_1 U + B_0}}, \quad (12)$$

здесь

$$L_1 = \frac{E_3}{C_2}, \quad L_0 = \frac{E_2 C_2 - E_3 C_1}{C_2^2}, \quad S_1 = E_1 - C_0 L_1 - C_1 L_0, \quad S_0 = E_0 - C_0 L_0.$$

Если $B_1^2 - 4B_2B_0 = 0$ и $-B_1/(2B_2) \in [-h/2, h/2]$, то в силу [10, п. 2] точка x лежит на маленьком кусочке проходящей через $y(u_n, v_m)$ касательной плоскости, по которому ведётся интегрирование в каноническом интеграле после линеаризации $y(u, v)$ вблизи $y(u_n, v_m)$. В данном случае в знаменателе канонического интеграла с такой линеаризацией возникает особенность в точке x . Однако если в числителе провести такую же линеаризацию, а нормаль приближённо заменить нормалью в точке $y(u_n, v_m)$, то числитель будет тождественно равен нулю, так как x лежит в касательной плоскости, проходящей через $y(u_n, v_m)$. С другой стороны, исходный канонический интеграл (без линеаризации) не имеет особенностей, так как $x \notin \Gamma$, и может быть оценён как $O(hH)$. Поэтому если выполняется условие $B_1^2 - 4B_2B_0 = 0$ и $-B_1/(2B_2) \in [-h/2, h/2]$, то будем считать, что $K_{nm}(x) \approx 0$. Далее предполагаем, что указанное условие не выполняется.

Так как $B_2 > 0$, интегралы J_{11} и J_2 находятся с помощью табличных интегралов 2.261 и 2.264 из справочника [14]:

$$J_{11} = \frac{L_1}{B_2} \sqrt{B_2 U^2 + B_1 U + B_0} +$$

$$+ \left(L_0 - \frac{L_1 B_1}{2B_2} \right) \frac{1}{\sqrt{B_2}} \ln \left| 2B_2 U + B_1 + 2\sqrt{B_2} \sqrt{B_2 U^2 + B_1 U + B_0} \right| \Big|_{U=-h/2}^{U=h/2},$$

$$J_2 = \frac{F_1}{B_2} \sqrt{B_2 U^2 + B_1 U + B_0} +$$

$$+ \left(F_0 - \frac{F_1 B_1}{2B_2} \right) \frac{1}{\sqrt{B_2}} \ln \left| 2B_2 U + B_1 + 2\sqrt{B_2} \sqrt{B_2 U^2 + B_1 U + B_0} \right| \Big|_{U=-h/2}^{U=h/2}.$$

Остаётся вычислить интеграл J_{12} . Способ вычисления интеграла зависит от знака дискриминанта квадратного трёхчлена $C_2U^2 + C_1U + C_0$, стоящего в знаменателе подынтегральной функции.

Первый случай: $C_1^2 - 4C_2C_0 > 0$. В п. 2.1 показано, что квадратный трёхчлен, обозначенный через a , неотрицательный; следовательно, его дискриминант неположительный: $C_1^2 - 4C_2C_0 \leq 0$, поэтому первый случай не реализуется.

Второй случай: $C_1^2 - 4C_2C_0 = 0$. В этом случае $C_2U^2 + C_1U + C_0 = C_2(U - U_1)^2$, где $U_1 = -C_1/(2C_2)$ – корень этого трёхчлена. Если $U_1 \in [-h/2, h/2]$, то в силу (10) точка x лежит на маленьком кусочке проходящей через $y(u_n, v_m)$ касательной плоскости, по которому ведётся интегрирование в каноническом интеграле после линейаризации $y(u, v)$ вблизи $y(u_n, v_m)$. В данном случае в знаменателе канонического интеграла с такой линейаризацией возникает особенность в точке x . Однако если в числителе провести такую же линейаризацию, а нормаль приближённо заменить нормалью в точке $y(u_n, v_m)$, то числитель будет тождественно равен нулю, так как x лежит в касательной плоскости, проходящей через $y(u_n, v_m)$. С другой стороны, исходный канонический интеграл (без линейаризации) не имеет особенности, так как $x \notin \Gamma$, и может быть оценён как $O(hH)$. Поэтому если $U_1 \in [-h/2, h/2]$, то будем считать, что $K_{nm}(x) \approx 0$.

Пусть $U_1 \notin [-h/2, h/2]$. Воспользовавшись равенством

$$\frac{S_1U + S_0}{(U - U_1)^2} = \frac{S_1(U - U_1) + S_0 + U_1S_1}{(U - U_1)^2} = \frac{S_1}{U - U_1} + \frac{S_0 + U_1S_1}{(U - U_1)^2},$$

получим

$$J_{12} = \int_{-h/2}^{h/2} \frac{S_1 dU}{C_2(U - U_1)\sqrt{B_2U^2 + B_1U + B_0}} + \\ + \frac{S_0 + U_1S_1}{C_2} \int_{-h/2}^{h/2} \frac{dU}{(U - U_1)^2\sqrt{B_2U^2 + B_1U + B_0}}.$$

Сделав в интеграле J_{12} замену $t = 1/(U - U_1)$, будем иметь [12, ч. 1, гл. 7, § 10, п. 5]:

$$t_1 = \frac{2}{h - 2U_1}, \quad t_2 = -\frac{2}{h + 2U_1} \text{ — пределы интегрирования,} \\ J_{12} = -\frac{S_1}{C_2} \int_{t_2}^{t_1} \frac{\operatorname{sgn}(t) dt}{\sqrt{(U_1^2B_2 + U_1B_1 + B_0)t^2 + (2B_2U_1 + B_1)t + B_2}} - \\ - \frac{S_0 + U_1S_1}{C_2} \int_{t_2}^{t_1} \frac{\operatorname{sgn}(t)t dt}{\sqrt{(U_1^2B_2 + U_1B_1 + B_0)t^2 + (2B_2U_1 + B_1)t + B_2}} = \\ = -\frac{S_1}{C_2} \operatorname{sgn}(C_1) \int_{t_2}^{t_1} \frac{dt}{\sqrt{\omega_2t^2 + \omega_1t + B_2}} - \frac{S_0 + U_1S_1}{C_2} \operatorname{sgn}(C_1) \int_{t_2}^{t_1} \frac{t dt}{\sqrt{\omega_2t^2 + \omega_1t + B_2}},$$

где $\omega_1 = 2B_2U_1 + B_1$, $\omega_2 = U_1^2B_2 + U_1B_1 + B_0$. Как показано выше, $a \geq 0$, поэтому и $\omega_2 \geq 0$. Если $\omega_2 > 0$, то с помощью табличных интегралов 2.261 и 2.264 из справочника [14] находим

$$J_{12} = \frac{-S_1 \operatorname{sgn}(C_1)}{C_2\sqrt{\omega_2}} \ln |2\omega_2t + \omega_1 + 2\sqrt{\omega_2}\sqrt{\omega_2t^2 + \omega_1t + B_2}| \Big|_{t_2}^{t_1} -$$

$$-\frac{S_0 + U_1 S_1}{C_2} \operatorname{sgn}(C_1) \left(\frac{1}{\omega_2} \sqrt{\omega_2 t^2 + \omega_1 t + B_2} - \frac{\omega_1}{2\omega_2} \frac{1}{\sqrt{\omega_2}} \ln |2\omega_2 t + \omega_1 + 2\sqrt{\omega_2} \sqrt{\omega_2 t^2 + \omega_1 t + B_2}| \right) \Big|_{t_2}^{t_1}.$$

Если $\omega_2 = 0$, а $\omega_1 \neq 0$, то, пользуясь интегралами 1.2.18.5, 1.2.18.6 из справочника [13], получаем

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

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

$$J_{12} = -\frac{S_1 \operatorname{sgn}(C_1)}{C_2} \frac{1}{\sqrt{B_2}} t \Big|_{t_2}^{t_1} - \frac{S_0 + U_1 S_1}{C_2} \frac{\operatorname{sgn}(C_1)}{2\sqrt{B_2}} t^2 \Big|_{t_2}^{t_1}.$$

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

Третий случай: $C_1^2 - 4C_2 C_0 < 0$. В этом случае многочлен $C_2 U^2 + C_1 U + C_0$ неприводим над $\mathbb{R}$. Рассмотрим различные варианты вычисления интеграла J_{12} (см. (12)), воспользовавшись методом, предложенным в [12, ч. 1, гл. 7, § 10, п. 5, (7.75)] или в [14, раздел 2.25].

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

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

и

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

где

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

Вводя обозначения

$$\begin{aligned} J_{121} &= \frac{S_1}{2C_2} \int_{t=t_2}^{t=t_1} \frac{dt^2}{[t^2 + \sigma_1] \sqrt{B_2 t^2 + \sigma_2}}, \\ J_{122} &= \frac{1}{C_2 \sqrt{B_2}} \left(-\frac{S_1 C_1}{2C_2} + S_0 \right) \int_{t_2}^{t_1} \frac{dt}{[t^2 + \sigma_1] \sqrt{t^2 + \sigma_2 / B_2}}, \end{aligned}$$

получаем

$$J_{12} = J_{121} + J_{122}. \quad (13)$$

Воспользовавшись табличным интегралом 2.246 из справочника [14], находим интеграл J_{121} в явном виде:

1₁) если $\sigma_1 B_2 - \sigma_2 > 0$, то

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

2₁) если $\sigma_1 B_2 - \sigma_2 < 0$, то

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

3₁) если $\sigma_1 B_2 - \sigma_2 = 0$, то

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

Используя табличные интегралы 1.2.43.17, 1.2.45.10, 1.2.45.13 и 1.2.11.10 из справочника [13], находим интеграл J_{122} в явном виде:

1₂) если $\sigma_2 > 0$ и $\sigma_2 - B_2\sigma_1 < 0$, то

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

2₂) если $\sigma_2 - B_2\sigma_1 > 0$, то

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

3₂) если $\sigma_1 - \sigma_2/B_2 = 0$, то

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

4₂) если $\sigma_2 = 0$ и $|B_1/(2B_2)| > h/2$, то

$$J_{122} = \frac{1}{C_2\sqrt{B_2}} \left(-\frac{S_1 C_1}{2C_2} + S_0 \right) \operatorname{sgn}(B_1) \frac{1}{2\sigma_1} \ln \frac{t^2}{|t^2 + \sigma_1|} \Big|_{t_2}^{t_1}.$$

Итак, в этом варианте интеграл J_{12} вычисляется явно по формуле (13).

Вариант 2. Пусть $B_1 \neq B_2 C_1/C_2$. В [10, п. 2] показано, что в (12) квадратный трёхчлен под корнем квадратным неотрицателен (этот результат вытекает также и из приведённых в п. 2.1 выкладок), поэтому его дискриминант неположителен, т.е. $B_1^2/(4B_2^2) - B_0/B_2 \leq 0$. Положим $\chi_1^2 = B_0/B_2 - B_1^2/(4B_2^2) \geq 0$, где без нарушения общности $\chi_1 \geq 0$. Поэтому могут представиться только две возможности: $\chi_1 > 0$ и $\chi_1 = 0$. Рассмотрим их по отдельности.

Вариант 2а: $\chi_1 > 0$. Запишем квадратный трёхчлен в виде

$$B_2 U^2 + B_1 U + B_0 = B_2 \left[\left(U + \frac{B_1}{2B_2} \right)^2 + \chi_1^2 \right] = B_2 \chi_1^2 \left[\left(U + \frac{B_1}{2B_2} \right)^2 / \chi_1^2 + 1 \right] = B_2 \chi_1^2 (\operatorname{sh}^2 t + 1),$$

где сделана гиперболическая замена переменных

$$\operatorname{sh} t = \left(U + \frac{B_1}{2B_2} \right) / \chi_1, \quad U = \chi_1 \operatorname{sh} t - \frac{B_1}{2B_2}, \quad t = \operatorname{arcsch} \left[\left(U + \frac{B_1}{2B_2} \right) / \chi_1 \right].$$

Теперь рассмотрим в (12) второй квадратный трёхчлен в знаменателе (учитывая условие, которым выделяется третий случай) и линейную функцию в числителе – соответственно

$$C_2 U^2 + C_1 U + C_0 = C_2 \left(\chi_1^2 \operatorname{sh}^2 t - \frac{\chi_1 B_1}{B_2} \operatorname{sh} t + \frac{B_1^2}{4B_2^2} \right) + C_1 \chi_1 \operatorname{sh} t + C_0 - \frac{B_1 C_1}{2B_2} =$$

$$= C_2 \chi_1^2 \operatorname{sh}^2 t + \left(C_1 \chi_1 - \frac{\chi_1 B_1 C_2}{B_2} \right) \operatorname{sh} t + \frac{B_1^2 C_2}{4B_2^2} - \frac{B_1 C_1}{2B_2} + C_0 = \nu_2 \operatorname{sh}^2 t + \nu_1 \operatorname{sh} t + \nu_0 > 0$$

и

$$S_1 U + S_0 = S_1 \chi_1 \operatorname{sh} t - \frac{B_1 S_1}{2B_2} + S_0 = \varepsilon_1 \operatorname{sh} t + \varepsilon_0,$$

где $\nu_2 = C_2 \chi_1^2 > 0$, $\nu_1 = C_1 \chi_1 - \chi_1 B_1 C_2 / B_2$, $\nu_0 = B_1^2 C_2 / (4B_2^2) - B_1 C_1 / (2B_2) + C_0$, $\varepsilon_1 = S_1 \chi_1$, $\varepsilon_0 = S_0 - B_1 S_1 / (2B_2)$. Так как $\nu_2 \operatorname{sh}^2 t + \nu_1 \operatorname{sh} t + \nu_0 > 0$, то дискриминант этого квадратного относительно $\operatorname{sh} t$ трёхчлена отрицательный, т.е.

$$\nu_1^2 - 4\nu_2 \nu_0 < 0. \quad (14)$$

Кроме того, $\nu_1 \neq 0$ в силу условий, принятых в варианте 2 и в варианте 2а.

Учитывая, что $dU = \chi_1 \operatorname{ch} t dt$ и $\operatorname{sh}^2 t + 1 = \operatorname{ch}^2 t$, получаем

$$J_{12} = \frac{1}{\sqrt{B_2}} \int_{t_-}^{t_+} \operatorname{ch} t dt \frac{\varepsilon_1 \operatorname{sh} t + \varepsilon_0}{(\nu_2 \operatorname{sh}^2 t + \nu_1 \operatorname{sh} t + \nu_0) \operatorname{ch} t}, \quad t_{\pm} = \operatorname{arcsch} \left[\left(\pm \frac{h}{2} + \frac{B_1}{2B_2} \right) / \chi_1 \right].$$

Разобьём знаменатель на произведение двух линейных функций от $\operatorname{sh} t$ с комплексными коэффициентами

$$J_{12} = \frac{1}{\sqrt{B_2}} \int_{t_-}^{t_+} dt \frac{\varepsilon_1 \operatorname{sh} t + \varepsilon_0}{\nu_2 (\operatorname{sh} t - G_-)(\operatorname{sh} t - G_+)} = \frac{1}{\nu_2 \sqrt{B_2}} \int_{t_-}^{t_+} dt \left(\frac{\lambda_-}{\operatorname{sh} t - G_-} + \frac{\lambda_+}{\operatorname{sh} t - G_+} \right),$$

где G_- , G_+ – комплексные корни уравнения $\nu_2 \operatorname{sh}^2 t + \nu_1 \operatorname{sh} t + \nu_0 = 0$ относительно $\operatorname{sh} t$:

$$G_- = \frac{-\nu_1 - \sqrt{\nu_1^2 - 4\nu_2 \nu_0}}{2\nu_2} = g_1 - ig_2, \quad G_+ = \frac{-\nu_1 + \sqrt{\nu_1^2 - 4\nu_2 \nu_0}}{2\nu_2} = g_1 + ig_2, \\ g_1 = -\frac{\nu_1}{2\nu_2} \neq 0, \quad g_2 = \frac{\sqrt{|\nu_1^2 - 4\nu_2 \nu_0|}}{2\nu_2} > 0, \quad g_1, g_2 \in \mathbb{R}; \quad G_+ = \overline{G_-}. \quad (15)$$

Отметим, что $g_2 \neq 0$ в силу (14). Коэффициенты $\lambda_{\pm}$ зависят от $G_{\pm}$:

$$\lambda_- = \frac{\varepsilon_1 G_- + \varepsilon_0}{G_- - G_+} = \frac{\varepsilon_1 G_- + \varepsilon_0}{-2g_2 i}, \quad \lambda_+ = \frac{\varepsilon_1 G_+ + \varepsilon_0}{G_+ - G_-} = \frac{\varepsilon_1 G_+ + \varepsilon_0}{2g_2 i}. \quad (16)$$

Так как справедливы равенства

$$\frac{\lambda_{\pm}}{\operatorname{sh} t - G_{\pm}} = \frac{\lambda_{\pm}}{(e^t - e^{-t})/2 - G_{\pm}} \frac{e^t}{e^t} = \frac{2\lambda_{\pm} e^t}{e^{2t} - 2G_{\pm} e^t - 1},$$

то, перейдя к замене $\zeta = e^t$, $d\zeta = e^t dt$, $dt = d\zeta/\zeta$,

$$\zeta_1 = \exp \left(\operatorname{arcsch} \left[\left(-\frac{h}{2} + \frac{B_1}{2B_2} \right) / \chi_1 \right] \right), \quad \zeta_2 = \exp \left(\operatorname{arcsch} \left[\left(\frac{h}{2} + \frac{B_1}{2B_2} \right) / \chi_1 \right] \right), \quad (17)$$

будем иметь

$$J_{12} = \frac{1}{\nu_2 \sqrt{B_2}} \int_{t_-}^{t_+} dt \left(\frac{2\lambda_- e^t}{e^{2t} - 2G_- e^t - 1} + \frac{2\lambda_+ e^t}{e^{2t} - 2G_+ e^t - 1} \right) = \frac{1}{\nu_2 \sqrt{B_2}} \int_{\zeta_1}^{\zeta_2} \frac{d\zeta}{\zeta} \left(\frac{2\lambda_- \zeta}{\zeta^2 - 2G_- \zeta - 1} + \right. \\ \left. + \frac{2\lambda_+ \zeta}{\zeta^2 - 2G_+ \zeta - 1} \right) = \frac{2}{\nu_2 \sqrt{B_2}} \int_{\zeta_1}^{\zeta_2} d\zeta \left(\frac{\lambda_-}{\zeta^2 - 2G_- \zeta - 1} + \frac{\lambda_+}{\zeta^2 - 2G_+ \zeta - 1} \right).$$

В данном случае нельзя применить стандартные формулы для нахождения первообразной от вещественных подынтегральных функций, поскольку в нашем случае коэффициенты комплексные. Введём обозначения: $Z_{1-} = G_- - \sqrt{G_-^2 + 1}$, $Z_{1+} = G_- + \sqrt{G_-^2 + 1}$,

$$Z_{2-} = G_+ - \sqrt{G_+^2 + 1}, \quad Z_{2+} = G_+ + \sqrt{G_+^2 + 1}, \quad (18)$$

$$\gamma_- = \frac{1}{2\sqrt{G_-^2 + 1}}, \quad \gamma_+ = \frac{1}{2\sqrt{G_+^2 + 1}}. \quad (19)$$

Знаменатель в дробях (19) не обращается в нуль, поскольку $g_1 \neq 0$, как отмечено выше. В результате

$$\begin{aligned} J_{12} &= \frac{2}{\nu_2 \sqrt{B_2}} \int_{\zeta_1}^{\zeta_2} d\zeta \left(\frac{\lambda_-}{(\zeta - Z_{1-})(\zeta - Z_{1+})} + \frac{\lambda_+}{(\zeta - Z_{2-})(\zeta - Z_{2+})} \right) = \\ &= \frac{-2}{\nu_2 \sqrt{B_2}} \left[\lambda_- \gamma_- \int_{\zeta_1}^{\zeta_2} d\zeta \left(\frac{1}{\zeta - Z_{1-}} - \frac{1}{\zeta - Z_{1+}} \right) + \lambda_+ \gamma_+ \int_{\zeta_1}^{\zeta_2} d\zeta \left(\frac{1}{\zeta - Z_{2-}} - \frac{1}{\zeta - Z_{2+}} \right) \right]. \end{aligned}$$

Принимая во внимание выражения (15), (16), (18), (19), получаем $\lambda_- = \overline{\lambda_+}$, $Z_{1-} = \overline{Z_{2-}}$, $Z_{1+} = \overline{Z_{2+}}$, $\gamma_+ = \overline{\gamma_-}$. Тогда

$$J_{12} = \frac{-2}{\nu_2 \sqrt{B_2}} \int_{\zeta_1}^{\zeta_2} d\zeta \left[\overline{\lambda_+ \gamma_+ \left(\frac{1}{\zeta - Z_{2-}} - \frac{1}{\zeta - Z_{2+}} \right)} + \lambda_+ \gamma_+ \left(\frac{1}{\zeta - Z_{2-}} - \frac{1}{\zeta - Z_{2+}} \right) \right].$$

Так как $\zeta_1, \zeta_2 \in \mathbb{R}$, то под интегралом находится сумма двух комплексно сопряжённых функций, поэтому интеграл J_{12} вещественный:

$$J_{12} = \frac{-4}{\nu_2 \sqrt{B_2}} \operatorname{Re} \left[\lambda_+ \gamma_+ \int_{\zeta_1}^{\zeta_2} d\zeta \left(\frac{1}{\zeta - Z_{2-}} - \frac{1}{\zeta - Z_{2+}} \right) \right].$$

В итоге приходим к равенству

$$J_{12} = \frac{-4}{\nu_2 \sqrt{B_2}} \operatorname{Re} [\lambda_+ \gamma_+ (\ln(Z_{2-} - \zeta_2) - \ln(Z_{2-} - \zeta_1) - \ln(Z_{2+} - \zeta_2) + \ln(Z_{2+} - \zeta_1))]. \quad (20)$$

Логарифмы с комплексными аргументами $Z_{2\pm} - \zeta_l$, где $l = 1, 2$, преобразуются по формуле

$$\ln(Z_{2\pm} - \zeta_l) = \ln|Z_{2\pm} - \zeta_l| + i \arg(Z_{2\pm} - \zeta_l), \quad l = 1, 2. \quad (21)$$

Рассмотрим величины, входящие в обозначения (18), (19). Так как $g_2 = \operatorname{Im} G_+ > 0$, то можно считать, что $\arg G_+ \in (0, \pi)$, тогда $\arg(G_+^2 + 1) \in (0, 2\pi)$ и

$$\arg \sqrt{G_+^2 + 1} = \frac{1}{2} \arg(G_+^2 + 1) \in (0, \pi),$$

поэтому $\operatorname{Im} \sqrt{G_+^2 + 1} > 0$. Более того, если $\arg G_+ \in (0, \pi/2)$, то $\arg \sqrt{G_+^2 + 1} \in (0, \pi/2)$, а если $\arg G_+ \in (\pi/2, \pi)$, то $\arg \sqrt{G_+^2 + 1} \in (\pi/2, \pi)$. Следовательно,

$$\operatorname{sgn} \left(\operatorname{Re} \sqrt{G_+^2 + 1} \right) = \operatorname{sgn} \operatorname{Re} G_+ = \operatorname{sgn} g_1.$$

Положим

$$G_+^2 + 1 = g_1^2 - g_2^2 + 1 + 2g_1g_2i = |G_+^2 + 1| \exp(i\Psi), \quad \Psi = \arg(G_+^2 + 1),$$

$$\cos \Psi = \frac{g_1^2 - g_2^2 + 1}{|G_+^2 + 1|}, \quad |G_+^2 + 1| = \sqrt{(g_1^2 - g_2^2 + 1)^2 + 4g_1^2g_2^2}. \quad (22)$$

Используя тригонометрические формулы [15, с. 109], получаем

$$\sqrt{G_+^2 + 1} = \sqrt{|G_+^2 + 1|} \exp(i\Psi/2) = \sqrt{|G_+^2 + 1|} \left(\cos \frac{\Psi}{2} + i \sin \frac{\Psi}{2} \right),$$

$$\cos \frac{\Psi}{2} = \operatorname{sgn}(g_1) \sqrt{\frac{1 + \cos \Psi}{2}} = \operatorname{sgn}(g_1) \sqrt{\frac{|G_+^2 + 1| + g_1^2 - g_2^2 + 1}{2|G_+^2 + 1|}},$$

$$\sin \frac{\Psi}{2} = \sqrt{\frac{1 - \cos \Psi}{2}} = \sqrt{\frac{|G_+^2 + 1| - (g_1^2 - g_2^2 + 1)}{2|G_+^2 + 1|}}. \quad (23)$$

Поскольку $\arg \sqrt{G_+^2 + 1} = \Psi/2 \in (0, \pi)$, то $(-\Psi/2) \in (-\pi, 0)$. Следовательно,

$$\gamma_+ = \frac{1}{2\sqrt{G_+^2 + 1}} = \frac{\exp(-i\Psi/2)}{2\sqrt{|G_+^2 + 1|}},$$

$$\operatorname{Re} \gamma_+ = \frac{\cos(\Psi/2)}{2\sqrt{|G_+^2 + 1|}} = \operatorname{sgn}(g_1) \frac{\sqrt{|G_+^2 + 1| + g_1^2 - g_2^2 + 1}}{2\sqrt{2}|G_+^2 + 1|},$$

$$\operatorname{Im} \gamma_+ = -\frac{\sin(\Psi/2)}{2\sqrt{|G_+^2 + 1|}} = -\frac{\sqrt{|G_+^2 + 1| - (g_1^2 - g_2^2 + 1)}}{2\sqrt{2}|G_+^2 + 1|}, \quad (24)$$

где величины g_1 и g_2 определены в (15), а $|G_+^2 + 1|$ – в (22).

Лемма. Пусть $G_+ = g_1 + ig_2$, где g_1, g_2 – вещественные числа и $g_2 > 0$, тогда

$$g_2 = \operatorname{Im} G_+ > |\operatorname{Im} \sqrt{G_+^2 + 1}|.$$

Доказательство. Используя обозначения и соотношения из (22) и (23), находим

$$\operatorname{Im} \sqrt{G_+^2 + 1} = \sqrt{|G_+^2 + 1|} \sin \frac{\Psi}{2} =$$

$$= \sqrt{|G_+^2 + 1|} \sqrt{\frac{|G_+^2 + 1| - (g_1^2 - g_2^2 + 1)}{2|G_+^2 + 1|}} = \sqrt{\frac{|G_+^2 + 1| - (g_1^2 - g_2^2 + 1)}{2}}. \quad (25)$$

Так как $g_2 > 0$, то несложно проверить, что

$$|G_+^2 + 1| = \sqrt{(g_1^2 - g_2^2 + 1)^2 + 4g_1^2g_2^2} < g_1^2 + g_2^2 + 1,$$

поэтому справедлива оценка

$$\frac{1}{2}(|G_+^2 + 1| - (g_1^2 - g_2^2 + 1)) < g_2^2.$$

Из (25) вытекает равенство

$$\left(\operatorname{Im} \sqrt{G_+^2 + 1}\right)^2 = \frac{1}{2}(|G_+^2 + 1| - (g_1^2 - g_2^2 + 1)).$$

Следовательно, $(\operatorname{Im} \sqrt{G_+^2 + 1})^2 < g_2^2$, откуда $|\operatorname{Im} \sqrt{G_+^2 + 1}| < g_2$. Лемма доказана.

Следствие. Если выполнены условия леммы, то

$$\operatorname{Im}(Z_{2\pm} - \zeta_l) = \operatorname{Im}\left(G_+ \pm \sqrt{G_+^2 + 1}\right) > 0, \quad l = 1, 2.$$

Так как $Z_{2\pm} = G_+ \pm \sqrt{G_+^2 + 1}$ в силу (18), а ζ_1 и ζ_2 – вещественные числа, то

$$Z_{2\pm} - \zeta_l = \operatorname{Re}(Z_{2\pm} - \zeta_l) + i \operatorname{Im}(Z_{2\pm} - \zeta_l),$$

$$\operatorname{Re}(Z_{2\pm} - \zeta_l) = g_1 \pm \sqrt{|G_+^2 + 1|} \cos \frac{\Psi}{2} - \zeta_l, \quad \operatorname{Im}(Z_{2\pm} - \zeta_l) = g_2 \pm \sqrt{|G_+^2 + 1|} \sin \frac{\Psi}{2},$$

$$|Z_{2\pm} - \zeta_l| = \sqrt{\left(g_1 \pm \sqrt{|G_+^2 + 1|} \cos \frac{\Psi}{2} - \zeta_l\right)^2 + \left(g_2 \pm \sqrt{|G_+^2 + 1|} \sin \frac{\Psi}{2}\right)^2}, \quad l = 1, 2. \quad (26)$$

Согласно следствию к лемме справедливо неравенство $\operatorname{Im}(Z_{2\pm} - \zeta_l) > 0$, поэтому можно считать, что $\arg(Z_{2\pm} - \zeta_l) \in (0, \pi)$, где $l = 1, 2$. Следовательно,

$$\arg(Z_{2\pm} - \zeta_l) = \arccos \frac{\operatorname{Re}(Z_{2\pm} - \zeta_l)}{|Z_{2\pm} - \zeta_l|} = \arccos \frac{g_1 \pm \sqrt{|G_+^2 + 1|} \cos(\Psi/2) - \zeta_l}{|Z_{2\pm} - \zeta_l|}, \quad l = 1, 2. \quad (27)$$

В соотношениях (26), (27) используются обозначения из (15), (17), (22), (23).

Из равенств (16) вытекает, что

$$\lambda_+ = \frac{\varepsilon_1 G_+ + \varepsilon_0}{G_+ - G_-} = \frac{\varepsilon_1(g_1 + ig_2) + \varepsilon_0}{2ig_2} = \frac{\varepsilon_1}{2} - \frac{\varepsilon_1 g_1 + \varepsilon_0}{2g_2} i,$$

поэтому

$$\operatorname{Re} \lambda_+ = \frac{\varepsilon_1}{2}, \quad \operatorname{Im} \lambda_+ = -\frac{\varepsilon_1 g_1 + \varepsilon_0}{2g_2}.$$

Положим

$$\lambda_+ \gamma_+ = \Lambda_1 + i\Lambda_2, \quad \Lambda_1 = \frac{\varepsilon_1}{2} \operatorname{Re} \gamma_+ + \frac{\varepsilon_1 g_1 + \varepsilon_0}{2g_2} \operatorname{Im} \gamma_+, \quad \Lambda_2 = \frac{\varepsilon_1}{2} \operatorname{Im} \gamma_+ - \frac{\varepsilon_1 g_1 + \varepsilon_0}{2g_2} \operatorname{Re} \gamma_+,$$

где $\operatorname{Re} \gamma_+$ и $\operatorname{Im} \gamma_+$ однозначно определяются в (24) с использованием (15), (22). Применяя формулу (26), находим сумму логарифмов

$$\begin{aligned} s_1 &= \ln |Z_{2-} - \zeta_2| - \ln |Z_{2-} - \zeta_1| - \ln |Z_{2+} - \zeta_2| + \ln |Z_{2+} - \zeta_1| = \\ &= -\frac{1}{2} \sum_{q=1}^2 \sum_{l=1}^2 (-1)^{q+l} \ln \left[\left(g_1 + (-1)^q \sqrt{|G_+^2 + 1|} \cos \frac{\Psi}{2} - \zeta_l \right)^2 + \left(g_2 + (-1)^q \sqrt{|G_+^2 + 1|} \sin \frac{\Psi}{2} \right)^2 \right]. \end{aligned}$$

Используя формулу (27), вычисляем сумму аргументов

$$s_2 = \arg(Z_{2-} - \zeta_2) - \arg(Z_{2-} - \zeta_1) - \arg(Z_{2+} - \zeta_2) + \arg(Z_{2+} - \zeta_1) =$$

$$= - \sum_{q=1}^2 \sum_{l=1}^2 (-1)^{q+l} \arccos \frac{g_1 + (-1)^q \sqrt{|G_+^2 + 1|} \cos(\Psi/2) - \zeta_l}{\sqrt{\left(g_1 + (-1)^q \sqrt{|G_+^2 + 1|} \cos \frac{\Psi}{2} - \zeta_l\right)^2 + \left(g_2 + (-1)^q \sqrt{|G_+^2 + 1|} \sin \frac{\Psi}{2}\right)^2}}.$$

В формулах для s_1 , s_2 использованы обозначения из (15), (17), (22), (23).

Согласно (20) и (21) получаем

$$J_{12} = \frac{-4}{\nu_2 \sqrt{B_2}} \operatorname{Re} [\lambda_+ \gamma_+ (s_1 + i s_2)] = \frac{-4}{\nu_2 \sqrt{B_2}} \operatorname{Re} [(\Lambda_1 + i \Lambda_2)(s_1 + i s_2)].$$

В итоге приходим к равенству

$$J_{12} = \frac{-4}{\nu_2 \sqrt{B_2}} (\Lambda_1 s_1 - \Lambda_2 s_2).$$

Таким образом, в варианте 2а интеграл J_{12} вычислен явно.

