

Том 64, Номер 8

ISSN 0044-4669
Август 2024



ФИЦ ИУ РАН

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ



НАУКА
— 1727 —

СОДЕРЖАНИЕ

Том 64, номер 8, 2024

Сергею Константиновичу Годунову посвящается

К.В. Брушлинский, Г.В. Демиденко, В.Т. Жуков

1337

ОБЩИЕ ЧИСЛЕННЫЕ МЕТОДЫ

Аппроксимация и сглаживание функции на основе регуляризации по Годунову

Э.А. Бибердорф, К.К. Абдишерипов

1342

Метод разделения матричного спектра относительно прямой и задача о флаттере бесконечной полосы

Э.А. Бибердорф, А.С. Рудометова, Ван Ли, А.Д. Жумабаев

1355

Эффективное устойчивое интегрирование по времени уравнений Кана-Хилларда: явные, неявные и явно-итерационные схемы

М.А. Бочев, И.А. Фахурдинов, Е.Б. Савенков

1366

Оценки p -норм решений и обратных матриц систем линейных уравнений с циркулянтной матрицей

Ю.С. Волков, В.В. Богданов

1388

Численное решение уравнений ионизации методом дробных шагов

М.Б. Гавриков, А.А. Таюрский

1398

Оценки решений одной биологической модели с бесконечным распределенным запаздыванием

Т.К. Искаков, М.А. Скворцова

1409

Оценки сходимости итерационных методов численного моделирования трехмерных процессов в магнитной гидродинамике

А.Ю. Круковский, И.В. Попов, Ю.А. Повещенко

1424

Об одном подходе к повышению точности и эффективности метода РОМБ для решения нестационарного уравнения теплопроводности

Н.Я. Моисеев

1437

УРАВНЕНИЯ В ЧАСТНЫХ ПРОИЗВОДНЫХ

Об условиях разрешимости задачи Коши для одной псевдогиперболической системы

Л.Н. Бондарь, С.Б. Мингноров

1445

Сферические сплайн-решения неоднородного бигармонического уравнения

В.Л. Васкевич

1456

Энергетические оценки для одного класса псевдогиперболических операторов с переменными коэффициентами

Г.В. Демиденко

1466

Структурированные псевдоспектры в задачах пространственной устойчивости пограничных слоев <i>К.В. Демьянко, Г.В. Засько, Ю.М. Нечепуренко</i>	1476
Оценки решений класса неавтономных систем нейтрального типа с сосредоточенным и распределенным запаздываниями <i>И.И. Матвеева</i>	1486

МАТЕМАТИЧЕСКАЯ ФИЗИКА

Модель градиента плотности в сферически-симметричной постановке и ее явно-неявная диссипативная дискретизация для исследования динамики межфазной границы <i>В.А. Балашов, Е.А. Павлишина, Е.Б. Савенков</i>	1500
Схема расчета нестационарных течений теплопроводного газа в трехтемпературном приближении <i>В.Т. Жуков, О.Б. Феодоритова</i>	1517
Расчет нестационарных закрученных течений в соплах и трубах с помощью новой локально неявной схемы годуновского типа <i>Н.А. Зайцев, Ю.Г. Рыков</i>	1529
Метод свободной границы для сопряженных задач динамики газа и твердого тела <i>И.С. Меньшов</i>	1546

ИНФОРМАТИКА

О спектральных портретах матриц инцидентности графов ближайших соседей <i>А.А. Кислицын</i>	1561
--	------

СЕРГЕЮ КОНСТАНТИНОВИЧУ ГОДУНОВУ ПОСВЯЩАЕТСЯ

© 2024 г. К.В. Брушлинский^{1,*}, Г.В. Демиденко^{2,3,**}, В.Т. Жуков^{1,***}

¹ 125047 Москва, Миусская пл., 4, Институт прикладной математики им. М.В. Келдыша РАН

² 630090 Новосибирск, пр. Академика Коптюга, 4, Институт математики им. С.Л. Соболева СО РАН

³ 630090 Новосибирск, ул. Пирогова, 1, Новосибирский государственный университет

*e-mail: brush@keldysh.ru

**e-mail: demidenk@math.nsc.ru

***e-mail: zhukov@kiam.ru

Поступила в редакцию 20.05.2024 г.

Сергей Константинович Годунов родился 17 июля 1929 г. в Москве. В 1946 г. после окончания спецшколы ВВС он поступил на мехмат Московского государственного университета, где прошел хорошую школу и получил прочный математический фундамент. Его наставниками были выдающиеся советские ученые и педагоги Б.Н. Делоне и И.Г. Петровский. В 1951 г. после окончания МГУ он был принят в Расчетное бюро Математического института им. В.А. Стеклова АН СССР. С первых же дней работы в Расчетном бюро, а затем в Отделении прикладной математики МИАН, С.К. Годунов принимает активное участие в работе большого коллектива ученых, созданного для решения важнейших практических задач по математическому моделированию и проведению расчетов различных процессов в области ядерной физики. С этого времени началось его беззаветное служение науке, которое продолжалось более 70 лет.



Фиг. 1. Курсант Московской спецшколы ВВС — советник РАН Института математики им. С.Л. Соболева

Выдающийся ученый академик Сергей Константинович Годунов являлся одним из наиболее ярких лидеров современной прикладной математики и механики. Его фундаментальный вклад в развитие вычислительной математики, численных методов механики сплошных сред, теории дифференциальных уравнений трудно переоценить. Его научные труды в значительной степени определили современный облик численного анализа, заложили новые перспективные направления его развития. С.К. Годунов — автор более 300 научных работ, в том числе 18 монографий, многие из которых переведены на другие языки и стали классическими. Описание полученных им результатов содержится в обзоре [1]. Приведем лишь несколько результатов, которые принесли С.К. Годунову всемирную известность.

Одним из первых наиболее значимых результатов Сергея Константиновича является «разностная схема Годунова». Схема была разработана им для расчетов разрывных решений задач одномерной газовой динамики.

Эта работа была инициирована М.В. Келдышем и И.М. Гельфандом, предложивших осенью 1953 г. молодому учёному Сергею Годунову создать вариант метода Дж. Неймана и Р. Рихтмайера для расчета уравнений газовой динамики с использованием искусственной вязкости. Однако в 1954 г. Сергей Годунов предложил свой метод расчёта, и с течением времени его схема приобрела всемирную известность и была названа его именем. В основу разработки схемы было положено требование сохранения монотонности численных решений при переходе с каждого расчетного временного слоя на следующий, реализованное с помощью точного решения задачи Римана о распаде разрыва. При работе над этой задачей проявились яркий талант С.К. Годунова, нестандартность его мышления, колоссальная работоспособность и стремление докопаться до истины. Его результаты были опубликованы в 1959 г. в статье [2]. Многолетний опыт применения схемы Годунова обобщен в монографии [3], вышедшей в 1976 г. Но особенно большой интерес к схеме Годунова проявился в 80-е годы прошлого века, когда возникла необходимость разработки методов высокой точности для гиперболических систем. На основе метода, предложенного С.К. Годуновым, было разработано множество различных модификаций, которые используются в разнообразных приложениях от медицины до астрофизики.

Схема Годунова оказала существенное влияние на последующее развитие численных методов. В настоящее время имеется большое число приближенных методов для расчета разрывных решений нелинейных уравнений, в частности, многие разностные схемы основаны на идеях С.К. Годунова. В связи с этим большой интерес представляют комментарии и пояснения самого Сергея Константиновича по схеме Годунова [4]. Огромное влияние на развитие современной вычислительной математики оказала также монография [5], написанная С.К. Годуновым в соавторстве с В.С. Рябеньким. Эта монография была первым отечественным изложением вычислительной математики 50–60 гг. Она стала учебником и настольной книгой для многих специалистов. В ней излагались важные теоретические вопросы и конкретные разностные методы на модельных примерах. Впоследствии эта монография неоднократно переиздавалась с дополнениями.

В своей работе Сергей Константинович всегда руководствовался принципом «идти от приложений». За время работы в Отделении прикладной математики (с 1966 г. — Институт прикладной математики АН СССР, а с 1978 года — ИПМ им. М.В. Келдыша) им было создано множество методов для приближенного решения различных конкретных задач, заложены основы новых направлений прикладной математики и механики. В частности, им был создан метод установления для анализа обтекания тел потоком газа. В дальнейшем этот метод стал универсальным приемом для приближенного построения стационарных решений с помощью нестационарной задачи. Под руководством С.К. Годунова была начата разработка первых алгоритмов для построения криволинейных разностных сеток. Эти алгоритмы используются для расчета разрывных решений двумерных задач механики сплошной среды. В настоящее время разработкой алгоритмов для построения разностных сеток в многомерных областях занимаются в различных научных организациях в нашей стране и за рубежом. С.К. Годуновым разработан метод ортогональной прогонки для решения краевых задач для систем обыкновенных дифференциальных уравнений. Этот метод применяется при решении большого числа прикладных задач, в частности, он использовался при расчетах атомных реакторов.

Решая различные прикладные задачи механики сплошных сред, С.К. Годунов получил важные теоретические результаты для квазилинейных дифференциальных уравнений математической физики. В частности, он проводил исследования проблемы единственности решений, проблемы построения обобщенных решений. В своих работах (см., например, [6, 7]) он впервые поставил вопросы о глубокой связи законов термодинамики и корректности систем дифференциальных уравнений для моделей механики сплошных сред, об обобщении понятия энтропии и закона ее возрастания, о термодинамических соотношениях. Полученные в этой области результаты составили содержание его докторской диссертации (1965 г.). Эти работы С.К. Годунова послужили началом нового направления теоретических и прикладных исследований, посвященных гиперболическим системам законов сохранения в математической физике. Эти исследования активно проводятся в нашей стране и за рубежом.

После переезда С.К. Годунова в Новосибирск, куда его пригласил М.А. Лаврентьев, он был назначен заведующим лабораторией в Вычислительном центре СО АН СССР и продолжил исследования связей фундаментальных вопросов математической физики и теории дифференциальных уравнений. Эти исследования были связаны с работами, проводимыми в лаборатории А.А. Дерибаса в Институте гидродинамики СО АН СССР по сварке металлов взрывом. Проводимые при этом численные расчеты не давали полного объяснения процессов, происходящих в металлах, в частности, образования волн на границе между свариваемыми пластинами. Это приводило С.К. Годунова к мысли о необходимости уточнения классических моделей, используемых для описания высокоэнергетических и высокоскоростных воздействий на металлы. Дело в том, что при больших высокоскоростных воздействиях металлы претерпевают упругие и пластические деформации, при этом могут вести себя, как жидкости. А классические модели это не учитывают! Изучая результаты многочисленных экспериментов, Сергею Константиновичу совместно с учениками и коллегами из Института гидродинамики удалось получить уравнения, которые адекватно описывали как упругую деформацию среды, так и ее течение

при интенсивных воздействиях (нелинейная релаксационная модель Максвелла). При этом полученная модель позволяла проводить расчеты реальных ударно-волновых процессов в металлах. Однако система дифференциальных уравнений нелинейной модели Максвелла не входила в класс квазилинейных систем, введенных С.К. Годуновым в работе [7], гиперболичность которых обеспечивалась выполнением законов термодинамики. Это стало отправной точкой для многолетних исследований, которые привели к появлению класса симметрических гиперболических термодинамически согласованных систем законов сохранения [8, 9]. В настоящее время теория таких систем широко используется как для создания новых моделей сложных сплошных сред, так и для разработки новых численных методов [10].

Наряду с изучением гиперболических систем квазилинейных уравнений, возникающих в задачах сплошных сред, С.К. Годунов вместе со своими учениками проводил исследования линейных гиперболических уравнений и систем. В середине 70-х годов в Новосибирском государственном университете начал работать семинар по гиперболическим уравнениям под руководством С.К. Годунова. На нем, в частности, был поставлен вопрос: всегда ли уравнение высокого порядка, гиперболическое по Петровскому, может быть сведено к симметрической системе первого порядка, гиперболической по Фридрихсу? В одномерном случае, а также для уравнений второго порядка, такое сведение было известно. Однако сведение гиперболического уравнения с двумя пространственными переменными к симметрической гиперболической системе оказалось очень сложной задачей. Ее решение было получено С.К. Годуновым и В.И. Костиным в 1979 г. Затем В.И. Костиным и Т.Ю. Михайловой было установлено аналогичное сведение гиперболических уравнений, инвариантных относительно вращений. Но ответ на вопрос для произвольного гиперболического уравнения оказался отрицательным. В 1983 г. В.В. Иванов построил неожиданный пример гиперболического уравнения четвертого порядка с четырьмя пространственными переменными, которое не сводилось к симметрической системе. Однако вопрос о сведении произвольных гиперболических уравнений с тремя пространственными переменными к гиперболическим системам по Фридрихсу до сих пор остается открытым.

На семинаре по гиперболическим уравнениям рассматривались также вопросы о корректных постановках краевых задач. С.К. Годуновым была поставлена задача изучения корректности общих смешанных задач для гиперболических уравнений на основе «метода симметризации», т.е. метода, позволяющего сводить краевую задачу для гиперболического уравнения к краевым задачам для систем, гиперболических по Фридрихсу, с диссипативными граничными условиями. В этом направлении был получен ряд новых результатов, в частности, для краевых задач для векторного волнового уравнения были установлены существенно более сильные результаты по сравнению с результатами, полученными японскими математиками.

Осознавая исключительную важность применения аппарата линейной алгебры к задачам уравнений математической физики и механики сплошных сред, Сергей Константинович уделял большое внимание правильным постановкам задач вычислительной линейной алгебры. Он внёс огромный вклад в создание надежных алгоритмов их решения на компьютере. Интерес к этим вопросам у С.К. Годунова возник еще в 50-е годы при решении конкретных прикладных задач, связанных с ядерной физикой. Он был одним из первых математиков, кто стал давать правильные объяснения вычислительным парадоксам, подчеркивая, что при выполнении компьютерных расчетов, необходимо учитывать не только погрешности входных данных, но и ошибки округления, неизбежно возникающие в процессе вычислений.

В своей научной деятельности С.К. Годунов всегда мыслил независимо от устоявшихся парадигм. Широкий математический кругозор, научная смелость и необыкновенная интуиция позволяли ему «чувствовать задачу», отделять главное от второстепенных деталей. Отчетливо осознавая, что классический аппарат линейной алгебры не был ориентирован на компьютерные вычисления, Сергей Константинович ввёл фундаментальное понятие «гарантированная точность» вычислений на компьютере, им были определены новые понятия: ε -спектр, спектральный портрет, расслоение спектра, критерий качества дихотомии и др. [11–13]. Большая часть исследований по вычислительной линейной алгебре была выполнена в Институте математики Сибирского отделения, куда в 1980 г. по приглашению С.Л. Соболева перешел С.К. Годунов вместе со своей лабораторией. Именно в Институте математики в рамках новой концепции гарантированной точности С.К. Годуновым и его учениками был детально изучен ряд спектральных задач для конечномерных операторов, разработаны новые алгоритмы решения систем линейных алгебраических уравнений, решения задач устойчивости и дихотомии для обыкновенных дифференциальных уравнений, детально исследован метод ортогональной прогонки, получивший широчайшее распространение. Эти исследования также нашли многочисленных последователей и используются при создании методов и алгоритмов для суперкомпьютерных вычислений.

Выдающийся ученый С.К. Годунов являлся прекрасным педагогом. Он преподавал в Московском и Новосибирском государственных университетах. В 1977–1990 гг. являлся заведующим кафедрой дифференциальных уравнений в НГУ, основателем которой был академик С.Л. Соболев. Им были прочитаны оригинальные лекционные курсы “Уравнения математической физики”, “Механика сплошной среды”, “Дифференциальные уравнения”, “Современные аспекты линейной алгебры”. Широко известны его научно-образовательные семинары

“Гиперболические уравнения” (НГУ) и “Математика в приложениях” (ИМ СО РАН), которыми он руководил на протяжении многих лет. За время работы в Новосибирском государственном университете, в Вычислительном центре СО АН СССР, а затем в Институте математики, носящем сейчас имя С.Л. Соболева, Сергей Константинович воспитал немало учеников, которые и сейчас продолжают активно работать в науке.

Научные результаты С.К. Годунова получили высокую оценку со стороны государства и мирового научного сообщества. За выполнение специальных заданий Правительства и решение важных задач новой оборонной техники он был удостоен звания Лауреата Ленинской премии (1959 г.). С.К. Годунов награжден многими орденами: «Знак Почёта» (1954, 1981), «Трудового Красного Знамени» (1956, 1975), Почёта (2010), Александра Невского (2023). За цикл работ по исследованию процессов, сопутствующих сварке металлов взрывом, С.К. Годунову с соавторами присуждена премия им. А.Н. Крылова АН СССР (1972 г.), за книгу “Элементы механики сплошной среды” — премия им. М.А. Лаврентьева РАН (1993 г.). В 2022 г. «за выдающиеся результаты в области вычислительной математики» С.К. Годунов был награжден Золотой медалью им. Л. Эйлера РАН. Свидетельством мирового признания научных заслуг С.К. Годунова являются проходящие за рубежом международные конференции, посвященные его методам. В 1997 г. ему была присуждена степень почетного доктора Мичиганского университета США. В 2020 г. С.К. Годунову была присуждена премия им. Рамеша Агарвала в области вычислительной гидродинамики (SAE/Ramesh Agarwal Computational Fluid Dynamics Award).

* * *

Сергей Константинович Годунов был беззаветно предан науке. Он всегда был полон сил и творческих замыслов. Его идейная убежденность, высокая культура, широкая эрудиция всегда притягивала к нему коллег, учеников и последователей. Он был из тех людей, общение с которыми духовно обогащает. Светлая память о Сергее Константиновиче навсегда сохранится в сердцах его друзей, учеников, коллег и всех, кто его знал.

СПИСОК ЛИТЕРАТУРЫ

1. *Белых В.Н., Брушлинский К.В., Васкевич В.Л., Киселёв С.П., Крайко А.Н., Куликовский А.Г., Мали В.И., Пухначёв В.В., Роменский Е.И., Рябенский В.С.* Сергею Константиновичу Годунову 85 лет // Успехи мат. наук, т. 70, вып. 3, 2015.
2. *Годунов С.К.* Разностный метод численного расчета разрывных решений уравнений гидродинамики // Математический сборник, т. 47, № 3, 1959.
3. *Годунов С.К., Забродин А.В., Иванов М.Я., Крайко А.Н., Прокопов Г.П.* Численное решение многомерных задач газовой динамики. М.: Наука, 1976.
4. *Годунов С.К.* Воспоминания о разностных схемах. Новосибирск: Научная книга, 1997.
5. *Годунов С.К., Рябенский В.С.* Введение в теорию разностных схем. М.: Физматгиз, 1962.
6. *Годунов С.К.* Термодинамика газов и дифференциальные уравнения // Успехи мат. наук, т. 14, вып. 5, 1959.
7. *Годунов С.К.* Интересный класс квазилинейных систем // Докл. АН СССР, т. 139, № 3, 1961.
8. *Годунов С.К.* Элементы механики сплошной среды. М.: Наука, 1978.
9. *Годунов С.К., Роменский Е.И.* Элементы механики сплошных сред и законы сохранения. Новосибирск: Научная книга, 1998.
10. *Годунов С.К., Киселев С.П., Куликов И.М., Мали В.И.* Моделирование ударно-волновых процессов в упругопластических материалах на различных (атомный, мезо и термодинамический) структурных уровнях. Москва; Ижевск: Институт компьютерных исследований, 2014.
11. *Годунов С.К., Антонов А.Г., Кирилюк О.П., Костин В.И.* Гарантированная точность решения систем линейных уравнений в евклидовых пространствах. Новосибирск: Наука, Сиб. Отд-ние, 1988.
12. *Годунов С.К.* Современные аспекты линейной алгебры. Новосибирск: Научная книга, 1997.
13. *Годунов С.К.* Лекции по современным аспектам линейной алгебры. Новосибирск: Научная книга, 2002.

DEDICATED TO SERGEI KONSTANTINOVICH GODUNOV

K.V. Brushlinsky^{a,*}, G.V. Demidenko^{b,c,**}, V.T. Zhukov^{a,***}^a*Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences, Miusskaya sq., 4, Moscow, 125047 Russia*^b*Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Acad. Koptug Ave., 4, Novosibirsk, 630090 Russia,*^c*Novosibirsk State University, Pirogov St., 1, Novosibirsk, 630090 Russia*^{*}*e-mail: brush@keldysh.ru*^{**}*e-mail: demidenk@math.nsc.ru*^{***}*e-mail: zhukov@kiam.ru*

Received 20 May, 2024

Abstract. Sergei Konstantinovich Godunov was born on July 17, 1929, in Moscow. In 1946, after completing an Air Force special school, he enrolled in the Mechanics and Mathematics Faculty at Moscow State University, where he received a solid mathematical foundation and excellent training. His mentors included prominent Soviet scientists and educators B.N. Delone and I.G. Petrovsky. In 1951, after graduating from Moscow State University, he joined the Computational Bureau of the Steklov Mathematical Institute of the USSR Academy of Sciences. From his very first days working in the Computational Bureau, and later in the Department of Applied Mathematics at MIAS, S.K. Godunov became an active member of a large group of scientists dedicated to addressing critical practical tasks in mathematical modeling and calculations across various nuclear physics processes. This marked the beginning of his unwavering dedication to science, which continued for more than 70 years.

АППРОКСИМАЦИЯ И СГЛАЖИВАНИЕ ФУНКЦИИ НА ОСНОВЕ РЕГУЛЯРИЗАЦИИ ПО ГОДУНОВУ¹⁾

© 2024 г. Э. А. Бибердорф^{1,*}, К. К. Абдишерипов²

¹630090 Новосибирск, пр. Академика Коптюга, 4, Институт математики им. С.Л. Соболева СО РАН, Россия

²630090 Новосибирск, ул. Пирогова, 1, Новосибирский Государственный Университет, Новосибирск, Россия

*e-mail: math@biberdorf.ru

Поступила в редакцию 02.04.2024 г.

Переработанный вариант 02.04.2024 г.

Принята к публикации 02.05.2024 г.

Представлен новый подход к аппроксимации функций, основанный на идеях С.К. Годунова о регуляризации плохо обусловленных систем. Представленный метод позволяет определить значения функций в узлах мелкой сетки, исходя из данных на более крупной сетке, обеспечивая при этом контроль над гладкостью получаемой функции. Обоснованы оценки сходимости и гладкости, а также приведены результаты вычислительных экспериментов, иллюстрирующие эффективность предложенного метода. Библ. 4. Фиг. 6.

Ключевые слова: аппроксимация, сглаживание функций, регуляризация плохо обусловленных СЛАУ.

DOI: 10.31857/S0044466924080017, **EDN:** YBJDNW

ВВЕДЕНИЕ

Благодаря работам по разностным схемам имя Сергея Константиновича Годунова стало известно математикам и физикам всего мира. Несколько менее известны его статьи и книги по вычислительной линейной алгебре, однако именно в них наиболее отчетливо проявляются особенности его научного подхода к решению математических проблем. Обратимся к методу регуляризации по Годунову плохообусловленных систем линейных алгебраических уравнений (СЛАУ). Его возникновение связано с решением СЛАУ, представляющей собой дискретизированное интегральное уравнение Фредгольма первого рода. Известно, что такие СЛАУ, как правило, плохо обусловлены. С.К. Годунов отметил, что не все известные свойства искомого решения, которые присутствуют в исходной физической постановке задачи, нашли отражение в ее математической постановке. В частности, это касалось гладкости решения. Таким образом, получалось, что добавление в СЛАУ уравнений разделенных разностей делало более точным соответствие между физической и математической задачами. Оказалось, что при таком расширении системы значительно уменьшается число обусловленности, а специальный выбор параметра «склейки» двух подсистем приводил к тому, что регуляризация не увеличивает уровень погрешности. Описание этого метода можно найти в монографиях [1–3].