Вариант 2б: $\chi_1 = 0$. В этом случае

$$B_2 U^2 + B_1 U + B_0 = B_2 \left[\left(U + \frac{B_1}{2B_2} \right)^2 + \chi_1^2 \right] = B_2 \left(U + \frac{B_1}{2B_2} \right)^2 = B_2 (U + \Omega)^2,$$

где $\Omega = B_1/(2B_2)$. Как отмечено выше, считаем, что $\Omega \notin [-h/2, h/2]$. Тогда

$$\begin{aligned} J_{12} &= \int_{-h/2}^{h/2} dU \frac{S_1 U + S_0}{(C_2 U^2 + C_1 U + C_0) \sqrt{B_2 U^2 + B_1 U + B_0}} = \\ &= \frac{1}{\sqrt{B_2}} \int_{-h/2}^{h/2} \frac{(S_1 U + S_0) \operatorname{sgn}(U + \Omega) dU}{(C_2 U^2 + C_1 U + C_0)(U + \Omega)} = \frac{\operatorname{sgn}(\Omega)}{\sqrt{B_2}} \int_{-h/2}^{h/2} \frac{(S_1 U + S_0) dU}{(C_2 U^2 + C_1 U + C_0)(U + \Omega)} = \\ &= \frac{\operatorname{sgn}(\Omega)}{\sqrt{B_2}} \left[p_1 \int_{-h/2}^{h/2} \frac{U dU}{C_2 U^2 + C_1 U + C_0} + p_2 \int_{-h/2}^{h/2} \frac{dU}{C_2 U^2 + C_1 U + C_0} + p_3 \int_{-h/2}^{h/2} \frac{dU}{U + \Omega} \right], \quad (28) \end{aligned}$$

где

$$p_1 = -\frac{C_2(S_1 \Omega - S_0)}{C_1 \Omega - C_0 - C_2 \Omega^2}, \quad p_2 = \frac{S_0}{\Omega} - \frac{C_0(S_1 \Omega - S_0)}{\Omega(C_1 \Omega - C_0 - C_2 \Omega^2)}, \quad p_3 = \frac{S_1 \Omega - S_0}{C_1 \Omega - C_0 - C_2 \Omega^2}.$$

Интегралы в (28) табличные. Воспользуемся формулами 1.2.8.19 и 1.2.8.13 из [13]. Учитывая, что $C_1^2 - 4C_2 C_0 < 0$, находим

$$\begin{aligned} J_{12} &= \frac{\operatorname{sgn}(\Omega)}{\sqrt{B_2}} \left(p_3 \ln |U + \Omega| + \frac{p_1}{2C_2} \ln |C_2 U^2 + C_1 U + C_0| - \right. \\ &\left. - \frac{p_1 C_1}{C_2 \sqrt{4C_2 C_0 - C_1^2}} \operatorname{arctg} \left(\frac{2C_2 U + C_1}{\sqrt{4C_2 C_0 - C_1^2}} \right) + \frac{2p_2}{\sqrt{4C_2 C_0 - C_1^2}} \operatorname{arctg} \left(\frac{2C_2 U + C_1}{\sqrt{4C_2 C_0 - C_1^2}} \right) \right) \Big|_{U=-h/2}^{U=h/2}. \end{aligned}$$

Таким образом, в варианте 2б интеграл J_{12} также вычислен явно.

3. Основной результат. Сформулируем основной результат этой работы.

Теорема. Пусть Γ – простая гладкая замкнутая поверхность класса C^2 , ограничивающая объёмно-односвязную внутреннюю область, или простая гладкая ограниченная разомкнутая ориентированная поверхность класса C^2 , содержащая свои предельные точки. Пусть Γ

допускает параметризацию (1) со свойством (3) и $\mu(y) \in C^0(\Gamma)$. Тогда для гармонического потенциала двойного слоя (4) при $x \notin \Gamma$ имеет место квадратурная формула

$$\mathcal{W}_0[\mu](x) \approx -\frac{1}{4\pi} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \mu_{nm} K_{nm}(x), \quad (29)$$

где интегралы $K_{nm}(x)$ вычислены в явном виде в п. 2.

4. Стандартная квадратурная формула. Квадратурная формула (29) является альтернативой стандартной квадратурной формуле для потенциала двойного слоя вне поверхности Γ , используемой в инженерных расчётах [4, гл. 2]:

$$\mathcal{W}_0[\mu](x) \approx -\frac{1}{4\pi} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \mu_{nm} \frac{hH}{|x - y(u_n, v_m)|^3} \sum_{j=1}^3 \eta_j(y(u_n, v_m))(y_j(u_n, v_m) - x_j). \quad (30)$$

Стандартная квадратурная формула получается из формулы (6) заменой канонического интеграла из (7) на следующее его приближённое значение:

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

Пусть для упрощения Γ – замкнутая поверхность. Значение, вычисленное по стандартной квадратурной формуле (30), стремится к бесконечности, если точка x стремится изнутри или извне Γ к одной из точек $y(u_n, v_m) \in \Gamma$, хотя сам потенциал двойного слоя в наших предположениях ограничен на Γ и непрерывно продолжим на Γ как изнутри, так и извне. Другими словами, если точка x стремится к точке на Γ изнутри или извне, то потенциал двойного слоя стремится к конечному пределу. Стандартная квадратурная формула не сохраняет важнейшие свойства потенциала двойного слоя, а именно, ограниченность на Γ и непрерывную продолжимость на Γ извне и изнутри. Стандартная квадратурная формула расходится вблизи границы.

5. Численные тесты. Тестирование улучшенной (29) и стандартной (30) квадратурных формул проведено в случае, когда поверхность Γ является сферой единичного радиуса, которая задана параметрически уравнениями:

$$y_1(u, v) = \cos u \sin v, \quad y_2(u, v) = \sin u \sin v, \quad y_3(u, v) = \cos v, \quad (31)$$

причём $(u, v) \in [0, 2\pi] \times [0, \pi]$. Отметим, что в данном случае $|\eta(y(u, v))| = \sin v$ и $|\eta(y(u, 0))| = |\eta(y(u, \pi))| = 0$ для всех $u \in [0, 2\pi]$. Иначе говоря, $|\eta(y)| = 0$ на полюсах сферы при такой параметризации, но условия теоремы выполняются.

Согласно [16, § 27.6] плотность потенциала двойного слоя на поверхности Γ можно найти по формуле

$$\mu(x)|_{\Gamma} = \mathcal{W}_0[\mu](x)|_{\Gamma^-} - \mathcal{W}_0[\mu](x)|_{\Gamma^+}.$$

Здесь поверхность Γ рассматривается как двусторонняя, через Γ^- обозначена сторона, которую мы видим, глядя навстречу вектору нормали $\mathbf{n}_y$, а через Γ^+ – противоположная сторона. В формуле берутся предельные значения потенциала двойного слоя на разных сторонах Γ . Отметим, что направление единичной нормали $\mathbf{n}_y$ совпадает с направлением нормали η , так как вектор $\mathbf{n}_y$ получается из η в результате нормировки. Пусть теперь Γ – единичная сфера, заданная параметризацией (31), тогда формулы (2) для нормали η определяют внутреннюю нормаль на сфере, а значит, Γ^- – внутренняя сторона единичной сферы, а Γ^+ – её внешняя сторона.

В рассматриваемых тестовых примерах для потенциала двойного слоя с заданной на единичной сфере плотностью известно явное выражение во всём пространстве, поэтому точные значения потенциала можно сравнить с приближёнными, вычисленными по квадратурным

формулам. Во всех тестах приближённое значение потенциала двойного слоя вычислялось по стандартной квадратурной формуле (30) и по улучшенной квадратурной формуле (29) в некоторых точках на вспомогательных сферах, имеющих центры в начале координат и радиусы, равные $1 \pm \Delta R$. Тем самым вспомогательные сферы находятся либо внутри, либо снаружи сферы единичного радиуса, на которой задана плотность потенциала, на расстоянии ΔR от неё. Затем были рассчитаны значения абсолютных погрешностей в этих точках, и для каждой вспомогательной сферы определялись максимумы значений этих погрешностей.

Координаты точек, которые использовались для оценки максимальной абсолютной погрешности, выбирались следующими:

$$x_j^{ql} = Ry_j(u_q, v_l), \quad j = 1, 2, 3, \\ u_q = \frac{2\pi}{2N}q, \quad q = \overline{0, 2N}; \quad v_l = \frac{\pi}{2M}l, \quad l = \overline{1, 2M-1}, \quad (32)$$

где величина $y_j(u, v)$ определяется формулами (31), R – радиус вспомогательной сферы. В частности, эти точки расположены над и под центрами участков разбиения единичной сферы, серединами границ между такими участками и пересечениями этих границ. Отметим, что эти точки распределены по всей сфере.

Вычисления проводились для различных значений M и N . Величины шагов выбирались по формулам $h = 2\pi/N$, $H = \pi/M$. Если $N = M = 25$, то $h \approx 0.25$, $H \approx 0.13$; если $N = M = 50$, то $h \approx 0.126$, $H \approx 0.063$; если $N = M = 100$, то $h \approx 0.063$, $H \approx 0.031$.

В таблицах приведены рассчитанные максимальные значения абсолютных погрешностей. В левом столбце указано отличие радиуса вспомогательной сферы от единицы: для внутренних сфер радиус равен $1 - \Delta R$, для внешних $1 + \Delta R$. В верхней строке указаны значения M, N . Первое число в ячейках таблицы – максимальная погрешность для стандартной квадратурной формулы на данной вспомогательной сфере, а число после точки с запятой – максимальная погрешность на данной сфере для улучшенной формулы.

Тест 1. В данном тесте использовалась плотность потенциала $\mu(y(u, v)) = 1$; тогда гармонический потенциал двойного слоя имеет вид [16, § 27.2]

$$\mathcal{W}_0[\mu](x) = \begin{cases} 1 & \text{при } |x| < 1, \\ 0 & \text{при } |x| > 1. \end{cases}$$

В табл. 1 приведены рассчитанные максимальные значения абсолютных погрешностей.

Таблица 1. Максимальная абсолютная погрешность квадратурных формул в тесте 1

ΔR	$M = N = 25$	$M = N = 50$	$M = N = 100$
<i>Внутренние сферы</i>			
0.1	0.096; 0.021	0.0079; 0.0045	0.0020; 0.0011
0.06	0.44; 0.041	0.054; 0.0083	0.0055; 0.0018
0.03	2.42; 0.094	0.45; 0.022	0.056; 0.0041
<i>Внешние сферы</i>			
0.1	0.11; 0.021	0.0081; 0.0039	0.0020; 0.00092
0.06	0.47; 0.048	0.061; 0.0080	0.0055; 0.0016
0.03	2.46; 0.15	0.47; 0.023	0.060; 0.0041

Тест 2. В данном тесте использовалась плотность потенциала $\mu(y(u, v)) = \cos v$; тогда гармонический потенциал двойного слоя имеет вид

$$\mathcal{W}_0[\mu](x) = \begin{cases} \frac{2|x| \cos \vartheta}{3} & \text{при } |x| < 1, \\ -\frac{\cos \vartheta}{3|x|^2} & \text{при } |x| > 1, \end{cases}$$

где ϑ – зенитный угол в сферических координатах с центром в начале координат. В табл. 2 приведены рассчитанные максимальные значения абсолютных погрешностей.

Таблица 2. Максимальная абсолютная погрешность квадратурных формул в тесте 2

ΔR	$M = N = 25$	$M = N = 50$	$M = N = 100$
<i>Внутренние сферы</i>			
0.1	0.032; 0.011	0.0078; 0.0030	0.0020; 0.00076
0.06	0.18; 0.017	0.020; 0.0044	0.0055; 0.0012
0.03	1.14; 0.043	0.19; 0.0086	0.020; 0.0021
<i>Внешние сферы</i>			
0.1	0.040; 0.0076	0.0077; 0.0016	0.0020; 0.00042
0.06	0.20; 0.017	0.020; 0.0030	0.0055; 0.00080
0.03	1.16; 0.050	0.20; 0.0086	0.020; 0.0017

Тест 3. В данном тесте использовалась плотность потенциала $\mu(y(u, v)) = \cos u \sin v$; тогда гармонический потенциал двойного слоя имеет вид

$$\mathcal{W}_0[\mu](x) = \begin{cases} \frac{2|x| \cos \varphi \sin \vartheta}{3} & \text{при } |x| < 1, \\ -\frac{\cos \varphi \sin \vartheta}{3|x|^2} & \text{при } |x| > 1, \end{cases}$$

где ϑ и φ – зенитный и азимутальный углы в сферических координатах с центром в начале координат. В табл. 3 приведены рассчитанные максимальные значения абсолютных погрешностей.

Таблица 3. Максимальная абсолютная погрешность квадратурных формул в тесте 3

ΔR	$M = N = 25$	$M = N = 50$	$M = N = 100$
<i>Внутренние сферы</i>			
0.1	0.096; 0.023	0.0055; 0.0049	0.000028; 0.0012
0.06	0.44; 0.043	0.054; 0.0088	0.0021; 0.0019
0.03	2.42; 0.096	0.45; 0.022	0.056; 0.0043
<i>Внешние сферы</i>			
0.1	0.12; 0.019	0.0082; 0.0036	0.000069; 0.00085
0.06	0.48; 0.045	0.061; 0.0077	0.0029; 0.0016
0.03	2.46; 0.15	0.47; 0.023	0.060; 0.0040

Тест 4. В данном тесте использовалась плотность потенциала $\mu(y(u, v)) = (3 \cos^2 v - 1)/2$; тогда гармонический потенциал двойного слоя имеет вид

$$\mathcal{W}_0[\mu](x) = \begin{cases} \frac{3|x|^2(3 \cos^2 \vartheta - 1)}{10} & \text{при } |x| < 1, \\ -\frac{(3 \cos^2 \vartheta - 1)}{5|x|^3} & \text{при } |x| > 1, \end{cases}$$

где ϑ – зенитный угол в сферических координатах с центром в начале координат. В табл. 4 приведены рассчитанные максимальные значения абсолютных погрешностей.

Закключение. Из таблиц видно, что улучшенная квадратурная формула обеспечивает более высокую точность вычислений вблизи границы Γ , чем стандартная. Кроме того, значения, вычисленные по стандартной формуле, стремятся к бесконечности при приближении к границе. Тесты показывают, что улучшенная формула даёт хорошую точность вычислений для

всех точек, расположенных на расстоянии H и более от границы Γ . В этом случае улучшенная квадратурная формула имеет второй порядок сходимости и обеспечивает погрешность вычислений порядка $O(hH)$. На расстояниях порядка hH до границы улучшенная формула даёт погрешность $O(H)$, а стандартная формула расходится.

Таблица 4. Максимальная абсолютная погрешность квадратурных формул в тесте 4

ΔR	$M = N = 25$	$M = N = 50$	$M = N = 100$
<i>Внутренние сферы</i>			
0.1	0.047; 0.012	0.0079; 0.0026	0.0020; 0.00063
0.06	0.22; 0.022	0.027; 0.0045	0.0055; 0.0011
0.03	1.21; 0.048	0.22; 0.011	0.028; 0.0022
<i>Внешние сферы</i>			
0.1	0.057; 0.010	0.0077; 0.0019	0.0020; 0.00044
0.06	0.24; 0.023	0.031; 0.0039	0.0055; 0.00080
0.03	1.23; 0.074	0.23; 0.012	0.03; 0.0020

СПИСОК ЛИТЕРАТУРЫ

1. Белоцерковский С.М., Лифанов И.К. Численные методы в сингулярных интегральных уравнениях. М., 1985.
2. Лифанов И.К. Метод сингулярных интегральных уравнений и численный эксперимент. М., 1995.
3. Сетуха А.В. Численные методы в интегральных уравнениях и их приложения. М., 2016.
4. Бреббия К., Теллес Ж., Вроубел Л. Методы граничных элементов. М., 1987.
5. Krutitskii P.A., Kwak D.Y., Hyon Y.K. Numerical treatment of a skew-derivative problem for the Laplace equation in the exterior of an open arc // J. of Eng. Math. 2007. V. 59. P. 25–60.
6. Крутицкий П.А., Колыбасова В.В. Численный метод решения интегральных уравнений в задаче с наклонной производной для уравнения Лапласа вне разомкнутых кривых // Дифференц. уравнения. 2016. Т. 52. № 9. С. 1262–1276.
7. Крутицкий П.А. Смешанная задача для уравнения Лапласа вне разрезов на плоскости // Дифференц. уравнения. 1997. Т. 33. № 9. С. 1181–1190.
8. Krutitskii P.A. The 2-dimensional Dirichlet problem in an external domain with cuts // Zeitschr. für Analysis und ihre Anwend. 1998. Bd. 17. № 2. S. 361–378.
9. Krutitskii P.A. The Dirichlet problem for the two-dimensional Laplace equation in a multiply connected domain with cuts // Proc. of the Edinburgh Math. Soc. 2000. V. 43. № 2. P. 325–341.
10. Крутицкий П.А., Федотова А.Д., Колыбасова В.В. Квадратурная формула для потенциала простого слоя // Дифференц. уравнения. 2019. Т. 55. № 9. С. 1269–1284.
11. Бутузов В.Ф., Крутицкая Н.Ч., Медведев Г.Н., Шишкин А.А. Математический анализ в вопросах и задачах. М., 2000.
12. Ильин В.А., Позняк Э.Г. Основы математического анализа. Ч. 2. М., 1973.
13. Прудников А.П., Брычков Ю.А., Маричев О.И. Интегралы и ряды. М., 1981.
14. Градштейн И.С., Рыжик И.М. Таблицы интегралов, сумм, рядов и произведений. М., 1963.
15. Цыпкин А.Г., Цыпкин Г.Г. Математические формулы. М., 1985.
16. Владимиров В.С. Уравнения математической физики. М., 1981.

Институт прикладной математики
им. М.В. Келдыша РАН, г. Москва,
Московский государственный университет
им. М.В. Ломоносова

Поступила в редакцию 03.03.2021 г.
После доработки 03.03.2021 г.
Принята к публикации 27.04.2021 г.

ЧИСЛЕННЫЕ МЕТОДЫ

УДК 519.622+517.958:533.7

ОБ УСТОЙЧИВЫХ МЕТОДАХ РУНГЕ–КУТТЫ ДЛЯ РЕШЕНИЯ ГИПЕРБОЛИЧЕСКИХ УРАВНЕНИЙ РАЗРЫВНЫМ МЕТОДОМ ГАЛЁРКИНА

© 2021 г. В. В. Лукин, В. Н. Корчагова, С. М. Сауткина

Рассматриваются методы Рунге–Кутты, область устойчивости которых включает круг максимального диаметра для заданных количества стадий и порядка. Указанные методы применяются для решения начальных задач, полученных при аппроксимации гиперболических систем с использованием разрывного метода Галёркина. Предложены два трёхстадийных метода из рассматриваемого класса, для которых с помощью тестовых задач для уравнения переноса и для системы уравнений газовой динамики исследована способность сохранять устойчивость и монотонность численного решения при максимально возможных шагах сетки по времени.

DOI: 10.31857/S0374064121070086

Введение. Математическое моделирование скоростных течений газа, содержащих разрывы решения и/или области резкого и неустойчивого его изменения, является одной из важных областей приложения моделей газовой динамики. Успех численного моделирования в равной степени зависит от выбора способа аппроксимации соответствующей системы гиперболических уравнений как по времени, так и по пространству [1]. Сложность течений, описываемых разрывными решениями, предъявляет к выбору разностных схем высокие требования по точности, устойчивости и монотонности [2].

В последние годы особой популярностью при численном моделировании течений жидкости и газа пользуются схемы на основе разрывного метода Галёркина [3–9]. Достоинства метода определяются его высокими разрешающими возможностями, основанными на полиномиальной аппроксимации решения произвольного порядка внутри каждой ячейки сетки и допущении разрывов решения на границах ячеек. При использовании такого подхода очень важен выбор схемы аппроксимации уравнений по времени. Необходимо выбирать методы, обеспечивающие скорость вычислений, их устойчивость и обладающие свойством невозрастания паразитных осцилляций решения. К таким схемам относятся, например, явные методы Рунге–Кутты, удовлетворяющие свойству TVD – total variation diminishing [10]. При этом важным является использование методов с оптимальной для данного класса задач областью устойчивости [11–13]. В то же время требования к точности метода сталкиваются с проблемой потери порядка аппроксимации на разрывах [14].

Цель статьи – выбор и исследование методов Рунге–Кутты с увеличенной областью устойчивости, адаптированных для решения задач газовой динамики с использованием разрывного метода Галёркина.

1. Разрывный метод Галёркина. Рассмотрим скалярный закон сохранения

$$\frac{\partial U}{\partial t} + \operatorname{div} F(U) = 0, \quad (1)$$

где $U(x, t)$ – неизвестная функция, $F(U) = (F_1(U), \dots, F_M(U))^T$ – вектор потоков, N – размерность пространства. Будем считать, что в пространственной области Ω на временном отрезке $[0, T]$ задана начальная задача

$$U(x, 0) = U_0(x), \quad x \in \Omega. \quad (2)$$

Введём в области Ω сетку $\Omega_h = \{I_j\}_{j=1}^n$. На каждой ячейке I_j зададим пространство $P^m(I_j)$ полиномов степени не выше m и выберем в нём ортонормированный в $L^2(I_j)$ базис $\{\varphi_j^{(s)}(x)\}_{s=0}^m$. Вне ячейки I_j положим $\varphi_j^{(s)}(x) = 0$, $\varphi = \overline{1, n}$, $s = \overline{0, m}$. Будем искать при-

ближённое решение U_h задачи (1), (2) в виде линейной комбинации выбранных базисных функций

$$U_h(x, t) = \sum_{j=1}^n \sum_{s=0}^m U_j^{(s)}(t) \varphi_j^{(s)}(x). \quad (3)$$

Умножив уравнение (1) скалярно на базисную функцию $\varphi_j^{(s)}(x)$, получим

$$\int_{I_j} \frac{\partial U}{\partial t} \varphi_j^{(s)}(x) dV + \int_{I_j} \sum_{i=1}^N \frac{\partial F_i}{\partial x_i} \varphi_j^{(s)}(x) dV = 0,$$

Далее преобразуем второе слагаемое:

$$\int_{I_j} \frac{\partial U}{\partial t} \varphi_j^{(s)}(x) dV + \int_{I_j} \sum_{i=1}^N \frac{\partial (F_i \varphi_j^{(s)}(x))}{\partial x_i} dV - \int_{I_j} \sum_{i=1}^N F_i \frac{\partial \varphi_j^{(s)}(x)}{\partial x_i} dV = 0,$$

откуда, наконец, получим с использованием формулы Остроградского–Гаусса

$$\int_{I_j} \frac{\partial U}{\partial t} \varphi_j^{(s)}(x) dV + \oint_{\partial I_j} \sum_{i=1}^N F_i n_i \varphi_j^{(s)}(x) dS - \int_{I_j} \sum_{i=1}^N F_i \frac{\partial \varphi_j^{(s)}(x)}{\partial x_i} dV = 0, \quad (4)$$

где $n = (n_1, \dots, n_N)^T$ – вектор внешней нормали к поверхности ∂I_j , ограничивающей ячейку I_j .

Подставляя в равенство (4) вместо функции U представление (3) для функции U_h и учитывая ортонормированность базиса $\{\varphi_j^{(s)}(x)\}_{s=0}^m$, приходим к следующей системе обыкновенных дифференциальных уравнений для нахождения коэффициентов $U_j^{(s)}$ разложения приближённого решения:

$$\frac{dU_j^{(s)}}{dt} = - \oint_{\partial I_j} \sum_{i=1}^N F_i n_i \varphi_j^{(s)}(x) dS + \int_{I_j} \sum_{i=1}^N F_i \frac{\partial \varphi_j^{(s)}(x)}{\partial x_i} dV, \quad j = \overline{1, n}, \quad s = \overline{0, m},$$

или

$$\frac{dU_j^{(s)}}{dt} = L_h(\dots, U_\beta^{(\alpha)}, \dots), \quad j = \overline{1, n}, \quad s = \overline{0, m}. \quad (5)$$

Проецируя равенство (2) (с заменой функции U функцией U_h) на базисные функции $\varphi_j^{(s)}$, получим начальные условия задачи Коши

$$U_j^{(s)}(0) = \int_{I_j} U_0(x) \varphi_j^{(s)}(x) dV, \quad j = \overline{1, n}, \quad s = \overline{0, m}. \quad (6)$$

Описанный подход определяет разрывный метод Галёркина [3]. Он позволяет строить высокоточные разностные схемы с использованием компактного шаблона аппроксимации даже для методов высокого порядка, а также добиваться хорошей разрешающей способности алгоритма на сильных разрывах.

2. Выбор метода интегрирования по времени. Как правило, в силу нелинейности оператора L_h для решения задачи Коши (5), (6) используются явные методы Рунге–Кутты. В то же время при использовании этих методов нужно для обеспечения устойчивости накладывать существенные ограничения на шаг по времени. Целесообразно выбирать такие методы,

которые при необходимом порядке точности имели бы как можно большую область устойчивости. При этом выбор области устойчивости, вытянутой вдоль действительной оси, не всегда является оптимальным [13].

Рассмотрим одномерное линейное уравнение переноса

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0 \quad (7)$$

и будем решать его на отрезке $[0, L] = \Omega$. Введём на нём равномерную сетку $\Omega_h = \{x_i = ih, i = \overline{0, n}, h = L/n\}$. В каждом узле зададим функцию $u_i(t)$, представляющую собой приближённое значение решения. Аппроксимируя пространственную производную разностью назад, получим систему ОДУ для функций u_i , соответствующих внутренним узлам:

$$\frac{du_i}{dt} = -c \frac{u_i - u_{i-1}}{h}. \quad (8)$$

Рассмотрим задачу на собственные значения $-c(u_i - u_{i-1}) = \lambda h u_i$ с периодическими граничными условиями $u_0 = u_n$. Её решение имеет вид

$$\mu_k = \exp(-i\tilde{2}\pi k/n) - 1,$$

где $\mu_k = \lambda_k h/c$, $\tilde{i} = \sqrt{-1}$, $k = \overline{0, n-1}$. На рис. 1 точками на комплексной плоскости изображены значения μ_k . Видим, что область устойчивости метода, применяемого для интегрирования системы (8), не стоит выбирать вытянутой вдоль действительной оси, как это характерно, например, для методов Рунге–Кутты–Чебышёва [15, гл. IV, § 2]. Эта область должна включать в себя круг максимально возможного диаметра, касающийся мнимой оси в точке $(0, 0)$. Очевидно, что в случае более сложного, нелинейного вида вектора потоков $F(U)$ расположение собственных чисел линейной части оператора L_h будет иным. Тем не менее приведённый пример демонстрирует общий принцип выбора методов по области устойчивости для гиперболических уравнений.

В работе [13] рассмотрен подход к построению методов Рунге–Кутты с областью устойчивости, включающей в себя круг максимального диаметра (условие МД), касающийся начала координат. Для методов первого и второго порядка точности авторы получили вид функции устойчивости, обеспечивающей выполнение условия МД, для произвольного количества стадий s . В частности, для метода второго порядка такая функция имеет вид

$$R(\mu) = \frac{s-1}{s} \left(\frac{\mu}{s-1} + 1 \right)^s + \frac{1}{s}. \quad (9)$$

Соответствующие области устойчивости для количества стадий $s = 2, 3, 4, 5, 6$ приведены на рис. 2. При этом задание вида функции устойчивости не задаёт единственным образом сам метод.

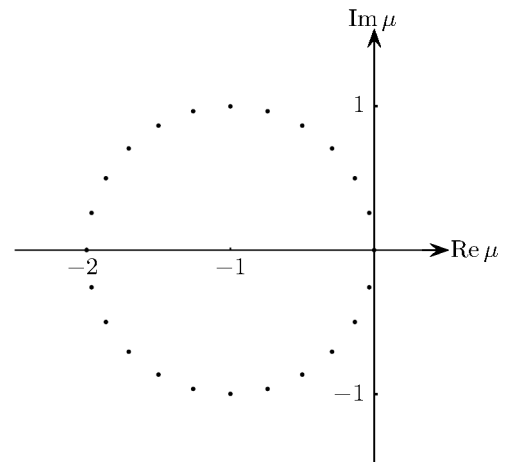


Рис. 1. Приведённые собственные числа μ_k оператора разностной производной назад.

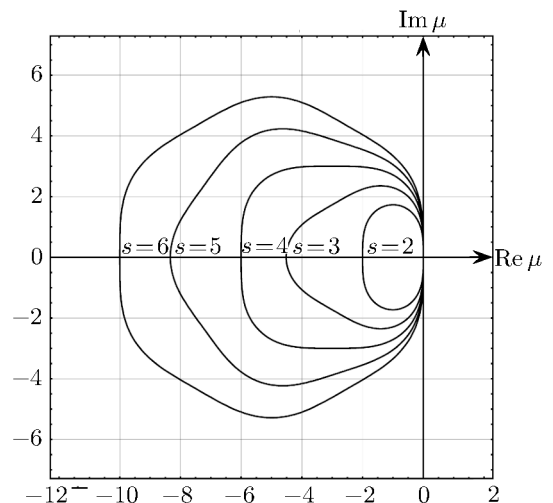


Рис. 2. Области устойчивости, соответствующие функции устойчивости (9), для количества стадий $s = 2, 3, 4, 5, 6$.

2.1. Интерполяционное представление методов Рунге–Кутты. Для автономной системы ОДУ $u' = f(u)$ явный метод Рунге–Кутты вычисления значения приближённого решения $y_{l+1} \approx u(t_l + \tau)$ по значению $y_l \approx u(t_l)$ имеет вид [16, гл. II, § 2]

$$g_i = y_l + \tau \sum_{j=1}^{i-1} a_{i,j} k_j, \quad k_i = f(g_i), \quad i = \overline{1, s}, \quad (10)$$

$$y_{l+1} = y_l + \tau \sum_{j=1}^s b_j k_j. \quad (11)$$

Далее будем опираться на следующее представление метода.

Утверждение 1. Метод Рунге–Кутты (10), (11) в случае, если $a_{i,i-1} \neq 0$ для всех $i = \overline{2, s}$, может быть представлен в виде

$$g_1 = y_l, \quad g_i = g_{i-1} + \tau a_{i,i-1} k_{i-1} + \sum_{j=1}^{i-1} \eta_j^i g_j, \quad i = \overline{2, s}, \quad (12)$$

$$g_{s+1} = g_s + \tau \xi k_s, \quad y_{l+1} = \sum_{j=1}^{s+1} \sigma_j g_j, \quad (13)$$

где σ_j , $j = \overline{1, s}$, и η_j^i , $i = \overline{2, s}$, $j = \overline{1, i-1}$, – некоторый набор постоянных коэффициентов, вычисляемых через коэффициенты $a_{i,j}$, $i = \overline{2, s}$, $j = \overline{1, i-1}$, и b_j , $j = \overline{1, s}$, при этом коэффициент $\xi \neq 0$ может быть выбран произвольно, и $\sigma_{s+1} = b_s/\xi$.

Доказательство. Для первой и второй стадий представление, очевидно, верно: $g_1 = y_l$, $k_1 = f(g_1)$, $g_2 = y_l + \tau a_{2,1} k_1$. Предположим, что оно верно и для всех g_j , $j = \overline{2, i-1}$, $i \leq s$. Тогда

$$k_j = \frac{1}{\tau a_{j+1,j}} \left(g_j + \tau a_{j+1,j} k_j + \sum_{\alpha=1}^j \eta_{\alpha}^{j+1} g_{\alpha} - g_j - \sum_{\alpha=1}^j \eta_{\alpha}^{j+1} g_{\alpha} \right) = \frac{1}{\tau a_{j+1,j}} \left(g_{j+1} - g_j - \sum_{\alpha=1}^j \eta_{\alpha}^{j+1} g_{\alpha} \right)$$

и

$$\begin{aligned} g_i &= y_l + \tau \sum_{j=1}^{i-1} a_{i,j} k_j = y_l + \tau \sum_{j=1}^{i-2} a_{i,j} k_j + \tau a_{i,i-1} k_{i-1} + \tau \sum_{j=1}^{i-2} (a_{i,j} - a_{i-1,j}) k_j = \\ &= g_{i-1} + \tau a_{i,i-1} k_{i-1} + \sum_{j=1}^{i-2} \frac{a_{i,j} - a_{i-1,j}}{a_{j+1,j}} \left(g_{j+1} - g_j - \sum_{\alpha=1}^j \eta_{\alpha}^{j+1} g_{\alpha} \right) = g_{i-1} + \tau a_{i,i-1} k_{i-1} + \sum_{j=1}^{i-1} \eta_j^i g_j. \end{aligned}$$

Из последнего равенства несложно вытекают формулы для вычисления η_j^i , но в силу громоздкости этих формул приводить их здесь не будем. Полученное представление позволяет в ходе расчётов хранить лишь значения g_i и не хранить k_i .

Аналогично получаем представление

$$y_{l+1} = y_l + \tau \sum_{j=1}^s b_j k_j = g_1 + \sum_{j=1}^{s-1} \frac{b_j}{a_{j+1,j}} \left(g_{j+1} - g_j - \sum_{\alpha=1}^j \eta_{\alpha}^{j+1} g_{\alpha} \right) + \tau b_s k_s.$$

Выберем $\xi \neq 0$ и обозначим $g_{s+1} = g_s + \tau \xi k_s$. Тогда $k_s = (g_{s+1} - g_s)/\tau \xi$ и

$$y_{l+1} = y_l + \tau \sum_{j=1}^s b_j k_j = \sum_{j=1}^{s+1} \sigma_j g_j,$$

причём $\sigma_{s+1} = b_s/\xi$. Значения коэффициентов σ_j , $j = \overline{1, s}$ могут быть получены явно. Утверждение доказано.

2.2. Свойство TVD. Согласно теореме Годунова [17] монотонные разностные схемы высокого порядка для решения гиперболических систем уравнений вида $u_t = L(u)$, где L – дифференциальный оператор, содержащий производные только по пространственным переменным, не могут быть линейными. В случае нелинейных систем гиперболических уравнений для введения понятия монотонной схемы часто используют TVD-свойство [1, с. 101]. В работе [10] понятие TVD-свойства рассматривается и для конструирования методов Рунге–Кутты, применяемых для интегрирования по времени систем гиперболических уравнений. Свойство TVD в этом случае определяется следующим образом.

Пусть разностный оператор L_h аппроксимирует оператор L на пространственной сетке Ω_h . Пусть $Y_l = \{y_l^i\}$ – приближённое решение на временном слое l и решение Y_{l+1} на следующем временном слое вычисляется с использованием метода (10), (11). Этот метод обладает свойством TVD, если при всех l имеет место неравенство $TV(Y_{l+1}) \leq TV(Y_l + \tau L_h(Y_l))$, где $TV(Y) = \sum_j |y^{j+1} - y^j|$ – полная вариация сеточной функции $Y = \{y^i\}$ и суммирование выполняется по всей сетке Ω_h .

В работе [10] показано, что справедливо

Утверждение 2. Метод Рунге–Кутты обладает свойством TVD, если для коэффициентов его представления (12), (13) выполняются неравенства

$$a_{i,i-1} > 0, \quad i = \overline{2, s}, \quad \delta_{j,i-1} + \eta_j^i \geq 0, \quad i = \overline{2, s}, \quad j = \overline{1, i-1},$$

где $\delta_{i,j}$ – символ Кронекера, и существует хотя бы одно такое значение ξ , при котором все σ_j неотрицательны, $j = \overline{1, s+1}$.

2.3. Примеры методов. Среди методов, наиболее часто использующихся при реализации разрывного метода Галёркина, можно выделить следующие два [10]. Оба не обладают свойством МД, но обладают свойством TVD.

Метод 1. Двухстадийный метод второго порядка:

$$k_1 = f(y_n), \quad k_2 = f(y_n + \tau k_1), \quad y_{n+1} = y_n + \frac{\tau}{2} k_1 + \frac{\tau}{2} k_2.$$

В интерполяционном представлении при выборе $\xi = 1$ этот метод принимает вид

$$g_1 = y_l, \quad g_2 = g_1 + \tau f(g_1), \quad (14)$$

$$g_3 = g_2 + \tau f(g_2), \quad y_{l+1} = \frac{1}{2} g_1 + \frac{1}{2} g_3. \quad (15)$$

Его функцией устойчивости является $R(\mu) = 1 + \mu + \mu^2/2$, а область устойчивости изображена на рис. 3, а.

В дальнейшем метод (14), (15) будем называть RK2TVD.

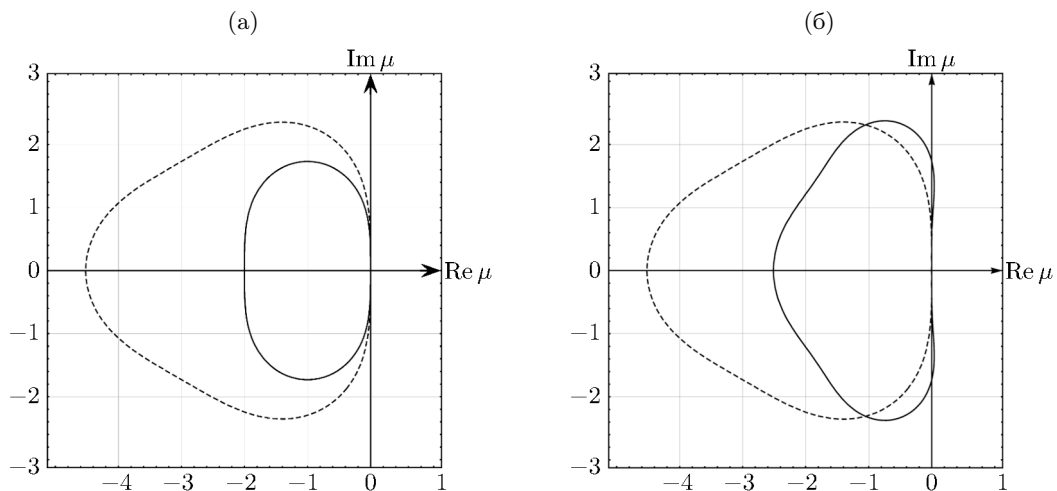


Рис. 3. Области устойчивости методов RK2TVD ((а), сплошная линия), RK3TVD ((б), сплошная линия) и RK2MD-TVD (штриховая линия).

Метод 2. Трёхстадийный метод третьего порядка (используется способ записи [10]):

$$y^* = y_l + \tau f(y_l),$$

$$y^{**} = \frac{3}{4}y_l + \frac{1}{4}y^* + \frac{1}{4}\tau f(y^*),$$

$$y_{l+1} = \frac{1}{3}y_l + \frac{2}{3}y^{**} + \frac{2}{3}\tau f(y^{**}).$$

В интерполяционном представлении при выборе $\xi = 1$ этот метод принимает вид

$$g_1 = y_l, \quad g_2 = g_1 + \tau f(g_1), \quad (16)$$

$$g_3 = g_2 + \frac{\tau}{4}f(g_2) + \frac{3}{4}g_1 - \frac{3}{4}g_2, \quad (17)$$

$$g_4 = g_3 + \tau f(g_3), \quad y_{l+1} = \frac{1}{3}g_1 + \frac{2}{3}g_4. \quad (18)$$

Его функцией устойчивости является $R(\mu) = 1 + \mu + \mu^2/2 + \mu^3/6$, а область устойчивости изображена на рис. 3, 6.

В дальнейшем метод (16)–(18) будем называть RK3TVD.

Теперь построим примеры трёхстадийных методов Рунге–Кутты второго порядка, удовлетворяющих условию МД. Для таких методов их функцией устойчивости является

$$R(\mu) = 1 + \mu + \frac{\mu^2}{2} + \frac{\mu^3}{12}.$$

Тогда условия второго порядка [16, гл. II, § 1] дают следующие уравнения относительно коэффициентов метода (10), (11):

$$b_1 + b_2 + b_3 = 1, \quad a_{2,1}b_2 + a_{3,1}b_3 + a_{3,2}b_3 = \frac{1}{2}, \quad a_{2,1}a_{3,2}b_3 = \frac{1}{12}. \quad (19)$$

Из доказательства утверждения 1 видно, что удобно выбрать $a_{i,j} = a_{i-1,j}$, $i = \overline{3, s}$, $j = \overline{1, i-1}$, поскольку тогда второе равенство (12) сводится к уравнению $g_i = g_{i-1} + \tau a_{i,i-1}k_{i-1}$, $i = \overline{3, s}$. В данном случае выберем $a_{3,1} = a_{2,1} = q$ и обозначим $a_{3,2} = r$. Дополнительно положим $q = r$. Тогда система (19) примет вид

$$b_1 + b_2 + b_3 = 1, \quad q(b_2 + 2b_3) = \frac{1}{2}, \quad q^2b_3 = \frac{1}{12}.$$

Варьируя параметр q , можно получать разные варианты методов.

Метод 3. Справедливо следующее

Утверждение 3. Метод Рунге–Кутты вида

$$g_1 = y_l, \quad k_1 = f(g_1),$$

$$g_2 = y_l + \frac{\tau}{2}k_1, \quad k_2 = f(g_2),$$

$$g_3 = y_l + \frac{\tau}{2}k_1 + \frac{\tau}{2}k_2, \quad k_3 = f(g_3),$$

$$y_{l+1} = y_l + \frac{\tau}{3}(k_1 + k_2 + k_3)$$

имеет второй порядок и удовлетворяет условиям МД и TVD.

Доказательство. Действительно, условия порядка (19), включающие требования условия МД, выполнены. В интерполяционной форме записи будем иметь

$$\begin{aligned} g_1 &= y_l, & g_2 &= g_1 + \frac{\tau}{2}f(g_1), & g_3 &= g_2 + \frac{\tau}{2}f(g_2), \\ g_4 &= g_3 + \tau\xi f(g_3), & y_{l+1} &= \frac{1}{3}g_1 + \frac{2-1/\xi}{3}g_3 + \frac{1}{3\xi}g_4 \end{aligned} \quad (20)$$

и, выбрав $\xi = 1/2$, получим

$$g_4 = g_3 + \frac{\tau}{2}f(g_3), \quad y_{l+1} = \frac{1}{3}g_1 + \frac{2}{3}g_4. \quad (21)$$

В дальнейшем метод (20), (21) будем называть RK2MD-TVD.

Метод 4. Справедливо следующее

Утверждение 4. Метод Рунге–Кутты вида

$$\begin{aligned} g_1 &= y_l, & k_1 &= f(g_1), \\ g_2 &= y_l + \frac{\tau}{3}k_1, & k_2 &= f(g_2), \\ g_3 &= y_l + \frac{\tau}{3}k_1 + \frac{\tau}{3}k_2, & k_3 &= f(g_3), \\ y_{l+1} &= y_l + \frac{\tau}{4}k_1 + \frac{3\tau}{4}k_3 \end{aligned}$$

имеет второй порядок и удовлетворяет условию МД, но не TVD.

Доказательство. Действительно, условия порядка (19), включающие требования условия МД, выполнены. В интерполяционной форме записи получим

$$\begin{aligned} g_1 &= y_l, & g_2 &= g_1 + \frac{\tau}{3}f(g_1), & g_3 &= g_2 + \frac{\tau}{3}f(g_2), \\ g_4 &= g_3 + \tau\xi f(g_3), & y_{l+1} &= \frac{1}{4}g_1 + \frac{3}{4}g_2 + \frac{3}{4\xi}(g_4 - g_3). \end{aligned} \quad (22)$$

Очевидно, что никаким выбором значения ξ невозможно удовлетворить условию TVD. Выбирая, например, для симметрии записи $\xi = 1/3$, получаем

$$g_4 = g_3 + \frac{\tau}{3}f(g_3), \quad y_{l+1} = \frac{1}{4}g_1 + \frac{3}{4}g_2 - \frac{9}{4}g_3 + \frac{9}{4}g_4. \quad (23)$$

В дальнейшем метод (22), (23) будем называть RK2MD.

3. Численные примеры. Рассмотрим результаты применения описанных методов для практических вычислений. Интерес представляет максимально достижимый для каждого метода шаг по времени при фиксированной пространственной сетке. Поскольку использовались аппроксимации пространственных операторов не выше второго порядка, то преимущество метода RK3TVD в порядке точности не принималось в расчёт.

Для сравнения методов между собой в описанных ниже тестовых примерах для разных методов вычислялись максимально достижимые временные шаги τ_m и в одномерном случае соответствующие им отношения $C_m = \tau_m/h$, где h – шаг пространственной сетки. За референсные значения принимались максимально достижимые шаги для метода RK2TVD, поскольку он имеет наименьшую область устойчивости.

В таблицах ниже использованы следующие обозначения:

$$T_{3-1} = \tau_m(\text{RK2MD-TVD})/\tau_m(\text{RK2TVD}), \quad T_{3-2} = \tau_m(\text{RK2MD-TVD})/\tau_m(\text{RK3TVD}),$$

$$T_{4-1} = \tau_m(\text{RK2MD})/\tau_m(\text{RK2TVD}), \quad T_{4-2} = \tau_m(\text{RK2MD})/\tau_m(\text{RK3TVD}),$$

$$T_{2-1} = \tau_m(\text{RK3TVD})/\tau_m(\text{RK2TVD}).$$

Поскольку метод RK2TVD в отличие от остальных является двухстадийным, наравне с коэффициентами T_{2-1} , T_{3-1} и T_{4-1} имеет смысл рассматривать величины $E_{J-1} = 2T_{J-1}/3$, $J = 2, 3, 4$, характеризующие сравнительную вычислительную сложность методов, связанную с количеством вычислений значений оператора L_h за один временной шаг.

3.1. Линейное одномерное уравнение переноса. Рассмотрена начальная задача для уравнения (7) при $c = 1$ на пространственном отрезке $[0, 2]$ и временном интервале $(0, 1)$. Использована пространственная сетка, состоящая из 200 ячеек, оператор $L \equiv -cu_x$ аппроксимировался с использованием левой конечно-разностной производной первого порядка и с использованием потока второго порядка с лимитером minmod [1, с. 92]. Начальные условия выбирались следующими:

$$u(x, 0) = \begin{cases} 3x - 1, & \text{если } 1/3 < x < 2/3, \\ 0 & \text{в противном случае.} \end{cases}$$

Результаты исследования максимальных шагов приведены в табл. 1. Отношения максимальных шагов для разных методов в случае использования левой разностной производной соответствуют тому, как относятся друг к другу радиусы вписанных кругов соответствующих областей устойчивости (ср. рис. 1 и 3).

Таблица 1. Соотношения максимальных шагов при решении линейного одномерного уравнения переноса

Аппроксимация оператора L_h	Параметры методов											
	RK2TVD	RK3TVD			RK2MD-TVD				RK2MD			
	C_m	C_m	T_{2-1}	E_{2-1}	C_m	T_{3-1}	E_{3-1}	T_{3-2}	C_m	T_{4-1}	E_{4-1}	T_{4-2}
Левая разность	1.0	1.245	1.26	0.84	2.0	2.0	1.33	1.6	2.0	2.0	1.33	1.6
minmod	0.95	1.15	1.21	0.81	1.7	1.79	1.19	1.48	1.45	1.525	1.02	1.26

В случае же нелинейной аппроксимации повышенной точности с использованием лимитера minmod наблюдаются как отличия от указанных соотношений для областей устойчивости, так и различия между результатами методов RK2MD-TVD и RK2MD. Оба они удовлетворяют свойству МД и имеют одну и ту же область устойчивости. Тем не менее отсутствие TVD-свойства у второго метода негативно сказывается на максимальной длине временного шага, что косвенно свидетельствует об изменении характера расположения собственных чисел линейной части пространственного оператора.

Как видно из таблицы, метод RK3TVD выигрывает у метода RK2TVD по устойчивости, но существенно проигрывает ему по вычислительной сложности. Этот разрыв оправдан в тех случаях, когда более высокий порядок метода играет существенную роль.

3.2. Одномерная газовая динамика. Для тестирования рассмотренных методов Рунге–Кутты использовался программный комплекс RKDG-1D [5], позволяющий решать одномерные задачи динамики совершенного невязкого газа. Такие задачи описываются системой уравнений [2, с. 31–33]

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \rho \vec{v} = 0, \quad (24)$$

$$\frac{\partial \rho \vec{v}}{\partial t} + \vec{\nabla} \cdot (\rho \vec{v} \vec{v} + p \hat{I}) = 0,$$

$$\frac{\partial e}{\partial t} + \vec{\nabla} \cdot (\vec{v}(e + p)) = 0, \quad (25)$$

где ρ – плотность, $\vec{v} = (u, v, w)^T$ – скорость, p – давление, e – полная энергия единицы объёма, $\hat{I}$ – единичный тензор. Дополняет систему уравнение состояния

$$p = \rho\varepsilon(\gamma - 1),$$

где ε – удельная внутренняя энергия, γ – показатель адиабаты.

В одномерном случае система (24), (25) сводится к системе

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} = 0, \quad (26)$$

где вектор U консервативных переменных и вектор $F(U)$ потоков следующие:

$$U = (\rho, \rho u, \rho v, \rho w, e)^T, \quad F(U) = (\rho u, p + \rho u^2, \rho uv, \rho uw, e)^T.$$

Программный комплекс, описанный в работе [5], основан на разрывном методе Галёркина. Он имеет модульную структуру и обладает широкими возможностями добавления новых методов и алгоритмов, включая методы интегрирования систем обыкновенных дифференциальных уравнений. Для решения тестовой задачи использована линейная аппроксимация приближённого решения на ячейках разностной сетки, численный поток HLLC, индикатор Harten, лимитеры WENO_S [18] и HWENO_SC [19].

В качестве тестовой задачи использован вариант задачи Римана о распаде разрыва – задача Сода [1, с. 227]. Она успешно применяется для тестирования численных схем, поскольку содержит в своем решении три типа разрывов – ударную волну, волну разрежения и контактный разрыв. Задача решалась на отрезке $[0, 1]$ с начальным положением разрыва в точке $x^* = 0.5$ до момента времени $T = 0.2$ с начальными условиями

$$(\rho, u, v, w, p) = \begin{cases} (1, 0, 0, 0, 1), & \text{если } x \leq 0.5; \\ (0.125, 0, 0, 0, 0.1), & \text{если } x > 0.5. \end{cases}$$

Результаты исследования максимальных шагов приведены в табл. 2. Как видим, тенденции, характерные для нелинейной аппроксимации пространственного оператора, подтверждаются. Метод RK2MD-TVD в реальном расчёте оказывается более устойчивым, чем RK2MD за счёт меньших искажений спектра пространственного оператора при выполнении свойства TVD. При этом на спектр оказывает влияние и используемый лимитер. При схожих параметрах T и E для всех методов Рунге–Кутты лимитер HWENO_SC позволяет считать с шагом, более чем в полтора раза большим, чем лимитер WENO_S. Однако существенно выше и вычислительная сложность этого лимитера, поскольку он предполагает переход от консервативных к характеристическим переменным задачи.

Таблица 2. Соотношения максимальных шагов при решении одномерной задачи газовой динамики разрывным методом Галёркина

Лимитер	Параметры методов											
	RK2TVD	RK3TVD			RK2MD-TVD				RK2MD			
	C_m	C_m	T_{2-1}	E_{2-1}	C_m	T_{3-1}	E_{3-1}	T_{3-2}	C_m	T_{4-1}	E_{4-1}	T_{4-2}
WENO_S	0.195	0.225	1.15	0.77	0.34	1.74	1.16	1.51	0.32	1.64	1.09	1.42
HWENO_SC	0.31	0.36	1.16	0.77	0.55	1.77	1.18	1.53	0.5	1.61	1.075	1.39

3.3. Двумерная газовая динамика. Для системы уравнений (24), (25) рассмотрена также пространственно двумерная постановка задачи. С использованием разрывного метода Галёркина разработан параллельный программный комплекс для решения задач газовой динамики на смешанных неструктурированных сетках [20]. Как и в одномерном случае, программный комплекс благодаря своей модульной структуре позволяет использовать различные

комбинации численных потоков, индикаторов и лимитеров. При решении тестовых задач использовались два лимитера – BJVertex [21] и WENO_SC [18, 19], а также два численных потока – Local Lax–Friedrichs и HLL [1, с. 75].

В качестве тестовой выбрана задача о набегающей вертикальной ударной волны с числом Маха $M_s = 10$ на клин, расположенный по отношению к горизонтали под углом $\varphi = 30^\circ$. В результате соударения волны и клина формируется сложная ударно-волновая структура, называемая двойным маховским отражением. Параметры набегающего потока $\rho = 8$, $u = 8.25$, $p = 116.518$, начальные условия $\rho = 1.4$, $u = v = w = 0$, $p = 1$, показатель адиабаты $\gamma = 1.4$.

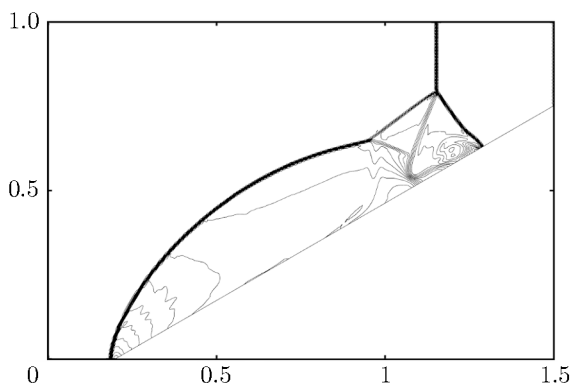


Рис. 4. Линии уровня плотности в задаче о двойном маховском отражении.

В расчётах использованы неструктурированные треугольные сетки из 36 тыс. и 74 тыс. ячеек. Средние длины сторон в указанных сетках равны примерно 0.01 и 0.005 соответственно. Расчёты проводились до момента времени $t^* = 0.1$ с постоянным в течение расчёта шагом по времени.

Характерный вид решения изображён на рис. 4. При постепенном увеличении шага по времени численное решение искажается в силу проявлений немонотонности схемы.

Прежде всего искажения возникают в треугольнике между контактным разрывом и ударной волной. Затем паразитные осцилляции появляются в области постоянства решения. Наконец, в численном решении возникают нефизические волны, расчёт становится неустойчивым и аварийно останавливается до достижения t^* .

При анализе результатов расчётов делалось различие между шагом, при котором проявляются существенные искажения решения, и шагом, при котором расчёт преждевременно прерывается.

Результаты исследования максимальных шагов для грубой и подробной сеток приведены в табл. 3 и 4 соответственно. Исследованы три сочетания численных потоков и лимитеров: поток Lax–Friedrichs и лимитер BJVertex; поток Lax–Friedrichs и лимитер WENO_SC; поток HLL и лимитер BJVertex.

Таблица 3. Соотношения максимальных шагов при решении задачи двойного маховского отражения разрывным методом Галёркина, сетка из 36 тыс. ячеек

Метод	Параметры методов											
	RK2TVD	RK3TVD			RK2MD-TVD				RK2MD			
	$\tau_m \cdot 10^4$	$\tau_m \cdot 10^4$	T_{2-1}	E_{2-1}	$\tau_m \cdot 10^4$	T_{3-1}	E_{3-1}	T_{3-2}	$\tau_m \cdot 10^4$	T_{4-1}	E_{4-1}	T_{4-2}
<i>Появление видимых искажений решения</i>												
BJVertex+LF	0.77	1.00	1.30	0.87	1.60	2.08	1.39	1.60	1.60	2.08	1.39	1.60
WENO_SC+LF	0.79	1.00	1.27	0.84	1.80	2.28	1.52	1.80	1.60	2.03	1.35	1.60
BJVertex+HLL	1.00	1.30	1.30	0.87	1.70	1.70	1.13	1.31	1.70	1.70	1.13	1.31
<i>Аварийный останов расчёта</i>												
BJVertex+LF	1.60	1.90	1.19	0.79	3.30	2.06	1.38	1.74	2.30	1.44	0.96	1.21
WENO_SC+LF	1.60	1.90	1.19	0.79	2.70	1.69	1.13	1.42	2.30	1.44	0.96	1.21
BJVertex+HLL	1.80	1.80	1.00	0.67	3.50	1.94	1.30	1.94	2.50	1.39	0.93	1.39

Во всех случаях метод RK2MD-TVD даёт возможность считать с бóльшим шагом по времени, чем остальные методы. Его преимущество над методом RK2TVD по количеству вычислений оператора L_h в подавляющем большинстве расчётов составляет не менее 13%, а часто – около 30%. Вычислительное преимущество этого метода над RK3TVD составляет не менее 25%. При этом не обладающий TVD-свойством метод RK2MD хотя и близок к методу RK2MD-TVD по характерным шагам возникновения немонотонности решения, но существенно уступает ему в общей устойчивости, иногда по вычислительным затратам проигрывая и

RK2TVD. Это означает, что данный метод оказывается менее устойчив к возникающим артефактам численного решения, вызванным немонотонностью.

Таблица 4. Соотношения максимальных шагов при решении задачи двойного маховского отражения разрывным методом Галёркина, сетка из 74 тыс. ячеек

Метод	Параметры методов											
	RK2TVD	RK3TVD			RK2MD-TVD				RK2MD			
	$\tau_m \cdot 10^4$	$\tau_m \cdot 10^4$	T_{2-1}	E_{2-1}	$\tau_m \cdot 10^4$	T_{3-1}	E_{3-1}	T_{3-2}	$\tau_m \cdot 10^4$	T_{4-1}	E_{4-1}	T_{4-2}
Появление видимых искажений решения												
BJVertex+LF	0.51	0.60	1.18	0.78	1.00	1.96	1.31	1.67	1.00	1.96	1.31	1.67
WENO_SC+LF	0.44	0.58	1.32	0.88	1.00	2.27	1.52	1.72	1.00	2.27	1.52	1.72
BJVertex+HLL	0.65	0.80	1.23	0.82	1.00	1.54	1.03	1.25	1.10	1.69	1.13	1.38
Аварийный останов расчёта												
BJVertex+LF	0.66	0.79	1.20	0.80	1.30	1.97	1.31	1.65	1.20	1.82	1.21	1.52
WENO_SC+LF	0.66	0.79	1.20	0.80	1.30	1.97	1.31	1.65	1.20	1.82	1.21	1.52
BJVertex+HLL	0.81	0.84	1.04	0.69	1.50	1.85	1.23	1.79	1.20	1.48	0.99	1.43

Оба метода, обладающие свойством МД, позволяют считать с существенно бóльшим временным шагом, чем метод RK3TVD. Одновременно максимальный шаг, обеспечиваемый методом RK2TVD, как правило, лишь немногим меньше, чем у RK3TVD. В то же время все три указанных метода имеют лишь второй порядок точности, а RK3TVD – третий. Это важно как при численном исследовании гладких решений, так и для получения высокого разрешения разрывных решений при использовании, впрочем, соответствующей по порядку аппроксимации оператора L_h .

Возможность увеличения шага по времени после фиксации искажений решения существенно зависит от используемых элементов разностной схемы, прежде всего, численных потоков. Так, например, при использовании потока HLL минимальный шаг, при котором расчёт неустойчив, мало отличается от длины шага, когда возникают видимые искажения. При этом, в среднем, схема с потоком HLL оказалась более устойчивой, чем с потоком Lax–Friedrichs, как относительно немонотонности решения, так и относительно аварийных остановов.

Отметим, что выбор лимитера без изменения численного потока больше влияет на проявления немонотонности решения, чем на устойчивость метода. В частности, расчёты с лимитером BJVertex позволяют получать корректную картину решения при большем шаге, чем с лимитером WENO_SC.

Заключение. Изучены методы Рунге–Кутты, область устойчивости которых включает круг максимального для заданных количества стадий и порядка диаметра. Такие методы востребованы при аппроксимации гиперболических систем уравнений с использованием разрывного метода Галёркина. Получены условия, при которых такие методы удовлетворяют условию TVD. Рассмотрены два метода рассматриваемого класса, их свойства исследованы на численных примерах. С помощью тестовых задач для линейного одномерного уравнения переноса и для системы уравнений газовой динамики в одномерной и двумерной пространственных постановках исследована способность полученных методов сохранять устойчивость и монотонность численного решения при максимально возможных шагах по времени.

Исследование В.В. Лукина выполнено при финансовой поддержке Российского научного фонда (проект 17-79-20445).

СПИСОК ЛИТЕРАТУРЫ

1. Куликовский А.Г., Погорелов Н.В., Семенов А.Ю. Математические вопросы численного решения гиперболических систем уравнений. М., 2001.
2. Самарский А.А., Попов Ю.П. Разностные методы решения задач газовой динамики. М., 1992.
3. Cockburn B., Shu C.-W. TVB Runge–Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework // Math. of Comp. 1989. V. 52. № 186. P. 411–435.

4. Cockburn B., Shu C.-W. Runge–Kutta discontinuous Galerkin methods for convection-dominated problems // J. Sci. Comput. 2001. V. 16. P. 2440–2463.
5. Галенова В.Д., Лукин В.В., Марчевский И.К., Фуфаев И.Н. Сравнительное исследование лимитеров семейства WENO и Hermite WENO для расчета одномерных течений газа методом RKDG // Препринты ИПМ им. М.В. Келдыша. 2017. № 131.
6. Dolejsi V., May G., Rangarajan A. A continuous hp-mesh model for adaptive discontinuous Galerkin schemes // Appl. Numer. Math. 2018. V. 124. P. 1–21.
7. Klein B., Müller B., Kummer F., Oberlack M. A high-order discontinuous Galerkin solver for low-Mach number flows // Int. J. Num. Meth. Fluids. 2015. V. 81. P. 489–520.
8. Galanin M.P., Lukin V.V., Chechetkin V.M. 3D hydrodynamical simulation of accretion disk in binary star system using RKDG CFD solver // J. Phys. Conf. Ser. 2018. V. 1103. P. 012019.
9. Галанин М.П., Савенков Е.Б., Токарева С.А. Решение задач газовой динамики с ударными волнами RKDG-методом // Мат. моделирование. 2008. Т. 20. № 11. С. 55–66.
10. Shu C.-W., Osher. S. Efficient implementation of essentially non-oscillatory shock capturing schemes // J. Comp. Phys. 1988. V. 77. P. 439–471.
11. Gottlieb S., Ketcheson D., Shu C.-W. High order strong stability preserving time discretizations // J. Sci. Comput. 2009. V. 38. P. 251–289.
12. Toulorge T., Desmet W. Optimal Runge–Kutta schemes for discontinuous Galerkin space discretizations applied to wave propagation problems // J. Comp. Phys. 2011. V. 231. P. 2067–2091.
13. Jeltsch R., Torrilhon M. Flexible stability domains for explicit Runge–Kutta methods // Some Topics in Industrial and Applied Mathematics / Eds. R. Jeltsch, D. Li, I.H. Sloan. Beijing, 2007.
14. Ладонкина М.Е., Неклюдова О.А., Остапенко В.В., Тишкин В.Ф. Исследование точности разрывного метода Галеркина при расчете решений с ударными волнами // Препринты ИПМ им. М.В. Келдыша. 2018. № 195.
15. Хайрер Э., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Жесткие и дифференциально-алгебраические задачи. М., 1999.
16. Хайрер Э., Нерсетт С., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Нежесткие задачи. М., 1990.
17. Годунов С.К. Разностный метод численного расчета разрывных решений уравнений гидродинамики // Мат. сб. 1959. Т. 47. Вып. 3. С. 271–306.
18. Zhong X., Shu C.-W. A simple weighted essentially nonoscillatory limiter for Runge–Kutta discontinuous Galerkin methods // J. Comp. Phys. 2013. V. 232. P. 397–415.
19. Zhu J., Zhong X., Shu C.-W., Qiu J.-X. Runge–Kutta discontinuous Galerkin method with a simple and compact Hermite WENO limiter // Comm. in Comp. Phys. 2016. V. 19. P. 944–969.
20. Корчагова В.Н., Сауткина С.М., Лукин В.В., Фуфаев И.Н. Экспериментальная оценка эффективности распараллеливания RKDG-метода для решения двумерных газодинамических задач // XIV междунар. конф. “Параллельные вычислительные технологии” (ПаВТ–2020). Пермь, 31 марта – 2 апреля 2020 г. Короткие статьи и описания плакатов. Челябинск, 2020. С. 142–151.
21. Giuliani A., Krivodonova L. Analysis of slope limiters on unstructured triangular meshes // J. of Comp. Phys. 2018. V. 374. P. 1–26.

Институт прикладной математики
им. М.В. Келдыша РАН, г. Москва,
Институт системного программирования
им. В.П. Иванникова РАН, г. Москва,
Московский государственный технический
университет им. Н.Э. Баумана

Поступила в редакцию 11.02.2021 г.
После доработки 11.02.2021 г.
Принята к публикации 27.04.2021 г.

УДК 519.633

КОМПАКТНЫЕ РАЗНОСТНЫЕ СХЕМЫ НА ТРЁХТОЧЕЧНОМ ШАБЛОНЕ ДЛЯ ГИПЕРБОЛИЧЕСКИХ УРАВНЕНИЙ ВТОРОГО ПОРЯДКА

© 2021 г. П. П. Матус, Хоанг Тхи Киеу Ань

Для уравнений Клейна–Гордона с постоянными и переменными коэффициентами на трёхточечном шаблоне по пространству рассматриваются компактные разностные схемы $4+2$ порядка аппроксимации. Предложены новые компактные схемы для одного типа квазилинейных гиперболических уравнений второго порядка. В случае постоянных коэффициентов доказана сильная устойчивость разностного решения по отношению к малому возмущению начальных условий, правой части и коэффициентов уравнения. Получены априорные оценки устойчивости и сходимости разностного решения в сильных сеточных нормах.

DOI: 10.31857/S0374064121070098

1. Введение. Повышение точности вычислительного метода для решения задач математической физики на минимальных шаблонах всегда представляло собой актуальную задачу численного анализа (см., например, [1–4]). Среди методов построения разностных схем повышенного порядка аппроксимации особое место занимают так называемые компактные схемы, которые пишутся на шаблоне, несущественно отличающемся от традиционных для данного уравнения [5]. Основопологающей работой по этой тематике для классических уравнений математической физики с самосопряжённым эллиптическим оператором является работа А.А. Самарского [6], опубликованная им более 50 лет назад. Для других классов уравнений, включая уравнения типа конвекции–диффузии, нелинейные уравнения без смешанных производных и задачи аэрогидродинамики, компактные разностные схемы построены в [7–9].

В настоящей работе для различных типов уравнения Клейна–Гордона исследуются на обычном трёхточечном шаблоне компактные разностные схемы $4+2$ порядка точности. Это уравнение играет важную роль в математической физике; в частности, оно используется при изучении солитонов и в физике конденсированного вещества [10]. Компактные разностные схемы для такого уравнения строятся и изучаются, например, в [9, 11]. Некоторые результаты по данной тематике также анонсированы в [12, 13]. Для этого уравнения, несмотря на линейность дифференциальной и разностной задач с переменными коэффициентами, для получения соответствующих априорных оценок не удаётся применить известные результаты по теории Самарского устойчивости трёхслойных операторно-разностных схем [1, гл. VI, § 3]. В работе с использованием метода энергетических неравенств для компактных разностных схем, аппроксимирующих уравнения Клейна–Гордона с переменными коэффициентами, получены априорные оценки устойчивости и сходимости разностного решения в сеточных нормах $L_2(\omega_h)$, $W_2^1(\omega_h)$, $C(\omega_h)$ или $L_\infty(\omega_h)$. На примере вычислительного эксперимента для квазилинейного уравнения, которое является следствием системы уравнений газовой динамики, показывается, как использовать правило Рунге для определения разных порядков скорости сходимости решения разностной схемы в случае двух независимых переменных.

1. Необходимые и достаточные условия устойчивости двух и трёхслойных операторно-разностных схем. При исследовании компактных разностных схем, аппроксимирующих линейные уравнения Клейна–Гордона, естественно воспользоваться общей теорией Самарского [1, гл. VI, § 3] операторно-разностных схем. Ниже предлагается использовать другие канонические формы операторно-разностных схем, для которых условия устойчивости формулируются значительно проще и не содержат условий на связь между операторами.