Позднее уже в 2000-х годах в личных беседах С.К. Годунов высказывал мысль, что аналогичный подход мог бы применяться для сглаживания функций, заданных на дискретной сетке. В настоящей работе, развивая эту идею, мы представляем новый метод обработки данных, который сочетает в себе свойства аппроксимации и сглаживания.

Аппроксимация используется во многих прикладных задачах и состоит, как правило, из двух этапов. Первый этап — вычисление коэффициентов аппроксимирующей функции по заданным значениям исходной функции, второй — вычисление значений аппроксимирующей функции в интересующих точках. Основные сложности численной реализации методов аппроксимации связаны с вычислительной трудоемкостью и/или вычислительной неустойчивостью. Особенно очевидны эти проблемы при использовании аппроксимации степенными функциями. Первый этап в этом случае состоит из решения линейной алгебраической системы, которая с ростом степени аппроксимирующего полинома стремительно становится плохо обусловленной.

В настоящей работе предлагается метод аппроксимации, обладающий двумя основными отличительными чертами. Первая из них заключается в том, что метод позволяет вычислять значения аппроксимирующей функции сразу же в процессе решения специальной СЛАУ. Таким образом, два этапа аппроксимации сливаются в

¹⁾ Работа выполнена в рамках государственного задания Института математики им. С.Л. Соболева СО РАН (проект № FWNF-2022-0008), фонд «El-yurt umidi», Узбекистан.

Очевидно, что эта система является недоопределенной и не может быть однозначно решена. Для ее корректной разрешимости необходимо добавить соотношения на производные:

$$g^{(p+1)}(\xi_k) = 0, \quad k = 0, 1, \dots, nm, \quad (5)$$

или разделенные разности порядка $p + 1$:

$$D_{p+1}g(\xi_k) = 0, \quad k = 0, 1, \dots, nm. \quad (6)$$

Для того чтобы система уравнений относительно значений $g(\xi_k)$ была замкнутой, из равенств (6) выбираем

$$D_{p+1}g(\xi_k) = 0, \quad k = \left[\frac{p}{2}\right], \dots, \left[nm - \frac{p+1}{2}\right]. \quad (7)$$

Совокупность равенств (7) образует систему линейных алгебраических уравнений $D_{p+1}g = 0$, где строчки матрицы D_{p+1} содержат коэффициенты разделенных разностей.

Теперь объединим обе системы:

$$Ag + \begin{pmatrix} r \\ \dots \\ \rho \end{pmatrix} = \begin{pmatrix} f \\ \dots \\ 0 \end{pmatrix}, \quad A = \begin{pmatrix} B \\ \dots \\ D_{p+1} \end{pmatrix}. \quad (8)$$

В результате получена переопределенная система, которая может быть несовместной. Поэтому в уравнение вводится невязка. Далее предполагается решать эту систему методом наименьших квадратов, который невязку минимизирует. Заметим, что минимизация невязки ρ , в частности, означает, что аппроксимация обладает также сглаживающим эффектом.

Таким образом, для приближенного вычисления значений аппроксимирующего полинома достаточно решить систему (8).

2. АППРОКСИМАЦИЯ КУСОЧНО-ПОСТОЯННОЙ ФУНКЦИЕЙ

Будем рассматривать случай $p = 0$, когда в качестве дополнительных уравнений берутся разделенные разности первого порядка. Таким образом, матрица D_1 имеет вид

$$D_1 = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 \\ 0 & 1 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 & -1 \end{pmatrix}. \quad (9)$$

2.1. Применение метода Лагранжа

Для определения свойств такой аппроксимации можно переформулировать задачу как задачу нахождения условного минимума и воспользоваться методом Лагранжа. С этой целью введем дополнительные обозначения (см. фиг. 2):

$$\begin{aligned} \tilde{f}_j &= g(x_j), \quad j = 1, \dots, m, \\ \delta_{j,k} &= g(\xi_{jn+k}) - g(\xi_{jn+k-1}), \quad j = 1, \dots, m-1, \quad k = 1, \dots, n. \end{aligned} \quad (10)$$

В этих обозначениях метод наименьших квадратов, примененный к системе (8), приводит к вычислению минимума функции

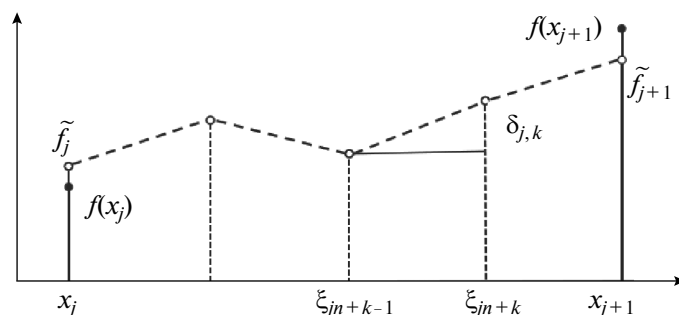
$$F(\tilde{f}_0, \dots, \tilde{f}_m; \delta_{0,1}, \dots, \delta_{m-1,n}) = \sum_{j=0}^m (f_j - \tilde{f}_j)^2 + \sum_{j=0}^{m-1} \sum_{k=1}^n \delta_{j,k}^2$$

при выполнении условий

$$\tilde{f}_{j+1} - \tilde{f}_j = \sum_{k=1}^n \delta_{j,k}, \quad j = 0, \dots, m-1.$$

Для применения метода Лагранжа составляем вспомогательную функцию

$$\Phi(\lambda_0, \dots, \lambda_{m-1}; \delta_{1,1}, \dots, \delta_{l-1,n}) = \sum_{j=0}^m (f_j - \tilde{f}_j)^2 + \sum_{j=0}^{m-1} \sum_{k=1}^n \delta_{j,k}^2 + \sum_{j=0}^{m-1} \lambda_j \left(\tilde{f}_{j+1} - \tilde{f}_j - \sum_{k=1}^n \delta_{j,k} \right).$$



Фиг. 2. Обозначения для аппроксимации в случае $p = 0$.

Приравниваем к нулю производные функции Φ по всем параметрам:

$$\begin{aligned}\Phi_{\lambda_j} &= \tilde{f}_{j+1} - \tilde{f}_j - \sum_{k=1}^n \delta_{j,k} = 0, \\ \Phi_{\delta_{j,k}} &= 2\delta_{j,k} - \lambda_j = 0, \\ \Phi_{\tilde{f}_0} &= -2(f_0 - \tilde{f}_0) - \lambda_0 = 0, \\ \Phi_{\tilde{f}_1} &= -2(f_1 - \tilde{f}_1) + \lambda_0 - \lambda_1 = 0, \\ \Phi_{\tilde{f}_j} &= -2(f_j - \tilde{f}_j) + \lambda_{j-1} - \lambda_j = 0, \\ \Phi_{\tilde{f}_m} &= -2(f_m - \tilde{f}_m) + \lambda_{m-1} = 0.\end{aligned}\quad (11)$$

Из первых двух равенств получаем, что

$$\delta_{j,k} = \frac{\lambda_j}{2}, \quad (12)$$

$$\tilde{f}_{j+1} - \tilde{f}_j = \sum_{k=1}^n \delta_{j,k} = n\delta_{j,k} = \frac{n}{2}\lambda_j. \quad (13)$$

Замечание. Равенство (12) означает, что в промежутках между точками x_j аппроксимация производится *линейно*. Следовательно, внутри каждого интервала (x_{j-1}, x_j) производная аппроксимирующей функции равна $(\tilde{f}_j - \tilde{f}_{j-1})/\Delta$.

Подставим выражение (13) в равенства для производных Φ_{f_j} :

$$\begin{aligned}\Phi_{\tilde{f}_0} &= -2(f_0 - \tilde{f}_0) - \lambda_0 = -2(f_0 - \tilde{f}_0) - \frac{2}{n}(\tilde{f}_1 - \tilde{f}_0) = 0, \\ \Phi_{\tilde{f}_j} &= -2(f_j - \tilde{f}_j) + \lambda_j - \lambda_{j+1} = -2(f_j - \tilde{f}_j) - \frac{2}{n}(\tilde{f}_{j+1} - \tilde{f}_j) + \frac{2}{n}(\tilde{f}_j - \tilde{f}_{j-1}) = 0, \\ \Phi_{\tilde{f}_m} &= -2(f_m - \tilde{f}_m) + \lambda_{m-1} = -2(f_m - \tilde{f}_m) + \frac{2}{n}(\tilde{f}_m - \tilde{f}_{m-1}) = 0.\end{aligned}$$

В матрично-векторном виде эти равенства принимают вид

$$f = \left(I + \frac{1}{n}T_1\right)\tilde{f}, \quad (14)$$

где

$$T_1 = \begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{bmatrix}.$$

Таким образом,

$$f - \tilde{f} = \frac{1}{n}T_1\tilde{f}.$$

Так как $\|T_1\| \leq 4$, то очевидна оценка

$$\|f - \tilde{f}\| \leq \frac{4}{n}\|\tilde{f}\|.$$

Здесь и далее обозначение $\|\cdot\|$, примененное к вектору, означает евклидову норму, а примененное к матрице — спектральную норму.

Из соотношения

$$\left(I + \frac{1}{n}T_1\right)^{-1}f = \tilde{f}$$

можно вывести и априорную оценку. Так как при $n \geq 5$

$$\left(I + \frac{1}{n}T_1\right)^{-1} = I - \frac{1}{n}T_1 \sum_{j=0}^{\infty} \left(\frac{-1}{n}T_1\right)^j = I + \Omega, \quad \|\Omega\| \leq \frac{4/n}{1 - 4/n},$$

то

$$\|f - \tilde{f}\| \leq \frac{4/n}{1 - 4/n} \|f\|. \quad (15)$$

Тем самым доказана следующая теорема.

Теорема 1. Пусть функция $f(x)$ задана в точках (1), матрицы A, B, D_1 имеют вид (8), (4) и (9) соответственно. Если при этом вектор g является обобщенным нормальным решением системы (8), а вектор $\tilde{f}$ задан равенством (10), то для векторов f и $\tilde{f}$ верна оценка (15), откуда следует, что данная аппроксимация имеет первый порядок точности.

2.2. Гладкость аппроксимации

Пусть матрица R_1 размера $m \times (m+1)$ имеет вид

$$R_1 = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 \\ 0 & 1 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 & -1 \end{pmatrix},$$

т.е. она отличается от матрицы D_1 размером. Тогда сглаживающий эффект данной аппроксимации описывается в следующей лемме.

Лемма 1. Для разностей элементов векторов f и $\tilde{f}$ имеет место неравенство

$$\|R_1 f\| \geq \left(1 + \frac{\lambda_{\min}(R_1 R_1^*)}{n}\right) \|R_1 \tilde{f}\|.$$

Для доказательства достаточно умножить обе части уравнения (14) на R_1 и заметить, что $T_1 = R_1^* R_1$, в результате чего получается следующая цепочка:

$$\|R_1 f\| = \|R_1 \tilde{f} + \frac{1}{n} R_1 R_1^* R_1 \tilde{f}\| = \left\| \left(I + \frac{1}{n} R_1 R_1^*\right) R_1 \tilde{f} \right\| \geq \left(1 + \frac{\lambda_{\min}(R_1 R_1^*)}{n}\right) \|R_1 \tilde{f}\|.$$

Таким образом, имеет место убывание разностных производных.

Усилить или ослабить эффект сглаживания можно путем добавления в систему весового параметра η :

$$A = \begin{pmatrix} B \\ \eta D_1 \end{pmatrix}.$$

Очевидно, что чем меньше вес, тем точнее аппроксимирующая функция, и тем меньше ее гладкость, и наоборот.

Повторив описанный в п. 2.1 процесс для этого случая, получим оценки, учитывающие множитель η :

$$\|f - \tilde{f}\| \leq \frac{4\eta^2}{n} \|\tilde{f}\|,$$

$$\|f - \tilde{f}\| \leq \frac{4\eta^2/n}{1 - 4\eta^2/n} \|f\|,$$

а также

$$\|R_1 f\| \geq \left(1 + \frac{4\eta^2}{n} \lambda_{\min}(R_1 R_1^*)\right) \|R_1 \tilde{f}\|.$$

Таким образом, при $p = 0$, несмотря на то что ожидается аппроксимация константой, т.е. полиномом нулевого порядка, в реальности имеет место кусочно-линейная аппроксимация. При этом такая аппроксимация более трудоемка, чем обычная кусочно-линейная интерполяция, когда линейными отрезками соединяются известные значения функции $f(x_j)$. Однако при этом данный подход к аппроксимации позволяет автоматически сгладить исходную функцию.

2.3. Обусловленность системы

Оценку числа обусловленности системы (8) можно получить на основе следующей теоремы (доказательство см. в [4]).

Теорема 2 (об окаймлении). Если $N \times N$ -матрицу A окаймить справа дополнительным столбцом из N элементов, то сингулярные числа $\sigma_1(B), \sigma_2(B), \dots, \sigma_N(B)$ матрицы B размера $N \times (N + 1)$, полученной в результате окаймления, связаны с сингулярными числами $\sigma_1(A), \sigma_2(A), \dots, \sigma_N(A)$ матрицы A условиями перемежаемости

$$\sigma_N(B) \geq \sigma_N(A) \geq \sigma_{N-1}(B) \geq \dots \geq \sigma_1(B) \geq \sigma_1(A).$$

Дополним матрицу D_1 последней строкой матрицы B и обозначим полученную квадратную $(mn + 1) \times (mn + 1)$ матрицу $\hat{D}_1$:

$$\hat{D}_1 = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 \\ 0 & 1 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 & -1 \\ 0 & \dots & 0 & 0 & -1 \end{pmatrix}.$$

Так как матрица A , определенная равенством (8), может рассматриваться как окаймление матрицы $\hat{D}_1$, то, применяя теорему 2, получаем неравенства

$$\sigma_{\min}(\hat{D}_1) \leq \sigma_{\min}(A) \leq \sigma_{\max}(\hat{D}_1) \leq \sigma_{\max}(A).$$

Отсюда следует оценка для числа обусловленности матрицы

$$\text{cond}(A) \leq \frac{\sigma_{\max}(A)}{\sigma_{\min}(\hat{D}_1)} = \|A\| \|\hat{D}_1^{-1}\|.$$

При этом имеют место очевидные ограничения на нормы

$$\|A\| \leq 2, \quad \|\hat{D}_1^{-1}\| \leq mn + 1.$$

Тем самым доказана следующая лемма.

Лемма 2. Пусть матрицы B и D_1 размеров $(m + 1) \times (mn + 1)$ и $mn \times (mn + 1)$ соответственно заданы формулами (4) и (9). Тогда число обусловленности матрицы

$$A = \begin{pmatrix} B \\ D_1 \end{pmatrix}$$

не превосходит $2(mn + 1)$.

3. ТОЧНОСТЬ АППРОКСИМАЦИИ КУСОЧНО-ЛИНЕЙНОЙ ФУНКЦИЕЙ

В данном разделе некоторые обозначения будут изменены для удобства выкладок.

Аппроксимация линейной функцией предполагает, что в качестве дополнительных условий используется равенство нулю вторых производных или разделенных разностей. Это означает, что регуляризирующая матрица D_2 размера $(mn - 1) \times (mn + 1)$ имеет вид

$$D_2 = \begin{pmatrix} 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \end{pmatrix}. \quad (16)$$

Введем дополнительные обозначения:

$$d_j = g(\xi_j) - g(\xi_{j-1}), \quad 1 \leq j \leq nm,$$

$$\varepsilon_j = d_{j+1} - d_j = g(\xi_{j+1}) - 2g(\xi_j) + g(\xi_{j-1}), \quad 1 \leq j \leq nm - 1,$$

$$\delta_j = g(x_j) - f_j, \quad 0 \leq j \leq m.$$

Решение системы (8) методом наименьших квадратов означает минимизацию нормы вектора невязки

$$(\delta_0, \dots, \delta_m, \varepsilon_1, \dots, \varepsilon_{nm-1})^T.$$

Чтобы сформулировать дополнительные условия, заметим, что

$$d_{nj+k} = d_{nj+1} + \sum_{i=1}^{k-1} \varepsilon_{nj+i}, \quad g(\xi_{n(j+1)}) - g(\xi_{nj}) = \sum_{i=1}^n d_{nj+i}, \quad 0 \leq j \leq m-1.$$

Подставим первое из этих равенств во второе и получим следующие соотношения:

$$nd_{nj+1} + \sum_{i=1}^{n-1} (n-i)\varepsilon_{nj+i} = f_{j+1} - f_j + \delta_{j+1} - \delta_j, \quad (17)$$

$$nd_{nj} - \sum_{i=1}^{n-1} i\varepsilon_{n(j-1)+i} = f_j - f_{j-1} + \delta_j - \delta_{j-1}. \quad (18)$$

Из равенства (17) вычтем (18), используя то, что $d_{nj+1} - d_{nj} = \varepsilon_{nj}$. Получаем при $j = 1, \dots, m-1$

$$f_{j-1} - 2f_j + f_{j+1} + \delta_{j-1} - 2\delta_j + \delta_{j+1} - \sum_{i=1}^n i\varepsilon_{n(j-1)+i} - \sum_{i=1}^{n-1} (n-i)\varepsilon_{nj+i} = 0. \quad (19)$$

Теперь можно применить метод Лагранжа для вычисления условного минимума. Функция Лагранжа имеет вид

$$\begin{aligned} \Phi = & \sum_{j=0}^m \delta_j^2 + \sum_{i=1}^{nm-1} \varepsilon_i^2 + \\ & + \sum_{j=1}^{m-1} \lambda_j \left(f_{j-1} - 2f_j + f_{j+1} + \delta_{j-1} - 2\delta_j + \delta_{j+1} - \sum_{i=1}^n i\varepsilon_{n(j-1)+i} - \sum_{i=1}^{n-1} (n-i)\varepsilon_{nj+i} \right). \end{aligned}$$

Приравниваем к нулю частные производные функции Лагранжа:

$$\begin{aligned} \Phi_{\varepsilon_j} &= 2\varepsilon_j - \lambda_1 i = 0, \\ \Phi_{\varepsilon_{nj+i}} &= 2\varepsilon_{nj+i} - \lambda_j(n-i) - \lambda_{j+1}i = 0, & i = 1, \dots, n, \\ \Phi_{\varepsilon_{n(m-1)+i}} &= 2\varepsilon_{n(m-1)+i} - \lambda_{m-1}(n-i) = 0, & j = 1, \dots, m-1; \end{aligned} \quad (20)$$

$$\begin{aligned} \Phi_{\delta_0} &= 2\delta_0 + \lambda_1 = 0, \\ \Phi_{\delta_1} &= 2\delta_1 - 2\lambda_1 + \lambda_2 = 0, \\ \Phi_{\delta_j} &= 2\delta_j + \lambda_{j-1} - 2\lambda_j + \lambda_{j+1} = 0, & j = 2, \dots, m-3. \\ \Phi_{\delta_{m-1}} &= 2\delta_{m-1} + \lambda_{m-3} - 2\lambda_{m-2} = 0, \\ \Phi_{\delta_m} &= 2\delta_m + \lambda_{m-2} = 0, \end{aligned} \quad (21)$$

Производные функции Φ по параметрам λ_j составляют систему (19), которую можно записать в виде

$$\Delta + \varphi = E, \quad (22)$$

где при $1 \leq j \leq mn-1$

$$\Delta_j = \delta_{j-1} - 2\delta_j + \delta_{j+1}, \quad E_j = \sum_{i=1}^n i\varepsilon_{n(j-1)+i} + \sum_{i=1}^{n-1} (n-i)\varepsilon_{nj+i}, \quad \varphi_j = f_{j-1} - 2f_j + f_{j+1}.$$

Обращение равенств (20) приводит к выражениям

$$\begin{aligned} \varepsilon_i &= \frac{i\lambda_1}{2}, \quad i = 1, \dots, n, \\ \varepsilon_{nj+i} &= \frac{(n-i)\lambda_j + i\lambda_{j+1}}{2}, \quad i = 1, \dots, n, \quad j = 1, \dots, m-1, \\ \varepsilon_{n(m-1)+i} &= \frac{(n-i)\lambda_{m-1}}{2}, \quad i = 1, \dots, n-1. \end{aligned} \quad (23)$$

Аналогичным образом из (21) выводим выражения

$$\begin{aligned}\delta_0 &= -\frac{\lambda_1}{2}, \\ \delta_1 &= -\frac{-2\lambda_1 + \lambda_2}{2}, \\ \delta_j &= -\frac{\lambda_{j-1} - 2\lambda_j + \lambda_{j+1}}{2}, \quad j = 2, \dots, m-3, \\ \delta_{m-1} &= -\frac{\lambda_{m-2} - 2\lambda_{m-1}}{2}, \\ \delta_m &= -\frac{\lambda_{m-1}}{2},\end{aligned}\tag{24}$$

которые можно представить в виде

$$\delta = -\frac{1}{2}R^*\lambda.\tag{25}$$

Здесь матрица R размера $(m-1) \times (m+1)$ имеет трехдиагональный вид

$$R_2 = \begin{pmatrix} 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \end{pmatrix}.$$

Отсюда следует, что вектор Δ из (22) можно выразить через λ_j :

$$\Delta = \begin{pmatrix} \delta_0 - 2\delta_1 + \delta_2 \\ \vdots \\ \delta_{m-2} - 2\delta_{m-1} + \delta_m \end{pmatrix} = \begin{pmatrix} 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} \delta_0 \\ \delta_1 \\ \vdots \\ \delta_m \end{pmatrix} = -\frac{1}{2}R_2R_2^* \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{m-1} \end{pmatrix}.$$

Далее можно показать, что

$$E = \begin{pmatrix} \sum_{i=1}^n i\varepsilon_i + \sum_{i=1}^{n-1} (n-i)\varepsilon_{n+i} \\ \sum_{i=1}^n i\varepsilon_{n+i} + \sum_{i=1}^{n-1} (n-i)\varepsilon_{2n+i} \\ \vdots \\ \sum_{i=1}^n i\varepsilon_{n(m-2)+i} + \sum_{i=1}^{n-1} (n-i)\varepsilon_{n(m-1)+i} \end{pmatrix} = \frac{n}{12}(n^2S + T_2) \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{m-1} \end{pmatrix}.$$

где S и T – симметричные трехдиагональные матрицы

$$S = \begin{pmatrix} 4 & 1 & & & \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & & 1 & 4 \end{pmatrix}, \quad T_2 = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & & -1 & 2 \end{pmatrix}.$$

Действительно, из (23) следует

$$\begin{aligned}E &= \frac{1}{2} \begin{pmatrix} \lambda_1 \left(\sum_{i=1}^n i^2 + \sum_{i=1}^{n-1} (n-i)^2 \right) + \lambda_2 \left(\sum_{n=1}^{n-1} i(n-i) \right) \\ \lambda_1 \left(\sum_{n=1}^{n-1} i(n-i) \right) + \lambda_2 \left(\sum_{i=1}^n i^2 + \sum_{i=1}^{n-1} (n-i)^2 \right) + \lambda_3 \left(\sum_{n=1}^{n-1} i(n-i) \right) \\ \vdots \\ \lambda_{m-2} \left(\sum_{n=1}^{n-1} i(n-i) \right) + \lambda_{m-1} \left(\sum_{i=1}^n i^2 + \sum_{i=1}^{n-1} (n-i)^2 \right) \end{pmatrix} = \\ &= \frac{n}{12} \begin{pmatrix} \lambda_1(4n^2 + 2) + \lambda_2(n^2 - 1) \\ \lambda_2(4n^2 + 2) + (\lambda_1 + \lambda_3)(n^2 - 1) \\ \vdots \\ \lambda_{m-1}(4n^2 + 2) + \lambda_{m-2}(n^2 - 1) \end{pmatrix} = \frac{n}{12} \begin{pmatrix} n^2(4\lambda_1 + \lambda_2) + 2\lambda_1 - \lambda_2 \\ n^2(\lambda_1 + 4\lambda_2 + \lambda_3) - \lambda_1 + 2\lambda_1 - \lambda_3 \\ \vdots \\ n^2(4\lambda_{m-1} + \lambda_{m-2}) + 2\lambda_{m-1} - \lambda_{m-2} \end{pmatrix}.\end{aligned}$$

В результате система (22) превращается в СЛАУ

$$\left[\frac{1}{2} R_2 R_2^* + \frac{n}{12} (n^2 S + T_2) \right] \lambda = \varphi,$$

где $\lambda = (\lambda_1, \dots, \lambda_{m-1})^T$. Оценим решение этой системы:

$$\lambda = \frac{12}{n^3} \left[I + \frac{1}{n^2} S^{-1} \left(T_2 + \frac{6}{n} R_2 R_2^* \right) \right]^{-1} S^{-1} \varphi. \quad (26)$$

Простая структура матриц R_2 , S и T_2 позволяет провести следующие оценки:

$$\|T_2\| = 4, \quad \|R_2\| = 4, \quad \|R_2 R_2^*\| \leq 16,$$

$$\lambda_{\min}(S) \geq 4 - 1 - 1 = 2, \quad \|S^{-1}\| = (\lambda_{\min}(S))^{-1} \leq \frac{1}{2}.$$

Введем обозначение

$$\Omega = -\frac{1}{n^2} S^{-1} \left(T_2 + \frac{6}{n} R_2 R_2^* \right),$$

тогда

$$\|\Omega\| \leq \frac{1}{n^2} \|S^{-1}\| \cdot \|T_2 + \frac{6}{n} R_2 R_2^*\| \leq \frac{2}{n^2} \left(1 + \frac{24}{n} \right).$$

Обратим внимание, что при $n \geq 5$

$$\|\Omega\| \leq \frac{1}{2}.$$

Следовательно, при $n \geq 5$ может быть получена также следующая оценка:

$$\| [I + \frac{1}{n^2} S^{-1} (T_2 + \frac{6}{n} R_2 R_2^*)]^{-1} \| = \| (I - \Omega)^{-1} \| \leq 1 + \frac{\|\Omega\|}{1 - \|\Omega\|} \leq 1 + \frac{4}{n^2} \left(1 + \frac{24}{n} \right) \leq 2.$$

Отсюда для (26) и (25) мы получаем

$$\|\lambda\| \leq \frac{12}{n^3} \|\varphi\|, \quad \|\delta\| \leq \frac{24}{n^3} \|\varphi\| \leq \frac{96}{n^3} \|f\|.$$

Таким образом, была доказана теорема о точности описанной аппроксимации.