Пусть задано действительное конечномерное евклидово пространство H и сетка по времени $\overline{\omega}_\tau = \{t_n = n\tau, \quad n = \overline{0, N_0}, \quad \tau N_0 = T\} = \omega_\tau \cup \{0\}$. Скалярное произведение в H обозначим через $(\cdot, \cdot)$, и пусть $A, B, D : H \rightarrow H$ – линейные операторы, не зависящие от τ и t_n .

Рассмотрим задачу Коши для двухслойной операторно-разностной схемы

$$By_t + Ay^{(0.5)} = \varphi(t), \quad t \in \omega_\tau, \quad (1)$$

$$y_0 = u_0, \quad (2)$$

где $y^n = y(t_n) \in H$ – искомая вектор-функция, а $\varphi^n = \varphi(t_n)$, u_0 заданы, $y^{(\sigma)} = \sigma y^{n+1} + (1 - \sigma)y^n$, $y^n \in H$. Тогда критерий Самарского (см. [1, с. 333]) при $\varphi(t) \equiv 0$ можно сформулировать следующим образом.

Теорема 1. *Условия*

$$B = B(t) \geq 0, \quad A = A^* > 0, \quad A - \text{постоянный оператор,}$$

необходимы и достаточны для устойчивости решения разностной задачи (1), (2) в H_A по начальным данным, т.е. для выполнения оценки

$$\|y^n\|_A \leq \|u_0\|_A, \quad n = \overline{1, N_0},$$

где, как обычно, $\|v\|_A = (Av, v)$ для любого $v \in H$.

Для трёхслойных операторно-разностных схем будем использовать следующую каноническую форму:

$$Dy_{tt} + A_1y^{(0.5, 0.5)} = \varphi(t), \quad 0 < t \in \omega_\tau, \quad (3)$$

$$y_0 = u_0, \quad y_1 = u_1, \quad (4)$$

где $y^{(\sigma_1, \sigma_2)} = \sigma_1 y^{n+1} + (1 - \sigma_1 - \sigma_2)y^n + \sigma_2 y^{n-1}$, $0 \leq \sigma_1, \sigma_2 \leq 1$.

В работе используются обозначения из [1, 2]. Далее будем предполагать, что D, A_1 – положительные, самосопряжённые и постоянные операторы:

$$D = D^* > 0, \quad A_1^* = A_1 > 0. \quad (5)$$

Тогда имеет место следующая

Теорема 2. *Разностная схема (3), (4) устойчива в H_{A_1} по начальным данным и правой части, и имеет место априорная оценка*

$$Q^{n+1} \leq Q^1 + \sum_{k=1}^n \tau \|\varphi(t_k)\|_{D^{-1}}.$$

Здесь $Q^n = \{\|y_t\|_D^2 + (\|y\|_{A_1}^2 + \|\check{y}\|_{A_1}^2)/2\}^{1/2}$.

Доказательство. Умножая уравнение (3) скалярно в H на $2\tau y_t^\circ = \tau(y_t + y_t)$, приходим к энергетическому соотношению

$$Q_{n+1}^2 - Q_n^2 = 2\tau(y_t^\circ, \varphi) \leq \tau(\|y_t\|_D + \|y_t\|_D)\|\varphi\|_{D^{-1}} \leq \tau(Q_{n+1} + Q_n)\|\varphi(t_n)\|_{D^{-1}}.$$

Отсюда и следует требуемая оценка.

2. Уравнение Клейна–Гордона с постоянными коэффициентами.

2.1. Постановка задачи и разностная схема. В области $\overline{Q}_T = \{(x, t) : 0 \leq x \leq l, \quad 0 \leq t \leq T\}$ рассмотрим начально-краевую задачу для уравнения Клейна–Гордона с постоянными коэффициентами

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} - mu + f(x, t), \quad m = \text{const} > 0, \quad (6)$$

$$u(x, 0) = u_0(x), \quad \frac{\partial u}{\partial t}(x, 0) = \overline{u}_0(x), \quad (7)$$

$$u(0, t) = \mu_1(t), \quad u(l, t) = \mu_2(t). \quad (8)$$

Отметим, что уравнение (6) является обобщением волнового уравнения и используется для описания быстро движущихся частиц, имеющих массу покоя. Здесь и далее относительно решения дифференциальной задачи будем предполагать, что оно существует, единственно и обладает всеми непрерывными в $\overline{Q}_T$ производными, необходимыми по ходу изложения.

На равномерной сетке узлов $\overline{\omega} = \overline{\omega}_h \times \overline{\omega}_\tau = \{(x_i, t_n) \in \overline{Q}_T\}$, $\overline{\omega}_h = \{x_i = ih, \quad 0 \leq i \leq N, \quad h = l/N\} = \omega_h \cup \{0, l\}$, $\overline{\omega}_\tau = \{t_n = n\tau, \quad 0 \leq n \leq N_0, \quad \tau = T/N_0\} = \omega_\tau \cup \{0\}$ дифференциальную задачу заменим разностной:

$$y_{\bar{t}t} = \Lambda y^{(\sigma, \sigma)} - m \left(y + \frac{h^2}{12} \Lambda y \right) + \varphi, \quad (x, t) \in \omega_h \times \omega_\tau, \quad (9)$$

$$y(x, 0) = u_0(x), \quad x \in \overline{\omega}_h, \quad y_t(x, 0) = u_1(x), \quad x \in \omega_h, \quad (10)$$

$$y(0, t) = \mu_1(t), \quad y(l, t) = \mu_2(t), \quad t \in \omega_\tau, \quad (11)$$

в которой

$$\hat{y} = y^{n+1}, \quad \check{y} = y^{n-1}, \quad y_i^n = y(x_i, t_n),$$

$$\Lambda y = y_{\bar{x}x}, \quad \varphi = f + \frac{h^2}{12} \Lambda f, \quad \sigma = \bar{\sigma} - \frac{h^2}{12\tau^2},$$

$$u_1(x) = \bar{u}_0(x) + \frac{\tau}{2} [u_0''(x) - m u_0(x) + f(x, 0)], \quad x \in \omega_h.$$

Как и в монографии [1, с. 309], нетрудно показать, что для невязки

$$\psi = -u_{\bar{t}t} + \Lambda u^{(\sigma, \sigma)} - m \left(u + \frac{h^2}{12} \Lambda u \right) + \varphi$$

и погрешности аппроксимации второго начального условия имеют место априорные оценки

$$\|\psi\| \leq M(h^4 + \tau^2), \quad M = \text{const} > 0, \quad (12)$$

$$\|\overset{\circ}{\psi}\| = \|u_1 - u_t^0\| \leq M_1 \tau^2, \quad M_1 = \text{const} > 0, \quad (13)$$

т.е. разностная схема (9)–(11) аппроксимирует исходную дифференциальную задачу с четвёртым порядком по пространству и вторым по времени.

2.2. Устойчивость по начальным данным и правой части. Для исследования этих вопросов в линейном случае обычно применяют теорию трёхслойных операторно-разностных схем, разработанную А.А. Самарским [1, гл. VI, § 3]. Требование самосопряжённости основного пространственного оператора

$$(Ay)_i = -(\Lambda y)_i, \quad i = \overline{1, N-1}, \quad (Ay)_0 = 0, \quad (Ay)_N = 0, \quad (14)$$

приводит к жёстким ограничениям на однородность граничных условий. Чтобы избежать этого, рассмотрим возмущённое решение $\tilde{y}$, полученное по разностной схеме (9)–(11) с возмущённой правой частью $\tilde{f}$ и возмущёнными начальными условиями $\tilde{u}_0, \tilde{u}_1$. Тогда задача для возмущения $\overline{y} = \tilde{y} - y$ может быть записана в операторном виде:

$$D\overline{y}_{\bar{t}t} + A_1\overline{y} = \overline{\varphi}, \quad \overline{\sigma} = 1, \quad (15)$$

$$\overline{y}(0) = \overline{u}_0, \quad \overline{y}_t(0) = \overline{u}_1, \quad (16)$$

$$D = E + \sigma\tau^2 A, \quad A_1 = mE + \left(1 - \frac{mh^2}{12}\right) A.$$

Здесь $\overline{u}_0 = \tilde{u}_0 - u_0$, $\overline{u}_1 = \tilde{u}_1 - u_1$, $\overline{\varphi} = \tilde{\varphi} - \varphi$.

Оператор A , определённый соотношением (14), является постоянным, положительным и самосопряжённым, т.е. $0 < A^* = A : H \rightarrow H$, здесь H – пространство сеточных функций, заданных на $\bar{\omega}_h$ и равных нулю при $x = 0$ и $x = l$. Поэтому операторы D и A_1 также являются положительными, постоянными и самосопряжёнными: $D = D^* > E$, $A_1 = A_1^* > A$.

Отметим следующие хорошо известные свойства оператора A [1, гл. II, § 3]:

$$\lambda_1 E \leq A \leq \lambda_2 E, \quad A^{-1} \leq \frac{1}{\delta} E, \quad \delta = \frac{8}{l^2}, \quad (17)$$

$$\lambda_1 = \frac{4}{h^2} \sin^2 \frac{\pi h}{2l} \geq \delta, \quad \lambda_2 = \frac{4}{h^2} \cos^2 \frac{\pi h}{2l} < \frac{4}{h^2}. \quad (18)$$

В дальнейшем нам понадобится следующая

Лемма 1 [1, с. 373]. Пусть в канонической форме (15), (16) операторы D и A_1 являются постоянными, положительными и самосопряжёнными в H и, кроме того, выполнено неравенство

$$D \geq \frac{1+\varepsilon}{4} \tau^2 A_1, \quad \text{где } \varepsilon > 0 \text{ – некоторое число.} \quad (19)$$

Тогда для решения схемы (15), (16) имеет место априорная оценка

$$\|\bar{y}^{n+1}\|_D \leq \sqrt{\frac{1+\varepsilon}{\varepsilon}} \left(\|\bar{y}(0)\|_D + \|D\bar{y}_t(0)\|_{A_1^{-1}} + \sum_{s=1}^n \tau \|\bar{\varphi}_s\|_{A_1^{-1}} \right). \quad (20)$$

В частности, если $\varepsilon = 1$, $D \geq E$, $A_1^{-1} < A^{-1} \leq \delta^{-1} E$, оценка (20) примет вид

$$\|\bar{y}^{n+1}\| \leq \sqrt{2} \left(\|\bar{y}(0)\|_D + \frac{1}{\delta} \|D\bar{y}_t(0)\| + \frac{1}{\delta} \sum_{s=1}^n \tau \|\bar{\varphi}_s\| \right). \quad (21)$$

Применим оценку (21) к исследованию устойчивости компактной разностной схемы (9)–(11). Далее будем предполагать, что

$$\frac{h}{\sqrt{6}} \leq \tau \leq \frac{1}{\sqrt{m}}, \quad \sigma = 1 - \frac{h^2}{12\tau^2}. \quad (22)$$

Тогда условие (19) выполнено. Действительно,

$$D - \frac{\tau^2}{2} A_1 = \frac{1}{2} E + \left(\frac{1}{2} - \frac{m\tau^2}{2} \right) E + \tau^2 \left[\sigma - \frac{1}{2} \left(1 - \frac{mh^2}{12} \right) \right] A \geq \frac{1}{2} E.$$

Итак, имеет место следующая

Теорема 3. Пусть выполнено условие (22). Тогда разностная схема с неоднородными граничными условиями (9)–(11) условно устойчива по начальным данным и правой части, а для её решения имеет место априорная оценка

$$\|\tilde{y}^{n+1} - y^{n+1}\| \leq \sqrt{2} \left(\|\tilde{u}_0 - u_0\|_D + \frac{1}{\delta} \|D(\tilde{u}_1 - u_1)\| + \frac{1}{\delta} \sum_{s=1}^n \tau \|\tilde{\varphi}_s - \varphi_s\| \right).$$

Замечание. В линейном случае после замены

$$u = v + \frac{x}{l} \mu_2(t) + \frac{l-x}{l} \mu_1(t)$$

исходная дифференциальная задача сводится к задаче для v с однородными граничными условиями. При этом устойчивость по граничным условиям и правой части решения разностных схем для этих задач имеет место одновременно для обеих задач. Однако для нелинейного уравнения Клейна–Гордона такая замена может не дать желаемого результата.

Итак, мы доказали устойчивость решения разностной схемы в слабой энергетической норме $L_2(\omega_h)$ при весьма слабых ограничениях на вес функции σ . Такие оценки полезны при исследовании сходимости разностных схем с обобщёнными решениями [14]. Чтобы получить априорные оценки в более сильных нормах $W_2^1(\omega_h)$ и $C(\omega_h)$, воспользуемся теоремой 2 и вложением (см. [1, с. 107])

$$\|\tilde{y} - y\|_C \leq \frac{\sqrt{l}}{2} \|\tilde{y}_x - y_x\|. \quad (23)$$

Запишем разностную схему для возмущения $\bar{y} = \tilde{y} - y$ в каноническом виде (3) при $h/\sqrt{6} \leq \tau \leq 1/\sqrt{m}$, $\sigma = 1 - h^2/(12\tau^2)$,

$$D\bar{y}_{tt} + A_1\bar{y}^{(0.5,0.5)} = \bar{\varphi}, \quad t \in \omega_\tau, \quad (24)$$

$$\bar{y}(0) = \bar{u}_0, \quad \bar{y}_t(0) = \bar{u}_1, \quad (25)$$

где

$$D = E + \tau^2(\sigma - 0.5)A_1 = D^* \geq \frac{1}{2}E, \quad D^{-1} \leq 2E, \quad A_1^* = A_1 = mE + \left(1 - \frac{mh^2}{12}\right)A > A \geq \delta E.$$

Тогда имеет место следующая

Теорема 4. *Решение разностной схемы (24), (25) при $h/\sqrt{6} \leq \tau \leq 1/\sqrt{m}$, $\sigma = 1 - h^2/(12\tau^2)$ устойчиво по начальным данным и правой части, и для всех $n = \overline{2, N_0}$ имеют место априорные оценки*

$$\begin{aligned} \frac{1}{\sqrt{2}}(\|\bar{y}_x^n\|^2 + \|\bar{y}_x^{n-1}\|^2)^{1/2} &\leq Q^1 + 2 \sum_{k=1}^{n-1} \tau \|\tilde{\varphi}(t_k) - \varphi(t_k)\|, \\ \|\tilde{y}^n - y^n\|_C &\leq \sqrt{\frac{l}{2}} \left\{ Q^1 + 2 \sum_{k=1}^{n-1} \tau \|\tilde{\varphi}(t_k) - \varphi(t_k)\| \right\}. \end{aligned} \quad (26)$$

Доказательство. При $h/\sqrt{6} \leq \tau \leq 1/\sqrt{m}$, $\sigma = 1 - h^2/(12\tau^2)$ выполнено условие (5), поэтому в силу теоремы 2 имеем

$$\frac{1}{\sqrt{2}}(\|\bar{y}^n\|_{A_1}^2 + \|\bar{y}^{n-1}\|_{A_1}^2)^{1/2} \leq Q^1 + \sum_{k=1}^{n-1} \tau \|\bar{\varphi}(t_k)\|_{D^{-1}}.$$

Так как $A_1 > A$ и $D^{-1} \leq 2E$, то справедлива оценка

$$\frac{1}{\sqrt{2}}(\|\bar{y}_x^n\|^2 + \|\bar{y}_x^{n-1}\|^2)^{1/2} \leq Q^1 + 2 \sum_{k=1}^{n-1} \tau \|\bar{\varphi}(t_k)\|.$$

Следовательно, из леммы и вложения (23) вытекает, что

$$\|\tilde{y}^n - y^n\|_C \leq \sqrt{\frac{l}{2}} \left\{ Q^1 + 2 \sum_{k=1}^{n-1} \tau \|\tilde{\varphi}(t_k) - \varphi(t_k)\| \right\}.$$

Теорема доказана.

2.3. Сильная устойчивость. При исследовании корректности разностных схем основное внимание уделяется устойчивости решения по начальным данным и правой части [1, 2]. Однако при численном решении дифференциальной задачи может оказаться, что коэффициенты уравнения заданы не точно, а приближённо. Это показывает, насколько важно изучение схем с

возмущёнными коэффициентами. Под *сильной устойчивостью* понимается устойчивость решения разностной задачи по отношению к малому возмущению начальных условий, правой части и коэффициентов уравнения [15].

Наряду с разностной схемой (9)–(11) рассмотрим соответствующую ей возмущённую задачу

$$\tilde{y}_{tt} = \Lambda \tilde{y}^{(\sigma, \sigma)} - \tilde{m} \left(\tilde{y} + \frac{h^2}{12} \Lambda \tilde{y} \right) + \tilde{\varphi}, \quad (x, t) \in \omega_h \times \omega_\tau, \quad (27)$$

$$\tilde{y}(x, 0) = \tilde{u}_0(x), \quad x \in \overline{\omega}_h, \quad \tilde{y}_t(x, 0) = \tilde{u}_1(x), \quad x \in \omega_h, \quad (28)$$

$$\tilde{y}(0, t) = \mu_1(t), \quad \tilde{y}(l, t) = \mu_2(t), \quad t \in \omega_\tau. \quad (29)$$

Вычитая из уравнений (27)–(29) соответствующие уравнения (9)–(11), получаем задачу для возмущения $\overline{y} = \tilde{y} - y$, которую запишем в операторном виде (15), (16):

$$D\overline{y}_{tt} + \tilde{A}_1\overline{y} = \overline{\varphi} - (\tilde{A}_1 - A_1)y,$$

$$\overline{y}(0) = \overline{u}_0, \quad \overline{y}_t(0) = \overline{u}_1,$$

$$\tilde{A}_1 = \tilde{m}E + \left(1 - \frac{\tilde{m}h^2}{12}\right)A.$$

Здесь $D = D^* > E$, $\tilde{A}_1 = \tilde{A}_1^* > A \geq \delta E$ – положительные, постоянные и самосопряжённые операторы.

Заметим, что при $h/\sqrt{6} \leq \tau \leq \sqrt{1/\tilde{m}}$ выполнено условие (19), поэтому на основании априорной оценки (21) для $\overline{y}$ получаем неравенство

$$\|\overline{y}^{n+1}\| \leq \sqrt{2} \left(\|\overline{y}(0)\|_D + \frac{1}{\delta} \|D\overline{y}_t(0)\| + \frac{1}{\delta} \sum_{s=1}^n \tau \{ \|\overline{\varphi}_s\| + \|(\tilde{A}_1 - A_1)y_s\| \} \right).$$

В силу соотношений (17), (18) имеем $\|A\| < 4/h^2$, и, следовательно, из (21) вытекает оценка

$$\|(\tilde{A}_1 - A_1)y_s\| = \left\| (\tilde{m} - m) \left(E - \frac{h^2}{12} A \right) y_s \right\| < \frac{4}{3} K_s |\tilde{m} - m|,$$

в которой

$$K_s = \sqrt{2} \left(\|u_0\|_D + \frac{1}{\delta} \|Du_1\| + \frac{1}{\delta} \sum_{r=1}^{s-1} \tau \|\varphi_r\| \right).$$

Итак, мы можем сформулировать теорему о сильной устойчивости.

Теорема 5. Пусть выполнено следующее условие на шаги сетки:

$$\frac{h}{\sqrt{6}} \leq \tau \leq \frac{1}{\sqrt{\tilde{m}}}, \quad \overline{m} = \max\{m, \tilde{m}\}.$$

Тогда решение разностной схемы (9)–(11) сильно устойчиво и для её возмущения имеет место априорная оценка

$$\|\tilde{y}^{n+1} - y^{n+1}\| \leq \sqrt{2} \left\{ \|\tilde{u}_0 - u_0\|_D + \frac{1}{\delta} \|D(\tilde{u}_1 - u_1)\| + \frac{1}{\delta} \sum_{s=1}^n \tau \left(\|\tilde{\varphi}_s - \varphi_s\| + \frac{4}{3} K_s |\tilde{m} - m| \right) \right\}, \quad n = \overline{1, N_0 - 1}.$$

2.4. Теорема о сходимости. Обозначим через $z = y - u$ погрешность метода. Заменяя в разностных уравнениях (9)–(11) y на $z + u$, получаем задачу для z :

$$z_{tt} = \Lambda z^{(\sigma, \sigma)} - m \left(z + \frac{h^2}{12} \Lambda z \right) + \psi, \quad (x, t) \in \omega_h \times \omega_\tau, \quad (30)$$

$$z(x, 0) = 0, \quad x \in \bar{\omega}_h, \quad z_t(x, 0) = \overset{\circ}{\psi}, \quad x \in \omega_h, \quad (31)$$

$$z(0, t) = 0, \quad z(l, t) = 0, \quad t \in \omega_\tau. \quad (32)$$

Так как задачи (9)–(11) и (30)–(32) идентичны, то мы можем воспользоваться теоремой 4 для оценки погрешности метода.

Теорема 6. Пусть выполнены условия теоремы 4. Тогда решение разностной задачи (9)–(11) сходится к точному решению дифференциальной задачи (6)–(8) в сеточной норме $C(\omega_h)$ и для её решения имеет место оценка точности вида

$$\max_{t \in \bar{\omega}_\tau} \|y^n - u^n\|_C \leq M_2(h^4 + \tau^2), \quad M_2 = \text{const} > 0.$$

Доказательство. Действительно, из априорных оценок (12), (13) и (26) следует неравенство

$$\|y^n - u^n\|_C \leq \sqrt{\frac{l}{2}} \left\{ Q(z^1) + 2 \sum_{k=1}^{n-1} \tau \|\psi(t_k)\| \right\} \leq M_2(h^4 + \tau^2).$$

Следовательно, разностное решение сходится к точному решению с четвёртым порядком по пространству и вторым по времени.

3. Уравнение Клейна–Гордона с переменными коэффициентами.

3.1. Постановка задачи и разностная схема. В области $\bar{Q}_T = \{(x, t) : 0 \leq x \leq l, 0 \leq t \leq T\}$ рассмотрим начально-краевую задачу для уравнения Клейна–Гордона с переменными коэффициентами

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x} \left(k(x, t) \frac{\partial u}{\partial x} \right) - mu + f(x, t), \quad m = \text{const} > 0, \quad (33)$$

$$u(x, 0) = u_0(x), \quad \frac{\partial u}{\partial t}(x, 0) = \bar{v}_0(x), \quad (34)$$

$$u(0, t) = \mu_1(t), \quad u(l, t) = \mu_2(t), \quad (35)$$

где $0 < k_1 \leq k(x, t) \leq k_2$, $u(x, t) \in C^{4,6}(\bar{Q}_T)$, $p \in C^{0,5}(\bar{Q}_T)$, $f \in C^{0,4}(\bar{Q}_T)$.

На построенной сетке $\bar{\omega}$ исходную дифференциальную задачу аппроксимируем разностной схемой вида

$$y_{\bar{t}t} = \Lambda y^{(\sigma, \sigma)} - \frac{h^2}{12} \Lambda(p y_{\bar{t}t}) - m \left[y^{(\sigma, \sigma)} + \frac{h^2}{12} \Lambda(p y^{(\sigma, \sigma)}) \right] + \varphi, \quad (x, t) \in \omega_h \times \omega_\tau, \quad (36)$$

$$y(x, 0) = u_0(x), \quad x \in \bar{\omega}_h, \quad y_t(x, 0) = u_1(x), \quad x \in \omega_h, \quad (37)$$

$$y(0, t) = \mu_1(t), \quad y(l, t) = \mu_2(t), \quad t \in \omega_\tau. \quad (38)$$

Здесь

$$\Lambda y = (a(x, t_n) y_{\bar{x}})_x, \quad \sigma = 0.5, \quad \varphi = f + \frac{h^2}{12} \Lambda(pf), \quad p(x, t) = \frac{1}{k(x, t)},$$

$$a(x, t) = 6 \left[p(x - h, t) + 4p\left(x - \frac{h}{2}, t\right) + p(x, t) \right]^{-1}, \quad 0 < c_1 \leq a(x, t) \leq c_2,$$

$$u_1(x) = \bar{v}_0(x) + \frac{\tau}{2} [(k(x, 0) u'(x, 0))' - m u(x, 0) + f(x, 0)], \quad x \in \omega_h.$$

Следуя работам [5, 6, 8], нетрудно показать, что разностная схема (36)–(38) аппроксимирует исходную задачу (33)–(35) с четвёртым порядком по пространству и вторым по времени, т.е. для её невязки

$$\psi = -u_{\bar{t}t} + \Lambda u^{(\sigma, \sigma)} - \frac{h^2}{12} \Lambda(p u_{\bar{t}t}) - m \left[u^{(\sigma, \sigma)} + \frac{h^2}{12} \Lambda(p u^{(\sigma, \sigma)}) \right] + \varphi$$

и второго начального условия имеют место априорные оценки

$$\|\psi\| \leq M(h^4 + \tau^2), \quad M = \text{const} > 0,$$

$$\|\overset{\circ}{\psi}\| = \|u_1 - u_1^0\| \leq M_1 \tau^2, \quad M_1 = \text{const} > 0.$$

3.2. Устойчивость. Чтобы избежать громоздких выкладок, ограничимся случаем зависимости коэффициента $k = k(x)$ только от пространственной переменной. Рассмотрим возмущённое решение $\tilde{y}$, полученное по разностной схеме (36)–(38) с возмущённой правой частью $\tilde{f}$ и возмущёнными начальными условиями $\tilde{u}_0, \tilde{u}_1$. Тогда задача для возмущения $\bar{y} = \tilde{y} - y$ примет вид

$$\bar{y}_{tt} = \Lambda \bar{y}^{(\sigma, \sigma)} - \frac{h^2}{12} \Lambda(p \bar{y}_{tt}) - m \left[\bar{y}^{(\sigma, \sigma)} + \frac{h^2}{12} \Lambda(p \bar{y}^{(\sigma, \sigma)}) \right] + \bar{\varphi}, \quad (x, t) \in \omega_h \times \omega_\tau, \quad (39)$$

$$\bar{y}(x, 0) = \bar{u}_0(x), \quad x \in \bar{\omega}_h, \quad \bar{y}_t(x, 0) = \bar{u}_1(x), \quad x \in \omega_h, \quad (40)$$

$$\bar{y}(0, t) = 0, \quad \bar{y}(l, t) = 0, \quad t \in \omega_\tau. \quad (41)$$

Здесь $\bar{u}_0 = \tilde{u}_0 - u_0$, $\bar{u}_1 = \tilde{u}_1 - u_1$, $\bar{\varphi} = \tilde{\varphi} - \varphi$.

К сожалению, несмотря на линейность разностной задачи, к ней неприменима теория Самарского трёхслойных операторно-разностных схем [1]. При использовании метода энергетических неравенств в дальнейшем кроме некоторых известных фактов из теории разностных схем: первая разностная формула Грина, неравенство Коши–Буняковского с ε [1, гл. II, § 3], разностный аналог леммы Гронуолла [16, гл. III, § 1], нам понадобится также следующая

Лемма 2. *Выражение*

$$Q^n = \|\bar{y}_t\|^2 + \frac{1}{2}(a, \bar{y}_x^2 + \check{y}_x^2) + \frac{m}{2}(\|\bar{y}\|^2 + \|\check{y}\|^2) - \frac{h^2}{12}(ap_{(-1)}, \bar{y}_{xt}^2) - \frac{mh^2}{24}(ap_{(-1)}, \bar{y}_x^2 + \check{y}_x^2),$$

где $p_{(-1)} = p_{i-1}$, при выполнении условий

$$h \leq h_0, \quad h_0 = \sqrt{\frac{3k_1}{m}}, \quad \tau \geq \frac{2h}{\sqrt{3k_1}} \quad (42)$$

неотрицательно: $Q^n \geq 0$.

Доказательство. Достаточно показать, что выражение

$$I_1 = \frac{1}{4}(a, \bar{y}_x^2 + \check{y}_x^2) - \frac{h^2}{12}(ap_{(-1)}, \bar{y}_{xt}^2) - \frac{mh^2}{24}(ap_{(-1)}, \bar{y}_x^2 + \check{y}_x^2) \quad (43)$$

неотрицательно. С учётом очевидных неравенств

$$-\frac{h^2}{12}(ap_{(-1)}, \bar{y}_{xt}^2) \geq -\frac{h^2}{6k_1 \tau^2}(a, \bar{y}_x^2 + \check{y}_x^2) \quad \text{и} \quad -\frac{mh^2}{24}(ap_{(-1)}, \bar{y}_x^2 + \check{y}_x^2) \geq -\frac{mh^2}{24k_1}(a, \bar{y}_x^2 + \check{y}_x^2)$$

для выражения (43) имеет место оценка

$$I_1 \geq c_3(a, \bar{y}_x^2 + \check{y}_x^2), \quad c_3 = \frac{1}{4} - \frac{h^2}{6k_1 \tau^2} - \frac{mh^2}{24k_1}.$$

При первом из условий (42) выполнено неравенство

$$\frac{1}{4} - \frac{mh^2}{24k_1} \geq \frac{1}{8}.$$

Следовательно,

$$c_3 \geq -\frac{h^2}{6k_1\tau^2} + \frac{1}{8} \geq 0$$

при выполнении второго из условий (42). Лемма доказана.

Чтобы получить априорную оценку для $\bar{y}$, умножим разностное уравнение (39) скалярно на $2\tau y_{\circ}$ и применим первую разностную формулу Грина. Получим следующее энергетическое соотношение:

$$Q^{n+1} - \frac{h^2}{12}(ap_{\bar{x}}(\bar{y}_t - \bar{y}_{\bar{t}}), \bar{y}_{\bar{x}t} + \bar{y}_{\bar{x}\bar{t}}] = Q^n + \frac{m\tau h^2}{12}(ap_{\bar{x}}(\hat{\bar{y}} + \check{\bar{y}}), \bar{y}_{\circ_{t\bar{x}}}] + 2\tau(\bar{\varphi}, \bar{y}_{\circ_t}). \quad (44)$$

Рассмотрим в (44) слагаемые, отличные от Q^n и Q^{n+1} . Применяя неравенство Коши–Буняковского с ε , легко получаем следующие оценки:

$$-\frac{h^2}{12}(ap_{\bar{x}}(\bar{y}_t - \bar{y}_{\bar{t}}), \bar{y}_{\bar{x}t} + \bar{y}_{\bar{x}\bar{t}}] \geq -ch(\|\bar{y}_t\|^2 + \|\bar{y}_{\bar{t}}\|^2), \quad (45)$$

$$\frac{m\tau h^2}{12}(ap_{\bar{x}}(\hat{\bar{y}} + \check{\bar{y}}), \bar{y}_{\circ_{t\bar{x}}}] \leq ch\left[\frac{m}{2}(\|\hat{\bar{y}}\|^2 + \|\bar{y}\|^2) + \frac{m}{2}(\|\bar{y}\|^2 + \|\check{\bar{y}}\|^2)\right], \quad (46)$$

$$2\tau(\bar{\varphi}, \bar{y}_{\circ_t}) \leq \varepsilon\tau\|\bar{\varphi}\|^2 + \frac{\tau}{2\varepsilon}(\|\bar{y}_t\|^2 + \|\bar{y}_{\bar{t}}\|^2), \quad (47)$$

где $c > 0$ – константа, зависящая от m , ε , $\max_{x \in \bar{\omega}_h} |p_{\bar{x}}(x)|$ и в каждом конкретном случае своя.

Учитывая неравенства (45)–(47) в (44), при выполнении условий (42) приходим к рекуррентному соотношению

$$Q^{n+1} \leq (1 + \tau c)Q^n + \tau c\|\bar{\varphi}\|^2 \leq e^{c\tau}Q^n + \tau c\|\bar{\varphi}\|^2. \quad (48)$$

Итак, имеет место следующая

Теорема 7. Пусть выполнено условие

$$\tau \geq \max\left\{1, \sqrt{\frac{2}{3k_1}}\right\}h.$$

Тогда имеет место оценка

$$Q^{n+1} \leq e^{c\tau n}\left(Q^1 + c\sum_{k=1}^n \tau\|\bar{\varphi}^k\|^2\right), \quad (49)$$

означающая ρ -устойчивость решения разностной схемы (36)–(38) по начальным данным и правой части в сеточных нормах $L_2(\omega_h)$, $W_2^1(\omega_h)$, $C(\bar{\omega}_h)$.

Доказательство теоремы следует из неравенства (48), леммы Гронуолла и вложения [1, с. 107]

$$\|\tilde{y} - y\|_C \leq \frac{\sqrt{l}}{2}\|\tilde{y}_{\bar{x}} - y_{\bar{x}}\|.$$

3.3. Сходимость разностной схемы в сеточной норме $C(\bar{\omega}_h)$. Заменяя в разностных уравнениях (36)–(38) y на $z + u$, где u – решение задачи (33)–(35), получаем для погрешности z задачу

$$z_{tt} = \Lambda z^{(\sigma, \sigma)} - \frac{h^2}{12}\Lambda(pz_{tt}) - m\left[z^{(\sigma, \sigma)} + \frac{h^2}{12}\Lambda(pz^{(\sigma, \sigma)})\right] + \psi, \quad (x, t) \in \omega_h \times \omega_{\tau}, \quad (50)$$

$$z(x, 0) = 0, \quad x \in \bar{\omega}_h, \quad z_t(x, 0) = \overset{\circ}{\psi}, \quad x \in \omega_h, \quad \overset{\circ}{\psi} = O(\tau^2), \quad (51)$$

$$z(0, t) = 0, \quad z(l, t) = 0, \quad t \in \omega_\tau. \quad (52)$$

Задачи (50)–(52) и (39)–(41) идентичны. Поэтому можно применить теорему 7 для оценки погрешности метода. Тогда в соответствии с неравенством (49) получаем оценку

$$\|z\|_C^2 \leq M_1 \left\{ \|\psi\|^\circ{}^2 + \frac{4c_2}{c^2} \|\psi_{\bar{x}}\|^\circ{}^2 + \frac{4m}{c^2} \|\psi\|^\circ{}^2 + cT \max_{t \in \omega_\tau} \|\psi(t)\|^2 \right\},$$

где $M_1 = \text{const} > 0$.

Итак, мы можем сформулировать теорему о сходимости.

Теорема 8. Пусть выполнены условия теоремы 7. Тогда решение разностной схемы (36)–(38) сходится к точному решению дифференциальной задачи (33)–(35) в сеточной норме $C(\overline{\omega}_h)$ и для её решения имеет место оценка точности

$$\|y^n - u^n\|_C \leq M_2(h^4 + \tau^2), \quad n = \overline{0, N_0},$$

где $M_2 = \text{const} > 0$.

4. Квазилинейные уравнения Клейна–Гордона. Недостатком предложенных А.А. Самарским компактных разностных схем для уравнений с переменными коэффициентами является невозможность их обобщения на случай квазилинейных уравнений, так как соответствующий шаблонный функционал должен вычисляться в несуществующей для квазилинейного случая полущелой точке. Тем не менее компактные схемы порядка $4 + 2$, аналогичные схемам для случая постоянных коэффициентов, можно строить и для квазилинейных уравнений Клейна–Гордона

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 \phi(u)}{\partial x^2} - m f_1(u) + f(x, t), \quad m = \text{const} > 0,$$

$$u(x, 0) = u_0(x), \quad \frac{\partial u}{\partial t}(x, 0) = \overline{u}_0(x),$$

$$u(0, t) = \mu_1(t), \quad u(l, t) = \mu_2(t),$$

с условием $\phi'_u = k(u) \geq k_1 > 0$.

Разностная схема $4 + 2$ порядка аппроксимации на стандартном шаблоне имеет вид

$$y_{\bar{t}t} = [\phi(y)]_{\bar{x}x}^{(\sigma, \sigma)} - m \overline{f}_1(y) + \overline{f} - \frac{h^2}{12} y_{\bar{t}t\bar{x}x}, \quad (x, t) \in \omega_h \times \omega_\tau,$$

$$y(x, 0) = u_0(x), \quad x \in \overline{\omega}_h, \quad y_t(x, 0) = u_1(x), \quad x \in \omega_h,$$

$$y(0, t) = \mu_1(t), \quad y(l, t) = \mu_2(t), \quad t \in \omega_\tau,$$

где

$$\overline{v} = v + \frac{h^2}{12} v_{\bar{x}x} = \frac{5}{6} v + \frac{1}{12} (v_{+1} + v_{-1}), \quad 0 < \sigma \leq 1,$$

$$u_1(x) = \overline{u}_0(x) + \frac{\tau}{2} [\phi''(u_0(x)) - m f_1(u_0(x)) + f(x, 0)], \quad x \in \omega_h.$$

Для реализации этой схемы необходимо использовать итерационный метод Ньютона.

5. Вычислительный эксперимент. Ниже приводятся результаты численных расчётов при решении начально-краевой задачи для уравнения вида

$$\frac{\partial^2 u}{\partial t^2} = -\frac{1}{\gamma} \frac{\partial^2 u^{-\gamma}}{\partial x^2} \quad (53)$$

с выбранными параметрами $\gamma = 5$, $l = 2$, $T = 2$. Начальные и краевые условия определяются из точного решения

$$u(x, t) = \left(\frac{t+1}{x+1} \right)^{2/(1+\gamma)}.$$

Уравнение (53) является следствием системы уравнений газовой динамики в переменных Лагранжа с уравнением состояния для политропного газа, в котором $u = \eta$ – удельный объём.

Разностная схема 4+2 порядка аппроксимации для уравнения (53) имеет вид

$$y_{tt} = \sigma(\phi(\hat{y}))_{\bar{x}x} + (1 - 2\sigma)(\phi(y))_{\bar{x}x} + \sigma(\phi(\check{y}))_{\bar{x}x} - \frac{h^2}{12} y_{t\bar{t}\bar{x}x}, \quad (54)$$

где $\sigma = 1$, $\phi(u) = -u^{-\gamma}/\gamma$.

Для нахождения решения разностной схемы (54) применяется итерационный метод Ньютона:

$$\hat{y}^{(k+1)} - 2y + \check{y} = \tau^2 [\phi(\hat{y}^{(k)}) + \phi'_y(\hat{y}^{(k)})(\hat{y}^{(k+1)} - \hat{y}^{(k)})]_{\bar{x}x} - \tau^2 [\phi(y)]_{\bar{x}x} + \tau^2 [\phi(\check{y})]_{\bar{x}x} - \frac{h^2}{12} (\hat{y}^{(k+1)} - 2y + \check{y})_{\bar{x}x},$$

$$\hat{y}^{(0)} = 2y - \check{y}, \quad k = 0, 1, 2, \dots - \text{номер итерации.}$$

На каждом слое этот процесс будет останавливаться, когда при некотором M будет выполнено условие

$$\|\hat{y}^{(M+1)} - \hat{y}^{(M)}\|_C \leq \varepsilon, \quad \varepsilon = 10^{-7}.$$

Порядок сходимости по временной и пространственной переменным в норме $L_\infty = C$ определяется по следующим формулам:

$$p_\infty^h = \log_2 \frac{\|z(2h, \tau)\|_{L_\infty}}{\|z(h, \tau)\|_{L_\infty}}, \quad p_\infty^\tau = \log_2 \frac{\|z(h, 2\tau)\|_{L_\infty}}{\|z(h, \tau)\|_{L_\infty}}. \quad (55)$$

Так как разностное решение сходится к точному решению с четвёртым порядком по пространству и вторым по времени, то для проверки скорости сходимости по временной переменной мы выбираем такие шаги h и τ , чтобы выполнялось неравенство $h^4 \leq \tau^2$. Тогда получается схема $O(\tau^2)$ и мы работаем со вторым правилом Рунге (55).

Аналогично, при рассмотрении порядка по h в расчётах следим, чтобы выполнялись неравенство $h^4 \geq \tau^2$. Тогда можно применить первое правило Рунге (55).

В таблицах приведена скорость сходимости приближённого решения к точному.

Таблица 1. Скорость сходимости по пространственному направлению

h	τ	$\ z\ _{L_\infty}$	$p_{L_\infty}^h$	$\ z\ _{L_2}$	$p_{L_2}^h$	k
$h_0=0.5$	$\tau_0=0.25$	1.29E-02	–	1.05E-02	–	3
$h_0/2^1$	$\tau_0/4^1$	7.47E-04	4.10917	6.20E-04	4.08544	2
$h_0/2^2$	$\tau_0/4^2$	4.49E-05	4.05519	3.82E-05	4.02065	2
$h_0/2^3$	$\tau_0/4^3$	2.79E-06	4.00895	2.38E-06	4.00215	2
$h_0/2^4$	$\tau_0/4^4$	1.75E-07	3.99592	1.49E-07	3.9998	1
$h_0/2^5$	$\tau_0/4^5$	1.16E-08	3.91571	9.94E-09	3.90473	1

Таблица 2. Скорость сходимости по временному направлению

h	τ	$\ z\ _{L_\infty}$	$p_{L_\infty}^\tau$	$\ z\ _{L_2}$	$p_{L_2}^\tau$	k
$h_0=0.001$	$\tau_0=0.25$	1.32E-02	–	1.11E-02	–	3
h_0	$\tau_0/2^1$	3.34E-03	1.97629	2.59E-03	2.09106	3
h_0	$\tau_0/2^2$	7.66E-04	2.12479	6.25E-04	2.05417	2
h_0	$\tau_0/2^3$	1.79E-04	2.10033	1.54E-04	2.02276	2
h_0	$\tau_0/2^4$	4.49E-05	1.99439	3.82E-05	2.00736	2
h_0	$\tau_0/2^5$	1.12E-05	1.99691	9.55E-06	2.00212	2

Отсюда видим, что построенная разностная схема имеет четвёртый порядок точности по пространственной переменной и второй по временной.

Кроме того, на рис. 1, 2 по цветам и фигурам хорошо видна сходимость приближённого решения к точному решению при измельчении шагов сетки h и τ . Для удобства визуального наблюдения приведённые результаты получены в области $0 \leq x \leq 10$, $0 \leq t \leq 10$.

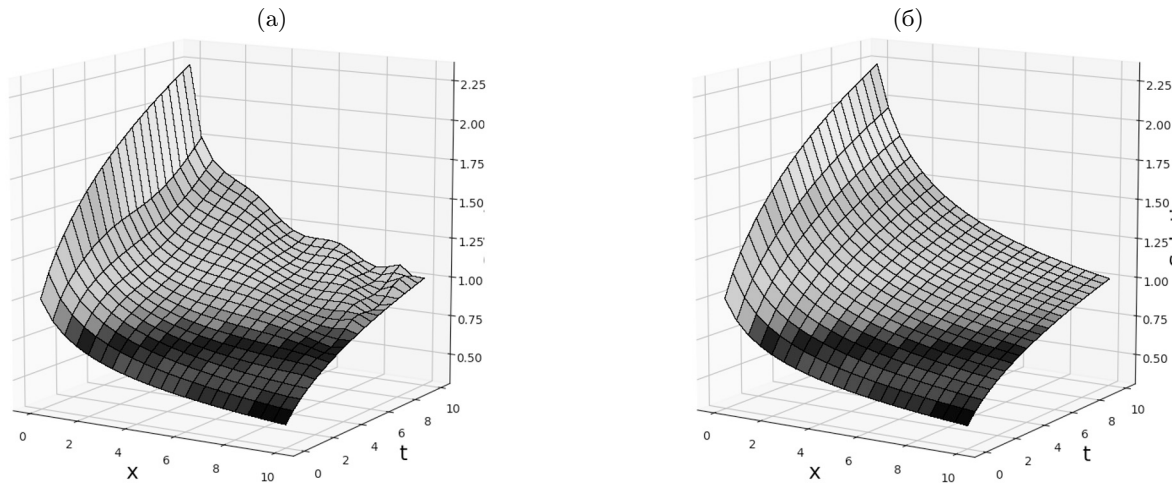


Рис. 1. Численное (а) и точное (б) решения при $h = 0.5$, $\tau = 0.5$.

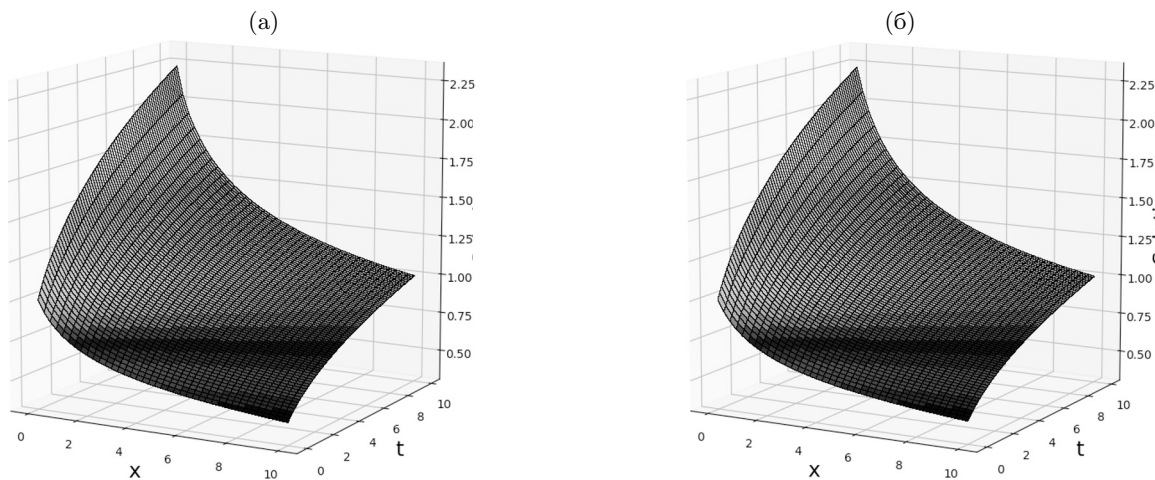


Рис. 2. Численное (а) и точное (б) решения при $h = 0.25$, $\tau = 0.125$.

Проведённый вычислительный эксперимент подтверждает наши теоретические выводы.

СПИСОК ЛИТЕРАТУРЫ

1. Самарский А.А. Теория разностных схем. М., 1989.
2. Самарский А.А., Вабичевич П.Н., Матус П.П. Разностные схемы с операторными множителями. Минск, 1998.
3. Lemeshevsky S., Matus P., Poliakov D. Exact Finite-Difference Schemes. De Gruyter, 2016.
4. Вабичевич П.Н. Двухслойные схемы повышенного порядка аппроксимации для нестационарных задач математической физики // Журн. вычислит. математики и мат. физики. 2010. Т. 50. № 1. С. 118–130.
5. Паасонен В.И. Компактные схемы для систем уравнений второго порядка с конвективными членами // Числ. методы механики сплошной среды. 1998. Т. 3. № 1. С. 55–66.
6. Самарский А.А. Схемы повышенного порядка точности для многомерного уравнения теплопроводности // Журн. вычислит. математики и мат. физики. 1963. Т. 3. № 5. С. 812–840.

7. Толстых А.И. Компактные разностные схемы и их применение в задачах аэрогидродинамики. М., 1990.
8. Паасонен В.И. Обобщение методов повышенной точности для нелинейных уравнений 2-го порядка в ортогональных системах координат // Числ. методы механики сплошной среды. 1977. Т. 8. № 2. С. 94–99.
9. Паасонен В.И. Диссипативные асимметричные компактные схемы для уравнения колебаний // Вычислит. технологии. Спец. выпуск. 2001. Т. 6. № 2. С. 475–479.
10. Caudreу P.J., Eilbeck J.C., Gibbon J.D. The sine-Gordon equation as a model classical field theory // *Nuovo Cimento B. Ser. 11*. 1975. V. 25. № 2. P. 497–512.
11. Luo Y., Li X., Guo C. Fourth-order compact and energy conservative scheme for solving nonlinear Klein–Gordon equation // *Numer. Methods Partial Differ. Equat.* 2017. V. 33. № 4. P. 1283–1304.
12. Матус П.П., Хоанг Тхи Киеу Ань. Компактные разностные схемы для уравнения Клейна–Гордона // Докл. НАН Беларуси. 2020. Т. 64. № 5. С. 526–533.
13. Матус П.П., Хоанг Тхи Киеу Ань. Компактные разностные схемы для уравнения Клейна–Гордона с переменными коэффициентами // Докл. НАН Беларуси. 2021. Т. 65. № 1. С. 25–32.
14. Москальков М.Н. О точности разностных схем, аппроксимирующих волновое уравнение с кусочно-гладкими решениями // Журн. вычислит. математики и мат. физики. 1974. Т. 14. № 2. С. 390–401.
15. Матус П.П., Панайотова Й.Н. Коэффициентная устойчивость трехслойных операторно-разностных схем // Журн. вычислит. математики и мат. физики. 2001. Т. 41. № 5. С. 722–731.
16. Самарский А.А., Гудин А.В. Устойчивость разностных схем. М., 1973.

Институт математики НАН Беларуси,
г. Минск,
Католический университет им. Иоанна-Павла II,
г. Люблин, Польша,
Белорусский государственный университет,
г. Минск

Поступила в редакцию 31.12.2020 г.
После доработки 31.12.2020 г.
Принята к публикации 27.04.2021 г.

ЧИСЛЕННЫЕ МЕТОДЫ

УДК 519.622.2

АЛГОРИТМ АДАПТИВНОЙ ИНТЕРПОЛЯЦИИ НА РАЗРЕЖЕННЫХ СЕТКАХ ДЛЯ ЧИСЛЕННОГО ИНТЕГРИРОВАНИЯ СИСТЕМ ОБЫКНОВЕННЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ С ИНТЕРВАЛЬНЫМИ НЕОПРЕДЕЛЁННОСТЯМИ

© 2021 г. А. Ю. Морозов, Д. Л. Ревизников

Рассматриваются теоретические аспекты обобщения алгоритма адаптивной интерполяции на случай большого количества интервальных неопределённостей с использованием разреженных сеток. Суть классического алгоритма адаптивной интерполяции заключается в построении адаптивного иерархического разбиения области неопределённости на подобласти, каждой из которых соответствует полиномиальная функция некоторой степени, интерполирующая зависимость решения задачи от точечных значений интервальных параметров с заданной точностью. Основным недостатком алгоритма является его экспоненциальная зависимость от количества интервальных параметров. Уже при наличии пяти-шести интервальных неопределённостей алгоритм становится практически не применим. В связи с этим предлагается вместо интерполяции на регулярной сетке использовать интерполяцию на адаптивных разреженных сетках, что позволяет в ряде случаев существенно расширить область применения алгоритма. Получена оценка глобальной погрешности алгоритма на разреженных сетках. Показана линейная зависимость глобальной погрешности алгоритма от уровня сетки. На представительном ряде задач продемонстрирована эффективность предложенного подхода.

DOI: 10.31857/S0374064121070104

Введение. Для многих прикладных и исследовательских задач характерно отсутствие точной информации о параметрах изучаемого явления или процесса [1]. По этой причине подобные задачи обычно формулируются в виде систем обыкновенных дифференциальных уравнений (ОДУ) с интервальными параметрами и интервальными начальными условиями. Решением таких систем является интервальная оценка значений фазовых переменных в зависимости от времени. Существует несколько подходов к решению данного класса задач: одни основаны на использовании классической интервальной арифметики [2–4], другие – на аппроксимации множества решений с помощью геометрических примитивов [5–7], а в основе третьих лежит определение явной зависимости решения задачи от значений интервальных параметров [8–11]. В общем случае наблюдается некоторый баланс между вычислительной сложностью методов и качеством получаемых интервальных оценок. Методы, основанные на классической интервальной арифметике, зачастую могут давать чрезмерно завышенные интервальные оценки решения, но при этом обладают небольшой вычислительной сложностью. С другой стороны, методы, восстанавливающие явную зависимость решения от интервальных параметров, имеют экспоненциальную сложность, но при этом получают близкие к истинным интервальные оценки. С практической точки зрения интерес представляют методы, которые не завышают оценки. В связи с этим актуальной является задача уменьшения вычислительной сложности соответствующих методов.

Ранее авторами был разработан и апробирован алгоритм адаптивной интерполяции для моделирования динамических систем с интервальными параметрами [11, с. 16–19]. Его суть заключается в адаптивном иерархическом разбиении области неопределённости на подобласти, для каждой из которых выполняется построение полиномиальной функции [12], интерполирующей зависимость решения задачи от точечных значений интервальных параметров. Все подобласти представляют собой прямоугольные параллелепипеды размерности, равной количеству интервальных параметров. В каждой подобласти строится регулярная сетка. Каждому

узлу в сетке соответствует решение исходной неинтервальной задачи при значениях параметров, определяемых положением узла в пространстве. Количество узлов в одной сетке оценивается как $(p + 1)^m$, где p – степень интерполяционного полинома по каждому измерению, а m – количество интервальных параметров. При увеличении m время работы алгоритма адаптивной интерполяции становится чрезвычайно большим, что ограничивает область его применения.

В работах [13, 14] дано теоретическое обоснование алгоритма и выполнено его тестирование как на модельных задачах, так и на прикладных задачах химической кинетики. В [15] рассмотрены различные аспекты разработки параллельной версии алгоритма под архитектуру графических процессоров с использованием технологии CUDA [16]. В [17] продемонстрирована возможность определения бифуркаций и хаоса в динамических системах с помощью анализа получающегося адаптивного разбиения области неопределённости в процессе работы алгоритма. В работе [18] выполнен анализ вычислительных затрат алгоритма и получены рекомендации по выбору оптимальной степени интерполяционного полинома p . В [19] показана возможность использования разреженных сеток в алгоритме адаптивной интерполяции для моделирования динамических систем, содержащих большое количество интервальных параметров, и продемонстрирована эффективность данного подхода. Настоящая работа посвящена теоретическим аспектам применения разреженных сеток для численного интегрирования систем обыкновенных дифференциальных уравнений с интервальными начальными условиями и параметрами.

Разреженные сетки основаны на иерархическом базисе и предназначены для решения задач представления, интегрирования и интерполяции многомерных функций [20]. За счёт незначительного ухудшения асимптотических свойств погрешности они позволяют существенно снизить “проклятие размерности” (экспоненциальный рост числа узлов при увеличении размерности). Классические разреженные сетки являются результатом оптимизации вычислительных затрат для аппроксимации функций с ограниченными смешанными производными [21, с. 3]. В [22] даётся введение в разреженные сетки и технику работы с ними, а также приводится программный код на языке программирования Python с их реализацией. Работа [23] посвящена алгоритмам распараллеливания, а работа [24] – различным приложениям. В [25] рассматриваются вопросы повышения эффективности разреженных сеток.

Существует адаптивный вариант разреженных сеток, который позволяет динамически учитывать особенности рассматриваемой функции и подстраиваться под них: в тех областях, в которых имеются резкие переходы, происходит уплотнение сетки, а в областях с плавными переходами – её разрежение. Кроме того, сетки учитывают разное влияние подмножеств переменных на конечный результат. Например, если исходная зависимость является линейной комбинацией функций от определённых подгрупп переменных, то построение сетки будет происходить только на подмножествах, соответствующих этим подгруппам.

Решение системы ОДУ с интервальными начальными условиями и параметрами в каждый момент времени можно рассматривать как некоторую многомерную вектор-функцию, которая в общем случае является “чёрным ящиком”. В начальный момент времени эта функция тривиальная, её компоненты являются либо константами, либо линейными функциями. Основная цель – построить для каждого момента времени соответствующий интерполянт. От шага к шагу интерполяционная сетка может изменяться, так как происходит её адаптация. Для того чтобы вычислить значение исходной вектор-функции в произвольной точке на некотором шаге, нет необходимости интегрировать систему ОДУ из начального состояния; вместо этого с помощью интерполяции вычисляется решение на предыдущем шаге, и затем оно переносится на текущий шаг. При этом важно понимать как в результате поведёт себя глобальная погрешность. Этот вопрос рассматривается в данной работе.

1. Постановка задачи. В общем виде рассматривается система ОДУ, состоящая из n_e уравнений с n_p параметрами:

$$\frac{du_i(t)}{dt} = g_i(u_1(t), u_2(t), \dots, u_{n_e}(t), v_1, v_2, \dots, v_{n_p}, t), \quad i = \overline{1, n_e}, \quad t \in [t_0, t_N],$$

$$u_i(t_0) \in [\underline{u}_i^0, \overline{u}_i^0], \quad i = \overline{1, m_e},$$

$$\begin{aligned} u_i(t_0) &= u_i^0, \quad i = \overline{m_e + 1, n_e}, \\ v_i &\in [\underline{v}_i, \overline{v}_i], \quad i = \overline{1, m_p}, \end{aligned} \quad (1)$$

где m_e – количество интервальных начальных условий, m_p – количество интервальных параметров. Вектор-функция $\mathbf{g} = (g_1, g_2, \dots, g_{n_e})^T$ удовлетворяет всем условиям, обеспечивающим существование и единственность решения при всех $u_i(t_0) \in [\underline{u}_i^0, \overline{u}_i^0]$, $i = \overline{1, m_e}$, и $v_i \in [\underline{v}_i, \overline{v}_i]$, $i = \overline{1, m_p}$.

Сгруппируем вместе интервальные неопределённости и выполним преобразование системы (1) к автономной системе ОДУ без параметров путём добавления в систему фиктивных уравнений:

$$\begin{aligned} \frac{dy_i(t)}{dt} &= f_i(y_1(t), y_2(t), \dots, y_n(t)), \quad i = \overline{1, n}, \\ y_i(t_0) &\in [\underline{y}_i^0, \overline{y}_i^0], \quad i = \overline{1, m}, \\ y_i(t_0) &= y_i^0, \quad i = \overline{m + 1, n}, \quad t \in [t_0, t_N], \end{aligned} \quad (2)$$

где $n = n_e + n_p + 1$ – количество уравнений, $m = m_e + m_p$ – количество интервальных начальных условий,

$$\begin{aligned} y_i(t) &= u_i(t), \quad \underline{y}_i^0 = \underline{u}_i^0, \quad \overline{y}_i^0 = \overline{u}_i^0; \quad f_i = g_i, \quad i = \overline{1, m_e}, \\ y_{m_e+i}(t) &= v_i, \quad \underline{y}_{m_e+i}^0 = \underline{v}_i^0, \quad \overline{y}_{m_e+i}^0 = \overline{v}_i^0; \quad f_{m_e+i} \equiv 0, \quad i = \overline{1, m_p}, \\ y_{m+i}(t) &= u_{m_e+i}(t), \quad y_{m+i}^0 = u_{m_e+i}^0; \quad f_{m+i} = g_{m_e+i}, \quad i = \overline{1, n_e - m_e}, \\ y_{n_e+m_p+i}(t) &= v_{m_p+i}, \quad y_{n_e+m_p+i}^0 = v_{m_p+i}^0; \quad f_{n_e+m_p+i} \equiv 0, \quad i = \overline{1, n_p - m_p}, \\ y_n(t) &= t, \quad y_n^0 = t_0; \quad f_n \equiv 1. \end{aligned}$$

Предположим, что решение системы (2) существует при всех указанных начальных данных на всём рассматриваемом отрезке и является функцией от независимой переменной t и точечных значений интервальных начальных условий:

$$\mathbf{y}(y_1^0, y_2^0, \dots, y_m^0, t), \quad y_i^0 \in [\underline{y}_i^0, \overline{y}_i^0], \quad i = \overline{1, m}.$$

Рассматривается $N + 1$ точка из интервала интегрирования $[t_0, t_N]$:

$$t_0 < t_1 < t_2 < \dots < t_k < \dots < t_N.$$

Требуется для каждого значения t_k , $k = \overline{0, N}$, построить вектор-функцию $\mathbf{P}^k(y_1^0, y_2^0, \dots, y_m^0) = (P_1^k(y_1^0, \dots, y_m^0), \dots, P_n^k(y_1^0, \dots, y_m^0))$, где $y_i^0 \in [\underline{y}_i^0, \overline{y}_i^0]$, $i = \overline{1, m}$, интерполирующую значение $\mathbf{y}^k(y_1^0, y_2^0, \dots, y_m^0) = \mathbf{y}(y_1^0, y_2^0, \dots, y_m^0, t_k)$.

2. Алгоритм адаптивной интерполяции на разреженных сетках. Выполним последовательное построение решения системы (2). В начальный момент времени t_0 компоненты $P_j^0(y_1^0, \dots, y_m^0)$, $j = \overline{1, n}$, вектор-функции $\mathbf{P}^0(y_1^0, y_2^0, \dots, y_m^0)$ будут являться линейными функциями

$$P_i^0(y_1^0, y_2^0, \dots, y_m^0) = y_i^0, \quad i = \overline{1, m},$$

и константами

$$P_i^0(y_1^0, y_2^0, \dots, y_m^0) = y_i^0, \quad i = \overline{m + 1, n}.$$

Предположим, что на k -м шаге известен интерполяционный полином $\mathbf{P}^k(y_1^0, y_2^0, \dots, y_m^0)$. В процессе построения $\mathbf{P}^{k+1}$ необходимо многократно вычислять исходную вектор-функцию

$\mathbf{y}^{k+1}$ в определённых точках $(\hat{y}_1^0, \hat{y}_2^0, \dots, \hat{y}_m^0)$. Для этого с помощью какого-либо численного метода интегрирования решается следующая система ОДУ:

$$\frac{d\hat{y}_i(t)}{dt} = f_i(\hat{y}_1(t), \hat{y}_2(t), \dots, \hat{y}_n(t)), \quad \hat{y}_i(t_k) = P_i^k(\hat{y}_1^0, \hat{y}_2^0, \dots, \hat{y}_m^0), \quad i = \overline{1, n}, \quad t \in [t_k, t_{k+1}], \quad (3)$$

и в качестве $\mathbf{y}^{k+1}$ берётся $\hat{\mathbf{y}}^{k+1}$.

На каждом шаге возникают две локальные погрешности: первая – от численного метода интегрирования системы ОДУ, а вторая – от интерполяции при вычислении $\mathbf{P}^k(\hat{y}_1^0, \hat{y}_2^0, \dots, \hat{y}_m^0)$. Будем предполагать, что первая несущественно мала по сравнению со второй.

Согласно постановке задачи интерполяции [12] интерполяционный полином обязательно должен проходить через все узловые точки. Следовательно, погрешность интерполяции в точках, соответствующих узлам, тождественно равна нулю. Поэтому при построении интерполяционной сетки на $(k+1)$ -м шаге погрешность от интерполяции будет возникать только в узлах, которых не было в сетке на k -м шаге.

Для того чтобы оценить глобальную погрешность необходимо оценить количество шагов, в которых происходило уплотнение сетки, а это, в свою очередь, зависит от свойств используемой интерполяции.

Далее приводится описание классических разреженных сеток согласно работам [21, 22].

В основе базиса лежит *функция шляпы*

$$\varphi(x) = \begin{cases} 1 - |x|, & \text{если } x \in [-1, 1], \\ 0 & \text{в противном случае.} \end{cases} \quad (4)$$

Сначала рассматривается построение интерполяционного полинома для гладкой функции одной переменной $f(x)$, $x \in [0, 1]$, $f(0) = f(1) = 0$.

Сетка уровня l определяется следующим образом:

$$G_l = \{x_{l,i} : x_{l,i} = i \cdot h_l, \quad 1 \leq i \leq 2^l - 1, \quad h_l = 2^{-l}\},$$

и её i -му узлу ставится в соответствие базисная функция $\varphi_{l,i}$, определяемая через функцию шляпы (4) равенством

$$\varphi_{l,i}(x) = \varphi\left(\frac{x - i \cdot h_l}{h_l}\right) = \varphi(2^l \cdot x - i).$$

Интерполяционный полином записывается в виде

$$P(x) = \sum_{i=1}^{2^l-1} a_{l,i} \varphi_{l,i}(x), \quad (5)$$

где $a_{l,i} = f(x_{l,i})$. Здесь мы имеем классическую кусочно-линейную интерполяцию.

Выполняется переход к иерархическому базису. Для этого базисные функции с чётными индексами уровня k выражаются через базисные функции уровня $k-1$:

$$\varphi_{k,2i}(x) = \varphi_{k-1,i}(x) - (\varphi_{k,2i-1}(x) + \varphi_{k,2i+1}(x))/2, \quad 1 \leq i \leq 2^{k-1} - 1. \quad (6)$$

Полученное выражение (6) подставляется в функцию (5):

$$P(x) = \sum_{k=1}^l \sum_{\substack{i=1 \\ i - \text{нечётное}}}^{2^k-1} a_{k,i} \varphi_{k,i}(x), \quad (7)$$

где $a_{k,i} = f(x_{k,i}) - (f(x_{k,i-1}) + f(x_{k,i+1}))/2$.

Коэффициенты $a_{k,i}$ характеризуют вторую производную функции $f(x)$ и при возрастании k к бесконечности стремятся к нулю как $O(2^{-2k})$. В отличие от (5) слагаемые, входящие в (7), не равнозначны: чем больше k , тем меньше вклад соответствующих слагаемых в результат.

D -мерный базис конструируется при помощи произведения одномерных базисов:

$$\varphi_{\mathbf{l},\mathbf{i}}(\mathbf{x}) = \prod_{j=1}^d \varphi_{l_j,i_j}(x_j), \quad 1 \leq i_j \leq 2^{l_j} - 1, \quad i_j - \text{нечётное},$$

где $\mathbf{l} = (l_1, l_2, \dots, l_d)$, $\mathbf{i} = (i_1, i_2, \dots, i_d)$ – мультиуровень и мультииндекс базисной функции, $\mathbf{x} = (x_1, x_2, \dots, x_d)$ – вектор переменных.

В зависимости от ограничений на мультиуровень $\mathbf{l}$ базисных функций, использующихся в интерполяционном полиноме, получается или полная сетка, или разреженная сетка. Для полной сетки уровня n_g имеем ограничение $\|\mathbf{l}\|_\infty \leq n_g$, где $\|\mathbf{l}\|_\infty = \max_{j=\overline{1,d}} l_j$, а для разреженной сетки уровня n_g – ограничение

$$\|\mathbf{l}\|_1 \leq n_g + d - 1,$$

где $\|\mathbf{l}\|_1 = \sum_{j=1}^d l_j$. Сравним асимптотические характеристики сеток. Обозначим $p = 2^{n_g} - 1$. Для полной сетки количество узлов оценивается как $O(p^d)$, а погрешность – $O(h_{n_g}^2)$; для разреженной сетки количество узлов – $O(p(\log_2 p)^{d-1})$, погрешность – $O(h_{n_g}^2 (\log_2 p)^{d-1})$ [21].

Интерполяционный полином:

$$P(\mathbf{x}) = \sum_{\mathbf{l},\mathbf{i}} a_{\mathbf{l},\mathbf{i}} \varphi_{\mathbf{l},\mathbf{i}}(\mathbf{x}), \quad \|\mathbf{l}\|_1 \leq n_g + d - 1,$$

$$1 \leq i_j \leq 2^{l_j} - 1, \quad i_j - \text{нечётное}, \quad j = \overline{1,d}, \quad (8)$$

$$a_{\mathbf{l},\mathbf{i}} = \sum_{\delta_1, \dots, \delta_d} (-1/2)^{\sum_{j=1}^d |\delta_j|} \times$$

$$\times f(x_{l_1, i_1 + \delta_1}, x_{l_2, i_2 + \delta_2}, \dots, x_{l_d, i_d + \delta_d}),$$

$$-1 \leq \delta_j \leq 1, \quad j = \overline{1,d}, \quad (9)$$

где $f(x_1, x_2, \dots, x_d)$ – интерполируемая гладкая функция, определённая на единичном d -мерном кубе и равная нулю на его границе.