Теорема 3. Пусть функция $f(x)$ задана в точках (1), матрицы A , B , D_2 имеют вид (8), (4) и (16) соответственно. Если при этом вектор g является обобщенным нормальным решением системы (8), вектор $\tilde{f}$ задан равенством $\tilde{f}_{j+1} = g_{nj+1}$ ($j = 1, \dots, m$), то для векторов f и $\tilde{f}$ верна оценка

$$\|f - \tilde{f}\| \leq \frac{96}{n^3} \|f\|$$

при $n \geq 5$, где $\tilde{f}_j = g(x_j)$.

Таким образом, данная аппроксимация имеет третий порядок по n .

3.1. Гладкость аппроксимации

По определению вектора δ имеет место равенство $\tilde{f} = \delta + f$, где $\tilde{f}_j = g(x_j)$. Умножим обе части этого равенства на R_2 :

$$\|R_2 \tilde{f}\| = \|R_2 \delta + R_2 f\| = \|R_2 f + \frac{1}{2} R_2 R_2^* \Theta R_2 f\| \leq \|I + \frac{1}{2} R_2 R_2^* \Theta\| \|R_2 f\|,$$

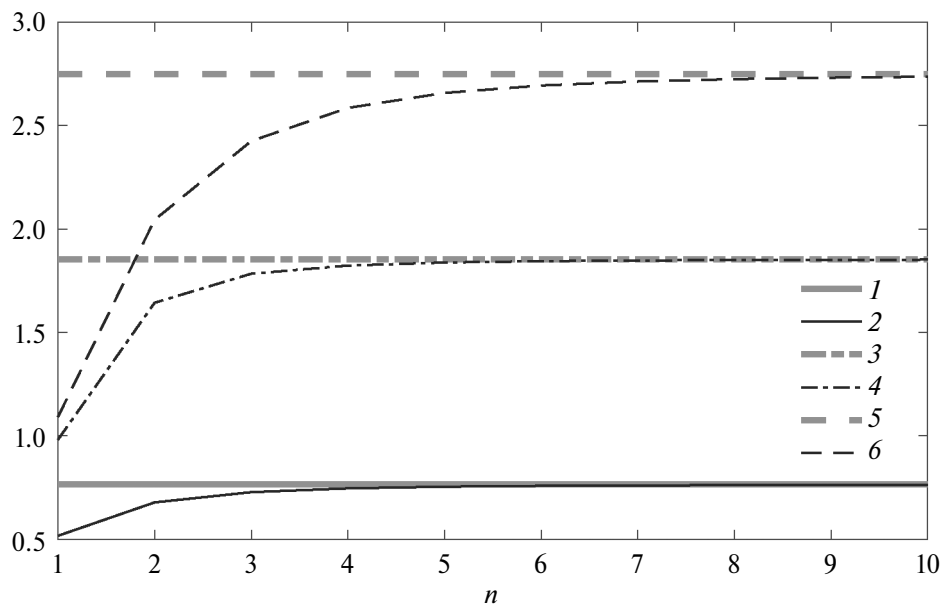
где

$$\Theta = \left(-\frac{1}{2} R_2 R_2^* - \frac{n}{12} (n^2 S + T_2) \right)^{-1}.$$

Используя выкладки предыдущего раздела, получаем оценку

$$\|\Theta\| = \left\| \frac{12}{n^3} (1 - \Omega)^{-1} \right\| \leq \frac{24}{n^3}$$

при $n \geq 5$. Отсюда следует лемма 3.



Фиг. 3. Сравнение $\|R_2 f\|$ и $\|R_2 \tilde{f}\|$, подробное описание см. в тексте.

Лемма 3. Для разностей элементов векторов f и $\tilde{f}$ имеет место неравенство

$$\|R_2 \tilde{f}\| < \left(1 + \frac{16 \cdot 24}{n^3}\right) \|R_2 f\|.$$

Это утверждение значительно более слабое, чем лемма 1, так как оно не показывает уменьшения производной, а значит, сглаживания. Однако проведенные многочисленные численные эксперименты убеждают, что сглаживание все-таки имеет место. На примере фиг. 3 можно видеть сравнение графиков $\|R_2 f\|$ и $\|R_2 \tilde{f}\|$ в зависимости от n при $m = 10$ для функций $\cos(2x)$ (кривые 1, 2 на фиг. 3), $\cos(x) + \sin(3x)$ (кривые 3, 4), $\cos(4x) + \sin(y)$ (кривые 5, 6). Очевидно, что вторые разности самих функций $\|R_2 f\|$ от параметра аппроксимации n не зависят, поэтому им на фиг. 3 соответствуют горизонтальные прямые. Видно, что к ним при растущем n кривые $\|R_2 \tilde{f}\|$ приближаются снизу.

Добавление в матрицу системы параметра η , с помощью которого можно увеличить или уменьшить вес производных,

$$A = \begin{pmatrix} B \\ \eta D_2 \end{pmatrix}$$

изменяет доказанные выше соотношения следующим образом:

$$\|f - \tilde{f}\| \leq \frac{24\eta}{n^3} \left(1 + \frac{2}{n^2} \left(1 + \frac{24\eta}{n}\right)\right) \|f\|.$$

3.2. Обусловленность систем

Для того чтобы оценить обусловленность расширенной матрицы в данном случае, воспользуемся следующими фактами и обозначениями. Обозначим $\tilde{D}_2$ квадратную матрицу размера $(mn+1) \times (mn+1)$, составленную из матрицы D_2 (16) и первой и последней строк матрицы B , и представим ее в блочном

$$\tilde{D}_2 = \begin{pmatrix} 1 & & & & & \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 1 & & & \vdots \\ 0 & \tilde{T} & & 0 \\ \vdots & & & 1 \\ 0 & 0 & \dots & 1 \end{pmatrix}, \quad \|\tilde{D}_2\| \leq 4.$$

При этом имеет место разложение

$$\tilde{T} = - \begin{pmatrix} 1 & -1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & 1 & -1 & 0 \\ 0 & \dots & 0 & 1 & -1 \\ 0 & \dots & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & 1 & -1 & 0 \\ 0 & \dots & 0 & 1 & -1 \\ 0 & \dots & 0 & 0 & 1 \end{pmatrix}^*.$$

Согласно теореме об окаймлении верна цепочка неравенств

$$\sigma_{\min}(\tilde{T}) \leq \sigma_{\min}(\tilde{D}_2) \leq \sigma_{\min}(A) \leq \sigma_{\max}(\tilde{T}) \leq \sigma_{\max}(\tilde{D}_2).$$

Отсюда следует оценка для числа обусловленности

$$\text{cond}(A) \leq \frac{\sigma_{\max}(A)}{\sigma_{\min}(\tilde{T})} \leq \|A\| \|\tilde{T}^{-1}\| \leq 4(mn+1)^2,$$

которую можно сформулировать в виде леммы.

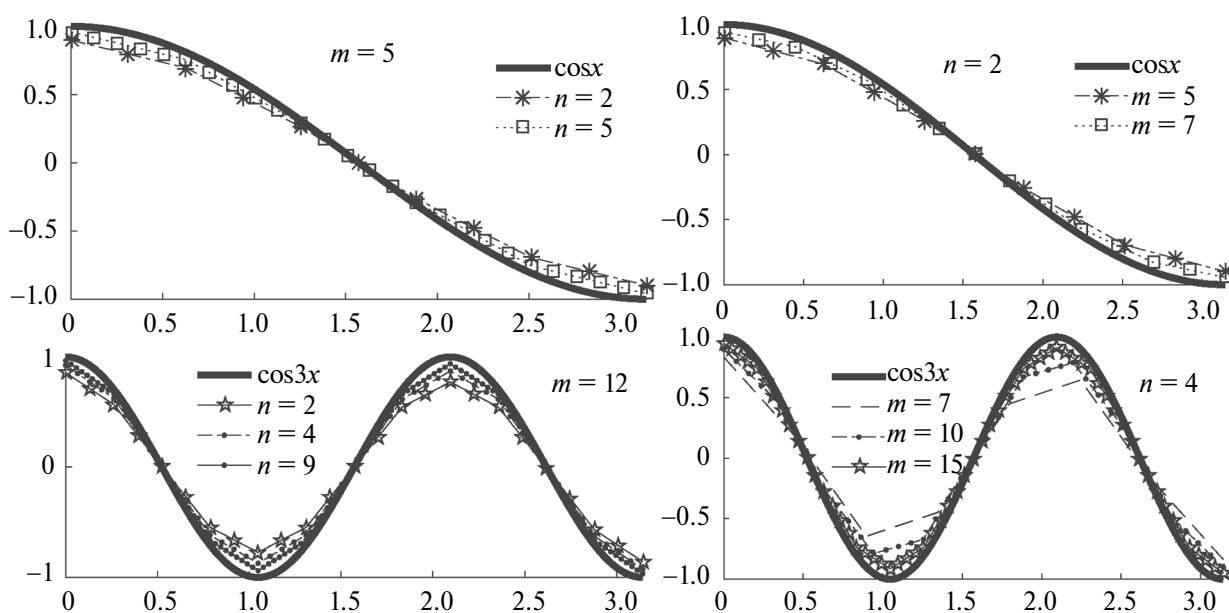
Лемма 4. Пусть матрицы B и D_2 размеров $(m+1) \times (mn+1)$ и $(mn-1) \times (mn+1)$ соответственно заданы формулами (4) и (16). Тогда число обусловленности матрицы

$$A = \begin{pmatrix} B \\ \dots \\ D_2 \end{pmatrix}$$

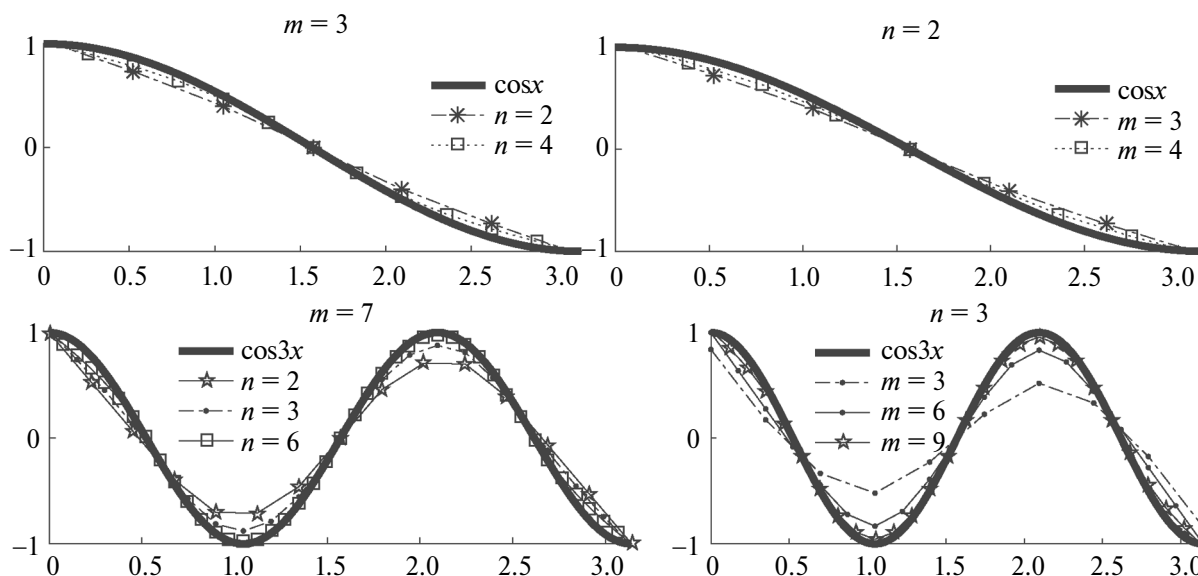
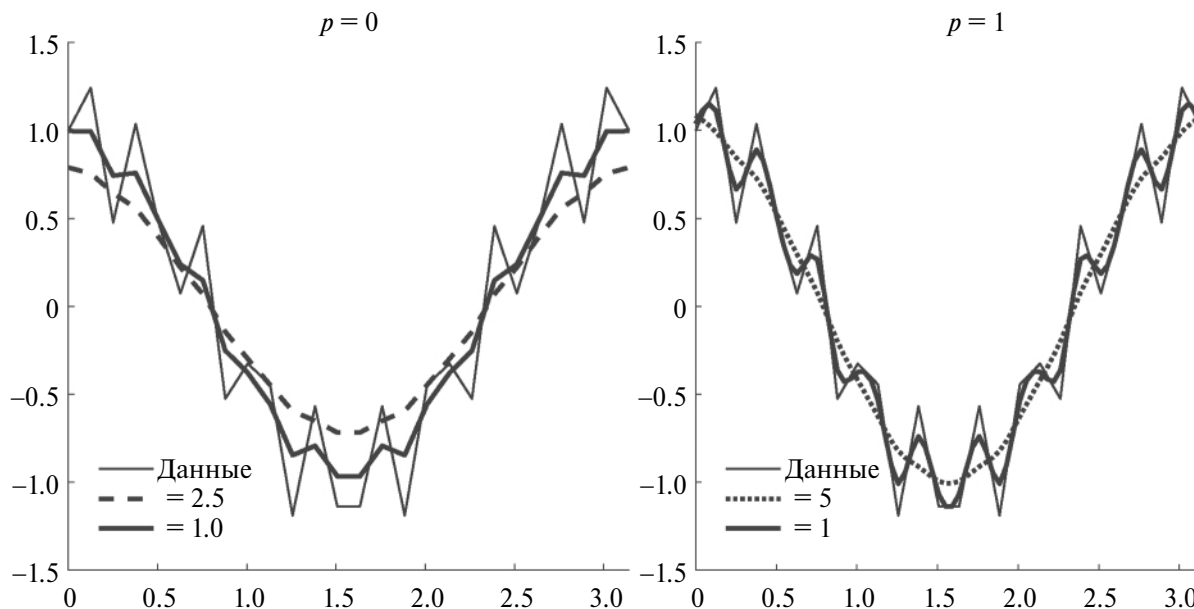
не превосходит $4(mn+1)^2$.

4. РЕЗУЛЬТАТЫ ВЫЧИСЛЕНИЙ

Используем представленный метод для аппроксимации функции $f(x) = \cos kx$ ($k = 1, 3$) на отрезке $[1, 3]$. На фиг. 4, 5 демонстрируется сходимость полученных аппроксимаций в зависимости от параметров m, n , а также от частоты осцилляций функции. На фиг. 6 представлены результаты влияния на сглаживание зашумленных данных весового параметра η . В обоих случаях использованы значения $m = 25, n = 1$.



Фиг. 4. Аппроксимация при $p = 0$.

Фиг. 5. Аппроксимация при $p = 1$.Фиг. 6. Сглаживание при $p = 0$.

ЗАКЛЮЧЕНИЕ

В рамках настоящей работы представлен гибридный метод обработки данных, объединяющий в себе свойства аппроксимации и сглаживания и основанный на оригинальной идее регуляризации плохо обусловленных систем, предложенной С.К. Годуновым. Интересной особенностью метода является то, что значения аппроксимации вычисляются без определения коэффициентов ее разложения по какому-либо базису.

Этот подход может применяться, например, для перехода с крупной сетки на более мелкую при численном решении начально-краевых задач с быстро меняющимися граничными условиями или в задачах, требующих сглаживания. Вероятно, в ходе дальнейших исследований будут выявлены и другие потенциальные возможности, заложенные в представленном методе.

СПИСОК ЛИТЕРАТУРЫ

1. Годунов С.К., Антонов А.Г., Кирилук О.П., Костин В.И. Гарантированная точность решения систем линейных уравнений в евклидовых пространствах. Новосибирск: Наука, 1988. С. 456.

2. Бибердорф Э.А., Попова Н.И. Гарантированная точность современных алгоритмов линейной алгебры. Новосибирск: Изд-во СО РАН, 2006. С. 319.
3. Кабанихин С.И. Обратные и некорректные задачи. Новосибирск: Сиб. науч. изд-во, 2009. С. 458.
4. Годунов С.К. Современные аспекты линейной алгебры. Новосибирск: Науч. книга, 1997. С. 388.

APPROXIMATION AND SMOOTHING OF A FUNCTION BASED ON GODUNOV REGULARIZATION

E. A. Biberdorf^{a,*}, K. K. Abdisheripov^b

^a*Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Acad. Koptyug Ave., 4,
Novosibirsk, 630090 Russia*

^b*Novosibirsk State University, Pirogov St., 1, Novosibirsk, 630090 Russia*

*e-mail: math@biberdorf.ru

Received 02 April, 2024

Revised 02 April, 2024

Accepted 02 May, 2024

Abstract. A new approach to function approximation is presented, based on S.K. Godunov's ideas on the regularization of ill-conditioned systems. The proposed method allows for determining function values at nodes of a finer grid from data on a coarser grid while ensuring control over the smoothness of the resulting function. Convergence and smoothness estimates are substantiated, and results from computational experiments illustrate the effectiveness of the proposed method.

Keywords: approximation, function smoothing, regularization of ill-conditioned systems of linear equations.

МЕТОД РАЗДЕЛЕНИЯ МАТРИЧНОГО СПЕКТРА ОТНОСИТЕЛЬНО ПРЯМОЙ И ЗАДАЧА О ФЛАТТЕРЕ БЕСКОНЕЧНОЙ ПОЛОСЫ¹⁾

© 2024 г. Э. А. Бибердорф^{1,*}, А. С. Рудометова¹, Ван Ли², А. Д. Жумабаев²

¹630090 Новосибирск, пр. Академика Коптюга, 4, Институт математики им. С.Л. Соболева СО РАН, Россия

²630090 Новосибирск, ул. Пирогова, 1, Новосибирский Государственный Университет, Новосибирск, Россия
*e-mail: math@biberdorf.ru

Поступила в редакцию 02.04.2024 г.

Переработанный вариант 02.04.2024 г.

Принята к публикации 02.05.2024 г.

Предлагается новый метод разделения матричного спектра относительно прямой, основанный на дробно-линейном преобразовании. Отмечается, что он имеет ряд преимуществ в сравнении с подходами, основанными на экспоненциальном преобразовании, а именно, его область применимости шире, а количество итераций, необходимых для его сходимости, значительно меньше. Предложенный метод используется для исследования задач о флаттере бесконечной полосы с различными условиями закрепления кромок, которые после подходящей дискретизации дифференциальных операторов сводятся к спектральным задачам для матриц. Исследование областей устойчивости методом дихотомии спектра относительно мнимой оси позволяет построить нейтральные кривые в плоскости параметров задачи о флаттере. Библ. 18. Фиг. 5.

Ключевые слова: дихотомия спектра, область устойчивости, флаттер, дискретизация дифференциального оператора, нейтральная кривая.

DOI: 10.31857/S0044466924080026, EDN: YBIPXW

ВВЕДЕНИЕ

Информация о расположении спектров линейных операторов используется для решения большого круга прикладных задач. В частности, спектральный критерий применяется в задачах устойчивости, к которым относится проблема флаттера. В задачах такого рода требуется информация о группах собственных значений, находящихся в областях устойчивости или неустойчивости. Следует отметить, что в несимметрическом случае спектр оператора может быть крайне чувствителен к возмущениям и вычислительным погрешностям. Это также необходимо учитывать при разработке вычислительных алгоритмов. С.К. Годунов был участником советского ядерного проекта, в котором к качеству и точности вычислений были особые требования. Сергей Константинович, как автор вычислительных методов, считал, что результат выполнения численного алгоритма должен сопровождаться оценкой его точности. В этом смысле ситуация с классической спектральной задачей в несимметричном случае была, очевидно, неудовлетворительной. Работа над этой проблемой под руководством С.К. Годунова в 80–90-е годы прошлого века привела к изменению самой постановки спектральной проблемы и созданию метода дихотомии матричного спектра, который впоследствии был им системно изложен в монографиях [1], [2].

Данный подход основан на идеях теории устойчивости А.М. Ляпунова, а также продолжает ряд других методов (см. [3]), которые позволяют по спектру специально сконструированной симметричной матрицы судить о расположении собственных значений заданной несимметричной матрицы A .