Для учёта граничных значений в базис добавляются базисные функции нулевого уровня. Если $l_j = 0$, то в (8) $i_j = 0, 1$, а в (9) $\delta_j = 0$. Учёт граничных значений эквивалентен построению разреженных сеток на всех гранях области определения функции меньшей размерности. В частности, для трёхмерного куба имеем 6 граней размерности 2, 12 граней (рёбер) размерности 1 и 8 граней (вершин) размерности 0. В общем случае для d -мерного куба количество граней меньшей размерности равняется $3^d - 1$.

На рис. 1 снизу продемонстрирован одномерный иерархический базис, а сверху показан пример интерполяции функции

$$f(x) = 0.5 + 5x - 4x^2$$

на интервале $[0, 1]$.

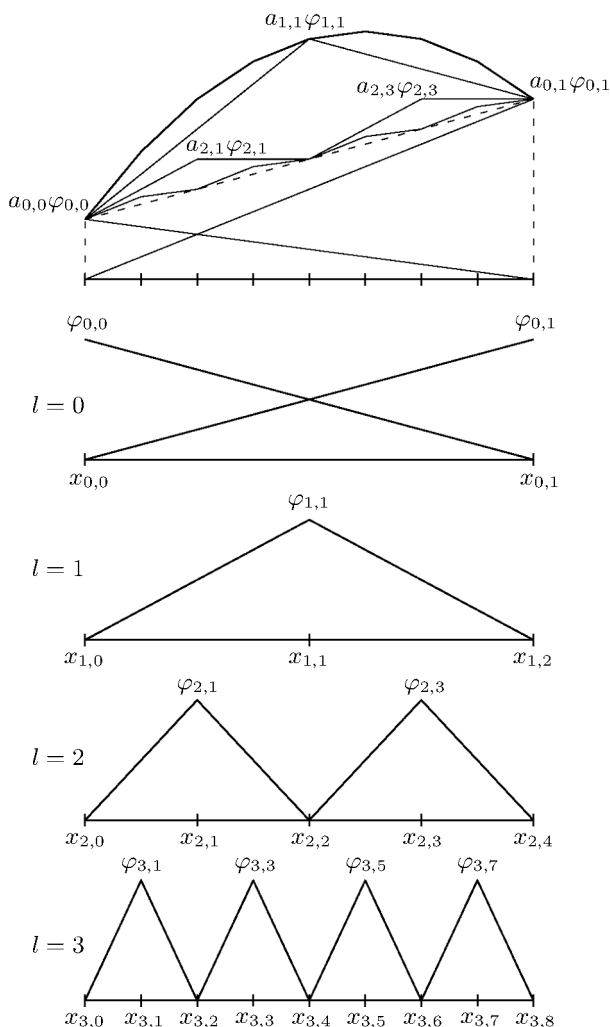


Рис. 1. Иерархический базис.

На рис. 2, а показаны узлы, соответствующие двумерным базисным функциям разного уровня, а на рис. 2, б – все узлы разреженной сетки одновременно.

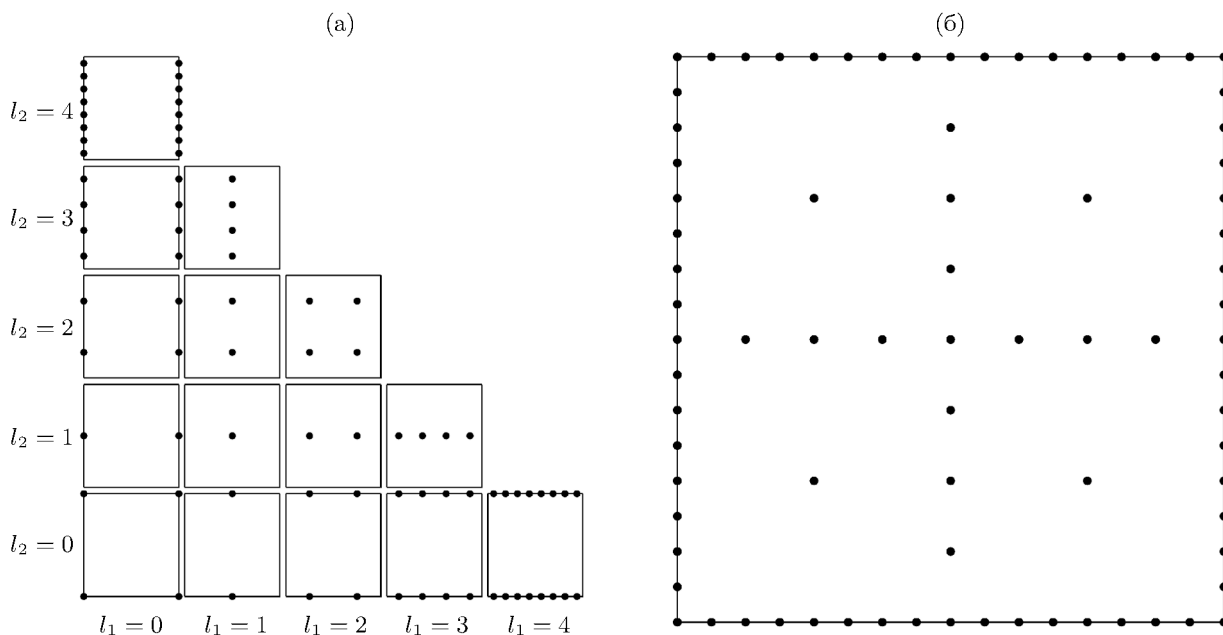


Рис. 2. Разреженная сетка третьего уровня: (а) узлы, соответствующие базисным функциям одного уровня по отдельности; (б) все узлы сетки одновременно.

Классические разреженные сетки являются результатом оптимизации вычислительных затрат для аппроксимации функций с ограниченными смешанными производными. Определим соответствующее пространство Соболева [21, 22]:

$$H_{\text{mix}}(\Omega) = \left\{ f : \Omega \rightarrow \mathbb{R} \mid \frac{\partial^{\|\alpha\|_1} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_d^{\alpha_d}} \in L_2(\Omega), \quad \|\alpha\|_\infty \leq 2, \quad f|_{\partial\Omega} = 0 \right\}.$$

Согласно работе [21, с. 3], если $f \in H_{\text{mix}}(\Omega)$, то

$$\|f(\mathbf{x}) - P(\mathbf{x})\|_2 = O(h_{n_g}^2 n_g^{d-1}).$$

Для построения решения исходной задачи область неопределённости с помощью линейных преобразований трансформируется в единичный m -мерный куб. Соотношения (8) и (9) записываются следующим образом:

$$\begin{aligned} \mathbf{P}^k(y_1^0, y_2^0, \dots, y_m^0) &= \sum_{\mathbf{l}, \mathbf{i}} \mathbf{a}_{\mathbf{l}, \mathbf{i}}^k \varphi_{\mathbf{l}, \mathbf{i}}(y_1^0, y_2^0, \dots, y_m^0), \\ |\mathbf{l}|_1 &\leq n_g + m - 1, \quad 1 \leq i_j \leq 2^{l_j} - 1, \quad i_j - \text{нечётное}, \quad j = \overline{1, m}, \\ \mathbf{a}_{\mathbf{l}, \mathbf{i}}^k &= \sum_{\delta_1, \dots, \delta_m} \left(-\frac{1}{2} \right)^{\sum_{j=1}^m |\delta_j|} \hat{\mathbf{y}}^k(y_{1, l_1, i_1 + \delta_1}^0, y_{2, l_2, i_2 + \delta_2}^0, \dots, y_{m, l_m, i_m + \delta_m}^0), \\ -1 &\leq \delta_j \leq 1, \quad j = \overline{1, m}. \end{aligned} \quad (10)$$

Предположим, что на k -м шаге решения системы (2) фиксируется локальная погрешность eps:

$$\max_{\substack{\hat{y}_i^0 \in [\underline{y}_i^0, \overline{y}_i^0] \\ i = \overline{1, m} \\ j = \overline{1, n}}} |P_j^k(\hat{y}_1^0, \hat{y}_2^0, \dots, \hat{y}_m^0) - \hat{y}_j^k(\hat{y}_1^0, \hat{y}_2^0, \dots, \hat{y}_m^0)| \leq \text{eps}, \quad (11)$$

На рис. 3, а показано множество решений системы (13) в различные моменты времени, а на рис. 3, б – узлы сетки. Точки на двух рисунках соответствуют друг другу. На рис. 3, в и 3, г представлены зависимости уровня сетки и погрешности от времени.

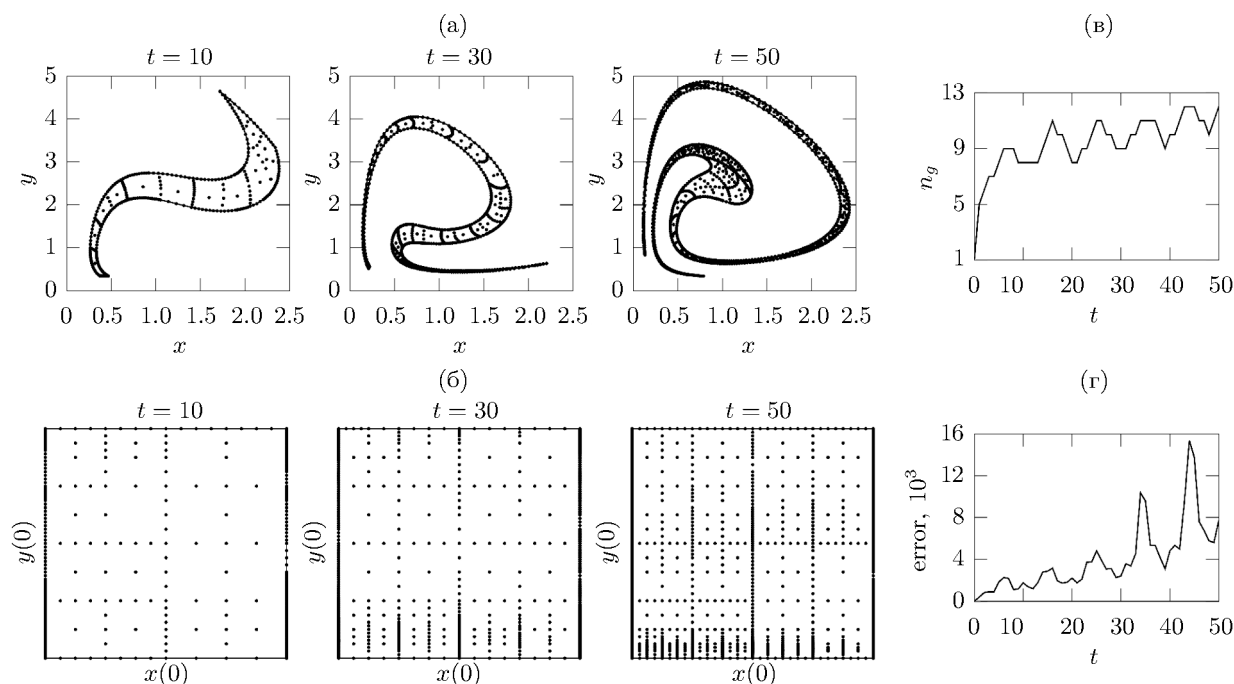


Рис. 3. Результаты решения системы (13): (а) – множество решений в различные моменты времени; (б) – узлы адаптивной разреженной сетки; (в) – зависимость уровня сетки от времени; (г) – зависимость глобальной погрешности от времени.

Здесь наблюдается корреляция между графиками, увеличение уровня приводит к увеличению погрешности, что соответствует ранее сформулированному утверждению. Для оценки погрешности в начальный момент времени случайным образом генерируется тестовое множество точек из области неопределённости параметров системы и строятся соответствующие решения, с которыми выполняется сравнение.

Далее рассмотрим автономные системы ОДУ, для которых решение может быть найдено аналитически. Первая система [11]:

$$\begin{aligned} x' &= a \left(\frac{\sqrt{3}}{2}x + \frac{1}{2}y \right), \quad y' = a \left(-\frac{1}{2}x + \frac{\sqrt{3}}{2}y \right), \quad t \in [0, 4], \\ x(0) &= x_0 \in [-1, 1], \quad y(0) = y_0 \in [-1, 1], \quad a \in [-1, 1]. \end{aligned} \quad (14)$$

Общее решение системы в (14) имеет следующее аналитическое представление:

$$\begin{aligned} x(x_0, y_0, a, t) &= \exp \left(\frac{\sqrt{3}}{2}at \right) \left(y_0 \sin \left(\frac{at}{2} \right) + x_0 \cos \left(\frac{at}{2} \right) \right), \\ y(x_0, y_0, a, t) &= \exp \left(\frac{\sqrt{3}}{2}at \right) \left(y_0 \cos \left(\frac{at}{2} \right) - x_0 \sin \left(\frac{at}{2} \right) \right). \end{aligned} \quad (15)$$

Проанализируем решение (15). Оно линейно зависит от начальных условий x_0 и y_0 и нелинейно – от параметра a . Старшие производные отличны от нуля только для параметра a , и, соответственно, коэффициенты (10) ненулевые только при $l_1 = 0$ и $l_2 = 0$. Следовательно, уплотнение сетки будет происходить на четырёх подмножествах:

$$\{-1\} \times \{-1\} \times a, \quad \{-1\} \times \{1\} \times a, \quad \{1\} \times \{-1\} \times a, \quad \{1\} \times \{1\} \times a.$$

На рис. 4, а и 4, б показаны множество решений системы (14) и адаптивная разреженная сетка в различные моменты времени. Уплотнение сетки происходит только на рёбрах, которые соответствуют параметру a . На рис. 4, в и 4, г представлены зависимости уровня сетки и погрешности от t . Видим, что аналогично предыдущей задаче увеличение уровня приводит к увеличению погрешности.

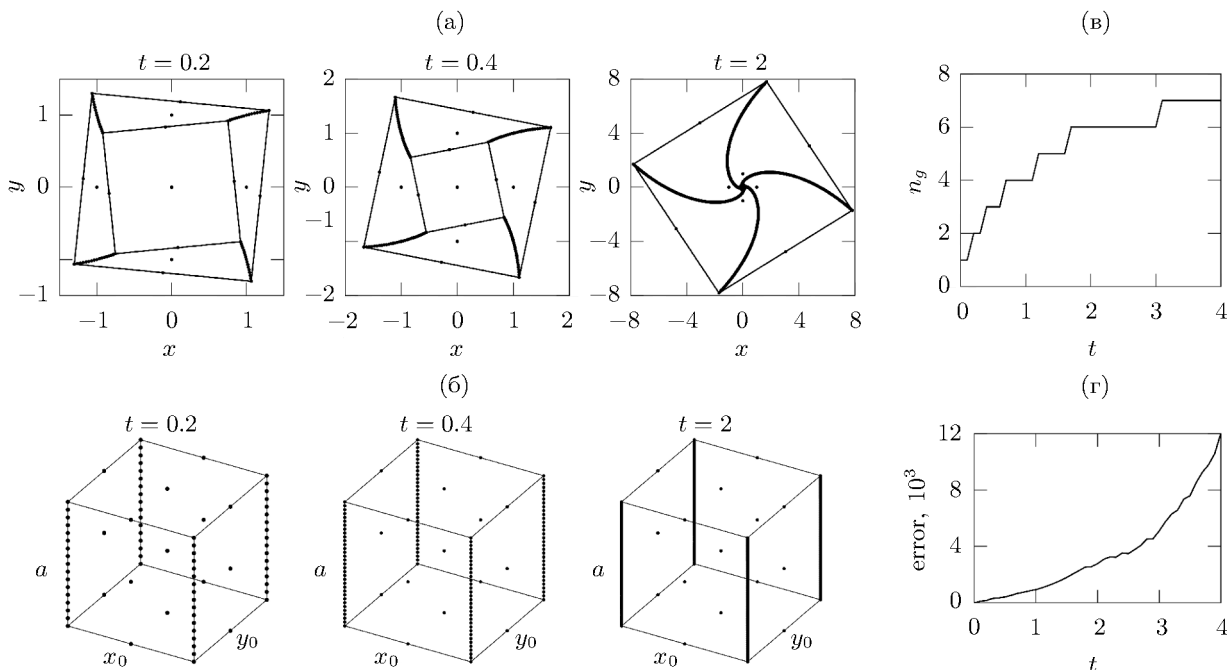


Рис. 4. Результаты решения системы (14): (а) – множество решений в различные моменты времени; (б) – узлы адаптивной разреженной сетки; (в) – зависимость уровня сетки от времени; (г) – зависимость глобальной погрешности от времени.

Введём в систему (14) ещё один интервальный параметр и получим вторую систему:

$$\begin{aligned} x' &= a(bx + \sqrt{1-b^2}y), \quad y' = a(-\sqrt{1-b^2}x + by), \quad t \in [0, 3], \\ x(0) &= x_0 \in [-1, 1], \quad y(0) = y_0 \in [-1, 1], \quad a \in [0, 1], \quad b \in [0, 1]. \end{aligned} \quad (16)$$

Общее решение системы в (16) аналитически записывается следующим образом:

$$\begin{aligned} x(x_0, y_0, a, b, t) &= \exp(abt)(y_0 \sin(a\sqrt{1-b^2}t) + x_0 \cos(a\sqrt{1-b^2}t)), \\ y(x_0, y_0, a, b, t) &= \exp(abt)(y_0 \cos(a\sqrt{1-b^2}t) - x_0 \sin(a\sqrt{1-b^2}t)). \end{aligned}$$

В отличие от предыдущего примера здесь старшие смешанные производные не равны нулю сразу по двум параметрам: a и b . Следовательно, уплотнение сетки будет происходить только на подмножествах, соответствующих этим параметрам:

$$\{-1\} \times \{-1\} \times a \times b, \quad \{-1\} \times \{1\} \times a \times b, \quad \{1\} \times \{-1\} \times a \times b, \quad \{1\} \times \{1\} \times a \times b.$$

Построение решения выполняется при $\epsilon = 5 \cdot 10^{-3}$. На рис. 5, а показано множество решений системы (16) в различные моменты времени. Решение системы в каждый момент времени принадлежит пространству $\mathbb{R}^4$, поэтому изображение на фазовой плоскости является лишь проекцией. На рис. 5, б показана адаптивная разреженная сетка, получающаяся в процессе интегрирования системы. Область неопределённости является четырёхмерным прямоугольным параллелепипедом, поэтому, по аналогии с рис. 5, а, здесь показана проекция.

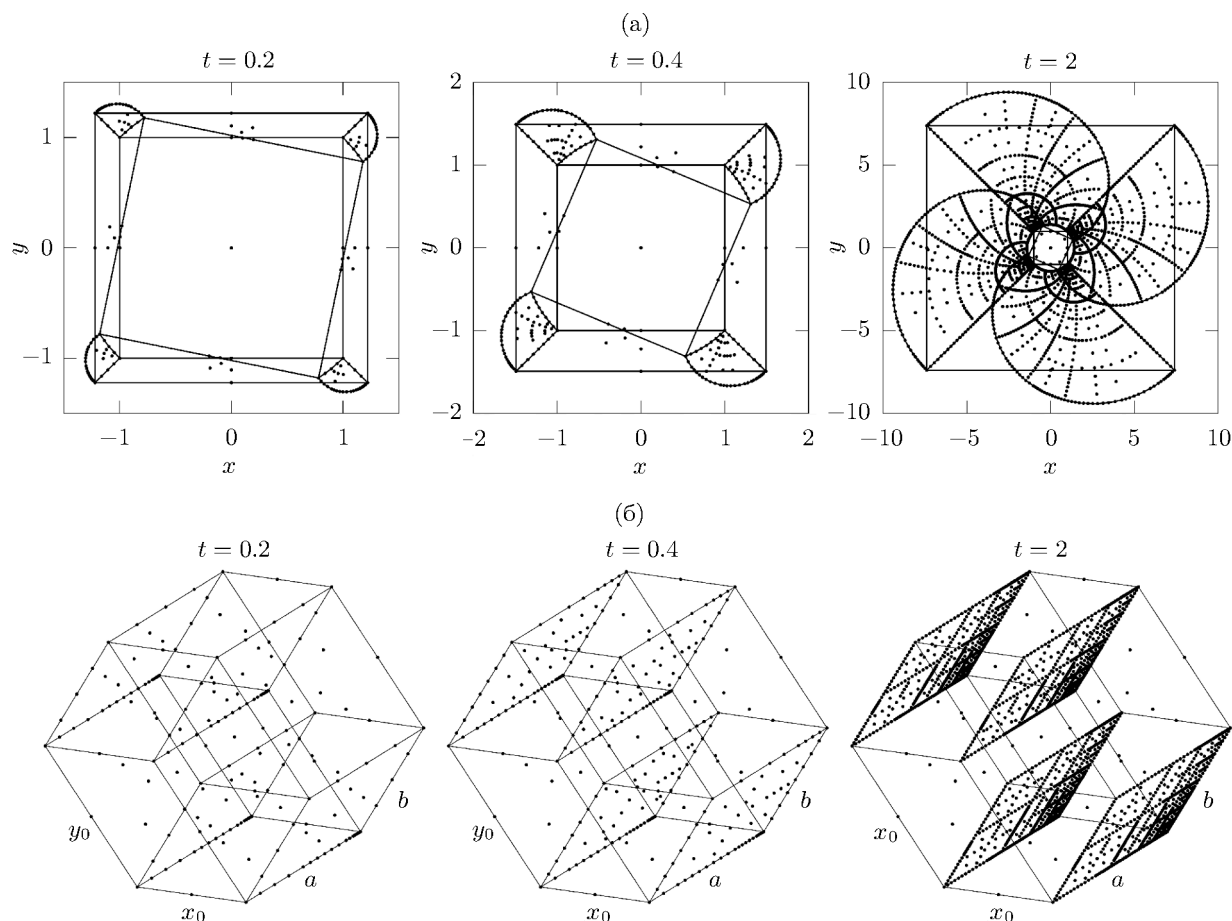


Рис. 5. Результаты решения системы (16): (а) – множество решений в различные моменты времени; (б) – узлы адаптивной разреженной сетки.

Выполним сравнение вычислительных затрат алгоритма адаптивной интерполяции на основе разреженных сеток с классическим вариантом алгоритма. В работе [13] определён критерий I , который характеризует вычислительные затраты и численно равен числу решённых неинтервальных систем (2) при конкретных точечных значениях интервальных параметров. В таблице приведены значения погрешности и критерия I для рассмотренных выше трёх систем ОДУ. Сравнение результатов показывает, что алгоритм адаптивной интерполяции на основе разреженных сеток для данных задач работает значительно быстрее, чем классический вариант алгоритма.

Таблица. Результаты работы классического алгоритма адаптивной интерполяции при $p = 3$ и алгоритма адаптивной интерполяции на основе разреженных сеток

Система ОДУ	Классический алгоритм		Алгоритм на основе разреженных сеток	
	errог	I	errог	I
(13)	$9.89 \cdot 10^{-3}$	1 560	$7.71 \cdot 10^{-3}$	820
(14)	$1.62 \cdot 10^{-2}$	1 318	$1.2 \cdot 10^{-2}$	521
(16)	$1.3 \cdot 10^{-1}$	14 549	$9.45 \cdot 10^{-2}$	1 968

Закключение. В работе рассмотрены теоретические аспекты обобщения алгоритма адаптивной интерполяции на случай большого числа интервальных параметров с использованием разреженных сеток. В процессе работы алгоритма на каждом шаге выполняется построение

функции, которая интерполирует зависимость решения от точечных значений интервальных параметров. В классическом варианте алгоритма используются регулярные сетки, на основе которых строится интерполяционный полином в форме Лагранжа. При этом количество узлов в сетке зависит экспоненциально от числа интервальных параметров. В данной статье изучен вопрос замены регулярных сеток на разреженные сетки. В процессе работы алгоритма происходит адаптация сетки к решению, при этом новые узлы создаются с помощью интерполяции и важно понимать, как в результате будет вести себя глобальная погрешность. Получена оценка глобальной погрешности и показано, что она линейно зависит от текущего уровня сетки. Проведена апробация алгоритма адаптивной интерполяции на разреженных сетках на нескольких системах ОДУ, демонстрирующая эффективность рассматриваемого подхода. Полученные результаты согласуются с теоретическими оценками.

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации (проект 075-15-2020-799).

СПИСОК ЛИТЕРАТУРЫ

1. Мозохина А.С., Мухин С.И. Некоторые точные решения задачи о течении жидкости в сокращающемся эластичном сосуде // Мат. моделирование. 2019. Т. 31. № 3. С. 124–140.
2. Moore R. Interval Analysis. New Jersey, 1966.
3. Moore R.E., Kearfott R.B., Cloud M.J. Introduction to Interval Analysis. Philadelphia, 2009.
4. Шарый С.П. Конечномерный интервальный анализ. Новосибирск, 2019.
5. Eijgenraam P. The Solution of Initial Value Problems Using Interval Arithmetic. Math. Centre Tracts. № 144. Stichting Math. Centrum. Amsterdam, 1981.
6. Lohner R.J. Enclosing the solutions of ordinary initial and boundary value problems // Comput. Arithmetic: Scientific Comput. and Program. Lang. 1987. P. 255–286.
7. Черноусько Ф.Л. Оценивание фазовых состояний динамических систем. Метод эллипсоидов. М., 1988.
8. Makino K., Berz M. Models and their applications // Numerical Software Verification 2017: conference. Heidelberg, Germany, July 22–23, 2017. Springer International Publishing AG, 2017. P. 3–13.
9. Nataraj P.S.V., Sondur S. The extrapolated Taylor Model // Reliable Computing. July, 2011. P. 251–278.
10. Розалев А.Н. Гарантированные методы решения систем обыкновенных дифференциальных уравнений на основе преобразования символьных формул // Вычислит. технологии. 2003. Т. 8. № 5. С. 102–116.
11. Морозов А.Ю., Ревизников Д.Л. Методы компьютерного моделирования динамических систем с интервальными параметрами. М., 2019.
12. Sauer T., Xu Y. On multivariate Lagrange interpolation // Math. of Comput. 1995. V. 64. № 211. P. 1147–1170.
13. Морозов А.Ю., Ревизников Д.Л. Алгоритм адаптивной интерполяции на основе kd-дерева для численного интегрирования систем обыкновенных дифференциальных уравнений с интервальными начальными условиями // Дифференц. уравнения. 2018. Т. 54. № 7. С. 963–974.
14. Морозов А.Ю., Ревизников Д.Л., Гудаснов В.Ю. Алгоритм адаптивной интерполяции на основе kd-дерева для решения задач химической кинетики с интервальными параметрами // Мат. моделирование. 2018. Т. 30. № 12. С. 129–144.
15. Morozov A.Yu., Reviznikov D.L. Modelling of dynamic systems with interval parameters on graphic processors // Program. Ingenieria. 2019. V. 10. № 2. P. 69–76.
16. Khoirudin, Shun-Liang J. Gpu application in cuda memory // Adv. Comput.: An Int. J. 2015. V. 6. № 2. P. 1–10.
17. Morozov A.Yu., Reviznikov D.L. Modeling of dynamic systems with interval parameters in the presence of singularities // Rus. J. Nonlin. Dyn. 2020. V. 16. № 3. P. 479–490.
18. Морозов А.Ю., Журавлев А.А., Ревизников Д.Л. Анализ и оптимизация алгоритма адаптивной интерполяции численного решения систем обыкновенных дифференциальных уравнений с интервальными параметрами // Дифференц. уравнения. 2020. Т. 56. № 7. С. 960–974.
19. Morozov A.Yu., Zhuravlev A.A., Reviznikov D.L. Sparse grid adaptive interpolation in problems of modeling dynamic systems with interval parameters // Mathematics. 2021. V. 9. № 4. P. 298.
20. Смоляк С.А. Квадратурные и интерполяционные формулы на тензорных произведениях некоторых классов функций // Докл. АН СССР. 1963. Т. 148. № 5. С. 1042–1045.

21. *Gerstner T., Griebel M.* Sparse grids // Encyclopedia of Quantitative Finance / Ed. R. Cont. New York, 2010.
22. *Garcke J.* Sparse grids in a nutshell. Sparse grids and applications // Lect. Not. in Comput. Sci. and Engin. V. 88. Berlin; Heidelberg, 2013. P. 57–80.
23. *Brumm J., Scheidegger S.* Using adaptive sparse grids to solve high-dimensional dynamic models // Econometrica. 2017. V. 85. № 5. P. 1575–1612.
24. *Bungartz H.-J., Griebel M.* Sparse grids // Acta Numerica. 2004. V. 13. № 1. P. 147–269.
25. *Judd K.L., Maliar L., Maliar S., Valero R.* Smolyak method for solving dynamic economic models: Lagrange interpolation, anisotropic grid and adaptive domain // J. of Economic Dyn. and Cont. 2014. V. 44. P. 92–123.
26. *Niesen J., Hall T.* On the Global Error of Discretization Methods for Ordinary Differential Equations. Cambridge, 2004.
27. *Арнольд В.И.* Обыкновенные дифференциальные уравнения. М., 2012.

Федеральный исследовательский центр
“Информатика и управление”
Российской академии наук, г. Москва,
Московский авиационный институт
(национальный исследовательский университет)

Поступила в редакцию 15.02.2021 г.
После доработки 15.02.2021 г.
Принята к публикации 27.04.2021 г.

ЧИСЛЕННЫЕ МЕТОДЫ

УДК 519.63+517.958:532.5

ЧИСЛЕННОЕ РЕШЕНИЕ УРАВНЕНИЙ
РЕЛАКСАЦИОННОЙ МОДЕЛИ БАЕРА–НУНЦИАТО
С ПОМОЩЬЮ РАЗРЫВНОГО МЕТОДА ГАЛЁРКИНА

© 2021 г. Р. Р. Тухватуллина, М. В. Алексеев, Е. Б. Савенков

Предлагается численный метод для решения задач двухфазной двухскоростной гидродинамики в рамках модели Баера–Нунциато с релаксацией. Уравнения модели решаются с помощью разрывного метода Галёркина с лимитером WENO-S, который применяется непосредственно к консервативным переменным модели. Релаксационные процессы моделируются с использованием неявного метода Рунге–Кутты второго порядка с адаптивным выбором шага интегрирования. Алгоритм включает в себя метод Ньютона для решения уравнений нелинейного метода Рунге–Кутты. Соответствующие матрицы Якоби вычисляются с применением численного дифференцирования. Приводятся результаты численных расчётов, демонстрирующих возможности предложенного алгоритма. Проводится сравнение результатов численных расчётов с аналитическим решением, а также с результатами, полученными другими авторами. Представлены результаты численных экспериментов с различными скоростями релаксации, в том числе рассмотрен случай “жёсткой” релаксации.

DOI: 10.31857/S0374064121070116

Введение. В настоящей работе рассматриваются вопросы численного решения двухфазной полностью неравновесной модели Баера–Нунциато (Б.–Н.) с релаксационными слагаемыми. Впервые данная модель была предложена в работе [1] для анализа процесса перехода дефлаграции в детонацию при моделировании динамики горения гранулированных взрывчатых веществ. В дальнейшем она применялась для решения широкого спектра задач и в настоящее время может рассматриваться как базовая модель для целого ряда обобщений [2–6].

Математическая формулировка модели приведена в п. 1 настоящей работы. Укажем здесь основные свойства системы уравнений Б.–Н., которые усложняют задачу построения вычислительных алгоритмов для её решения. Система уравнений Б.–Н. является гиперболической системой первого порядка и включает в себя уравнения как в дивергентной форме, так и в квазилинейной. Она не может быть записана в дивергентной форме. Система включает в себя релаксационные слагаемые, которые описывают процесс релаксации механических и термодинамических параметров фаз к равновесному значению. По крайней мере в ряде приложений характерные времена релаксации могут быть значительно меньше характерных времён протекания гидродинамических процессов. В этом смысле система уравнений Б.–Н. является “жёсткой”.

Вследствие сказанного численное решение уравнений Б.–Н. представляет собой сложную задачу, решению которой посвящено значительное число работ. Большая их часть направлена на преодоление указанных выше трудностей.

Из многочисленных методов решения системы уравнений Б.–Н. далее будем рассматривать только методы гоуновского типа. Основой методов этого класса является способ постановки и решения задачи Римана о распаде разрыва. Применительно к системе уравнений Б.–Н. основным вопросом является корректное определение обобщённого решения. Возникающая при этом трудность связана с тем, что, как сказано выше, эта система включает в себя уравнения как в дивергентной, так и в квазилинейной форме. Это делает классическое определение обобщённого решения неприменимым в рассматриваемом случае. Наиболее хорошо разработанным способом разрешения данной проблемы является теория, предложенная в фундаментальной работе [7]. Этот подход используется в настоящей работе.

Способам построения схем высокого порядка для модели Б.–Н. посвящено значительное число работ (см., например, [8–11]). Большинство из них основывается на применении методов

типа WENO. В настоящей работе применяется разрывный метод Галёркина/Рунге–Кутты (RK/DG, Runge–Kutta/Discontinuous Galerkin method [12]). Такой выбор обусловлен тем, что разрывный метод Галёркина допускает формальные обобщения на случай уравнений высокого порядка, к которым сводятся модели типа Б.–Н.

В рамках настоящей работы модель Б.–Н. рассматривается в качестве основы для дальнейших обобщений. В силу этого используются алгоритмы, наиболее полно отвечающие идее “physics-free”, т.е. минимально учитывающие особенности задачи, но, вместе с тем, пригодные для получения конечных результатов с приемлемой точностью. По этой причине в настоящей работе:

- используется максимально простой численный поток Русанова, гарантирующий устойчивость аппроксимаций;
- лимитирование решения осуществляется в терминах консервативных переменных, без перехода к характеристическим.

Для монотонизации решения используется лимитер WENO-S [13].