Результатом применения метода дихотомии к матричному пучку $A - \lambda B$ являются матрицы

$$H = \oint_{\gamma} (\lambda B - A)^{-1} C (\bar{\lambda} B^* - A^*)^{-1} |d\lambda| \quad \text{и} \quad P = \frac{1}{2\pi i} \oint_{\gamma} (\lambda B - A)^{-1} B d\lambda, \quad (0.1)$$

где $C = C^* > 0$ — произвольная матричная нормировка. Норма матрицы H позволяет оценить расстояние от спектра до кривой γ . При этом матрица P представляет собой проектор на правое приводящее подпространство

¹⁾Работа выполнена в рамках государственного задания Института математики им. С.Л. Соболева СО РАН (проект № FWNF-2022-0008).

пучка $A - \lambda B$ (инвариантное подпространство матрицы A в случае $B = I$), соответствующее собственным значениям, лежащим внутри контура γ , соответственно, след P равен числу таких собственных значений.

Итерационные алгоритмы дихотомии, не требующие вычисления контурных интегралов, разработаны для многих встречающихся на практике кривых: окружности (см. [4, 5]), мнимой оси (см. [6–9]), кривых второго порядка (см. [10–12]), угла (см. [13]), а также для разделения корней полинома относительно окружности (см. [14]). При этом алгоритм разделения спектра относительно единичной окружности можно назвать базовым, так как он является составной частью остальных существующих алгоритмов дихотомии. Один из его вариантов мы приведем в приложении.

Нужно отметить, что С.К. Годунов с соавторами ранее уже обращался к проблеме флаттера (см. [8]). Однако по причинам, которые мы описываем далее, существовавший на тот момент метод мог применяться только к максимально простой модели флаттера, описываемой обыкновенными дифференциальными уравнениями.

В настоящей работе рассматривается новый вариант дихотомии спектра относительно прямой. Его мы представляем в разд. 1. В разд. 2 этот модифицированный метод мы применяем к задаче флаттера для модели, предложенной в 1994 г. А.А. Ильюшиным и И.А. Кийко (см. [15]). С помощью дихотомии спектра дифференциального оператора модели находятся нейтральные кривые на плоскости параметров для бесконечной полосы с различными видами закрепления по краям.

1. ДИХОТОМИЯ МНИМОЙ ОСЬЮ

Подход к разделению спектра матрицы A относительно мнимой оси, применявшийся до сих пор, основан на свойствах экспоненциального отображения, которое переводит левую полуплоскость в единичный круг. Важно, что инвариантные подпространства матрицы, соответствующие собственным значениям, лежащим по разные стороны от мнимой оси, совпадают с инвариантными подпространствами матричной экспоненты, соответствующими частям ее спектра внутри и вне единичной окружности. Поэтому для решения задачи дихотомии мнимой осью достаточно применить алгоритм дихотомии единичной окружностью к матричному пучку $A_0 - \lambda B_0 = e^{\tau A} - \lambda I$. Выбор параметра τ , например, $\tau \approx 1/2\|A\|$, обусловлен не только стремлением обеспечить быструю сходимость при вычислении матричной экспоненты. Дело в том, что для матриц с достаточно большой нормой вычисление матричной экспоненты без предварительной нормировки невозможно в принципе. Однако, вследствие такой нормировки спектр исходной матрицы оказывается прижатым к мнимой оси и применение метода дихотомии приводит к орицательным результатам. В частности, по этой причине дихотомия мнимой осью, использующая экспоненциальное преобразование, не может применяться к матрицам, аппроксимирующим дифференциальные операторы.

Ранее в работе [9] эта проблема была решена за счет выбора нормировочного параметра τ в виде степени двойки $\tau = 2^{-K}$ и последующем выполнении K дополнительных итераций основного цикла алгоритма.

В настоящей работе в качестве альтернативы экспоненциальному преобразованию рассматривается дробно-линейное преобразование комплексной плоскости $\lambda \rightarrow \xi$, переводящее мнимую ось в единичную окружность:

$$\lambda = \frac{\xi + 1}{\xi - 1} \quad \text{или} \quad \xi = \frac{\lambda + 1}{\lambda - 1}, \quad (1.1)$$

где λ, ξ — комплексные числа. В результате этой замены исходная спектральная задача для пучка $A - \lambda B$ превращается в задачу о спектре пучка

$$A_0 - \xi B_0 = (A + B) - \xi(A - B). \quad (1.2)$$

Теорема 1. *Разделение спектра пучка $A - \lambda B$ относительно мнимой оси возможно тогда и только тогда, когда спектр пучка $A_0 - \xi B_0$ может быть разделен относительно единичной окружности. При этом их матричные критерии дихотомии (0.1) отличаются множителем 2, а проекторы на приводящие подпространства совпадают.*

Доказательство. Заметим, что если λ — чисто мнимое число, то соответствующее значение ξ (0.1) находится на единичной окружности. Выпишем интегральное представление критерия дихотомии для пучка $A - \lambda B$ и сделаем замену подинтегральной переменной:

$$\begin{aligned} H_\lambda &= \int_{Re(\lambda)=0} (\lambda B - A)^{-1} C (\bar{\lambda} B^* - A^*)^{-1} |d\lambda| = \int_{|\xi|=1} |\xi - 1|^2 (\xi B_0 - A_0)^{-1} C (\bar{\xi} B_0^* - A_0^*)^{-1} \left| d \frac{\xi + 1}{\xi - 1} \right| = \\ &= 2 \int_{|\xi|=1} (\xi B_0 - A_0)^{-1} C (\bar{\xi} B_0^* - A_0^*)^{-1} |d\xi| = 2H_\xi. \end{aligned}$$

То, что нормы матриц H_λ и H_ξ конечны или бесконечны одновременно, говорит о том, что спектры пучков $A - \lambda B$, $A_0 - \xi B_0$ одновременно пересекают мнимую ось и единичную окружность соответственно, тем самым доказано первое утверждение теоремы.

Далее предположим, что у исходного пучка $A - \lambda B$ отсутствуют собственные значения на мнимой оси. Тогда имеет место разложение

$$A - \lambda B = T \begin{pmatrix} \Lambda_- & \\ & \Lambda_+ \end{pmatrix} S^{-1} - \lambda T S^{-1}, \quad (1.3)$$

где спектры подматрицы Λ_- , Λ_+ лежит соответственно в левой и правой полуплоскостях. При этом матрицы $T = [T_1 | T_2]$ и $S = [S_1 | S_2]$ невырождены, а подматрицы T_1, T_2, S_1, S_2 представляет собой объединение базисных векторов в приводящих подпространствах, соответствующих собственным значениям в левой и правой полуплоскостях.

Для пучка, полученного после замены спектрального параметра, будет верно представление

$$\begin{aligned} A_0 - \xi B_0 &= T \begin{pmatrix} (\Lambda_- + I) - \xi(\Lambda_- - I) & \\ & (\Lambda_+ + I) - \xi(\Lambda_+ - I) \end{pmatrix} S^{-1} = \\ &= T_0 \begin{pmatrix} \Lambda_0 - \xi I & \\ & I - \xi \Lambda_\infty \end{pmatrix} S^{-1}, \end{aligned}$$

где

$$\begin{aligned} T_0 &= T \begin{pmatrix} \Lambda_- - I & \\ & \Lambda_+ + I \end{pmatrix}, \\ \Lambda_0 &= (\Lambda_- + I)(\Lambda_- - I)^{-1}, \quad \Lambda_\infty = (\Lambda_+ - I)(\Lambda_+ + I)^{-1}. \end{aligned}$$

То есть правая приводящая матрица сохраняется неизменной, а базисы левых приводящих подпространств подвергаются невырожденным преобразованиям

$$T_{01} = T_1(\Lambda_- - I), \quad T_{02} = T_2(\Lambda_+ + I),$$

а, значит, сами подпространства не меняются.

Из этого следует, что проекторы на правые и левые приводящие подпространства пучка $A - \lambda B$, соответствующие частям спектра, лежащим в левой и правой полуплоскостях, совпадают с проекторами на правые и левые приводящие подпространства пучка $A_0 - \xi B_0$, соответствующие собственным значениям внутри и вне единичной окружности. Например, проекторы на правые подпространства имеют вид

$$P = S \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} S^{-1}, \quad I - P = S \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} S^{-1}. \quad (1.4)$$

Теорема доказана.

Замечание. Следует обратить внимание на то, что среди собственных значений пучка могут быть бесконечные в случае, если матрица B вырождена. Мнимая ось проходит через бесконечную точку, а бесконечному значению λ соответствует $\xi = 1$. Следовательно, при наличии бесконечных собственных значений у пучка $A - \lambda B$ невозможна дихотомия его спектра мнимой осью, а спектра пучка $A_0 - \xi B_0$ единичной окружностью.

Итак, для разделения спектра пучка $A - \lambda B$ относительно мнимой оси нужно применить алгоритм дихотомии единичной окружностью к пучку (1.2). Полученные при этом матрицы H и P являются искомыми критерием дихотомии и проектором на правое приводящее подпространство пучка, соответствующее собственным значениям, лежащим в левой полуплоскости.

1.1. Пример

Продemonстрируем свойства нового подхода на простом примере с помощью спектрального портрета. Одномерные спектральные портреты — это эффективный способ визуализации спектра, основанный на методе дихотомии матричного спектра. Пусть $\gamma(x) = \{\lambda, \operatorname{Re} \lambda = x\}$ — параметризованное семейство прямых на комплексной плоскости, параллельных мнимой оси. Заметим, что разделение спектра пучка $A - \lambda B$ относительно прямой $\gamma(x)$ равносильно разделению спектра пучка $A - (\lambda + x)B = (A - xB) - \lambda B$ со сдвинутым параметром относительно мнимой оси. Таким образом, каждому значению параметра x можно сопоставить норму матричного критерия дихотомии $H(x)$, вычисленного для пучка $(A - xB) - \lambda B$. График функции $\log_{10} \|H(x)\|$ дает следующую информацию о расположении спектра. Если для фиксированного x значение этой функции ограничено, то на прямой $\gamma(x)$ отсутствуют собственные значения пучка $A - \lambda B$. Если в окрестности точки x функция $\log_{10} \|H(x)\|$ неограниченно растет (график функции образует «пик»), то вблизи прямой $\gamma(x)$ лежат точки

спектра пучка. В примере мы увидим, что бывают ситуации, когда вычисленный график $\log_{10} \|H(x)\|$ образует «полочку» вместо того, чтобы стремиться к бесконечности. Это происходит из-за того, что если в процессе итераций текущий параметр $\omega = \|H(x)\|$ превышает предельно допустимое значение, то алгоритм прерывает свою работу, выходным значением параметра дихотомии объявляется $\omega_{\max}$, а проектор $P(x)$ не вычисляется.

Дополнительную информацию о спектре можно получить из графика функции $\text{tr}P(x)$. Это кусочно-постоянная функция, значение которой в точке x совпадает с числом собственных значений, лежащих левее прямой $\gamma(x)$. Функция $\text{tr}P(x)$ меняет свое значение (график образует «ступеньку»), если на прямой $\gamma(x)$ лежат собственные значения пучка. Очевидно, что число собственных значений, лежащих вблизи $\gamma(x)$ совпадает с высотой «ступеньки».

Также для иллюстрации работы итерационного алгоритма дихотомии будем использовать функцию $k(x)$, значения которой будут совпадать с числом итераций, потребовавшихся для сходимости.

В качестве примера рассмотрим двухдиагональную матрицу A размера 5×5 , у которой на главной диагонали стоят числа 6, 7, 8, 9, 10, а на побочной — 7, 8, 9, 10. Для данной матрицы мы построим спектральные портреты, используя три варианта сведения задачи о разделении спектра относительно прямой к задаче о единичной окружности. Первый способ — использование экспоненциального преобразования $e^{\tau A}$ (см. [1], [6]). Вторым способом — использование экспоненты со специальной нормировкой $\tau = 2^{-K} \approx 1/2\|A\|$ и проведение K дополнительных итераций алгоритма дихотомии единичной окружностью (см. [9]). Третий способ — использование дробно-линейного преобразования.

На фиг. 1 изображены графики функций $\log_{10} \|H(x)\|$ и $k(x)$, полученные тремя способами. Из них видно, что новый метод дает наименьшие значения для обеих функций. Причем число итераций для нового метода меньше в разы. Если рассмотреть подробнее окрестность прямой, на которой находится собственное значение (см. фиг. 1, средний рисунок), можно видеть, что все три графика в окрестности прямой $x = 8$, на которой лежит собственное значение матрицы образуют полочку. Заметим, что наиболее широкую полочку имеет график, полученный с использованием экспоненциального преобразования. Остальные два графика довольно близки и имеют значительно более узкую полочку. Это означает, что метод, использующий экспоненциальное преобразование со специальной нормировкой и предложенный нами метод, основанный на дробно-линейном преобразовании, имеют более широкую область применимости и могут эффективно работать в окрестности прямых, на которых лежат собственные значения.

Отметим также, что во всех точках x , в которых вычислялся проектор $P(x)$, вне зависимости от использованного метода имеют место следующие относительные оценки

$$\frac{\|P^2(x) - P(x)\|}{\|P(x)\|} \leq 10^{-10}, \quad \frac{\|AP(x) - P(x)A\|}{\|A\|\|P(x)\|} \leq 10^{-13}.$$

2. ФЛАТТЕР БЕСКОНЕЧНОЙ ПОЛОСЫ

Рассмотрим математическую модель панельного флаттера (см. [15]). Пусть пластина постоянной толщины h из материала плотности ρ , занимающая в плоскости xy область G с кусочно-гладкой границей Γ обтекается с одной стороны потоком газа с вектором скорости $\vec{v} = (v_x, v_y) = (v \cos \theta, v \sin \theta)$. Уравнение в безразмерной форме для возмущенного состояния пластины $u(t, x, y)$ имеет вид

$$D\Delta^2 u + \gamma \vec{v} \nabla u + \gamma \frac{\partial u}{\partial t} + \rho h \frac{\partial^2 u}{\partial t^2} = 0. \quad (2.1)$$

Здесь Δ^2 — бигармонический оператор, $D = \frac{Eh^3}{12(1-\nu^2)}$ — цилиндрическая жесткость, E — модуль Юнга, ν — коэффициент Пуассона, γ — показатель адиабаты газа.

Будем искать решение (2.1) в виде $u(t, x, y) = \varphi(x, y)e^{\omega t}$. Возмущенное движение пластины будет устойчивым, если $\text{Re}(\omega) < 0$, и неустойчивым, если $\text{Re}(\omega) > 0$. Условие $\text{Re}(\omega) = 0$ таким образом будет определять критические параметры потока газа или пластины.

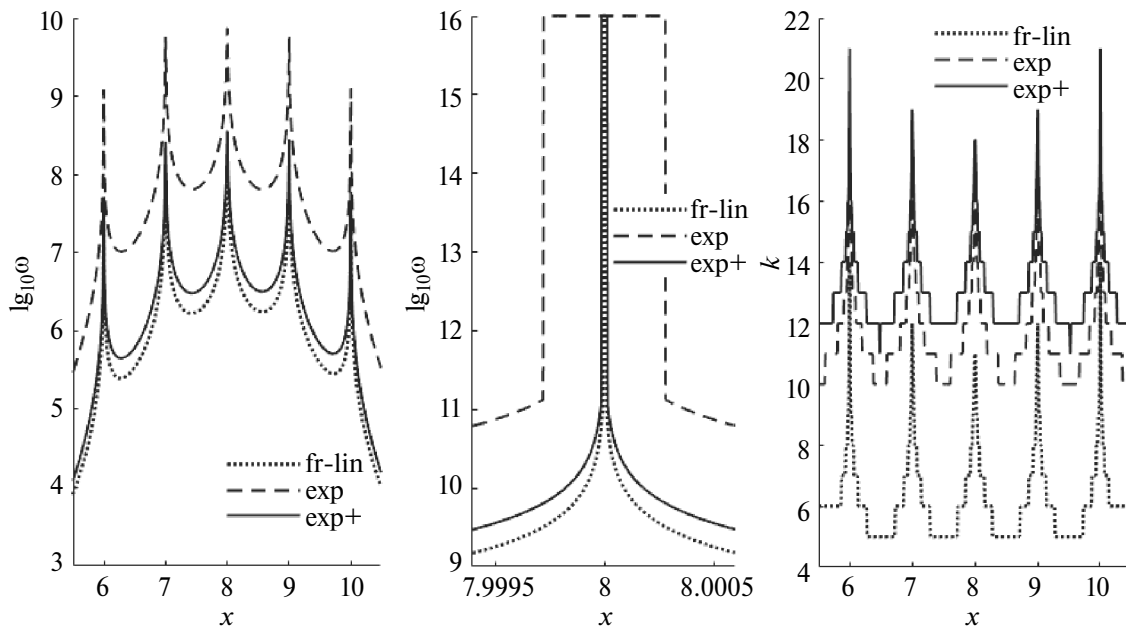
Для амплитуды колебания пластины $\varphi(x, y)$ получаем уравнение

$$D\Delta^2 \varphi + \gamma \vec{v} \nabla \varphi + (\gamma \omega + \rho h \omega^2) \varphi = 0, \quad (2.2)$$

которое вместе с подходящими условиями на границе

$$M_1(\varphi) = 0, \quad M_2(\varphi) = 0, \quad x, y \in \Gamma, \quad (2.3)$$

составляет задачу на собственные значения.



Фиг. 1. Линейный спектральный портрет матрицы A (слева и в центре) и число потребовавшихся итераций (справа). Использование экспоненциального отображения — сплошная линия, экспоненциального отображения со специальной нормировкой — штрих, дробно-рационального отображения — пунктир.

Далее в работе будем рассматривать пластину в виде упругой бесконечной полосы. Пусть в плоскости xy она занимает область $G : \{|x| < \infty, 0 \leq y \leq 1\}$. Решение (2.2), ограниченное на бесконечности, будем искать в виде $\varphi(x, y) = \psi(y)e^{-i\alpha x}$, где α — вещественный параметр. Тогда для $\psi(y)$ получаем уравнение

$$D \left(\frac{\partial^4 \psi}{\partial y^4} - 2\alpha^2 \frac{\partial^2 \psi}{\partial y^2} + \alpha^4 \psi \right) + \gamma v_y \frac{\partial \psi}{\partial y} - i\alpha \gamma v_x \psi = (\gamma \omega + \rho h \omega^2) \psi. \quad (2.4)$$

Поставим две краевые задачи, соответствующие разным типам закрепления кромок полосы при $y = 0$ и $y = 1$. В случае шарнирного закрепления кромок

$$\begin{aligned} y = 0 : \psi &= 0, \quad \frac{\partial^2 \psi}{\partial y^2} = 0, \\ y = 1 : \psi &= 0, \quad \frac{\partial^2 \psi}{\partial y^2} = 0, \end{aligned} \quad (2.5)$$

при жестком закреплении кромок полосы

$$\begin{aligned} y = 0 : \psi &= 0, \quad \frac{\partial \psi}{\partial y} = 0, \\ y = 1 : \psi &= 0, \quad \frac{\partial \psi}{\partial y} = 0. \end{aligned} \quad (2.6)$$

Для аппроксимации поставленных задач нужно провести дискретизацию дифференциальных операторов с учетом заданных граничных условий. В итоге получим задачу на собственные значения ω

$$(C + (\gamma \omega + \rho h \omega^2)I)\psi = 0, \quad (2.7)$$

где I — единичная матрица, а матрица C имеет вид

$$C = D[D_4 - 2\alpha^2 D_2 + \alpha^4 I] + \gamma v_y D_1 - i\alpha \gamma v_x I.$$

Заметим, что математическая формулировка условий жесткого закрепления совпадает с граничными условиями для уравнения Орра—Зоммерфельда, когда в качестве основного течения берется течение Пуазейля в бесконечном плоском канале. Поэтому для условий (2.6) мы используем аппроксимацию производных, которая применяется в таких случаях в многочисленных работах по гидродинамике, в том числе в исследованиях

спектров оператора Орра—Зоммерфельда. При этом аппроксимация решения производится в точках Гаусса—Лобатто $y_j = \cos(j\pi/N)$, $j = 0, \dots, N$, а матрицы D_1 , D_2 и D_4 представляют собой матричные коллокационные производные первого, второго и четвертого порядков, учитывающие однородные условия Дирихле и Неймана (подробное описание см. [17]). Можно отметить, что в гидродинамических задачах, а также в тестах на основе задачи Штурма—Лиувилля применение данной аппроксимации дифференциального оператора позволяет получить с высокой точностью собственные значения с небольшим модулем. Большие собственные значения вычисляются неточно. Однако в задачах устойчивости важна точность критических собственных значений, которые, как правило, невелики по модулю.

При шарнирном закреплении кромок используем метод дискретизации дифференциальных операторов, учитывающий условия вида (2.5), предложенный в работе [18]:

$$D_1 = \begin{pmatrix} u_1 & c_{21} & \cdots & c_{N1} \\ c_{12} & u_2 & \cdots & c_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1N} & c_{2N} & \cdots & u_N \end{pmatrix}, \quad D_2 = \begin{pmatrix} v_1 & b_{21} & \cdots & b_{N1} \\ b_{12} & v_2 & \cdots & b_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ b_{1N} & b_{2N} & \cdots & v_N \end{pmatrix}, \quad D_4 = D_2^2,$$

где

$$y_j = \frac{j\pi}{N+1}, \quad u_j = \frac{\pi}{2} \operatorname{ctg} y_j, \quad v_j = -\frac{(N+1)^2\pi^2}{3} - \frac{\pi^2}{6} + \frac{\pi^2}{2\sin^2 y_j}, \quad j = 1, \dots, N,$$

$$c_{ij} = (-1)^{i+j-1} \frac{\pi \sin y_i}{\cos y_j - \cos y_i}, \quad b_{ij} = (-1)^{i+j-1} \frac{2\pi^2 \sin y_j \sin y_i}{(\cos y_j - \cos y_i)^2}, \quad i, j = 1, \dots, N, \quad i \neq j.$$

Для проверки адекватности такой дискретизации воспользуемся тем, что для задачи о продольном обтекании (т.е. при $\vec{v} = (v_x, 0)$) полосы с шарнирно закрепленными кромками можно найти (см., например, [16]) точные собственные значения λ (связанные с частотой ω соотношением $\gamma\omega + \rho h\omega^2 = -\lambda$):

$$\lambda_n = (\alpha^2 + n^2\pi^2)^2 D - i\alpha\gamma v_x. \quad (2.8)$$

Определим относительную погрешность вычисленных собственных значений матрицы C по формуле

$$\Delta_n = \frac{|\lambda_n(C) - \lambda_n|}{|\lambda_n|}.$$

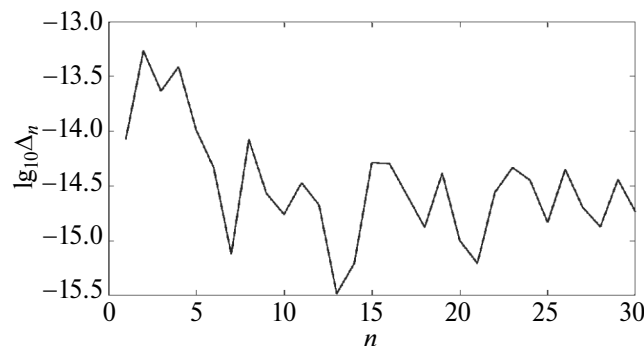
На фиг. 2 показано, что максимум погрешности собственных чисел дискретизованного оператора в сравнении с собственными числами дифференциального оператора, найденными в явном виде по формуле (2.8), приблизительно равен 10^{-13} . Это позволяет считать выбранный метод дискретизации дифференциального оператора для случая шарнирного закрепления достаточно точным.

Вспомним, что по исходной постановке задачи о флаттере нас интересует расположение частот ω относительно мнимой оси. Квадратичную спектральную задачу (2.7) можно привести к линейной относительно ω задаче

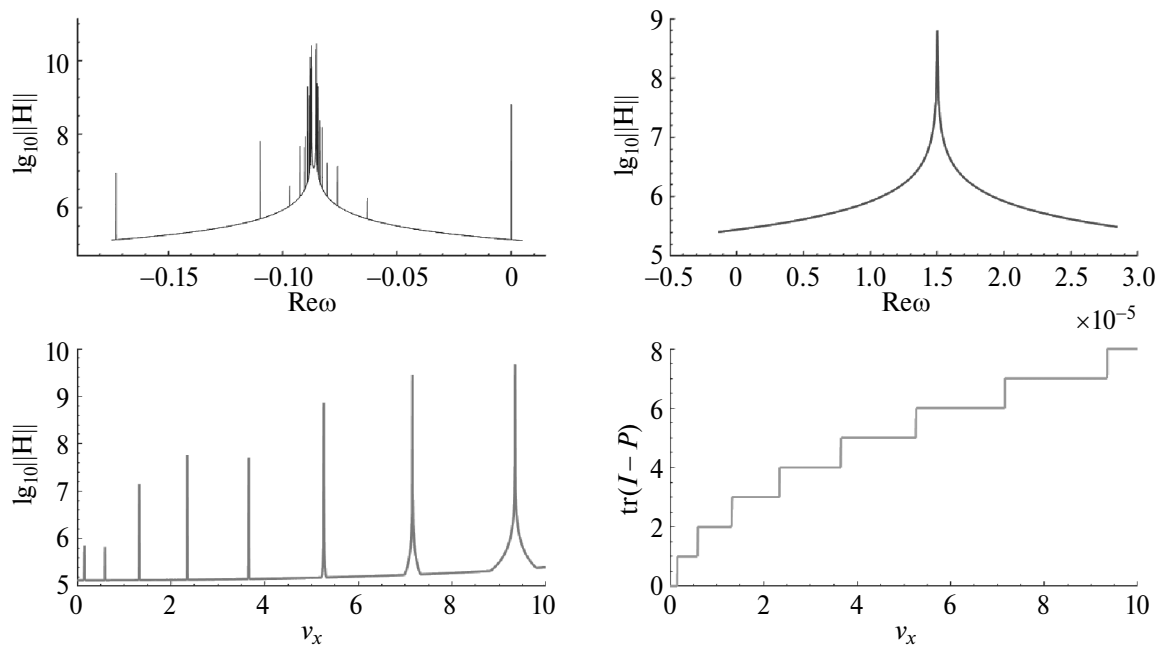
$$(A - \omega B)\varphi = 0 \quad (2.9)$$

с матрицами удвоенной размерности

$$A = \begin{pmatrix} A & \gamma I \\ 0 & I \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -\rho h I \\ I & 0 \end{pmatrix}.$$



Фиг. 2. График погрешности $\log_{10} \Delta_n$ в зависимости от номера собственного значения n , $N = 30$.



Фиг. 3. Спектральные портреты в случае шарнирного закрепления и продольного обтекания при $\alpha = 1$. Для верхних графиков скорость фиксирована $v_x = 0.1606$; $N = 10$.

Это возможно, так как собственные значения ω_j у задач (2.7) и (2.9) совпадают, а собственные векторы имеют вид

$$\Psi_j \quad \text{и} \quad \Phi_j = \begin{pmatrix} \Psi_j \\ \omega_j \Psi_j \end{pmatrix}$$

соответственно.

К матричному пучку $A - \omega B$ применим описанный в первом разделе алгоритм дихотомии собственных значений ω мнимой осью.

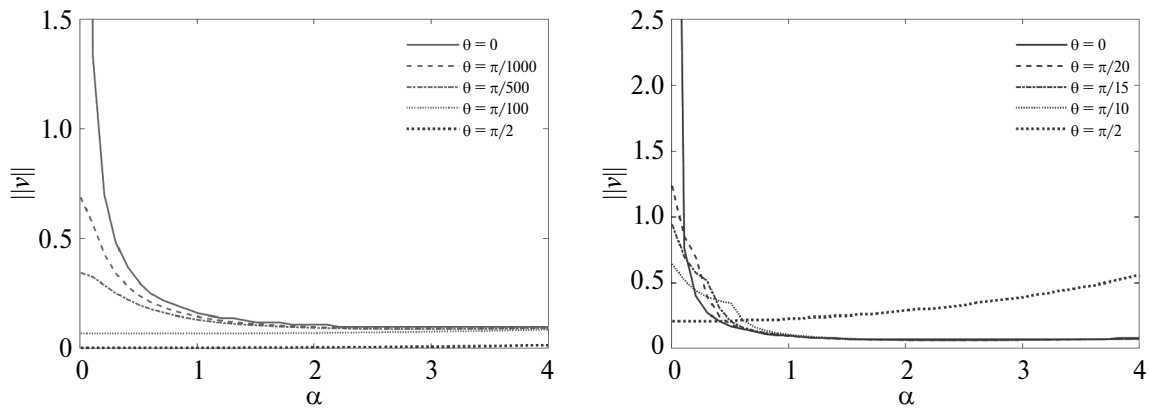
На фиг. 3 изображены одномерные спектральные портреты пучка $A - \omega B$ в случае шарнирного закрепления и продольного обтекания при $\alpha = 1$. На графиках в верхней части рисунка представлена зависимость нормы критерия дихотомии прямой, параллельной мнимой оси, от расположения прямой при фиксированной скорости потока $|v| = v_x = 0.1606$. При этом правый верхний рисунок представляет собой увеличенную часть левого верхнего рисунка. Видно, что при указанном значении скорости самый правый пик графика лежит справа от нуля, следовательно в правой полуплоскости находится как минимум одно собственное значение, что полностью соответствует теоретическим результатам (см. [16]).

В нижней части фиг. 3 изображена зависимость критерия дихотомии мнимой осью от модуля скорости также в случае шарнирного закрепления и продольного обтекания. Левый нижний график показывает, при каких значениях скорости собственные значения пучка $A - \omega B$ проходят через мнимую ось. На графике справа внизу изображен график следа проектора на подпространство, соответствующее собственным значениям пучка, лежащим в правой полуплоскости, расположение ступеней этого графика (т.е. значений скорости, при которых меняется количество собственных значений в правой полуплоскости) совпадает с положением пиков на левом графике.

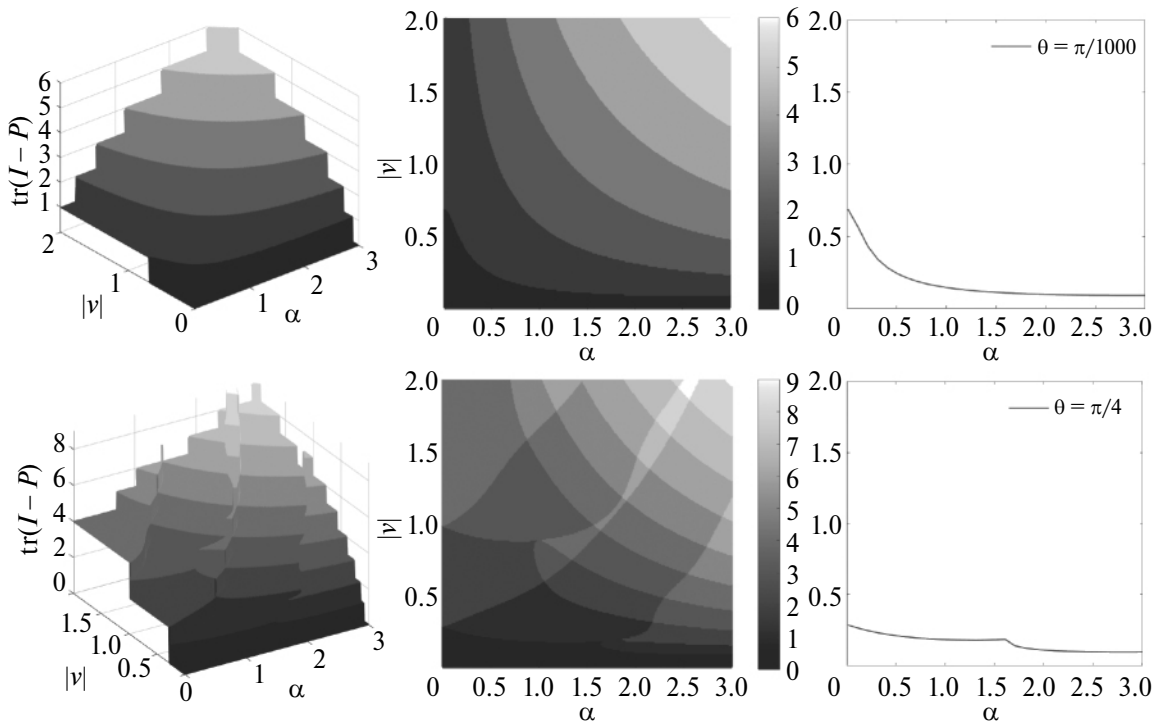
Заметим, что на плоскости параметров $\alpha, v = |\vec{v}|$ при фиксированном угле обтекания θ множества точек, при которых $\text{tr}(I - P)$ равен определенному натуральному значению, образуют области, в общем случае не односвязные. Так как нейтральная кривая является множеством таких α, v , при которых первое собственное значение переходит в правую полуплоскость, то ее можно представить как границу области на плоскости α, v , в которой $\text{tr}(I - P) = 0$. Заметим, что след проектора весьма удобно использовать как критерий в расчетах, так как на нейтральной кривой он меняется сразу на целое число.

На фиг. 4 изображены нейтральные кривые для разных углов обтекания. Видно, что характер кривых при шарнирном и жестком закреплении существенно различается.

Различие между шарнирным и жестким закреплении границ полосы также демонстрируется на фиг. 5. На нем разным цветом раскрашены области на плоскости параметров, в которых $\text{tr}(I - P)$ принимает различные значения. Видно, что в случае жесткого закрепления структура этих областей, а значит, и зависимость спек-



Фиг. 4. Нейтральные кривые, слева – шарнирное закрепление, справа – жесткое закрепление; $N = 40$.



Фиг. 5. Функция $\text{tr}(I - P)$ от α , $v = |\vec{v}|$, ее проекция и нейтральная кривая $\text{tr}(I - P) = 0$. Верхний ряд – шарнирное закрепление при $\theta = \pi/100$, нижний ряд – жесткое закрепление при $\theta = \pi/4$; $N = 40$.

тра задачи от параметров, существенно сложнее. Это, в частности, объясняет негладкость нейтральных кривых (см. фиг. 4, справа).

3. ЗАКЛЮЧЕНИЕ

В работе предложен новый способ разделения матричного спектра относительно прямой, основанный на дробно-линейном преобразовании. По сравнению с подходами, основанными на экспоненциальном преобразовании, он обладает несколькими преимуществами. Во-первых, для его реализации не требуется вычисление матричной экспоненты. Во-вторых, в случае его применения к матричному пучку нет необходимости вычислять обратную матрицу. На примерах было показано, что область его применимости шире, чем у метода, использующего экспоненту, а число итераций, требующихся для его сходимости, значительно меньше.

С помощью нового метода мы исследовали задачу о флаттере бесконечной полосы. В случае продольного обтекания полосы с шарнирно закрепленными кромками данные, полученные в ходе вычислений, хорошо согласуются с имеющимися аналитическими результатами. Также данный подход был использован для постро-

ения нейтральных кривых в области параметров задачи. Их изображения приведены при различных значениях угла между направлением набегающего потока и осью x .

4. ПРИЛОЖЕНИЕ

Ниже приведен один из существующих алгоритмов дихотомии единичной окружностью. Заметим, что условия сходимости основного цикла могут быть сформулированы по-разному. Можно следить за относительной сходимостью, как это сделано ниже, или за абсолютной. В условие можно также включать сходимость проектора P_k . Кроме того, опираясь на оценки сходимости (например, [14], теорема 5), можно априорно оценить число итераций, необходимых для сходимости алгоритма при условии $\|H\| \leq \omega_{\max}$. Эту оценку также можно использовать в качестве верхней границы числа итераций. При этом нужно принимать во внимание, что она, как правило, очень завышена, так как в обычной ситуации для сходимости с высокой точностью достаточно нескольких итераций.

Алгоритм дихотомии матричного спектра относительно единичной окружности

Дано: матричный пучок $A_0 - \lambda B_0$, ε_{it} — требуемая точность итерационного процесса, $\omega_{\max}$, $\mu_{\max}$ — максимальные значения критерия дихотомии и числа обусловленности матрицы.

Если $\text{cond}(A_0 - B_0) > \mu_{\max}$,

то дихотомия невозможна, конец расчетов.

иначе

$$H_0 = (A_0 - B_0)^{-1}(A_0 A_0^* + B_0 B_0^*)(A_0^* - B_0^*)^{-1}$$

Цикл пока $\|H_k - H_{k-1}\| > \varepsilon_{it} \|H_k\|$

Если $\|H_k\| \geq \mu_{\max}$ или $\text{cond}(A_k + B_k) > \mu_{\max}$,

то дихотомия невозможна, конец расчетов

иначе

$$V_{k+1} = (A_k + B_k)^{-1} A_k, \quad U_{k+1} = I_n - V_{k+1},$$

$$H_{k+1} = U_{k+1} H_k U_{k+1}^* + V_{k+1} H_k V_{k+1}^*,$$

$$qr \left(\begin{bmatrix} -B_k & A_k & 0 \\ A_k & 0 & -B_k \end{bmatrix} \right) = \begin{bmatrix} * & * & * \\ 0 & A_{k+1} & -B_{k+1} \end{bmatrix},$$

$$P_k = -(A_{k+1} - B_{k+1})^{-1} B_{k+1};$$

Конец цикла

Результат: критерий дихотомии $\omega = \|H_k\|$, проектор P_k .

Если $\|H_k\|$, вычисленная на шаге k , превосходит заданное значение $\omega_{\max}$, это значит, что окружность пересекает пятна ε -спектра для достаточно малого ε . Таким образом, алгоритм диагностирует ситуацию, когда внутренняя и внешняя части спектра не могут быть разделены достаточно надежно. В этом случае алгоритм прекращает свою работу, проектор не вычисляется.

О точности полученных результатов можно судить по величине невязки равенства $P^2 = P$, определяющего проектор. В случае задачи о спектре матрицы A (не пучка) для проверки точности можно использовать свойство перестановочности $AP = PA$ матрицы и проектора на инвариантное подпространство (см. [1]). Для матричного пучка $A - \lambda B$ аналогом этого равенства служат соотношения $QA = AP$ и $QB = BP$, где Q — проектор на левое приводящее подпространство. Как правило для решения задачи не требуется вычислять проектор Q , поэтому эти равенства для проверки точности обычно не используются.

В приведенных в статье примерах мы использовали следующие константы

$$\omega_{\max} = 10^{16}, \quad \mu_{\max} = 10^{10}, \quad \varepsilon_{it} = 10^{-5}.$$

СПИСОК ЛИТЕРАТУРЫ

1. Годунов С.К. Современные аспекты линейной алгебры. Новосибирск: Научная книга, 1997. С. 388.
2. Годунов С.К. Лекции по современным аспектам линейной алгебры. Новосибирск: Научная книга, 2002. С. 216.
3. Крейн М.Г., Неймарк М.А. Метод симметрических и эрмитовых форм в теории отделения корней алгебраических уравнений. ГНТИ Украины, Харьков, 1936. С. 40.

4. Годунов С.К. Круговая дихотомия матричного спектра// Сиб. матем. журн. 1986. Т. 27. № 5. С. 24–37.
5. Булгаков А.Я., Годунов С. К. Круговая дихотомия матричного спектра// Сиб. матем. журн. 1988. Т. 29. № 5. С. 59–70.
6. Малышев А.Н. Введение в вычислительную линейную алгебру. Новосибирск: Наука, Сибирское отделение, 1991. С. 228.
7. Godunov S.K., Sadkane M. Spectral Analysis of Symplectic Matrices with Application to the Theory of Parametric Resonance// SIAM J. on Matrix Analysis and Applications. 2006. V. 28. Iss. 4. P. 1083–1096.
8. Буньков В.Г., Годунов С.К., Курзин В.Б., Садкане М. Применение нового математического аппарата «Одномерные спектральные портреты матрицы» к решению проблемы аэроупругих колебаний решеток лопастей Ученые записки ЦАГИ. 2009. Т. 40. № 6. С. 3–13.
9. Бибердорф Э.А., Блинова М.А., Попова Н.И. Модификации метода дихотомии матричного спектра и их применение к задачам устойчивости// СибЖВМ. 2018. Т. 21. № 2. С. 139–153.
10. Godunov S.K., Sadkane M. Elliptic dichotomy of a matrix spectrum // Linear Algebra Appl. 1996. V. 248. P. 205–232.
11. Malyshev A.N., Sadkane M. On parabolic and elliptic spectral dichotomy// SIAM J. Matrix Anal. Appl. 1997. V. 18. P. 265–278.
12. Блинова М.А., Попова Н.И., Бибердорф Э.А. Приложение дихотомии матричного спектра к исследованию устойчивости течений // Марчуковские научные чтения - 2017. Тр. Междунар. науч. конф. 2017. С. 106–112.
13. Бибердорф Э.А. Алгоритм разделения матричного спектра относительно угла// Ж. вычисл. матем. и матем. физ. 2022. Т. 62. № 5. С. 742–756.
14. Бибердорф Э.А. Критерий дихотомии корней полинома единичной окружностью // СибЖИМ. 2000. Т. 3. № 1. С. 16–32.
15. Ильюшин А.А., Кийко И.А. Новая постановка задачи о флаттере пологой оболочки // ПММ. 1994. Т. 58. Вып. 3. С. 167–171.
16. Алгазин С.Д., Кийко И.А. Флаттер пластин и оболочек. Москва: Наука, 2006.
17. Trefethen L.N. Spectral Methods in MATLAB. SIAM. Philadelphia, 2000. P. 163.
18. Семисалов Б.В. Нелокальный алгоритм поиска решений уравнения Пуассона и его приложения // Ж. вычисл. матем. и матем. физ. 2014. Т. 54. № 7. С. 1111–1135.

A METHOD FOR SEPARATING THE MATRIX SPECTRUM BY A STRAIGHT LINE AND AN INFINITE STRIP FLUTTER PROBLEM

E. A. Biberdorf^{a,*}, A. S. Rudometova^a, Wang Li^b, A.D. Jumabaev^b

^a*Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Acad. Koptug Ave., 4,
Novosibirsk, 630090 Russia*

^b*Novosibirsk State University, Pirogov St., 1, Novosibirsk, 630090 Russia*

**e-mail: math@biberdorf.ru*

Received 02 April, 2024

Revised 02 April, 2024

Accepted 02 May, 2024

Abstract. A new method for separating the matrix spectrum relative to the straight line is proposed, based on a fractional-linear transformation. It is noted that it has a number of advantages over approaches based on an exponential transformation, namely, its applicability area is wider, and the number of iterations required for its convergence is significantly smaller. The proposed method is used to study the problems of infinite strip flutter with various edge fixing conditions, which after suitable discretization of differential operators are reduced to spectral problems for matrices. The study of stability regions by the method of spectrum dichotomy relative to the imaginary axis allows one to construct neutral curves in the plane of parameters of the flutter problem.

Keywords: spectrum dichotomy, stability region, flutter, differential operator discretization, neutral curve.

ЭФФЕКТИВНОЕ УСТОЙЧИВОЕ ИНТЕГРИРОВАНИЕ ПО ВРЕМЕНИ УРАВНЕНИЙ КАНА–ХИЛЛАРДА: ЯВНЫЕ, НЕЯВНЫЕ И ЯВНО-ИТЕРАЦИОННЫЕ СХЕМЫ¹⁾

© 2024 г. М. А. Бочев^{1,*}, И. А. Фахурдинов^{1,2,**}, Е. Б. Савенков^{1,***}

¹125047 Москва, Миусская пл., 4, ИПМ им. М.В. Келдыша РАН, Россия

²115409 Москва, Каширское ш., 31, НИЯУ МИФИ, Россия

*e-mail: botchev@kiam.ru

**e-mail: mv1451003@gmail.com

***e-mail: savenkov@keldysh.ru

Поступила в редакцию 02.04.2024 г.

Переработанный вариант 02.04.2024 г.

Принята к публикации 02.05.2024 г.

Предлагается новый алгоритм численного интегрирования по времени уравнения Кана–Хилларда, основанный на совместном применении метода расщепления Эйра и схемы локальных итераций (ЛИМ) для решения конечномерной задачи на каждом временном шаге. Предложенный метод является градиентно-устойчивым, допускает расчет с большими шагами по времени и имеет явный характер вычислений. Приведены результаты численных расчетов, демонстрирующие возможности предложенного метода и его сравнение с распространенными способами интегрирования по времени уравнения Кана–Хилларда. Библ. 65. Фиг. 5. Табл. 10.

Ключевые слова: уравнение Кана–Хилларда, градиентно-устойчивые схемы, расщепление Эйра, схема локальных итераций, ЛИМ.

DOI: 10.31857/S0044466924080034, EDN: YBFTGQ

1. ВВЕДЕНИЕ

Уравнение Кана–Хилларда было предложено в 1958 г. в оригинальной работе [1] для описания процесса разделения фаз в двухкомпонентных сплавах. В настоящее время оно имеет многочисленные приложения в самых разных областях науки, являясь одной из базовых моделей так называемой градиентной, или слабонелокальной, термомеханики сплошных сред (см. [2]). Как составная часть более сложных моделей, уравнение применяется в многофазной гидродинамике, материаловедении, задачах солидификации и теории фазовых переходов и многих других областях (см. [3], [4]). С теоретической точки зрения уравнение Кана–Хилларда является основой феноменологической теории спинодального распада (см. [5]) и имеет многочисленные применения в теоретической физике (см. [6]–[8] и ссылки там). Математической теории уравнения Кана–Хилларда посвящена обширная литература (см., например, [9]).

До недавнего времени уравнение Кана–Хилларда рассматривалось прежде всего как теоретическая модель соответствующих процессов. Однако в последние десятилетия оно стало активно применяться, независимо, либо как часть более сложных моделей, в качестве прикладного инструмента математического моделирования. Это привело к необходимости разработки эффективных вычислительных алгоритмов для его численного решения.

Возникающие при этом сложности связаны, прежде всего, с двумя основными факторами. Во-первых, уравнение Кана–Хилларда является нелинейным уравнением в частных производных, содержащим четвертые пространственные производные. Наличие нелинейностей приводит к тому, решения уравнения Кана–Хилларда эволюционируют на широком диапазоне пространственных и временных масштабов. Так, типичное решение задачи практически постоянно в пределах некоторых пространственных областей однородности, соответствующих «чистым» фазам системы. Эти области однородности, имеющие характерный диаметр d , отделены друг от друга тонкими слоями конечной толщины $\sim \epsilon \ll d$ («диффузными границами»), в пределах которых решение задачи является гладким, но меняется от своего минимального до своего максимального значения. Одновременно с этим, в начальный момент развития спинодального распада из случайного начального возмущения

¹⁾ Исследование Е.Б. Савенкова выполнено за счет гранта Российского научного фонда (проект № 2-11-00203).

(когда толщина ϵ диффузной границы сравнима с диаметром зон однородности, $\epsilon \sim d$), характерные времена эволюции решения имеют порядок ϵ^2 — в то время как в процессе дальнейшей эволюции, когда толщина диффузной границы становится существенно меньше диаметра областей однородности, $\epsilon \ll d$, характерные времена эволюции $\sim 1/\epsilon$ (см. [10], [11]).