Одной из основных трудностей при решении задач с использованием модели Б.–Н. является учёт релаксационных процессов. Более общая модель Б.–Н. с жёсткой релаксацией даёт результаты в соответствии с менее общими равновесными моделями [3–5], причём при этом делается значительно меньше предположений относительно протекающих физических процессов в многофазной среде. В этой модели релаксация скоростей фаз и их термодинамических параметров учитывается в виде источниковых слагаемых, при наличии которых система уравнений может становиться жёсткой. С вычислительной точки зрения операция интегрирования таких систем является сложной, поскольку характерные времена релаксации могут быть значительно меньше, чем шаг интегрирования по времени в вычислительной схеме.

Разработка эффективных методов аппроксимаций уравнений Б.–Н. в настоящее время не вполне завершена. Основные усилия направлены здесь на разработку методов интегрирования жёстких систем уравнений специального вида и алгоритмов решения полной задачи с учётом релаксационных слагаемых, в частности, разработку соответствующих схем расщепления по физическим процессам. Работы в этой области сосредоточены в основном в направлении развития так называемых асимптотически корректных разностных схем (“asymptotic preserving schemes”) (см. [14, 15]), которые, в частности, имеют равномерную по малому времени релаксации скорость сходимости численных аппроксимаций.

Целью настоящей работы является реализация и демонстрация возможностей алгоритма, который основывается на таких сравнительно простых средствах численных методов как:

- разрывный метод Галёркина для построения пространственных аппроксимаций уравнений модели;
- лимитирование консервативных переменных с использованием геометрического лимитера WENO-S;
- теория Dal Maso–Le Floch–Murat для формулировки задачи Римана и, как следствие, численных потоков для неконсервативной “части” задачи;
- простейшее расщепление по физическим процессам для учёта жёстких релаксационных слагаемых;
- неявный метод Рунге–Кутты для интегрирования уравнений модели.

Эффективность предложенной методики иллюстрируется результатами численных экспериментов в одномерном случае.

Структура работы следующая. Математическая модель Б.–Н. и её особенности представлены в п. 1. Разрывный метод Галёркина для гиперболических уравнений с неконсервативными слагаемыми кратко описан в пп. 2.1. Алгоритм вычисления релаксационных слагаемых изложен в пп. 2.2. Схема лимитирования представлена в пп. 2.3. Результаты численных экспериментов приведены в п. 3.

1. Модель Баера–Нунциато с релаксацией. Модель Б.–Н. с релаксационными слагаемыми, описывающая двухфазное течение, может быть представлена в виде системы гиперболических уравнений [5]

$$\frac{\partial \alpha_k}{\partial t} + \mathbf{u}_I \cdot \nabla \alpha_k = \nu(P_k - P_k^-), \quad (1.1a)$$

$$\frac{\partial \alpha_k \rho_k}{\partial t} + \nabla \cdot (\alpha_k \rho_k \mathbf{u}_k) = 0, \quad (1.1b)$$

$$\frac{\partial \alpha_k \rho_k \mathbf{u}_k}{\partial t} + \nabla \cdot (\alpha_k \rho_k \mathbf{u}_k \otimes \mathbf{u}_k) + \nabla (\alpha_k P_k) - P_I \nabla \alpha_k = \mu (\mathbf{u}_{\bar{k}} - \mathbf{u}_k), \quad (1.1c)$$

$$\frac{\partial \alpha_k \rho_k E_k}{\partial t} + \nabla \cdot (\alpha_k (\rho_k E_k + P_k) \mathbf{u}_k) - P_I \mathbf{u}_I \nabla \alpha_k = \mu (\mathbf{u}_{\bar{k}} - \mathbf{u}_k) \cdot \mathbf{u}_I + \nu (P_{\bar{k}} - P_k) P_I. \quad (1.1d)$$

Здесь α_k , ρ_k , $\mathbf{u}_k$, P_k , E_k – объёмная доля, плотность, поле скоростей, давление и полная энергия фазы $k = 1, 2$ соответственно; $\bar{k} = \{1, 2\} \setminus \{k\}$; ν , μ – релаксационные параметры. Для объёмных долей выполнено условие нормировки $\alpha_1 + \alpha_2 = 1$. Полная энергия k -й фазы определяется равенством

$$E_k = \mathcal{U}_k + \mathbf{u}_k \cdot \mathbf{u}_k / 2,$$

где $\mathcal{U}_k$ – внутренняя энергия фазы.

Уравнения модели (1.1) включают в себя уравнение динамики для объёмной доли (1.1a), законы сохранения массы (1.1b), импульса (1.1c) и энергии (1.1d).

Величины $\mathbf{u}_I$ и P_I являются так называемыми “интерфейсными” скоростью и давлением. Для замыкания модели необходимо указать их конкретную зависимость от переменных задачи. Она может быть определена целым рядом способов, причём выбор замыкающих соотношений непосредственно влияет на структуру волн и поведение фаз в модели (см. [1, 16–18]). В настоящей работе используется вариант, предложенный в основополагающей работе [1]:

$$\mathbf{u}_I = \mathbf{u}_1, \quad P_I = P_2,$$

где $k = 1$ соответствует менее плотной фазе (более сжимаемой), $k = 2$ – более плотной фазе (менее сжимаемой). Термодинамические свойства фаз определяются уравнениями состояния вида $\mathcal{U}_k = \mathcal{U}_k(P_k, \rho_k)$. В данной работе они определены равенствами

$$\mathcal{U}_k = \frac{P_k + \gamma_k P_{\infty, k}}{(\gamma_k - 1) \rho_k},$$

где $P_{\infty, k}$ и γ_k – параметры.

Отметим, что система уравнений (1.1) содержит нелинейные члены вида $P_I \nabla \alpha_k$ и не может быть записана в консервативном виде. Эта особенность является типичной для целого ряда многофазных многоскоростных моделей (см. [19]).

2. Вычислительные алгоритмы. Рассматриваемый алгоритм численного решения системы уравнений (1.1) основан на следующих подходах: использовании расщепления по физическим процессам для расчёта релаксационных правых частей системы; применении разрывного метода Галёркина для решения однородной системы (1.1) и неявного метода Рунге–Кутты для интегрирования “жёстких” правых частей системы (1.1). Далее описаны основные компоненты полного алгоритма решения задачи.

2.1. Разрывный метод Галёркина. В данном пункте представлена схема разрывного метода Галёркина [11] для одномерного варианта неоднородной гиперболической системы уравнений (1.1), записанной в абстрактной форме

$$\frac{\partial \mathbf{Q}(x, t)}{\partial t} + \mathcal{B}(\mathbf{Q}) \frac{\partial \mathbf{Q}}{\partial x} = \mathcal{S}(\mathbf{Q}), \quad (2.1)$$

где $x \in \Omega = [0, L] \subset \mathbb{R}$, $t \in [0, T] \subset \mathbb{R}$, $\mathbf{Q} = \mathbf{Q}(x, t) = (Q_1, \dots, Q_M)^T$ – решение уравнения (2.1), $M = 7$ – число компонент вектора $\mathbf{Q}$,

$$\mathcal{B}(\mathbf{Q}) = \partial \mathcal{F}(\mathbf{Q}) / \partial \mathbf{Q} + \mathcal{A}(\mathbf{Q}). \quad (2.2)$$

Здесь $\mathcal{A}$ – заданная матрица, зависящая от решения задачи и определяющая квазилинейную “часть” системы уравнений, $\mathcal{F}$ – заданный вектор потоков, $\mathcal{B}$ – матрица системы, записанной в квазилинейной форме, $\mathcal{S}$ – плотность источников. Далее считается, что матрицы $\mathcal{A}$, $\mathcal{B}$ и

вектор поток $\mathcal{F}$ являются непрерывными и имеют требуемое количество производных как функции вектора переменных $\mathbf{Q}$. Обобщённое решение данной задачи для консервативных систем (в случае, когда $\mathcal{A}(\mathbf{Q}) = 0$) может быть определено в классе обобщённых функций. Для неконсервативных систем (т.е. в случае, когда $\mathcal{A}(\mathbf{Q}) \neq 0$) обобщённое решение не может быть определено так же, как в случае консервативных систем. Для решения этой проблемы в настоящей работе используется распространённый подход, основанный на применении теории DLM (Dal Maso–Le Floch–Murat [7]). В рамках этого подхода для корректного определения “неконсервативного произведения” вида $\mathcal{A}(\mathbf{Q})$ вводится липшицево отображение $\Psi : [0, 1] \times \times \mathbb{R}^7 \rightarrow \mathbb{R}^7$ (путь), которое “соединяет” левое значение решения в точке разрыва с правым:

$$\Psi(\mathbf{Q}^+, \mathbf{Q}^-; 0) = \mathbf{Q}^-, \quad \Psi(\mathbf{Q}^+, \mathbf{Q}^-; 1) = \mathbf{Q}^+, \quad \Psi(\mathbf{Q}, \mathbf{Q}; s) = \mathbf{Q}.$$

Введём разбиение $\{\omega_i\}_{i=0}^{i=N}$ области Ω и обозначим ячейку сетки (конечный элемент)

$$\omega_i = [x_{i-1/2}, x_{i+1/2}], \quad 1 \leq i \leq N.$$

Обозначим через $\mathcal{V}_h^k(\Omega)$ подпространство в $\mathcal{L}^1(\Omega)$, состоящее из тех элементов, сужения (ограничения) которых на ячейки ω_i принадлежат векторному пространству $\mathcal{P}^k(\omega_i)$ полиномов степени не выше k :

$$\mathcal{V}_h^k = \{v \in \mathcal{L}^1(\Omega) : v|_{\omega_i} \in \mathcal{P}^k(\omega_i); \quad 1 \leq i \leq N\}.$$

Представим решение $\mathbf{Q}(x, t)$ в ячейке ω_i конечномерной аппроксимацией $\mathbf{Q}_h \in \mathcal{V}_h^k$:

$$\mathbf{Q}_h(x, t)|_{\omega_i} = \sum_{l=0}^k \psi_i^{(l)}(x) \mathbf{Q}_i^{(l)}(t), \quad (2.3)$$

где $\psi_i^{(l)}$ – полином Лежандра степени l . Здесь и в дальнейшем индекс i обозначает принадлежность i -й ячейке. Чтобы получить полудискретную систему уравнений для $\mathbf{Q}_h(x, t)$, умножим уравнение (2.1) на пробную функцию $v_h \in \mathcal{V}_h^k$ и проинтегрируем получившееся тождество по отрезку ω_i :

$$\int_{\omega_i} \frac{\partial \mathbf{Q}_h(x, t)}{\partial t} v_h(x) dx + \left\langle \left[\mathcal{B}(\mathbf{Q}(\cdot, t)) \frac{\partial \mathbf{Q}(\cdot, t)}{\partial x} \right]_{\Psi}, v_h \right\rangle = \int_{\omega_i} \mathcal{S}(\mathbf{Q}_h) v_h(x) dx.$$

Неконсервативное произведение может быть определено в обобщённом смысле следующим образом [20]:

$$\begin{aligned} \left\langle \left[\mathcal{B}(\mathbf{Q}) \frac{\partial \mathbf{Q}}{\partial x} \right]_{\Psi}, v_h \right\rangle &\stackrel{\text{def}}{=} \int_{\omega_i} \mathcal{B}(\mathbf{Q}(x, t)) \frac{\partial \mathbf{Q}(x, t)}{\partial x} v_h(x) dx + \\ &+ \sum_d \left(\int_0^1 \mathcal{B}(\Psi(\mathbf{Q}_d^+, \mathbf{Q}_d^-, s)) \frac{\partial \Psi}{\partial s}(\mathbf{Q}_d^+, \mathbf{Q}_d^-, s) ds \right) v_h(x_d), \end{aligned} \quad (2.4)$$

где индекс d нумерует точки разрыва решения $\mathbf{Q}$, при этом $\mathbf{Q}_d^+$ и $\mathbf{Q}_d^-$ – предельные значения решения соответственно справа и слева от разрыва в точке x_d в момент времени t . Заметим, что данное определение существенно зависит от выбранного пути Ψ . В случае, когда $\mathcal{A} = 0$ произведение (2.4) не зависит от выбранного пути и совпадает с классическим определением обобщённого произведения. Следствием определения (2.4) и системы (2.1) без релаксационных слагаемых является условие Гюгоньо, которому должно удовлетворять обобщённое решение в точках разрыва x_d :

$$\int_0^1 (\xi_d \mathcal{I} - \mathcal{B}(\Psi(\mathbf{Q}_d^-, \mathbf{Q}_d^+; s))) \frac{\partial \Psi}{\partial s}(\mathbf{Q}_d^-, \mathbf{Q}_d^+; s) ds = 0,$$

где ξ_d – скорость движения разрыва, $\mathcal{I}$ – единичная матрица. С учётом равенства (2.2) получим

$$\begin{aligned} \left\langle \left[\mathcal{B}(\mathbf{Q}) \frac{\partial \mathbf{Q}}{\partial x} \right]_{\Psi}, v_h \right\rangle &= \int_{\omega_i} \left(\mathcal{A}(\mathbf{Q}(x, t)) \frac{\partial \mathbf{Q}(x, t)}{\partial x} + \frac{\partial \mathcal{F}(\mathbf{Q})}{\partial x} \right) v_h(x) dx + \\ &+ \sum_d \left(\mathcal{F}(\mathbf{Q}_d^+) - \mathcal{F}(\mathbf{Q}_d^-) + \left(\int_0^1 \mathcal{A}(\Psi(\mathbf{Q}_d^+, \mathbf{Q}_d^-, s)) \frac{\partial \Psi}{\partial s}(\mathbf{Q}_d^+, \mathbf{Q}_d^-, s) ds \right) \right) v_h(x_d). \end{aligned}$$

Постановка задачи Римана для соответствующей неконсервативной системы представлена в [20]. На основе сделанных выше построений могут быть выведены соответствующие схемы типа Годунова. В частности, численная схема для разрывного метода Галёркина может быть записана в виде

$$\begin{aligned} \int_{\omega_i} \frac{\partial \mathbf{Q}_h}{\partial t} v_h(x) dx + \int_{\omega_i} \left(\mathcal{A}(\mathbf{Q}_h) \frac{\partial \mathbf{Q}_h}{\partial x} + \mathcal{F}(\mathbf{Q}_h) \right) v_h(x) dx + \\ + (v_h(x_{i+1/2})^- D_{i+1/2}^- + v_h(x_{i+1/2})^+ D_{i-1/2}^+) = \int_{\omega_i} \mathcal{S}(\mathbf{Q}_h) v_h(x) dx, \end{aligned}$$

где $D_{i\pm 1/2}^\mp$ и $v_h(x_{i\pm 1/2})^\mp$ определены на границах соответствующих ячеек,

$$\begin{aligned} D_{i+1/2}^- &= \mathcal{F}(\hat{\mathbf{Q}}_{i+1}) - \mathcal{F}(\mathbf{Q}_{i+1/2}^-) + \int_0^1 \mathcal{A}(\Psi(\mathbf{Q}_{i+1/2}^-, \hat{\mathbf{Q}}_{i+1}; s)) \frac{\partial \Psi}{\partial s}(\mathbf{Q}_{i+1/2}^-, \hat{\mathbf{Q}}_{i+1}; s) ds, \\ D_{i+1/2}^+ &= \mathcal{F}(\mathbf{Q}_{i+1/2}^+) - \mathcal{F}(\hat{\mathbf{Q}}_{i+1}) + \int_0^1 \mathcal{A}(\Psi(\hat{\mathbf{Q}}_{i+1}, \mathbf{Q}_{i+1/2}^+; s)) \frac{\partial \Psi}{\partial s}(\hat{\mathbf{Q}}_{i+1}, \mathbf{Q}_{i+1/2}^+; s) ds, \end{aligned}$$

здесь $\hat{\mathbf{Q}}_{i+1}$ – решение задачи Римана на границе ячеек.

Пусть выбран линейный путь $\Psi(\mathbf{Q}^-, \mathbf{Q}^+; s) = \mathbf{Q}_d^- + s(\mathbf{Q}_d^+ - \mathbf{Q}_d^-)$ и в качестве численного потока используется поток Русанова. Тогда в точках разрыва $x_d = x_{i\pm 1/2}$ будем иметь

$$D_d^\pm = \mathcal{F}(\mathbf{Q}_d^+) - \mathcal{F}(\mathbf{Q}_d^-) + A_d^\pm (\mathbf{Q}_d^+ - \mathbf{Q}_d^-), \quad A_d^\pm = \frac{1}{2} \int_0^1 (\mathcal{A}(\mathbf{Q}_d^- + s(\mathbf{Q}_d^+ - \mathbf{Q}_d^-)) \pm \mathcal{I} \Lambda) ds,$$

где $\Lambda = \lambda(s)$ – максимальные собственные значения матрицы $\mathcal{B}(\mathbf{Q}_d^- + s(\mathbf{Q}_d^+ - \mathbf{Q}_d^-))$. Отметим, что применение более простого потока Лакса–Фридрихса приводит, вообще говоря, к неустойчивой схеме (см. приложение 2).

Рассматривая произвольный полином $v_h(x) \in \text{span} \{\psi_i^{(l)}\}$, получаем следующую полудискретную систему уравнений относительно переменных $\{\mathbf{Q}_i^{(l)}\}$, определённых в представлении (2.3):

$$\mathbf{M} \frac{d\tilde{\mathbf{Q}}_i}{dt} = \mathbf{H}(\tilde{\mathbf{Q}}_i) + \mathbf{I}(\tilde{\mathbf{Q}}_i). \quad (2.5)$$

Здесь $\tilde{\mathbf{Q}}_i = (Q_{1,i}^{(0)}, \dots, Q_{M,i}^{(0)}, \dots, Q_{M,i}^{(k)})$ – вектор неизвестных. Его компоненты $Q_{m,i}^{(l)}$ в дальнейшем будем называть l -й гармоникой ($l = \overline{0, k}$) m -й компоненты вектора $\mathbf{Q}_i$. Матрица $\mathbf{M} \in \mathbb{R}^{M(k+1) \times M(k+1)}$ является матрицей Грама указанной системы базисных функций,

вектор $\mathbf{H}(\tilde{\mathbf{Q}}_i)$ представляет собой аппроксимацию дифференциального оператора в левой части системы уравнений (2.1), а вектор $\mathbf{I}(\tilde{\mathbf{Q}}_i)$ – аппроксимацию правой части (2.1),

$$I_m^{(l)} = \int_{\omega_i} S_m(\mathbf{Q}_h) \psi_i^{(l)} dx. \quad (2.6)$$

Одна из наиболее простых стратегий решения системы (2.5) – расщепление по процессам первого порядка. Разобьём временной интервал на части точками $\{t^n\}$. Далее на каждом временном шаге $t \in [t^n, t^{n+1}]$ сначала решается задача Коши для системы ОДУ

$$\mathbf{M} \frac{d\tilde{\mathbf{Q}}_i}{dt} = \mathbf{H}(\tilde{\mathbf{Q}}_i), \quad t \in (t^n, t^{n+1}], \quad (2.7)$$

с начальными данными $\tilde{\mathbf{Q}}_i^0 = \tilde{\mathbf{Q}}_i(t^n)$ для определения “промежуточного” решения $\tilde{\mathbf{Q}}_{i,H} = \tilde{\mathbf{Q}}_i(t^{n+1})$. Затем решается задача Коши для системы ОДУ

$$\mathbf{M} \frac{d\tilde{\mathbf{Q}}_i}{dt} = \mathbf{I}(\tilde{\mathbf{Q}}_i), \quad t \in (t^n, t^{n+1}], \quad (2.8)$$

с начальными данными $\tilde{\mathbf{Q}}_{i,H}$.

Для интегрирования по времени однородной системы (2.7) далее используется вариант метода Рунге–Кутты TVD/RK3 [12] с лимитированием консервативных переменных на каждом шаге метода. В настоящей работе используется лимитер WENO-S, описанный в п. 2.3. Для интегрирования по пространству применяется метод квадратур Гаусса–Лежандра.

Для интегрирования системы уравнений (2.8) используется неявный алгоритм Рунге–Кутты с автоматическим выбором шага интегрирования, описанный в п. 2.2.

2.2. Алгоритм расчёта релаксационных слагаемых. Задача Коши для ОДУ (2.8) решается неявным методом Рунге–Кутты второго порядка на временном интервале $t \in (t^n, t^{n+1}]$. Разобьём временной интервал на части точками $(t_0^n, t_1^n, \dots, t_{M-1}^n)$, где $t_{j+1}^n = t_j^n + \tau_j$, $t_0^n = t_n$, $t_{M-1}^n = t_{n+1}$. Таким образом, $\Delta t = t^{n+1} - t^n$ – шаг интегрирования газодинамической части задачи, а τ_j – шаг интегрирования релаксационной части. При этом шаг τ_j может быть переменным. Разностная схема решения задачи Коши для системы (2.8) имеет следующий вид:

$$\tilde{\mathbf{Q}}_i^0 = \tilde{\mathbf{Q}}_i(t^n); \quad \tilde{\mathbf{Q}}_i^{j+1} = \tilde{\mathbf{Q}}_i^j + \frac{1}{2} \tau_j (\mathbf{R}(\tilde{\mathbf{Q}}_i^{j+1}) + \mathbf{R}(\tilde{\mathbf{Q}}_i^j)), \quad j = 0, 1, \dots, \quad (2.9)$$

где $\mathbf{R} = \mathbf{M}^{-1}\mathbf{I}$. Аппроксимации релаксационных членов $\mathbf{I}$, определённые в (2.6), рассчитываются численно методом квадратур Гаусса–Лежандра.

Нелинейная система уравнений (2.9) решается численно методом Ньютона:

$$\partial \mathbf{F}(\tilde{\mathbf{Q}}_i^{j+1}) / \partial \tilde{\mathbf{Q}}_i^{j+1} |_{\tilde{\mathbf{Q}}_i^{j+1,s}} \cdot (\tilde{\mathbf{Q}}_i^{j+1,s+1} - \tilde{\mathbf{Q}}_i^{j+1,s}) = -\mathbf{F}(\tilde{\mathbf{Q}}_i^{j+1,s}), \quad (2.10)$$

где индекс s обозначает номер итерации, $\tilde{\mathbf{Q}}_i^{j+1,0} = \tilde{\mathbf{Q}}_i^j$ соответствует $s = 0$. Функция $\mathbf{F}$ имеет следующий вид:

$$\mathbf{F}(\tilde{\mathbf{Q}}_i^{j+1}) = \tilde{\mathbf{Q}}_i^{j+1} - \tilde{\mathbf{Q}}_i^j - 0.5 \tau_j (\mathbf{R}(\tilde{\mathbf{Q}}_i^{j+1}) + \mathbf{R}(\tilde{\mathbf{Q}}_i^j)).$$

Якобиан $\partial \mathbf{F}(\tilde{\mathbf{Q}}_i^{j+1}) / \partial \tilde{\mathbf{Q}}_i^{j+1}$ в (2.10) вычисляется с применением численного дифференцирования. Итерации метода Ньютона продолжаются до тех пор, пока величина $r = \max_m |r_m|$ не станет меньше заданного значения. Здесь r_m определены равенством

$$r_m = \begin{cases} (\tilde{\mathbf{Q}}_{i,m}^{j+1,s+1} - \tilde{\mathbf{Q}}_{i,m}^{j+1,s}) / \tilde{\mathbf{Q}}_{i,m}^{j+1,s+1}, & \text{если } |\tilde{\mathbf{Q}}_{i,m}^{j+1,s+1}| > 1, \\ \tilde{\mathbf{Q}}_{i,m}^{j+1,s+1} - \tilde{\mathbf{Q}}_{i,m}^{j+1,s} & \text{в противном случае.} \end{cases}$$

В реализованном алгоритме предусмотрен автоматический выбор шага интегрирования. В начальный момент $\tau_0 = 0.2\Delta t$. В случае, если превышено максимальное число $N_{\max}$ ньютоновских итераций (далее $N_{\max} = 20$) или если решение сходится к нефизичным результатам (отрицательное давление или объёмная доля и т.д.), то шаг интегрирования уменьшается. В случае, когда число итераций меньше минимального значения $N_{\min}$ (далее $N_{\min} = 6$), шаг интегрирования увеличивается.

2.3. Лимитирование переменных. Применение лимитера WENO-S состоит из двух шагов [13].

1. Идентификация ячеек, в которых решение подлежит лимитированию. В данной работе используется TVB-идентификатор, основанный на функции $\minmod$.

2. Применение лимитера WENO-S для реконструкции решения в отмеченных ячейках.

Идентификация ячеек, в которых решение подлежит лимитированию. Обозначим через $\bar{Q}_{m,i}$ усреднённое по объёму решение в ячейке ω_i , т.е.

$$\bar{Q}_{m,i} = \frac{1}{\Delta x_i} \int_{\omega_i} \tilde{Q}_m dx. \quad (2.11)$$

Здесь, как и выше, индекс i обозначает пространственную ячейку, а индекс m обозначает компоненту вектора $\tilde{Q}$, $m = \overline{1, M \cdot k}$.

Определим скачки численного решения в ячейках как $\tilde{Q}_{m,i}^+ = \tilde{Q}_{m,i+1/2}^- - \bar{Q}_{m,i}$, $\tilde{Q}_{m,i}^- = \bar{Q}_{m,i} - \tilde{Q}_{m,i-1/2}^+$. Рассмотрим в каждой ячейке функции

$$\begin{aligned} (\tilde{Q}_{m,i}^+)^{(\text{mod})} &= \minmod(\tilde{Q}_{m,i}^+, \bar{Q}_{m,i+1} - \bar{Q}_{m,i}, \bar{Q}_{m,i} - \bar{Q}_{m,i-1}), \\ (\tilde{Q}_{m,i}^-)^{(\text{mod})} &= \minmod(\tilde{Q}_{m,i}^-, \bar{Q}_{m,i+1} - \bar{Q}_{m,i}, \bar{Q}_{m,i} - \bar{Q}_{m,i-1}), \end{aligned} \quad (2.12)$$

где

$$\minmod(a_1, \dots, a_N) = \begin{cases} s \min(a_1, \dots, a_N), & \text{если } s = \text{sign}(a_k), \quad k = \overline{1, N}, \\ 0 & \text{в противном случае.} \end{cases}$$

Решение в ячейке ω_i подлежит лимитированию, если в (2.12) $(\tilde{Q}_{m,i}^\pm)^{(\text{mod})} \neq (\tilde{Q}_{m,i}^\pm)$, т.е. какая-либо из функций в (2.12) возвращает не первый аргумент.

Применение лимитера WENO-S состоит из следующих шагов.

1. Обозначим полиномы в ячейках ω_j , $j = i-1, i, i+1$, через $p_{-1}(x)$, $p_0(x)$ и $p_1(x)$ соответственно и модифицируем решения в соседних ячейках следующим образом:

$$p_{-1}^{\text{mod}}(x) = p_{-1} - \bar{p}_{-1} + \bar{p}_0, \quad p_1^{\text{mod}}(x) = p_1 - \bar{p}_1 + \bar{p}_0,$$

где $\bar{p}$ – усреднённое по объёму решение, представленное равенством (2.11).

2. Индикаторы гладкости β_i рассчитываем по формуле

$$\beta_i = \sum_{l=1}^k \int_{\omega_j} \Delta x_j^{2l-1} \left(\frac{\partial^l}{\partial x^l} p_j(x) \right)^2 dx,$$

где k – степень полинома $p_j(x)$.

3. Весы κ_j рассчитываем по формуле

$$\kappa_j = \bar{\kappa}_j / \sum_{n=-1,0,+1} \bar{\kappa}_n, \quad \bar{\kappa}_n = \frac{\gamma_n}{(\varepsilon + \beta_n)^r}, \quad n = -1, 0 + 1.$$

Здесь γ_n – линейный вес, $\varepsilon = 10^{-6}$ и $r = 2$. Линейные веса должны удовлетворять следующим требованиям: $\gamma_0 \gg \gamma_{\pm 1}$, $\gamma_{-1} + \gamma_0 + \gamma_{+1} = 1$. В настоящей работе $\gamma_0 = 0.998$, $\gamma_{\pm 1} = 0.001$.

4. Реконструкция решения в центральной ячейке ω_0 проводится по формуле

$$p_0^{\text{new}}(x) = \kappa_{-1} p_{-1}^{\text{mod}}(x) + \kappa_0 p_0(x) + \kappa_1 p_1^{\text{mod}}(x).$$

Применение лимитера в одномерном случае может быть записано в операторном виде $p_0^{\text{new}} = \Lambda_\alpha p_0$ для направления $\mathcal{O}_{x_\alpha}$. Для многомерного случая и декартовых сеток $p_0^{\text{new}} = \Lambda p_0$, где $\Lambda = \Lambda_z \Lambda_y \Lambda_x$.

Общая схема решения однородной гиперболической системы с применением описанного выше лимитера имеет приведённый ниже вид.

Для интегрирования по времени системы обыкновенных дифференциальных уравнений (2.7) используется TVD/RK3 метод Рунге–Кутты, представленный ниже. На k -й стадии метода Рунге–Кутты используется способ лимитирования переменных, описанный выше:

$$\begin{aligned}\tilde{Q}_i^1 &= \tilde{Q}_i^n + \Delta t \mathbf{P}(\tilde{Q}_i^n), & \overline{\tilde{Q}_i^1} &= \tilde{\Lambda} \tilde{Q}_i^1, \\ \tilde{Q}_i^2 &= \frac{3}{4} \tilde{Q}_i^n + \frac{1}{4} \tilde{Q}_i^1 + \frac{1}{4} \Delta t \mathbf{P}(\overline{\tilde{Q}_i^1}), & \overline{\tilde{Q}_i^2} &= \tilde{\Lambda} \tilde{Q}_i^2, \\ \tilde{Q}_i^{n+1} &= \frac{1}{3} \tilde{Q}_i^n + \frac{2}{3} \tilde{Q}_i^2 + \frac{2}{3} \Delta t \mathbf{P}(\overline{\tilde{Q}_i^2}), & \overline{\tilde{Q}_i^{n+1}} &= \tilde{\Lambda} \tilde{Q}_i^{n+1},\end{aligned}$$

где $\mathbf{P} = \mathbf{M}^{-1} \mathbf{H}$ (см. (2.5)).

3. Вычислительные эксперименты. В настоящем пункте рассмотрен ряд тестовых расчётов, которые демонстрируют возможности предложенного алгоритма. В тесте 1 проводится тестирование численного алгоритма решения системы ОДУ (2.8). Приводится сравнение с численными результатами, полученными другими авторами [21], а также с аналитическим решением (см. приложение 1). В тесте 2 численно исследуется метод для решения полной задачи (1.1). Рассматриваются численные расчёты задачи Римана с различными коэффициентами релаксации. Приводится сравнение с численными результатами, полученными конечно-объёмной схемой первого порядка с численным потоком HLLEM (соответствующие решения далее называются “решения HLLEM”). Все расчёты предложенным методом приведены для разрывного метода Галёркина с линейным восполнением решения в ячейках сетки, что соответствует $k = 1$ в аппроксимации (2.3).

Результаты расчётов представлены для примитивных переменных (фазовые давления и скорости).

Тест 1. Рассматривается задача Коши для уравнения (2.8).

Начальные условия имеют следующие вид: для нулевых гармоник

$$u_1 = -5, \quad u_2 = 5, \quad P_1 = 0.1, \quad P_2 = 20, \quad \alpha_1 = 0.9, \quad \rho_1 = 1.1111, \quad \rho_2 = 40,$$

для первых гармоник

$$\tilde{Q}_{2,3,6,7}^{(1)} = 0.01 \cdot \tilde{Q}_{2,3,6,7}^{(0)}, \quad \tilde{Q}_{1,4,5,8,9}^{(1)} = 0.$$

Параметры уравнения состояния: $\gamma_1 = 6$, $\pi_1 = 0$ Па, $\gamma_2 = 1.4$, $\pi_2 = 0$ Па. Значения параметров релаксации $\mu = 10^9$ кг/(м·с) и $\nu = 10$ Па⁻¹·с⁻¹. На рис. 1, а и рис. 1, б представлены результаты расчётов для нулевой и первой гармоник фазовых давлений соответственно с $\tau_0 = 2 \cdot 10^{-4}$ с. При расчёте с шагом $\tau_0 = 2 \cdot 10^{-4}$ с скорости фаз сразу выходят на стационарное значение 3 м/с (эти результаты не показаны на рис. 1). На рис. 1, в и рис. 1, г представлены результаты расчётов для нулевой и первой гармоник фазовых скоростей соответственно с $\tau_0 = 2 \cdot 10^{-9}$ с. Видим, что численный расчёт хорошо согласуется с результатами, полученными в [21] для нулевых гармоник (фазовые скорости и давления), а также хорошо согласуется с аналитическим решением для первых гармоник (фазовые скорости). Аналитическое решение для первых гармоник фазовых давлений в данной работе получено не было и сравнение с ним не приводится. Также видно, что нулевые и первые гармоники для фазовых давлений и скоростей релаксируют на одних и тех же временах. Резкое изменение давления (см. рис. 1, а и

рис. 1, б) в начальный момент времени связано с релаксацией скоростей, которая происходит на временах порядка 10 нс.