В совокупности это приводит к тому, что чисто явные алгоритмы интегрирования уравнения Кана—Хилларда по времени являются устойчивыми при временных шагах $\tau \sim h^4$, где h — шаг по пространству. Это ограничение хотя и отвечает физике задачи, в ряде случаев неприемлемо с точки зрения эффективного решения содержательных задач. Одновременно с этим, чисто неявные разностные схемы не гарантируют однозначную разрешимость уравнений конечномерной задачи при достаточно больших шагах по времени (см. [12], [13]). По этой причине построение разностных схем, которые были бы (а) однозначно разрешимы при произвольных шагах по времени; (б) устойчивыми и (в) консервативными — является сложной задачей. Базовые алгоритмы, удовлетворяющие этим критериям, были построены сравнительно недавно в цикле работ (см. [14], [15]). Они основаны на выпуклом расщеплении энергии системы, см. разд. 3. Отметим, что под «устойчивостью» в контексте численного решения уравнения Кана—Хилларда понимается так называемая «градиентная» или, что то же, «энергетическая» устойчивость, гарантирующая выполнение условие невозрастания конечномерного аналога свободной энергии системы (см. ниже разд. 2, 3). Это условие является сильным (выражение для свободной энергии системы включает в себя первые производные решения), но необходимо для обеспечения термодинамической согласованности конечномерных аппроксимаций решения уравнений Кана—Хилларда.

В настоящее время среди основных направлений исследований в области построения эффективных вычислительных алгоритмов интегрирования по времени для уравнения Кана—Хилларда можно выделить построение адаптивных алгоритмов (они необходимы в силу широкого диапазона характерных времен эволюции решения, см. выше), разработку энергетически—устойчивых схем интегрирования по времени (обеспечивающих термодинамическую корректность численного решения) и построение алгоритмов, с одной стороны, обеспечивающих приемлемый асимптотический запас устойчивости, а с другой — однозначную разрешимость конечномерной задачи и вычислительную эффективность.

Сколь-либо полный обзор работ по способам построения и исследования временных аппроксимаций уравнения Кана—Хилларда существенно выходит за рамки настоящей работы. Общее представление об основных направлениях разработки временных аппроксимаций для уравнения Кана—Хилларда дает обзор [16]. В качестве примера других работ отметим [17]–[23] (адаптивные алгоритмы и методы высокого порядка), [24]–[26] (схемы, допускающие «большие» шаги по времени), [14], [15], [27]–[33] (энергетически—устойчивые схемы).

Для построения пространственных аппроксимаций уравнений Кана—Хилларда могут быть использованы различные методы, среди которых классические методы конечных разностей, конечных объемов, спектральные методы, аппроксимации на основе классического и изогометрического метода конечных элементов (см. [34]–[47]).

Целью настоящей работы является численное исследование нового класса методов интегрирования по времени для решения уравнения Кана—Хилларда. Новые алгоритмы основаны на двух основных идеях: (а) использование явно—неявной аппроксимации по времени на основе выпуклого расщепления энергии системы в соответствии с работами [14], [15], обеспечивающего энергетическую устойчивость схемы, и (б) применение модифицированной схемы локальных итераций (ЛИМ), позволяющей сочетать явный характер вычислений с устойчивостью неявных схем. Метод расщепления Эйра (см. [14], [15]) в настоящее время является основой построения большинства аппроксимаций уравнения Кана—Хилларда, с одной стороны, обладающих свойством энергетической устойчивости, а с другой стороны — гарантирующих существование и единственность решения конечномерной задачи.

Методы локальных итераций (ЛИ) представляют собой явные схемы интегрирования по времени, где устойчивость достигается за счет применения многочленов Чебышёва. Итерации с многочленами Чебышёва используются в схемах интегрирования по времени по крайней мере с 50-х годов прошлого века (см. [48], [49]). Отметим, что в схемах ЛИ чебышёвские итерации применяются не для приближенного решения систем, возникающих в неявных схемах — в этом случае, как показано в 1952 г. Гельфандом и Локуциевским (см. [50] и [51, гл.10, § 4.12]), существенного выигрыша в вычислительной работе по сравнению с явной схемой получить невозможно. Принципиальной особенностью схем ЛИ является то, что итерации Чебышёва выбираются из соображений устойчивости и точности интегрирования по времени, а не из соображений скорейшей сходимости итераций к решению неявной схемы. Схемы ЛИ были предложены в работах [52]–[55] и получили развитие в статьях [56]–[58], см. также ссылки в [58]. Применение специальных чебышёвских итераций в схемах ЛИ позволяет использовать существенно большие шаги по времени, чем требуемые для устойчивого применения двухслойной явной схемы, сохраняя при этом число явный характер вычислений.

Первая схема ЛИ, представленная в [52]–[54], не обладает свойством монотонности (не сохраняет неотрицательности численного решения) и не является асимптотически устойчивым при больших временах $t \rightarrow \infty$.

Поэтому в данной работе нами используется модифицированная схема локальных итераций ЛИМ, разработанная в [55], которая обладает этими свойствами. Подробное описание и сравнение вариантов схем ЛИ можно найти в [58]. В настоящее время известен целый ряд примеров успешного применения ЛИМ для решения достаточно сложных задач (см., например, [59]–[61]). В работе [62] схема ЛИМ была применена для решения уравнений Кана–Хилларда в рамках численной реализации полной математической модели процессов кристаллизации металлов. В этой работе уравнение Кана–Хилларда линейризовывалось на каждом временном слое, и метод ЛИМ применялся для решения линейризованной задачи. Сравнение предложенного алгоритма с другими известными современными схемами интегрирования по времени не выполнялось, равно как не исследовались такие его свойства как энергетическая устойчивость.

В настоящей работе на примере модельного пространственно–одномерного уравнения Кана–Хилларда предложены новые варианты построения временных аппроксимаций, основываясь на методе ЛИМ и явно– неявных аппроксимациях по Эйру. Приводятся результаты численного исследования предложенного метода и его сравнение с рядом других разностных аппроксимаций.

В разд. 2 приведено краткое описание модели. В разд. 3 сначала приведено краткое описание известных алгоритмов, далее подробно рассматриваются новые методы. В разд. 4 представлены результаты численных расчетов. В разд. 5 дается обсуждение основных полученных результатов работы.

2. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

В рассматриваемом одномерном случае уравнение Кана–Хилларда имеет вид

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial x} \left(M \frac{\partial \mu}{\partial x} \right), \quad \mu(c) = F'(c) - \epsilon^2 \frac{\partial^2 c}{\partial x^2}, \quad F(c) = c^2(1 - c)^2, \quad (2.1)$$

где $c = c(x, t)$ — искомая функция, $M = M(c) > 0$ — коэффициент подвижности (далее в расчетах считаем $M \equiv 1$), $\mu = \mu(c, \partial c / \partial x)$ — химический потенциал, $\epsilon > 0$, $\epsilon = \text{const}$ — параметр, определяющий толщину диффузной границы, t — время, x — пространственная переменная.

Уравнение (2.1) решается в одномерной области $\Omega = (0, 1)$ при $t \in (0, T]$. На границе области заданы граничные условия

$$\frac{\partial c}{\partial x} = \frac{\partial \mu}{\partial x} = 0, \quad (2.2)$$

а при $t = 0$ — начальное условие

$$c(x, 0) = c^0(x). \quad (2.3)$$

С физической точки зрения уравнение (2.1) может быть получено следующим способом (см. [63]). Рассмотрим двухкомпонентную систему с концентрациями компонент $c_{1,2}$, $c_1 + c_2 = 1$. Будем считать, что свободная энергия такой системы имеет вид

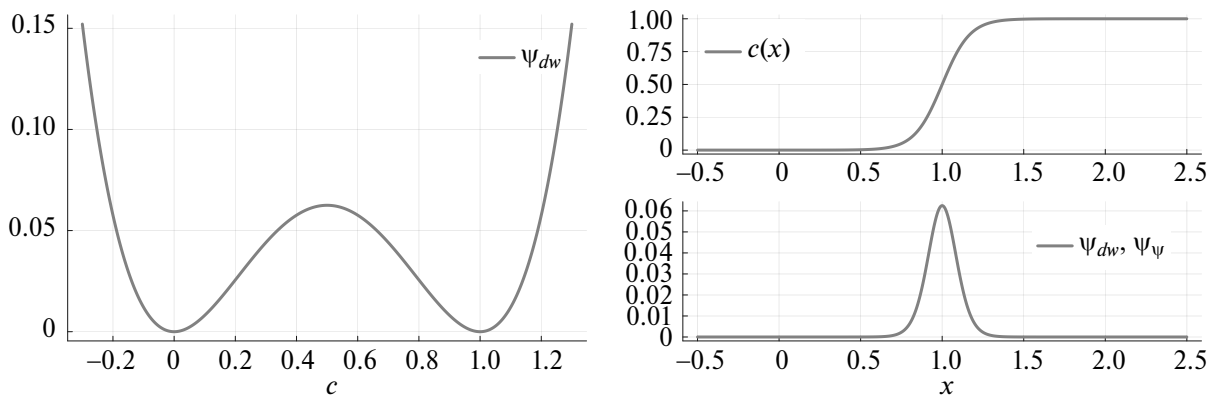
$$\Psi[c, \nabla c] = \int_{\Omega} \psi(c, \nabla c) dx, \quad (2.4)$$

где $c = c_1$,

$$\psi(c, \nabla c) = F(c) + \frac{\epsilon^2}{2} |\nabla c|^2, \quad F(c) = c^2(1 - c)^2, \quad (2.5)$$

где ψ — плотность свободной энергии системы, F — там называемый двухъямный потенциал, $\epsilon > 0$ — малый параметр. Первое слагаемое в (2.5) описывает «разделяющую» часть свободной энергии и обеспечивает несмешивание фаз (и, как следствие, областей постоянных значений концентраций $c_{1,2}$). Второе слагаемое в (2.5) позволяет отнести к диффузной границе раздела фаз заданную энергию. Входящий в него параметр ϵ определяет толщину диффузной границы, разделяющей «чистые» фазы. Учет градиентных членов позволяет учесть зависимость энергии системы не только от количества каждого компонента, но и от формы областей, которые они занимают. Такой вид свободной энергии является типичным и возникает во многих моделях слабонелокальной (или градиентной) термомеханики.

Отметим, что потенциал $F(c)$ в (2.5) является эмпирическим. Его характеристическим свойством является наличие двух минимумов при значениях $c = 0, 1$, которые отвечают «чистым» фазам. Состоянию $c = 1/2$ соответствует максимум $F(c)$; это состояние неустойчиво и препятствует образованию смеси фаз. Более точной и физически обоснованной моделью является так называемый логарифмический потенциал (см. [63]), удовлетворяющий тем же характеристическим условиям (указанный выше вид энергии может быть обоснованно получен как его аппроксимация).



Фиг. 1. Слева: вид разделяющей части ψ_{dw} свободной энергии. Справа: типичное решение задачи и распределение энергии.

Приведенное обсуждение иллюстрируется на фиг. 1. Слева на нем показан вид разделяющей части $\psi_{dw} \equiv F$ свободной энергии. Видно, что соответствующая зависимость имеет два минимума, соответствующие «чистым» фазам, то есть состояниям $c = c_1 = 0, c_2 = 1$ и $c = c_1 = 1, c_2 = 0$. Состояние системы в окрестности точки $c = 1/2$ является неустойчивым. На этом же рисунке справа показано типичное стационарное решение системы и графики различных членов свободной энергии ψ . Видно, что градиентная часть свободной энергии отлична от нуля лишь в пределах диффузной границы. Одновременно с этим, «разделяющая» часть свободной энергии равна нулю вне области, занятой диффузной границей.

Далее можно показать, что в случае, когда c является консервативной величиной (как в рассматриваемом здесь случае), кинетическое уравнение, описывающее эволюцию поля $c = c(x, t)$, имеет вид (см. [6], [63])

$$\frac{\partial c}{\partial t} = -\nabla \cdot (-M \nabla \mu), \quad \mu = \frac{\delta \psi[c, \nabla c]}{\delta c}, \quad (2.6)$$

где $\delta(\cdot)/\delta c$ — функциональная производная (производная Гато). При постоянном M в силу (2.5) последнее уравнение принимает вид

$$\frac{1}{M} \frac{\partial c}{\partial t} = \Delta \mu, \quad \mu = -\epsilon^2 \Delta c + F'(c),$$

где Δ — лапласиан, что в рассматриваемом одномерном случае в точности совпадает с (2.1).

Можно показать (см. [12], [63]), что решения уравнения (2.1) обладают следующими фундаментальными свойствами: на решениях $c = c(x, t)$ справедливо

$$\frac{d}{dt} \int_{\Omega} c = 0, \quad \frac{d}{dt} \Psi[c, \nabla c] \leq 0.$$

Другими словами, c является консервативной величиной, а свободная энергия невозрастающей функцией на решениях уравнений (2.1) или (2.6) с однородными граничными условиями (2.2).

Это сразу же накладывает соответствующие ограничения на численный алгоритм решения уравнения (2.1) — он должен быть консервативным (добиться этого достаточно легко) и энергетически — устойчивым в смысле выполнения дискретного эквивалента последнего неравенства (это более сложная задача). Помимо этого, конечномерная задача должна быть однозначна разрешима при заданном шаге по времени — как показала практика, добиться этого не так просто (это связано с тем, потенциал $F(c)$ не является выпуклым и, как следствие, конечномерное решение может «оказаться» в локальном минимуме, не являющимся глобальным и не достижимым из начального состояния).

3. ВЫЧИСЛИТЕЛЬНЫЕ АЛГОРИТМЫ

В настоящем разделе описаны вычислительные алгоритмы, применяемые далее для решения уравнения Кана-Хилларда. В первом разделе рассмотрены классические алгоритмы различной степени неясности для интегрирования уравнения Кана-Хилларда по времени. По этой причине рассмотрены только полудискретные аппроксимации. Второй раздел посвящен описанию полностью дискретных алгоритмов, с учетом конкретного способа аппроксимации по пространству. Наконец, в третьем разделе приводится описание новых алгоритмов на основе метода ЛИМ, предлагаемых в настоящей работе.

3.1. Классические алгоритмы

Рассмотрим ряд алгоритмов, которые далее будут принимать участие в сравнительном тестировании методов. Все они хорошо известны и описаны в многочисленной литературе. Для конкретики рассмотрим те из них, которые представлены в работе [64]. Все приведенные в настоящем разделе схемы аппроксимируют уравнение (2.1) и являются консервативными.

Всюду далее пространственные производные аппроксимируются методом конечных разностей. Будем считать, что для аппроксимации уравнения в области Ω используется равномерная сетка с шагом $h = 1/N$ и узлами $x_i = (i - 1/2)h$, где $i = \overline{0, N+1}$, а N — число ячеек сетки. Границы $x = 0$ и $x = 1$ области соответствуют полуцелым узлам $x_{1/2} = (x_0 + x_1)/2 = 0$ и $x_{N+1/2} = (x_N + x_{N+1})/2 = 1$. Узлы $x_1, x_2, \dots, x_N$ являются внутренними узлами области. Для аппроксимаций по времени используется равномерная сетка с шагом τ с узлами $t_n = n\tau$, $n = 0, 1, 2, \dots$. Таким образом, решение конечномерной задачи определено в точках (x_i, t_j) , $i = \overline{1, N}$, $j = 0, 1, \dots$. Соответствующие значения сеточных функций c_h будем обозначать как $c_i^n = c(x_i, t_n)$.

Всюду далее Δ_h обозначает стандартную трехточечную конечно-разностную аппроксимацию оператора Лапласа. В рассматриваемом пространственно-одномерном случае его значение в узлах $i = \overline{1, N-1}$ пространственной сетки определено как

$$\Delta_h c_i = \frac{1}{h^2} (c_{i+1} - 2c_i + c_{i-1}).$$

К «граничным» узлам $i = 0, N$ отнесены уравнения, аппроксимирующее однородное граничное условие Неймана (2.2) для аргумента оператора Лапласа,

$$c_0 - c_1 = 0, \quad c_N - c_{N-1} = 0. \quad (3.1)$$

Результирующий сеточный оператор, учитывающий дискретные граничные условия, далее будем также обозначать символом Δ_h .

Явная схема (explicit Euler) представляет собой аппроксимации уравнения (2.1) простейшего типа и имеет вид

$$\frac{1}{M} \frac{c_i^{n+1} - c_i^n}{\tau} = \Delta_h \mu_i^n, \quad \mu_i^n = F'(c_i^n) - \epsilon^2 \Delta_h c_i^n.$$

Схема имеет первый порядок аппроксимации по времени, второй — по пространству, является линейной и устойчивой при $\tau \sim h^4$, не обладает свойством энергетической устойчивости.

Неявная схема (implicit Euler) имеет вид

$$\frac{1}{M} \frac{c_i^{n+1} - c_i^n}{\tau} = \Delta_h \mu_i^{n+1}, \quad \mu_i^{n+1} = F'(c_i^{n+1}) - \epsilon^2 \Delta_h c_i^{n+1}.$$

Схема имеет первый порядок аппроксимации по времени, второй — по пространству, является нелинейной, однозначно разрешима и градиентно-устойчива только при достаточно маленьких шагах по времени, см. табл. 1.

Схема Кранка-Николсон (КН) имеет вид

$$\frac{1}{M} \frac{c_i^{n+1} - c_i^n}{\tau} = \Delta_h \mu_i^{n+1/2}, \quad \mu_i^{n+1/2} = \frac{1}{2} (\mu_i^n + \mu_i^{n+1}),$$

где μ_i^n, μ_i^{n+1} были определены выше. Схема имеет второй порядок аппроксимации по времени и пространству, является нелинейной, градиентно-устойчива и однозначно разрешима при достаточно маленьких шагах по времени, см. табл. 1.

Полунеявная схема (semi-implicit Euler) имеет вид

$$\frac{1}{M} \frac{c_i^{n+1} - c_i^n}{\tau} = \Delta_h \mu_i^{n+1}, \quad \mu_i^{n+1} = F'(c_i^n) - \epsilon^2 \Delta_h c_i^{n+1}.$$

Схема имеет первый порядок аппроксимации по времени, второй — по пространству, является линейной, однозначно разрешима, не обладает свойством энергетической устойчивости.

3.2. Расщепление Эйра

Приведенные выше схемы не обладают свойством энергетической устойчивости и/или не являются однозначно разрешимыми при достаточно больших шагах по времени (см. [12], [13]). Известный в настоящее время эффективный и общий способ построения разностных схем, как однозначно разрешимых при любых шагах по времени, так и энергетически устойчивых, был предложен в основополагающих работах Дэвида Эйра

Таблица 1. Свойства разностных схем по [12–15, 40] (* — при надлежащем выборе параметров расщепления, см. [40]).

Схема	Линейность	Град.-устойчивость	Разрешимость
Явная	Да	Нет	Да
Неявная	Нет	Условная, $\tau \leq \frac{1}{4}h^2$	$\tau \leq \frac{1}{18}h^2$
Кранка-Николсон	Нет	Условная	$\tau \leq \frac{1}{9}h^2$
Полунеявная	Да	Нет	Да
Нелин.-стаб.	Нет	Безусловная	Да
Линейно-стаб.	Да	Безусловная*	Да

(David J. Eyre) (см. [14], [15]). В силу того, что далее соответствующие математический аппарат будет использован при построении новых схем, рассмотрим этот вопрос подробнее.

Суть подхода Эйра заключается в следующем. Плотность свободной энергии ψ системы, см. (2.4), не является выпуклым функционалом. Это связано с тем, что разделяющая часть свободной энергии является «двухмянным» потенциалом и невыпукла. Предположим, тем не менее, что ψ можно представить в виде суммы двух функционалов так, что

$$\psi = \psi_c + \psi_e, \quad (3.2)$$

где ψ_c и $-\psi_e$ — *выпуклые* функции. Индексы «с» («contraction») и «е» («expansion») отражают физическую суть отдельных слагаемых. Здесь «выпуклость» функции f означает положительную определенность соответствующего гессиана, см. [13]; конкретно в рассматриваемом случае функция $f(x)$ выпукла, если $f''(x) \equiv d^2f/dx^2 \geq 0$. Тогда полудискретные (только по времени) аппроксимации вида

$$\frac{1}{M} \frac{c^{n+1} - c^n}{\tau} = \Delta [\mu_c(c^{n+1}) + \mu_e(c^n)], \quad (3.3)$$

где

$$\mu_{c,e} = \frac{\delta \psi_{c,e}}{\delta c}, \quad \mu = \mu_c + \mu_e,$$

обеспечивают выполнение следующего неравенства (означающего градиентную устойчивость схемы):

$$\psi(c^{n+1}) \leq \psi(c^n)$$

для любых шагов по времени $\tau > 0$. При этом выпуклость части ψ_c энергии, учитываемой в схеме неявно, гарантирует однозначную разрешимость (3.3).

Рассмотрим конкретный вид функционала энергии (2.4). С учетом того, что градиентное слагаемое в (2.4) относится к выпуклой части свободной энергии, расщепление ψ можно представить в виде

$$\psi_c = \frac{1}{2} |\nabla c|^2 + F_c, \quad \psi_e = F_e, \quad F_c + F_e = F.$$

Таким образом, построение расщепления энергии сводится к построению расщепления ее «разделяющей» части F так, чтобы выполнялось

$$F = F_c + F_e, \quad F_c'' \geq 0, \quad -F_e'' \geq 0. \quad (3.4)$$

Выражение для химического потенциала в этом случае принимает вид

$$\mu = \mu_c + \mu_e, \quad \mu_c = -\epsilon^2 \Delta c + F_c', \quad \mu_e = F_e',$$

а полудискретную схему (3.3) можно записать так:

$$\frac{1}{M} \frac{c^{n+1} - c^n}{\tau} = -\epsilon^2 \Delta^2 c + \Delta F_c'(c^{n+1}) + \Delta F_e'(c^n). \quad (3.5)$$

Расщепление вида (3.4) может быть построено разными способами. Один из них основан непосредственно на представлении разделяющей части F свободной энергии в виде двух слагаемых, удовлетворяющих соответствующим свойствам (градиентная часть свободной энергии (см. (2.5)) всегда относится к выпуклой части ψ

свободной энергии). Другой способ заключается в том, что выбирается некоторая регуляризирующая функция F_Γ , так что

$$F_c = F + F_\Gamma, \quad F_e = F - F_\Gamma.$$

Содержательно оба способа совпадают, однако технически, имея ввиду возможные обобщения, их удобнее различать.

Рассмотрим конкретное выражение для F , определяемое в настоящей работе в соответствии с (2.5). Для него справедливо

$$F(c) = c^2(1 - c^2), \quad F'(c) = 2c(2c^2 - 3c + 1), \quad F''(c) = 2(6c^2 - 6c + 1).$$

Корнями уравнения $F''(c) = 0$ являются $c_{1,2} = (3 \pm \sqrt{3})/6$, и, таким образом, $F''(c) < 0$ при $(3 - \sqrt{3})/6 < c < (3 + \sqrt{3})/6$; $F''(c) \geq 0$ при $c \leq (3 - \sqrt{3})/6$ и $c \geq (3 + \sqrt{3})/6$. Минимальное значение $F''(c)$ достигается в точке $c = 1/2$ и равняется -1 . Максимальное значение $F''(c)$ достигается при $c = 0$ и $c = 1$ и равняется 2. Таким образом, чтобы обеспечить выпуклость F_c и $-F_e$, достаточно положить

$$F_c = F(c) + \frac{1}{2}c^2, \quad F_e = -\frac{1}{2}c^2.$$

В этом случае приходим к схеме *нелинейно-стабилизированного расщепления* (НЛСР или non-linearly stabilized splitting, NLSS) вида:

$$\frac{1}{M} \frac{c_i^{n+1} - c_i^n}{\tau} = \Delta_h \mu_i^{n+1}, \quad \mu_i^{n+1} = [F'(c_i^{n+1}) + c_i^{n+1}] - c_i^n - \epsilon^2 \Delta_h c_i^{n+1}. \quad (3.6)$$

Ее также можно записать в виде

$$\frac{1}{M\tau} (c_i^{n+1} - c_i^n) = (\Delta_h - \epsilon^2 \Delta_h^2) c_i^{n+1} + \Delta_h F'(c_i^{n+1}) - \Delta_h c_i^n.$$

Схема имеет первый порядок аппроксимации по времени, второй — по пространству, является неявной, нелинейной, энергетически устойчива при произвольных шагах по времени.

Если же положить

$$F_c = c^2, \quad F_e = -c^2 + F(c),$$

то приходим к схеме *линейно-стабилизированного расщепления* (ЛСР или linearly stabilized splitting, LSS), которая имеет вид

$$\frac{1}{M} \frac{c_i^{n+1} - c_i^n}{\tau} = \Delta_h \mu_i^{n+1}, \quad \mu_i^{n+1} = 2c_i^{n+1} + [F'(c_i^n) - 2c_i^n] - \epsilon^2 \Delta_h c_i^{n+1}.$$

Для дальнейшего использования ее удобно записать в виде

$$\frac{1}{M\tau} (c_i^{n+1} - c_i^n) = (2\Delta_h - \epsilon^2 \Delta_h^2) c_i^{n+1} + \Delta_h F'(c_i^n) - 2\Delta_h c_i^n.$$

Схема имеет первый порядок аппроксимации по времени, второй — по пространству, является неявной, энергетически устойчива при произвольных шагах по времени при правильно выбранных параметрах расщепления (в общем случае, зависящих от решения задачи, см. [40]). В отличие от нелинейно-стабилизированной схемы (3.6) она является *линейной* относительно неизвестных на верхнем временном слое. Отметим, что в работе [13] анализируется целое трехпараметрическое семейство схем указанного вида.

Свойства приведенных выше схем в компактном виде приведены в табл. 1.

В полностью дискретном случае градиентная устойчивость схемы предполагает аппроксимацию функционала свободной энергии (2.4). В настоящей работе она согласована с используемой аппроксимацией уравнения и имеет вид

$$\Psi_h(\mathbf{c}) = h \sum_{i=1}^N F'(c_i) + \frac{1}{2} \epsilon^2 h \sum_{i=1}^N \frac{(c_{i+1} - c_i)^2}{h^2}. \quad (3.7)$$

Пусть далее матрица $\mathbf{A}$ размерности $N \times N$ соответствует указанной выше аппроксимации $-\Delta_h$ оператора $-\Delta$ с учетом однородных граничных условий Неймана (3.1). Дискретизация по пространству сводит исходную начально-краевую задачу (2.1)–(2.3) к задаче Коши

$$\frac{1}{M} \mathbf{c}' = -\mathbf{A} (F'(\mathbf{c}) + \epsilon^2 \mathbf{A} \mathbf{c}), \quad \mathbf{c}(0) = \mathbf{c}^0, \quad (3.8)$$

где компоненты вектор–функции $\mathbf{c}(t) : \mathbb{R} \rightarrow \mathbb{R}^N$ — это значения искомого решения на пространственной сетке, а вектор $\mathbf{c}^0$ содержит значения на сетке заданной функции $c^0(x)$.

Пусть $\mathbf{c}^n$ — вектор решения длины N на временном слое n . Тогда приведенные выше схемы могут быть компактно записаны следующим образом:

- явная (explicit Euler):

$$\frac{\mathbf{c}^{n+1} - \mathbf{c}^n}{\tau} = -\mathbf{A} (F'(\mathbf{c}^n) + \epsilon^2 \mathbf{A} \mathbf{c}^n); \quad (3.9)$$

- неявная (implicit Euler):

$$\frac{\mathbf{c}^{n+1} - \mathbf{c}^n}{\tau} = -\mathbf{A} (F'(\mathbf{c}^{n+1}) + \epsilon^2 \mathbf{A} \mathbf{c}^{n+1}); \quad (3.10)$$

- Кранка–Николсон (КН):

$$\frac{\mathbf{c}^{n+1} - \mathbf{c}^n}{\tau} = -\frac{1}{2} \mathbf{A} (F'(\mathbf{c}^{n+1}) + \epsilon^2 \mathbf{A} \mathbf{c}^{n+1} + F'(\mathbf{c}^n) + \epsilon^2 \mathbf{A} \mathbf{c}^n); \quad (3.11)$$

- полунявная:

$$\frac{\mathbf{c}^{n+1} - \mathbf{c}^n}{\tau} = -\mathbf{A} (F'(\mathbf{c}^n) + \epsilon^2 \mathbf{A} \mathbf{c}^{n+1}); \quad (3.12)$$

- линейно-стабилизированное расщепление (ЛСР):

$$\frac{\mathbf{c}^{n+1} - \mathbf{c}^n}{\tau} = -\mathbf{A} (F'(\mathbf{c}^n) - 2\mathbf{c}^n + 2\mathbf{c}^{n+1} + \epsilon^2 \mathbf{A} \mathbf{c}^{n+1}); \quad (3.13)$$

- нелинейно-стабилизированное расщепление (НЛСР):

$$\frac{\mathbf{c}^{n+1} - \mathbf{c}^n}{\tau} = -\mathbf{A} (F'(\mathbf{c}^{n+1}) + \mathbf{c}^{n+1} - \mathbf{c}^n + \epsilon^2 \mathbf{A} \mathbf{c}^{n+1}). \quad (3.14)$$

Описанные выше схемы имеют различную вычислительную сложность. Неявные нелинейные схемы (полностью неявная схема, схема типа Кранка–Николсон, нелинейно-стабилизированная схема) требуют решения на каждом временном слое системы нелинейных алгебраических уравнений, например, методом Ньютона; в свою очередь, на каждой ньютоновской итерации необходимо решить систему линейных алгебраических уравнений. Неявные, но линейные схемы (полунявная схема, линейно-стабилизированная схема) требуют решения только одной системы линейных алгебраических уравнений для определения решения на каждом временном слое. Наконец, явные схемы не требуют ни того, ни другого и имеют минимальную «стоимость» расчета решения на шаг по времени. Вместе с тем, неявные схемы, как правило, позволяют использовать существенно большие шаги по времени, чем явные. С другой стороны, решение системы линейных алгебраических уравнений в методе Ньютона, или независимо, является сложной и вычислительно затратной процедурой в том случае, если рассматриваются постановки с высокими сеточными размерностями. Это связано с тем, что итерационные методы решения систем линейных алгебраических уравнений, обеспечивающие высокую скорость сходимости итераций, обычно не допускают эффективную параллельную реализацию, особенно с использованием гибридных вычислительных систем с графическими ускорителями. Одновременно с этим, простые методы типа метода простой итерации или последовательной верхней релаксации, допускают эффективную параллельную реализацию, но сходятся слишком медленно для практического применения.

Целью настоящей работы является разработка эффективного вычислительного алгоритма, с одной стороны — градиентно-устойчивого и допускающего существенно большие временные шаги, чем чисто явная схема, а с другой — допускающего эффективную вычислительную реализацию в том числе с применением высокопроизводительных вычислительных систем с графическими ускорителями. Основной идеей алгоритма является использование линейно-стабилизированной схемы Эйра и использование модифицированной схемы локальных итераций ЛИМ для решения уравнений на временном слое. При этом расщепление Эйра позволяет использовать большие шаги по времени при сохранении градиентной устойчивости, а применение ЛИМ — обеспечивает вычислительную эффективность.

3.3. ЛИМ, модифицированная схема локальных итераций

Схему ЛИМ можно рассматривать как специальную явную схему, где используется многочлен Чебышёва такого порядка p , чтобы схема была устойчива для выбранного шага по времени τ и обладала свойством монотонности (неотрицательности решения). Сформулируем схему ЛИМ на основе схемы линейно-стабилизированного расщепления ЛСР (3.13), которую запишем в виде

$$\mathbf{c}^{n+1} = [\mathbf{I} + \tau \hat{\mathbf{A}}]^{-1} \hat{\mathbf{f}}^n, \quad \hat{\mathbf{A}} = \mathbf{A}(2\mathbf{I} + \epsilon^2 \mathbf{A}), \quad \hat{\mathbf{f}}^n = \mathbf{c}^n + \tau \mathbf{A}(2\mathbf{c}^n - F'(\mathbf{c}^n)), \quad (3.15)$$

где $\mathbf{I}$ — единичная матрица $N \times N$. Подчеркнем, что обратная матрица здесь, разумеется, не вычисляется, а решается линейная система. В схеме ЛИМ обратная матрицы заменяется специально выбранным матричным многочленом Чебышёва. Пусть λ_∞ — оценка наибольшего собственного значения $\hat{\mathbf{A}}$ (на практике можно положить $\lambda_\infty = \|\hat{\mathbf{A}}\|_1 = \max_j \sum_i |\hat{a}_{ij}|$). Определим порядок p многочлена Чебышёва, для которого будет выполняться устойчивость, так:

$$p = \left\lceil \frac{\pi}{4} \sqrt{\tau \lambda_\infty + 1} \right\rceil, \quad (3.16)$$

где, для $x \in \mathbb{R}$, $\lceil x \rceil$ обозначает наименьшее целое число, большее или равное x . Вычислим корни β_m многочлена Чебышёва степени p ,

$$\{\beta_m, m = 1, \dots, p\} = \left\{ \cos \pi \frac{2i-1}{2p}, i = 1, \dots, p \right\},$$

и упорядочим их в порядке, гарантирующем устойчивость, и так, что корень $\beta_1 = \cos(\pi/2p)$ является первым по порядку. Положим $z_1 = \beta_1$ и определим параметры чебышёвских итераций

$$a_m = \frac{\lambda_\infty}{1 + z_1} (z_1 - \beta_m), \quad m = 1, \dots, p.$$

Обозначим $\mathbf{y}^{(0)} = \mathbf{c}^n$ и получим решение $\mathbf{c}^{n+1}$ схемы ЛИМ на следующем слое, выполняя $2p - 1$ итераций Чебышёва следующим образом:

$$\begin{aligned} \mathbf{y}^{(m)} &= \frac{1}{1 + \tau a_m} \left(\mathbf{c}^n + \tau a_m \mathbf{y}^{(m-1)} + \tau (\hat{\mathbf{f}}^n - \hat{\mathbf{A}} \mathbf{y}^{(m-1)}) \right), \quad m = 1, \dots, p, \\ \mathbf{y}^{(p+m-1)} &= \frac{1}{1 + \tau a_m} \left(\mathbf{c}^n + \tau a_m \mathbf{y}^{(m-1)} + \tau (\hat{\mathbf{f}}^n - \hat{\mathbf{A}} \mathbf{y}^{(m-1)}) \right), \quad m = 2, \dots, p, \\ \mathbf{c}^{n+1} &= \mathbf{y}^{(2p-1)}. \end{aligned} \quad (3.17)$$

Заметим, что $a_1 = 0$ и, как нетрудно увидеть, решение $\mathbf{y}^{(1)}$ на первой итерации совпадает с решением на следующем шаге по времени явной схемы, сами же чебышёвские итерации начинаются с вычисления $\mathbf{y}^{(2)}$. Если в (3.17) ограничиться первой группой итераций, где вычисляются $\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(p)}$, то, положив $\mathbf{c}^{n+1} = \mathbf{y}^{(p)}$, получим обычную, немодифицированную схему локальных итераций (схему ЛИ). Как отмечалось выше, в этой, обычной, схеме ЛИ гарантируется устойчивость (при решении параболических задач), но свойство монотонности решения не гарантируется. Как видим, в схеме ЛИМ (3.17) при выполнении второй группы итераций, где вычисляются $\mathbf{y}^{(p+1)}, \dots, \mathbf{y}^{(2p-1)}$, чебышёвские итерации повторяются (с теми же параметрами a_m), но первый итерационный шаг с $a_1 = 0$ (дающий явную схему) не делается. Можно показать (см. [58]), что формулы (3.17) представляются в операторном виде

$$\mathbf{c}^{n+1} = (\mathbf{I} - \mathbf{F}_p^2) [\mathbf{I} + \tau \hat{\mathbf{A}}]^{-1} \hat{\mathbf{f}}^n, \quad (3.18)$$

где $\hat{\mathbf{A}}$ и $\hat{\mathbf{f}}^n$ определены в (3.15), а $\mathbf{F}_p$ — оператор многочлена Чебышёва,

$$\mathbf{F}_p = \prod_{m=p}^{m=1} \left(\mathbf{I} - \frac{1}{1 + \tau a_m} (\mathbf{I} + \tau \hat{\mathbf{A}}) \right).$$

Если в формуле (3.18) заменить $\mathbf{F}_p^2$ на $\mathbf{F}_p$, то получаем операторное представление обычной схемы ЛИ. Таким образом, как видим, в схеме ЛИМ вместо многочлена Чебышёва (как в обычной схеме ЛИ), используется квадрат этого многочлена, что и приводит к монотонности схемы.

Поскольку рассматриваемая схема ЛИМ (3.17) получена нами на основе линейно—стабилизированного расщепления (3.13), будем называть ее схемой ЛИМ–ЛСР. Чтобы отследить эффект применения расщепления Эйра в этой схеме, в представленные ниже численные тесты включим еще одну модификацию схемы ЛИМ, получаемую из линеаризации чисто неявной схемы (3.10). Проведем такую линеаризацию, аппроксимируя нелинейный неявный член в (3.10) следующим образом:

$$F'(\mathbf{c}^{n+1}) \approx F'(\mathbf{c}^n) + J_n(\mathbf{c}^{n+1} - \mathbf{c}^n), \quad (3.19)$$

где J_n — матрица Якоби отображения F' , вычисленная в $\mathbf{c}^n$. Подставка приближения (3.19) в (3.10) приводит к линеаризованной неявной (ЛН) схеме

$$\mathbf{c}^{n+1} = [\mathbf{I} + \tau \tilde{\mathbf{A}}_n]^{-1} \tilde{\mathbf{f}}^n, \quad \tilde{\mathbf{A}} = \mathbf{A}(J_n + \epsilon^2 \mathbf{A}), \quad \tilde{\mathbf{f}}^n = \mathbf{c}^n + \tau \mathbf{A}(J_n \mathbf{c}^n - F'(\mathbf{c}^n)), \quad (3.20)$$

где, так же как и в схеме (3.15), подразумевается решение линейной системы, так что обратная матрица не вычисляется. Сравнивая схемы (3.15) и (3.20), можно заметить, что заменяя J_n в (3.20) на $2\mathbf{I}$, получим схему (3.15).

Схема ЛИМ на основе схемы (3.20) формулируется в точности так же, как и на основе (3.15). Для этого достаточно повторить приведенные выше выкладки, заменяя $\tilde{\mathbf{A}}$ и $\tilde{\mathbf{f}}_n$ на, соответственно, $\tilde{\mathbf{A}}_n$ и $\tilde{\mathbf{f}}_n$. Назовём такую схему ЛИМ–ЛН (схема ЛИМ линеаризованная неявная).

Вычислительная эффективность схем ЛИМ следует из формулы числа итераций (3.16) и следующих соображений (см. [54], [58]). Если в решаемой задаче интервал по времени, на котором требуется просчитать задачу, увеличивается в s раз, то так же в s раз увеличиваются затраты при расчете явной схемой (в s раз возрастает требуемое число шагов во времени). Такое же увеличение затрат наблюдается и при увеличении в s раз верхней спектральной границы λ_∞ (поскольку шаг явной схемы ограничен условием устойчивости, обычно имеющего вид $\tau \leq 2/\lambda_\infty$). Напомним, что для уравнения Кана–Хилларда, вообще говоря, имеем $\lambda_\infty \sim h^{-4}$. Для схем ЛИМ вычислительные затраты растут медленнее: увеличение временного интервала или верхней спектральной границы λ_∞ в s раз означает увеличение затрат примерно в $\sqrt{s}$ раз (поскольку число чебышёвских итераций $p \sim (\tau \lambda_\infty)^{1/2}$, см. (3.16)). Из приведенных оценок ясно, что выигрыш по вычислительным затратам, достигаемый схемами ЛИМ, будет расти с ростом размерности задачи $N = 1/h$ (см. [54], [58]). Тем не менее, следует учитывать, что на практике потенциальный выигрыш может быть ограничен другими факторами, например, условиями точности расчета.

4. ВЫЧИСЛИТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ

4.1. Тесты на градиентную устойчивость

В представленных здесь тестах решается начально—краевая задача (2.1)–(2.3) для одномерного уравнения Кана–Хилларда. Для дискретизации по пространству используются стандартные конечные разности второго порядка точности на равномерной сетке, описанные в начале п. 3.1. В результате дискретизации по пространству из исходной начально—краевой задачи получаем задачу Коши (3.8). Ее численное интегрирование проводим схемами (3.9)–(3.14) и новыми схемами локальных итераций (3.17) на основе схем (3.13) и (3.20) (ЛИМ–ЛСР и ЛИМ–ЛН). Значения ϵ будем задавать следующими двумя способами:

$$\epsilon = \epsilon_4(h), \quad \epsilon_m(h) \equiv \frac{hm}{2\sqrt{2} \operatorname{arth}(9/10)}, \quad (4.1)$$

$$\epsilon = \epsilon_4(1/64), \quad (4.2)$$

т.е. ϵ задается либо в зависимости от шага пространственной сетки h (способ (4.1), либо используется фиксированное значение $\epsilon = \epsilon_4(1/64)$ с сетки $N = 64$ (способ (4.2)). Заметим, что $m = 4$ в формуле (4.1) определяет характерное число ячеек пространственной сетки, на которых стационарное решение меняется от минимального до максимального значения.

В табл. 2 для тестируемых схем представлены величины максимальных шагов по времени, при которых наблюдается градиентная устойчивость. При этом используем сеточно—зависимые значения ϵ (4.1) и считаем, что схема градиентно устойчива, если наблюдается пошаговое возрастание дискретной энергии (3.7) не более, чем на 1%, т.е. на каждом шаге по времени n выполняется условие

$$\psi_h(\mathbf{c}^{n+1}) \leq 1.01 \psi_h(\mathbf{c}^n). \quad (4.3)$$

Градиентная устойчивость тестировалась на нескольких векторах начальных значений $\mathbf{c}^0$, где каждая компонента вектора $\mathbf{c}^0$ бралась независимой одинаково распределенной случайной величиной на отрезке $[0, 1]$, округленной с точностью до сотых. В табл. 2 также включена линеаризованная неявная схема (3.20).

Таблица 2. Максимальные значения шагов по времени τ , при которых схемы градиентно устойчивы (т.е. выполняется условие (4.3)), для $\epsilon = \epsilon_4(h)$

h	1/32	1/64	1/128	1/256
Явная	8.8×10^{-5}	2.1×10^{-5}	5.4×10^{-6}	1.3×10^{-6}
Неявная	5.1×10^{-3}	1.3×10^{-3}	3.0×10^{-4}	8.1×10^{-5}
КН	2.4×10^{-3}	6.8×10^{-4}	1.7×10^{-4}	3.8×10^{-5}
Полунеявная	2.2×10^{-3}	5.7×10^{-4}	1.9×10^{-4}	4.0×10^{-5}
ЛСР	∞	∞	∞	∞
НЛСР	∞	∞	∞	∞
ЛИМ-ЛСР	∞	∞	∞	∞
ЛИМ-ЛН	9.9×10^{-3}	2.8×10^{-3}	5.8×10^{-4}	1.2×10^{-4}
ЛН	5.8×10^{-3}	1.6×10^{-3}	3.2×10^{-4}	8.3×10^{-5}

Как видим из таблицы, только схемы, основанные на расщеплении Эйра, являются безусловно градиентно устойчивыми. В тестах схема считалась безусловно устойчивой, если устойчивость выполнялась для $\tau \leq 500$. Далее отметим, что максимальный шаг для неявной схемы превосходит максимальный шаг явной схемы примерно в одно и то же число раз (≈ 60) для всех h . Сравнение схемы ЛН и неявной схемы говорит о том, что линеаризация не ухудшает градиентную устойчивость неявной схемы. Схема ЛИМ-ЛСР сохраняет безусловную градиентную устойчивость схемы ЛСР, а схема ЛИМ-ЛН — условную градиентную устойчивость схемы ЛН. Учитывая, что одна из целей данной работы — показать возможность замены неявных схем на явные устойчивые схемы типа ЛИМ, важно отметить, что обе схемы ЛИМ, как видно из табл. 2, работают успешно. При этом увеличение шага τ как в схеме ЛИМ-ЛСР, так в ЛИМ-ЛН, разумеется, означает увеличение числа чебышёвских итераций на шаг по времени. Заметим также, что представленные нами в табл. 2 значения близки к значениям в табл. 1 из работы [64].

4.2. Тесты на точность и эффективность

Цель обсуждаемых здесь тестов — проверить точность и эффективность рассмотренных схем и, в частности, выяснить, позволяют ли схемы ЛИМ получить выигрыш по вычислительным затратам по сравнению с явной схемой. Очевидно, что поскольку затраты на шаг по времени в явной схеме минимальны, повышение эффективности другими схемами возможно только при увеличении в них шага по времени. Однако увеличение шага по времени возможно, разумеется, лишь так, чтобы точность вычислений оставалась в допустимых пределах. Будем оценивать точность схем в конечный момент времени $t = T$ по референтному решению $\mathbf{c}_{\text{ref}}(T)$, которое на каждой пространственной сетке вычислялось явной схемой с очень маленьким шагом по времени ($\tau = 10^{-9}$). При этом вектор начального значения выбираем случайным образом так же, как и в предыдущей серии тестов, и затем используем на данной пространственной сетке для всех схем. Для каждой из схем достигнутую точность оцениваем по относительной норме ошибки

$$\frac{\|\mathbf{c}^n - \mathbf{c}_{\text{ref}}(T)\|}{\|\mathbf{c}_{\text{ref}}(T)\|}, \quad n = n_{\text{final}} = T/\tau, \quad (4.4)$$

где $\|\mathbf{c}\| = \sqrt{\mathbf{c}^T \mathbf{c}}$ — евклидова векторная норма.

Во всех тестах конечное время T выбираем равным $T = 0.2$. При таком T решение уже прошло начальную фазу формирования однородных областей на временах $t \approx \epsilon^2$, всем схемам требуется достаточно большое количество шагов по времени, но решение еще далеко от стационарного состояния.