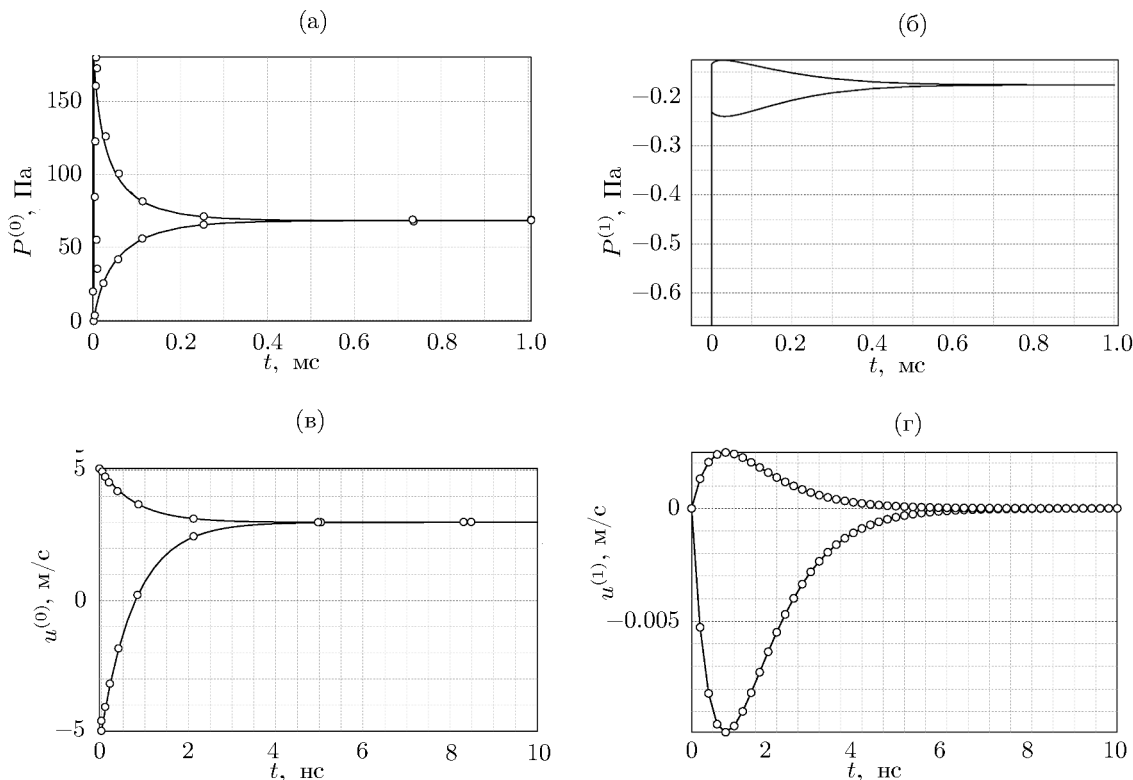


Рис. 1. Тест 1. Сплошные кривые – результаты численного расчёта, символы на (а) и (в) – результаты численного расчёта в [21], символы на (г) – аналитическое решение.

Тест 2. В данном тесте рассматривается задача Римана о распаде разрыва системы уравнений (1.1). В начальный момент времени разрыв находится в точке $x_0 = 0.6$ м. Длина расчётной области 1 м. Слева от разрыва $P_1 = P_2 = 1$, $\rho_1 = 1$, $\rho_2 = 0.2$, $u_1 = u_2 = 0$, $\alpha_1 = 0.55$, справа от разрыва $P_1 = P_2 = 0.1$, $\rho_1 = 0.125$, $\rho_2 = 2$, $u_1 = u_2 = 0$, $\alpha_1 = 0.45$. Параметры уравнения состояния

$$\gamma_1 = 2, \quad \pi_1 = 2 \text{ Па}, \quad \gamma_2 = 1.4, \quad \pi_2 = 0 \text{ Па}.$$

Число узлов сетки и шаг по времени задавались как $N = 500$ и $\Delta t = 10^{-6}$ с. Значение параметра релаксации $\mu = 10^6 \text{ кг}/(\text{м} \cdot \text{с})$, что соответствует времени релаксации меньшему, чем газодинамический шаг интегрирования по времени. Значения параметра релаксации $\nu = 1, 10$ и $100 \text{ Па}^{-1} \cdot \text{с}^{-1}$ (рис. 2–4). На графиках показаны только нулевые гармоники. Видим, что результаты численных расчётов хорошо согласуются с результатами HLLM для всех коэффициентов релаксации (см. рис. 2–4), а также что величина времени релаксации принципиально меняет структуру разрывов. Небольшая немонотонность результатов, представленных на рис. 4, по-видимому, связана с численным расчётом матрицы Якоби в методе Ньютона, а именно, с выбором шага численного дифференцирования.

Результаты расчётов при помощи конечно-объёмной схемы с потоком HLLM получены на сетке с $N = 10\,000$ шагами. Все остальные параметры не менялись.

Представленные результаты демонстрируют, что предложенный алгоритм обеспечивает устойчивый расчёт решения задачи при достаточно широком диапазоне изменения параметров релаксации, в том числе при жёсткой релаксации, и позволяет добиться сходного качества разрешения фронтов по сравнению с конечно-объёмной схемой первого порядка с потоками HLLM при условии одновременного использования: (а) значительно более грубых сеток; (б) наиболее простых численных потоков типа Русанова; (в) лимитирования консервативных переменных.

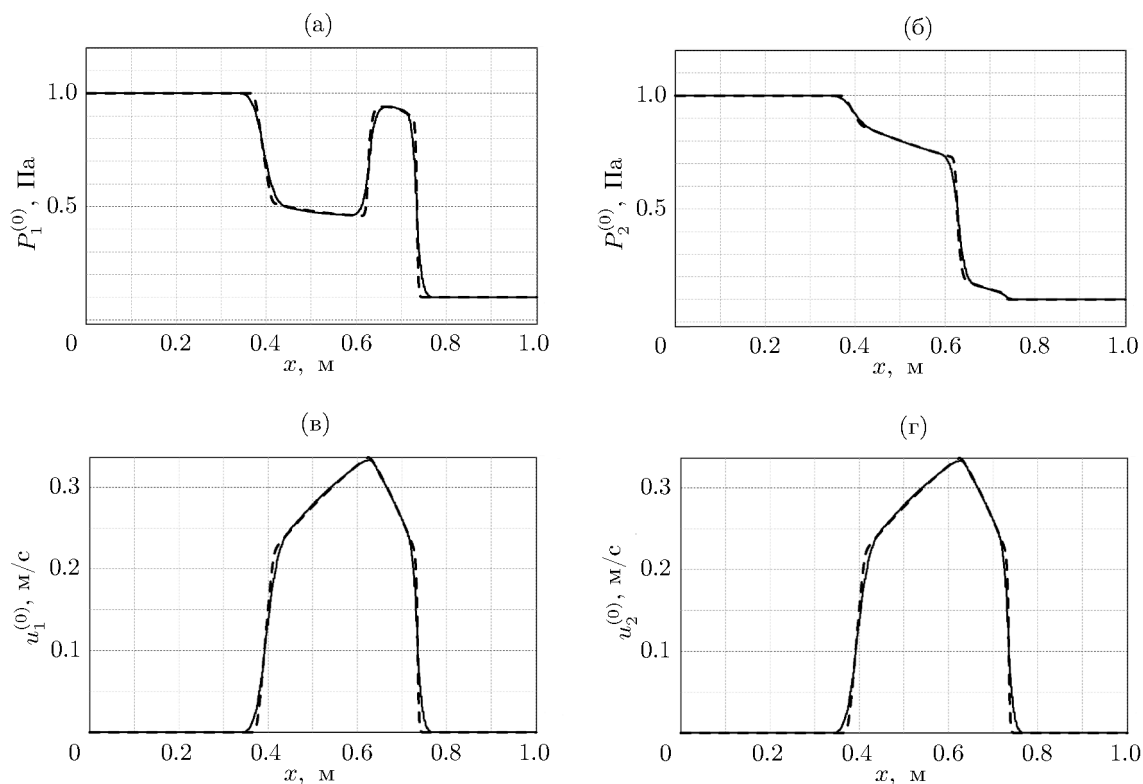


Рис. 2. Тест 2. Коэффициенты релаксации $\mu = 10^6$ кг/(м · с) и $\nu = 1$ Па⁻¹ · с⁻¹. Сплошные линии – результаты численного расчёта ($N = 500$), штриховые линии – результаты численного расчёта HLEM ($N = 10000$).

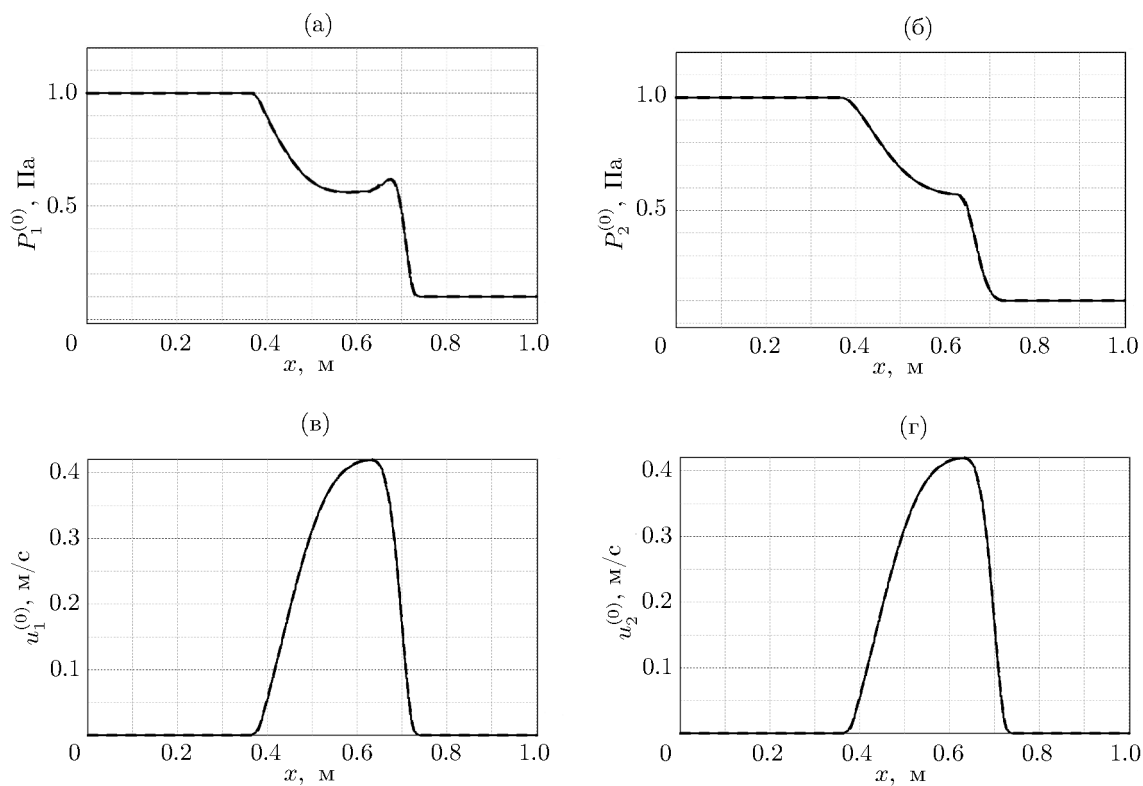


Рис. 3. Тест 2. Коэффициенты релаксации $\mu = 10^6$ кг/(м · с) и $\nu = 10$ Па⁻¹ · с⁻¹. Сплошные линии – результаты численного расчёта ($N = 500$), штриховые линии – результаты численного расчёта HLEM ($N = 10000$).

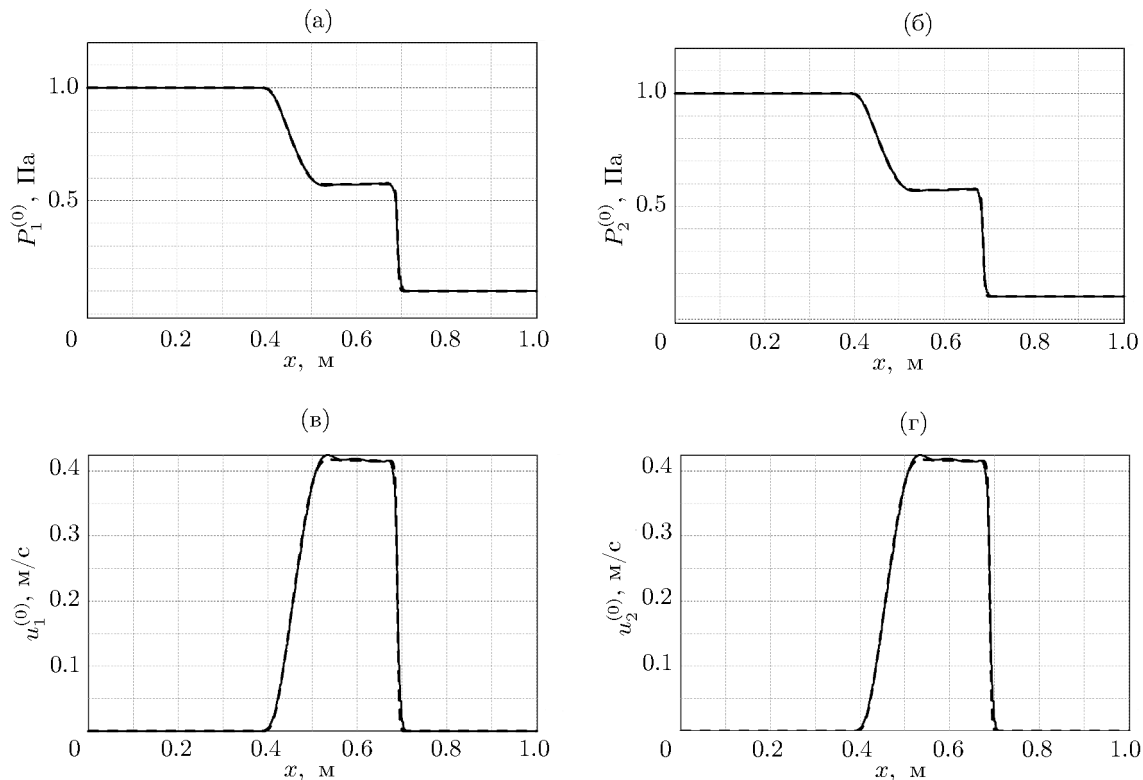


Рис. 4. Тест 2. Коэффициенты релаксации $\mu = 10^6 \text{ кг}/(\text{м} \cdot \text{с})$ и $\nu = 100 \text{ Па}^{-1} \cdot \text{с}^{-1}$. Сплошные линии – результаты численного расчёта ($N = 500$), штриховые линии – результаты численного расчёта HLLEM ($N = 10\,000$).

Заключение. В работе представлен подход к численному решению многофазной полностью неравновесной модели Баера–Нунциато с релаксационными слагаемыми в одномерном случае. В качестве численной схемы используется разрывный метод Галёркина с применением численного потока Русанова. Для монотонизации решения используется лимитер WENO-S, применяемый непосредственно к консервативным переменным модели. Учёт релаксационных слагаемых в численной схеме проводится неявным методом Рунге–Кутты второго порядка с адаптивным выбором шага интегрирования. Соответствующая нелинейная система уравнений решается методом Ньютона. Представленный комплексный подход применяется к одномерным тестовым задачам. Результаты соответствуют полученным ранее для численных схем с потоком HLLEM. В работе показано, что использование разрывного метода Галёркина позволяет добиться сходного разрешения волновых фронтов при значительно меньшем числе расчётных ячеек, чем в методах с использованием большей информации об исходной системе уравнений.

4. Приложения.

Приложение 1. Для уравнения релаксации скоростей (см. систему ОДУ (2.8)) можно выписать точное аналитическое решение задачи, которое имеет вид

$$\mathbf{W}(t) = \exp(\mathbf{A}t)\mathbf{W}(t_0), \quad t > t_0; \quad \tilde{Q}_{2,3}^{(l)}(t) = \tilde{Q}_{2,3}^{(l)}(t_0), \quad t = t_0,$$

где $l = 0, 1$, а вектор $\mathbf{W}$ и матрица $\mathbf{A}$ имеют следующий вид:

$$\mathbf{W} = \begin{pmatrix} \tilde{Q}_4^{(0)} \\ \tilde{Q}_4^{(1)} \\ \tilde{Q}_7^{(0)} \\ \tilde{Q}_7^{(1)} \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ b_{11} & b_{12} & b_{13} & b_{14} \\ -a_{11} & -a_{12} & -a_{13} & -a_{14} \\ -b_{11} & -b_{12} & -b_{13} & -b_{14} \end{pmatrix}.$$

Согласно (1.1) имеем $\tilde{Q}_2^{(0,1)} = (\alpha_1 \rho_1)^{(0,1)}$, $\tilde{Q}_3^{(0,1)} = (\alpha_2 \rho_2)^{(0,1)}$, $\tilde{Q}_4^{(0,1)} = (\alpha_1 \rho_1 u_1)^{(0,1)}$, $\tilde{Q}_7^{(0,1)} = (\alpha_2 \rho_2 u_2)^{(0,1)}$, где $u_{1,2}$ – скорость 1-й или 2-й фазы. Постоянные коэффициенты a_{ij} и b_{ij} имеют следующий вид:

$$\begin{aligned} a_{11} &= \frac{1}{2} \frac{\mu}{\tilde{Q}_2^{(1)}} \log \left(\frac{\tilde{Q}_2^{(0)} - \tilde{Q}_2^{(1)}}{\tilde{Q}_2^{(0)} + \tilde{Q}_2^{(1)}} \right), \quad a_{12} = -\frac{\mu}{\tilde{Q}_2^{(1)}} + \frac{1}{2} \frac{\mu \tilde{Q}_2^{(0)}}{(\tilde{Q}_2^{(1)})^2} \log \left(\frac{\tilde{Q}_2^{(0)} + \tilde{Q}_2^{(1)}}{\tilde{Q}_2^{(0)} - \tilde{Q}_2^{(1)}} \right), \\ a_{13} &= \frac{1}{2} \frac{\mu}{\tilde{Q}_3^{(1)}} \log \left(\frac{\tilde{Q}_3^{(0)} + \tilde{Q}_3^{(1)}}{\tilde{Q}_3^{(0)} - \tilde{Q}_3^{(1)}} \right), \quad a_{14} = \frac{\mu}{\tilde{Q}_3^{(1)}} + \frac{1}{2} \frac{\mu \tilde{Q}_3^{(0)}}{(\tilde{Q}_3^{(1)})^2} \log \left(\frac{\tilde{Q}_3^{(0)} - \tilde{Q}_3^{(1)}}{\tilde{Q}_3^{(0)} + \tilde{Q}_3^{(1)}} \right), \\ b_{11} &= -3 \frac{\mu}{\tilde{Q}_2^{(1)}} + \frac{3}{2} \frac{\mu \tilde{Q}_2^{(0)}}{(\tilde{Q}_2^{(1)})^2} \log \left(\frac{\tilde{Q}_2^{(0)} + \tilde{Q}_2^{(1)}}{\tilde{Q}_2^{(0)} - \tilde{Q}_2^{(1)}} \right), \quad b_{12} = 3 \frac{\mu \tilde{Q}_2^{(0)}}{(\tilde{Q}_2^{(1)})^2} + \frac{3}{2} \frac{\mu (\tilde{Q}_2^{(0)})^2}{(\tilde{Q}_2^{(1)})^3} \log \left(\frac{\tilde{Q}_2^{(0)} - \tilde{Q}_2^{(1)}}{\tilde{Q}_2^{(0)} + \tilde{Q}_2^{(1)}} \right), \\ b_{13} &= \frac{3\mu}{\tilde{Q}_3^{(1)}} + \frac{3}{2} \frac{\mu \tilde{Q}_3^{(0)}}{(\tilde{Q}_3^{(1)})^2} \log \left(\frac{\tilde{Q}_3^{(0)} - \tilde{Q}_3^{(1)}}{\tilde{Q}_3^{(0)} + \tilde{Q}_3^{(1)}} \right), \\ b_{14} &= -3 \frac{\mu \tilde{Q}_3^{(0)}}{(\tilde{Q}_3^{(1)})^2} - 4 \frac{\mu (\tilde{Q}_3^{(0)})^2}{(\tilde{Q}_3^{(1)})^3} \log \left(\frac{\tilde{Q}_3^{(0)} - \tilde{Q}_3^{(1)}}{\tilde{Q}_3^{(0)} + \tilde{Q}_3^{(1)}} \right). \end{aligned}$$

Приложение 2. В настоящем приложении показывается, что простейший поток Лакса–Фридрихса при его применении в рамках разрывного метода Галёркина с интегрированием по времени методом TVD/RK3 приводит к неустойчивой схеме.

Утверждение. Для линейного одномерного уравнения переноса

$$\frac{\partial q}{\partial t} + \frac{\partial f(q)}{\partial x} = 0; \quad f(q) = aq, \quad a = \text{const}, \quad (4.1)$$

с начальными условиями в виде кусочно-линейных функций

$$q_h(x, 0)|_{\omega_i} = Q_0 - \frac{2Q_1 x_i}{h} + \frac{2Q_1}{h} x, \quad Q_{0,1} = \text{const}, \quad (4.2)$$

заданными на каждом конечном элементе $\omega_i = [x_{i-1/2}, x_{i+1/2}]$, дискретное решение, полученное с использованием разрывного метода Галёркина с линейным восполнением решения в ячейках, интегрированием по времени с помощью схемы TVD/RK3, описанной в п. 2.3, и численного потока Лакса–Фридрихса вида

$$\hat{f}(q_l, q_r) = \frac{1}{2} [f(q_l) + f(q_r)] - \frac{h}{2\Delta t} (q_r - q_l), \quad (4.3)$$

удовлетворяет неравенству

$$|q_h^{n+1}| > |q_h^n|.$$

Доказательство. Считаем, что все ячейки сетки (конечные элементы) $\omega_i = [x_{i-1/2}, x_{i+1/2}]$ имеют одинаковую длину h . Каждая из них может быть отображена на канонический конечный элемент $\tilde{\omega}_i = [-1, +1]$. В канонических координатах базисные функции разрывного метода Галёркина имеют вид $\psi_0 = 0$, $\psi_1 = \xi$, а начальные условия – вид $q_h|_{\tilde{\omega}_i} = Q_0 + Q_1 \xi$.

В соответствии с п. 2.1 сначала получим полудискретное уравнение. Для этого умножим уравнение (4.1) на $\psi \in \text{span}(\psi_0, \psi_1)$ и проинтегрируем результат по $\tilde{\omega}_i$. Тогда получим

$$\frac{h}{2} \int_{-1}^1 \frac{\partial q}{\partial t} \psi(\xi) d\xi + \int_{-1}^1 \frac{\partial f(q)}{\partial \xi} \psi(\xi) d\xi = 0. \quad (4.4)$$

Полудискретное решение задачи в каждой ячейке сетки (конечном элементе) можно представить в виде

$$q_i := q_h(\xi, t)|_{\tilde{\omega}_i} = s_0^i(t) + s_1^i(t)\xi,$$

где $s_{0,1}^i = s_{0,1}^i(t)$ – гармоники решения в ячейке $\tilde{\omega}_i$.

Введём следующие обозначения:

$$q_i^- = q_h(-1, t)|_{\tilde{\omega}_i} = s_0^i - s_1^i, \quad q_i^+ = q_h(1, t)|_{\tilde{\omega}_i} = s_0^i + s_1^i.$$

После стандартных преобразований уравнения (4.4) получим

$$\frac{h}{2} \int_{-1}^1 \frac{\partial q_i}{\partial t} \psi(\xi) d\xi - \int_{-1}^1 f(q_i) \frac{\partial \psi(\xi)}{\partial \xi} d\xi + [\hat{f}(q_i) \psi|_{\xi=1} - \hat{f}(q_i) \psi|_{\xi=-1}] = 0,$$

где $\hat{f}$ – численный поток. Отсюда получаем следующие уравнения для определения гармоник решения в ячейке ω_i :

$$\frac{h}{2} \int_{-1}^1 \frac{d}{dt} (s_0^i(t) + s_1^i(t)\xi) d\xi + [\hat{f}(q_{i+1/2}) - \hat{f}(q_{i-1/2})] = 0,$$

$$\frac{h}{2} \int_{-1}^1 \frac{d}{dt} (s_0^i(t) + s_1^i(t)\xi)\xi d\xi + [\hat{f}(q_{i+1/2}) + \hat{f}(q_{i-1/2})] - \int_{-1}^1 f(q_i) d\xi = 0.$$

Тогда с учётом ортогональности полиномов ψ_0 и ψ_1 будем иметь

$$\frac{ds_0}{dt} = -\frac{1}{h}(\hat{f}(q_{i+1/2}) - \hat{f}(q_{i-1/2})),$$

$$\frac{ds_1}{dt} = -\frac{3}{h} \left(\hat{f}(q_{i+1/2}) + \hat{f}(q_{i-1/2}) - \int_{-1}^1 f(q_i) d\xi \right).$$

При использовании численного потока Лакса–Фридрихса (4.3) получаем

$$\hat{f}(q_{i+1/2}) = \frac{1}{2}[f(q_{i+1}^-) + f(q_i^+)] - \frac{h}{2\Delta t}(q_{i+1}^- - q_i^+) = \frac{a}{2}(q_{i+1}^- + q_i^+) - \frac{h}{2\Delta t}(q_{i+1}^- - q_i^+),$$

$$\hat{f}(q_{i-1/2}) = \frac{1}{2}[f(q_i^-) + f(q_{i-1}^+)] - \frac{h}{2\Delta t}(q_i^- - q_{i-1}^+) = \frac{a}{2}(q_i^- + q_{i-1}^+) - \frac{h}{2\Delta t}(q_i^- - q_{i-1}^+),$$

или, вводя число Куранта $\lambda = a\Delta t/h$,

$$\hat{f}(q_{i+1/2}) = \frac{h}{2\Delta t}[(\lambda - 1)q_{i+1}^- + (1 + \lambda)q_i^+], \quad \hat{f}(q_{i-1/2}) = \frac{h}{2\Delta t}[(\lambda - 1)q_i^- + (1 + \lambda)q_{i-1}^+].$$

Помимо этого в частном случае линейного уравнения переноса верно равенство

$$\int_{-1}^1 f(q_i) d\xi = a \int_{-1}^1 (s_0^i + s_1^i \xi) d\xi = 2as_0^i = \frac{2\lambda h s_0^i}{\Delta t}.$$

В результате система обыкновенных дифференциальных уравнений для определения гармоник решения в конечном элементе ω_i примет вид

$$\frac{ds_0^i}{dt} = \frac{1}{\Delta t}(-s_0^i - \lambda s_1^i) + \frac{1}{2\Delta t}[(1 + \lambda)q_{i-1}^+ + (1 - \lambda)q_{i+1}^-],$$

$$\frac{ds_1^i}{dt} = \frac{3}{\Delta t}(\lambda s_0^i - s_1^i) + \frac{3}{2\Delta t}[(1 - \lambda)q_{i+1}^- - (1 + \lambda)q_{i-1}^+].$$

Подставляя в последние уравнения выражения для численных потоков q_{i-1}^+ и q_{i+1}^- , получаем

$$\frac{ds_0^i}{dt} = \frac{1}{\Delta t}(-s_0^i - \lambda s_1^i) + \frac{1}{2\Delta t}[(1 + \lambda)(s_0^{i-1} + s_1^{i-1}) + (1 - \lambda)(s_0^{i+1} - s_1^{i+1})],$$

$$\frac{ds_1^i}{dt} = \frac{3}{\Delta t}(\lambda s_0^i - s_1^i) + \frac{3}{2\Delta t}[(1 - \lambda)(s_0^{i+1} - s_1^{i+1}) - (1 + \lambda)(s_0^{i-1} + s_1^{i-1})],$$

или

$$\frac{ds_0^i}{dt} = G_0(q_h), \quad \frac{ds_1^i}{dt} = G_1(q_h).$$

Учитывая начальные условия (4.2), находим, что

$$G_0(q_h) = 0, \quad G_1(q_h) = -\frac{6Q_1}{\Delta t}.$$

Непосредственной проверкой несложно убедиться в том, что для данной системы ОДУ применение аппроксимаций по времени вида TVD/RK3 приводит к следующим соотношениям:

$$s_1^{i,(1)} = Q_1 - 6Q_1 = -5Q_1,$$

$$s_1^{i,(2)} = \frac{3}{4}Q_1 + \frac{1}{4}(-5Q_1) + \frac{1}{4}(-6)(-5Q_1) = 7Q_1,$$

$$s_1^{i,new} = \frac{1}{3}Q_1 + \frac{2}{3}(7Q_1) + \frac{2}{3} - 6(7Q_1) = -23Q_1.$$

Поэтому значения решения на временном шаге $n + 1$ имеют вид

$$(s_0^i)^{n+1} = Q_0 \quad (s_1^i)^{n+1} = (-23)^n Q_1.$$

Таким образом, $|q_h^{n+1}| > |q_h^n|$, что доказывает утверждение.

Обратим внимание, что в рассмотренном выше случае алгоритм корректно пересчитывает нулевые гармоники дискретного решения в ячейках (которые соответствуют среднему значению решения в ячейках). При этом первые гармоники решения неограниченно возрастают, как только в начальном условии существуют соответствующие сколь угодно малые, но конечные возмущения. Одновременно с этим применение потока Лакса–Фридрихса для метода конечных объёмов первого порядка даёт устойчивое корректное решение.

Описанные теоретические построения подтверждаются результатами расчётов авторов (см. также работу [22]).

Исследование Р.Р. Тухватуллиной (введение, пп. 2.2, п. 3, приложение 1) выполнено при финансовой поддержке Российского научного фонда (проект 19-71-30004). Исследование М.В. Алексеева и Е.Б. Савенкова (введение, п. 1, пп. 2.1, 2.3, приложение 2) выполнено при поддержке Московского центра фундаментальной и прикладной математики (соглашение с Министерством науки и высшего образования Российской Федерации № 075-15-2019-1623).

СПИСОК ЛИТЕРАТУРЫ

1. Baer M., Nunziato J. A two-phase mixture theory for the deflagration-to-detonation transition (DDT) in reactive granular materials // Int. J. Multiph. Flow. 1986. № 12. P. 861–889.
2. Drew D., Passman S. Theory of Multicomponent Fluids. Springer, 2014.
3. Favrie N., Gavriluk S., Saurel R. Solid-fluid diffuse interface model in cases of extreme deformations // J. Comput. Phys. 2009. V. 228. № 16. P. 6037–6077.
4. Kapila A., Son S., Bdzil J., Menikoff R. Two-phase modeling of DDT: structure of the velocity-relaxation zone // Phys. Fluids. 1997. V. 9. № 12. P. 3885–3897.
5. Kapila A., Menikoff R., Bdzil J., Son S., Stewart S. Two-phase modeling of deflagration-to-detonation transition in granular materials: Reduced equations // Phys. Fluids. 2001. V. 13. № 10. P. 3002–3024.

6. *Murrone A., Guillard H.* A five-equation reduced model for compressible two phase flow problems // J. Comput. Phys. 2005. V. 202. № 2. P. 664–698.
7. *Dal Maso G., Le Floch P., Murat F.* Definition and weak stability of nonconservative products // J. Math. Pures Appl. 1995. V. 74. № 6. P. 483–548.
8. *Tokareva S., Toro E.* HLLC-type Riemann solver for the Baer–Nunziato equations of compressible two-phase flow // J. Comput. Phys. 2010. V. 229. № 10. P. 3573–3604.
9. *Dumbser M., Toro E.* A simple extension of the Osher Riemann solver to non-conservative hyperbolic systems // J. Sci. Comput. 2011. V. 48. P. 70–88.
10. *Franquet E., Perrier V.* Runge–Kutta discontinuous Galerkin method for the approximation of Baer and Nunziato type multiphase models // J. Comput. Phys. 2012. V. 291. P. 4096–4141.
11. *de Frahan H., Varadan S., Johnsen E.* A new limiting procedure for discontinuous Galerkin methods applied to compressible multiphase flows with shocks and interfaces // J. Comput. Phys. 2015. V. 280. P. 489–509.
12. *Cockburn B., Shu C.-W.* The Runge–Kutta local projection-discontinuous Galerkin finite element method for scalar conservation laws // ESAIM Math. Model. Numer. Anal. 1991. V. 25. № 3. P. 337–361.
13. *Zhong X., Shu C.-W.* A simple weighted essentially nonoscillatory limiter for Runge–Kutta discontinuous Galerkin methods // J. Comput. Phys. 2013. V. 232. № 1. P. 397–415.
14. *Jin S., Xin Z.* The relaxation schemes for systems of conservation laws in arbitrary space dimensions // Comm. Pure Appl. Math. 1995. V. 48. № 3. P. 0010–3640.
15. *Jin S.* Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations // SIAM J. Sci. Comput. 1999. V. 21. P. 441–454.
16. *Andrianov N., Warnecke G.* The Riemann problem for the Baer–Nunziato two-phase flow model // J. Comput. Phys. 2004. V. 195. № 2. P. 434–464.
17. *Daude F., Berry R., Galon P.* A Finite-volume method for compressible non-equilibrium two-phase flows in networks of elastic pipelines using the Baer–Nunziato model // Comput. Methods in Appl. Mech. Eng. 2019. V. 354. P. 820–849.
18. *Saurel R., Abgrall R.* A simple method for compressible multifluid flows // SIAM J. Sci. Comput. 1999. V. 21. № 3. P. 1115–1145.
19. *Nigmatulin R.* Dynamics of Multiphase Media. N.Y., 1990.
20. *Pares C.* Numerical methods for nonconservative hyperbolic systems: a theoretical framework // SIAM J. Numer. Anal. 2006. V. 44. № 1. P. 300–321.
21. *Chiocchetti S., Muller C.* A solver for stiff finite-rate relaxation in Baer–Nunziato two-phase flow model // Droplet Interactions and Spray Processes. 2020. P. 31–44.
22. *Rider W., Lowrie R.* The use of classical Lax–Friedrichs Riemann solvers with discontinuous Galerkin methods // Int. J. Numer. Methods Fluids. 2002. V. 40. P. 479–486.

Институт прикладной математики
им. М.В. Келдыша, г. Москва

Поступила в редакцию 01.03.2021 г.
После доработки 01.03.2021 г.
Принята к публикации 27.04.2021 г.