Значения верхних границ спектра λ_{∞} в зависимости от шага сетки h и соответствующие максимальные значения шага по времени τ явной схемы, при котором явная схема устойчива, приведены в табл. 3. В таблице максимальные значения τ определяются двумя способами:

(а) максимальный шаг τ , обеспечивающий градиентную устойчивость (невозрастание энергии (4.3)); (б) максимальный шаг τ , при котором схемы ЛИМ-ЛН и ЛИМ-ЛСР работают в режиме явной схемы, т.е. в соотношении (3.16) получаем $p = 1$, что означает, что нет необходимости выполнять чебышёвские итерации. Последнее условие означает устойчивость явной схемы в операторной евклидовой норме, что видно из следу-

Таблица 3. Верхняя граница спектра $\lambda_\infty = \|\hat{A}\|_1$ схемы ЛИМ-ЛСР и максимально возможные значения шага по времени $\tau_{\max}$, дающие градиентную устойчивость (строка 3 таблицы) и режим явной схемы в схеме ЛИМ-ЛСР (строки 4 и 6), в зависимости от h . Как видно из таблицы, $\lambda_\infty = \mathcal{O}(h^{-2})$ для $\epsilon = \epsilon_4(h)$, $\lambda_\infty = \mathcal{O}(h^{-4})$ для $\epsilon = \epsilon_4(1/64)$ и во всех случаях $\tau_{\max} = \mathcal{O}(\lambda_\infty^{-1})$

h	1/32	1/64	1/128	1/256	1/512
$\epsilon = \epsilon_4(h)$					
$\lambda_\infty = \ \hat{A}\ _1$	2.3×10^4	9.3×10^4	3.7×10^5	1.5×10^6	6.0×10^6
Градиентная устойчивость	8.8×10^{-5}	2.1×10^{-5}	5.4×10^{-6}	1.3×10^{-6}	3.3×10^{-7}
Режим явной схемы	2.6×10^{-5}	6.6×10^{-6}	1.6×10^{-6}	4.1×10^{-7}	1.0×10^{-7}
$\epsilon = \epsilon_4(1/64)$					
$\lambda_\infty = \ \hat{A}\ _1$	1.2×10^4	9.3×10^4	1.1×10^6	1.6×10^7	2.5×10^8
Режим явной схемы	5.1×10^{-5}	6.6×10^{-6}	5.6×10^{-7}	3.8×10^{-8}	2.4×10^{-9}

ющих соображений. Если в соотношении (3.16) для схемы ЛИМ-ЛН выполняется $p = 1$, то

$$\tau \lambda_\infty \leq \left(\frac{4}{\pi}\right)^2 - 1, \quad (4.5)$$

где $\lambda_\infty = \|\tilde{A}\|_1$, а $\tilde{A}$ определена в (3.20). Если $J_n = \partial F'(\mathbf{c}^n)/\partial \mathbf{c}$, то, учитывая приближение $F'(\mathbf{c}^n) \approx F'(\mathbf{0}) + J_n(\mathbf{c}^n - \mathbf{0}) = J_n \mathbf{c}^n$, для решения явной схемы $\mathbf{c}^{n+1}$ получаем

$$\|\mathbf{c}^{n+1}\| = \|\mathbf{c}^n - \tau A(F'(\mathbf{c}^n) + \epsilon^2 A \mathbf{c}^n)\| \approx \|(I - \tau \tilde{A})\mathbf{c}^n\| \leq \|I - \tau \tilde{A}\| \|\mathbf{c}^n\|.$$

Поскольку $\tilde{A}$ — симметричная неотрицательно определенная матрица, для евклидовой операторной нормы условие $\|I - \tau \tilde{A}\|_2 \leq 1$ равносильно условию $\tau \|\tilde{A}\|_2 \leq 2$, которое следует из (4.5) (заметим, что $\|\tilde{A}\|_2 \leq \|\tilde{A}\|_1 = \lambda_\infty$).

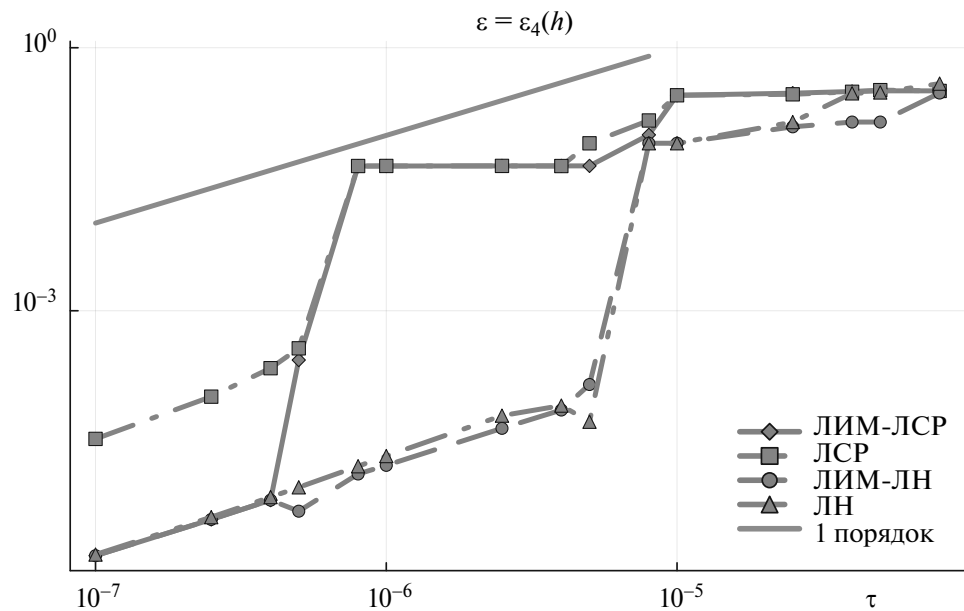
В табл. 3 зависимость верхней границы спектра λ_∞ от h дана для схемы ЛИМ-ЛСР, то есть $\lambda_\infty = \|\hat{A}\|_1$. В представленных здесь тестах во всех случаях значения $\|\hat{A}\|_1$ и $\|\tilde{A}\|_1$ оказываются настолько близкими, что число чебышёвских итераций в схемах ЛИМ-ЛСР и ЛИМ-ЛН практически во всех случаях одинаково. Значения в табл. 3 приведены для обоих способов задания ϵ , см. (4.1), (4.2). По данным таблицы легко проверить, что при выборе $\epsilon = \epsilon_4(h)$ матрица $\epsilon^2 A$ не зависит от h , так что, учитывая (3.15), получаем зависимость $\|\hat{A}\|_1 = \mathcal{O}(h^{-2})$. При втором способе задания ϵ (по формуле (4.2)) имеем $\|\hat{A}\|_1 = \mathcal{O}(h^{-4})$.

Первый порядок сходимости $\mathcal{O}(\tau)$ схем ЛИМ-ЛСР, ЛСР, ЛИМ-ЛН и ЛН подтверждается графиком на фиг. 2. Как видим из рисунка, обе схемы ЛИМ сохраняют свойства точности схем, на которых они основаны. Кроме того отметим, что схемы расщепления Эйра ЛСР и ЛИМ-ЛСР оказываются менее точными, чем схемы ЛИМ-ЛН и ЛН. В этих схемах градиентная устойчивость достигается за счет расщепления, приводящего к дополнительной ошибке.

В табл. 4–7 для явной схемы и схем ЛСР, ЛИМ-ЛСР, ЛИМ-ЛН представлены, для различных шагов по времени τ , вычислительные затраты этих схем и достигаемая ими точность. Как видим из таблиц, точность схем ЛСР и ЛИМ-ЛСР оказывается слишком низкой (будем считать значения ошибки (4.4) более 10^{-2} неприемлемо большими). Увеличение шага по времени по отношению к явной схеме, необходимое для компенсации увеличенных на шаг по времени затрат, приводит к недопустимо низкой точности. Схема ЛИМ-ЛН более точна и дает выигрыш по отношению к явной схеме на мелких сетках (в два раза на сетке $N = 512$). То, что низкая точность схем ЛСР и ЛИМ-ЛСР вызвана ошибкой расщепления Эйра, видно при переходе на схемы ЛН и ЛИМ-ЛН (значения ошибки для схемы ЛН в таблицах не показаны, они близки к значениям ошибки для схем ЛИМ-ЛН, см. фиг. 2).

Графики зависимости достигаемой точности от вычислительных затрат представлены на фиг. 3. Как видим из рисунка, ситуация для схем ЛИМ улучшается при переходе от сетки $N = 256$ к $N = 512$: на последней сетке возможно снижение затрат примерно в три раза.

В представленных тестах наблюдается довольно низкая точность всех схем кроме явной, и это обусловлено не только расщеплением Эйра. В наших тестах вектор начального решения представлял собой негладкую сеточную функцию, где каждая компонента вектора выбиралась случайным образом. Как обсуждалось выше



Фиг. 2. Сходимость (достигаемая точность в зависимости от шага по времени τ) для схем ЛИМ-ЛСР, ЛСР, ЛИМ-ЛН и ЛН на сетке $N = 256$, $\epsilon = \epsilon_4(h)$.

Таблица 4. Число матрично–векторных умножений (матвеков) и решений линейных систем (для неявных схем ЛСР и ЛН), достигнутая точность в зависимости от τ , $\epsilon = \epsilon_4(h)$, сетка по пространству $N = 64$

τ	Схема	Число матвеков / решений лин. систем $\times 10^3$	Ошибка (4.4)
5.0×10^{-5}	ЛИМ-ЛСР	12 / —	5.24×10^{-2}
	ЛСР	4 / 4	1.18×10^{-1}
	ЛИМ-ЛН	12 / —	8.34×10^{-3}
1.0×10^{-5}	явная	20 / —	1.04×10^{-4}
	ЛИМ-ЛСР	60 / —	2.52×10^{-2}
	ЛСР	20 / 20	2.92×10^{-2}
	ЛИМ-ЛН	60 / —	3.96×10^{-4}
1.0×10^{-6}	явная	200 / —	2.27×10^{-5}
	ЛИМ-ЛСР	200 / —	2.27×10^{-5}
	ЛСР	200 / 200	2.93×10^{-3}
	ЛИМ-ЛН	200 / —	2.27×10^{-5}

во введении, для таких решений наблюдается быстрое формирование областей однородности с характерными временами эволюции порядка ϵ^2 . Следовательно, для правильного отслеживания этого процесса шаг по времени должен выбираться так, чтобы $\tau \sim \mathcal{O}(\epsilon^2) = \mathcal{O}(h^2)$. Отсюда, учитывая, что $\epsilon^2 \|A\| \sim \epsilon^2 h^{-2} = \mathcal{O}(1)$, видим, что на начальной фазе интегрирования по времени шаг τ не может существенно превосходить шаг явной схемы. Он, в свою очередь, в данном случае определяется не условием устойчивости — а условием фактической точности расчета.

Чтобы протестировать устойчивость и точность рассматриваемых схем интегрирования при больших шагах по времени, проведем тесты с более гладким начальным условием. Будем задавать его интерполяцией на используемую сетку по пространству исходного негладкого начального решения с сетки $N = 64$ (на сетке $N = 64$ начальный вектор не меняется). При этом будет использоваться кусочно–кубическая эрмитова интерполяция,

Таблица 5. Число матрично–векторных умножений (матвеков) и решений линейных систем (для неявных схем ЛСР и ЛН), достигнутая точность в зависимости от τ , $\epsilon = \epsilon_4(h)$, сетка по пространству $N = 128$

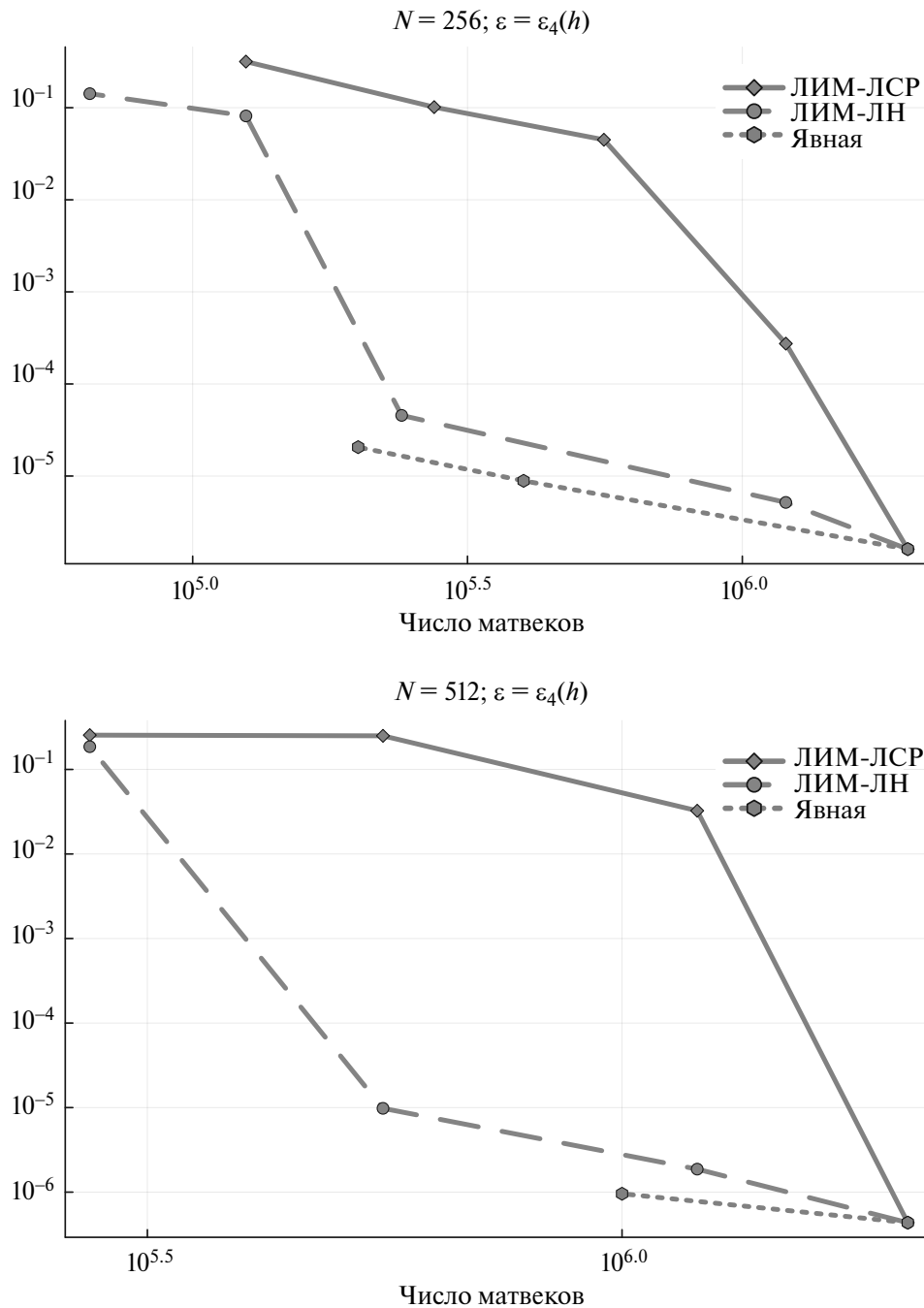
τ	Схема	Число матвеков / решений лин. систем $\times 10^3$	Ошибка (4.4)
2.5×10^{-5}	ЛИМ-ЛСР	40 / —	1.71×10^{-1}
	ЛСР	8 / 8	1.66×10^{-1}
	ЛИМ-ЛН	40 / —	9.06×10^{-2}
5.0×10^{-6}	явная	40 / —	3.37×10^{-4}
	ЛИМ-ЛСР	120 / —	3.60×10^{-2}
	ЛСР	40 / 40	1.19×10^{-2}
	ЛИМ-ЛН	120 / —	7.91×10^{-4}

Таблица 6. Число матрично–векторных умножений (матвеков) и решений линейных систем (для неявных схем ЛСР и ЛН), $\epsilon = \epsilon_4(h)$, сетка по пространству $N = 256$

τ	Схема	Число матвеков / решений лин. систем $\times 10^3$	Ошибка (4.4)
1.0×10^{-5}	ЛИМ-ЛН	140 / —	8.14×10^{-2}
5.0×10^{-6}	ЛИМ-ЛСР	200 / —	4.49×10^{-2}
	ЛСР	40 / 40	8.12×10^{-2}
	ЛИМ-ЛН	200 / —	1.44×10^{-4}
1.0×10^{-6}	явная	200 / —	2.06×10^{-5}
	ЛИМ-ЛСР	600 / —	4.48×10^{-2}
	ЛСР	200 / 200	4.48×10^{-2}
	ЛИМ-ЛН	600 / —	1.72×10^{-5}

Таблица 7. Число матрично–векторных умножений (матвеков) и решений линейных систем (для неявных схем ЛСР и ЛН), достигнутая точность в зависимости от τ , $\epsilon = \epsilon_4(h)$, сетка по пространству $N = 512$

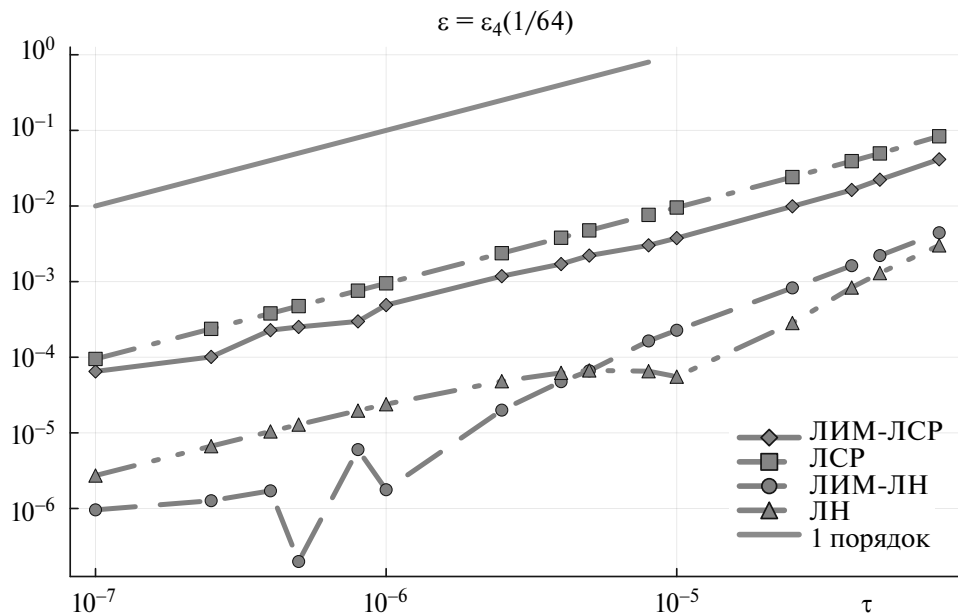
τ	Схема	Число матвеков / решений лин. систем $\times 10^6$	Ошибка (4.4)
2.0×10^{-6}	ЛИМ-ЛН	0.5 / —	8.25×10^{-6}
1.0×10^{-6}	ЛИМ-ЛСР	1 / —	3.25×10^{-2}
	ЛСР	0.2 / 0.2	3.25×10^{-2}
	ЛИМ-ЛН	1 / —	4.61×10^{-6}
2.0×10^{-7}	явная	1 / —	9.57×10^{-7}
	ЛИМ-ЛСР	3 / —	3.25×10^{-2}
	ЛСР	1 / 1	3.25×10^{-2}
	ЛИМ-ЛН	3 / —	7.35×10^{-7}



Фиг. 3. Зависимость достигаемой точности от числа матрично–векторных умножений (матвеков) для явной схемы и схем ЛИМ, сетка $N = 256$ (верхний график) и $N = 512$ (нижний график), $\varepsilon = \varepsilon_4(h)$. Увеличение τ для явной схемы невозможно из-за ограничений устойчивости.

исключающая появление новых экстремумов и дающая непрерывно–дифференцируемую функцию (в пакете octave такой метод интерполяции называется `rchip`). Кроме того, в тестах со сглаженными начальными условиями будем выбирать ε вторым способом (4.2), что позволит протестировать наши схемы в ситуации, когда норма оператора правой части растет как $\mathcal{O}(h^{-4})$, см. табл. 3.

Начнем с проверки первого порядка сходимости $\mathcal{O}(\tau)$ по графику на фиг. 4. Сравнивая его с графиком на фиг. 2, видим, что все схемы достигают гораздо более высокой точности, однако схемы расщепления Эйра ЛСР и ЛИМ–ЛСР по-прежнему менее точны, чем схемы ЛН и ЛИМ–ЛН. Небольшие осцилляции ошибки схемы ЛИМ–ЛН для $\tau \approx 10^{-6}$ вызваны изменением числа чебышёвских итераций на шаг по времени при данных τ .



Фиг. 4. Сходимость (достигаемая точность в зависимости от шага по времени τ) для схем ЛИМ-ЛСР, ЛСР, ЛИМ-ЛН и ЛН на сетке $N = 256$, $\epsilon = \epsilon_4(1/64)$

Таблица 8. Число матрично–векторных умножений (матвеков) и решений линейных систем (для неявных схем ЛСР и ЛН), достигнутая точность в зависимости от τ , $\epsilon = \epsilon_4(1/64)$, сглаженное начальное решение c^0 , сетка по пространству $N = 128$

τ	Схема	Число матвеков / решений лин. систем $\times 10^3$	Ошибка (4.4)
1.0×10^{-4}	ЛИМ-ЛСР	34 / —	7.46×10^{-2}
	ЛСР	2 / 2	1.37×10^{-1}
	ЛИМ-ЛН	34 / —	6.06×10^{-3}
5.0×10^{-6}	ЛИМ-ЛСР	200 / —	3.42×10^{-3}
	ЛСР	40 / 40	5.25×10^{-3}
	ЛИМ-ЛН	185 / —	1.09×10^{-4}
1.0×10^{-6}	явная	200 / —	3.02×10^{-5}
	ЛИМ-ЛСР	600 / —	8.14×10^{-4}
	ЛСР	200 / 200	1.04×10^{-3}
	ЛИМ-ЛН	600 / —	1.50×10^{-5}

Результаты тестов для сглаженных начальных условий и фиксированного $\epsilon = \epsilon_4(1/64)$ представлены в табл. 8–10 и на фиг. 5. Как видим, теперь обе схемы ЛИМ позволяют достичь существенного выигрыша по эффективности по сравнению с явной схемой. На сетке $N = 512$ достигается выигрыш примерно в 10 раз (см. нижний график на фиг. 5).

5. ЗАКЛЮЧЕНИЕ

Представленные результаты позволяют сделать следующие выводы.

1. Предложенные схемы (3.17) локальных итераций ЛИМ-ЛСР и ЛИМ-ЛН на основе схем неявных схем ЛСР (3.13) и ЛН (3.20) показали себя надежными в работе. Они сочетают в себе структурную простоту и параллелизуемость явных схем с устойчивостью неявных схем. Теоретические оценки эффективности для схем

Таблица 9. Число матрично–векторных умножений (матвеков) и решений линейных систем (для неявных схем ЛСР и ЛН), достигнутая точность в зависимости от τ , $\epsilon = \epsilon_4(1/64)$, сглаженное начальное решение $\mathbf{c}^0$, сетка по пространству $N = 256$

τ	Схема	Число матвеков / решений лин. систем $\times 10^6$	Ошибка (4.4)
1.0×10^{-5}	ЛИМ-ЛСР	0.38 / —	3.76×10^{-3}
	ЛСР	0.02 / 0.02	9.55×10^{-3}
	ЛИМ-ЛН	0.38 / —	2.27×10^{-4}
5.0×10^{-7}	ЛИМ-ЛСР	2 / —	2.51×10^{-4}
	ЛСР	0.4 / 0.4	4.75×10^{-4}
	ЛИМ-ЛН	2 / —	1.99×10^{-7}
1.0×10^{-7}	явная	2 / —	2.76×10^{-6}
	ЛИМ-ЛСР	6 / —	6.49×10^{-5}
	ЛСР	2 / 2	9.49×10^{-5}
	ЛИМ-ЛН	6 / —	9.59×10^{-7}

Таблица 10. Число матрично–векторных умножений (матвеков) и решений линейных систем (для неявных схем ЛСР и ЛН), достигнутая точность в зависимости от τ , $\epsilon = \epsilon_4(1/64)$, сглаженное начальное решение $\mathbf{c}^0$, сетка по пространству $N = 512$

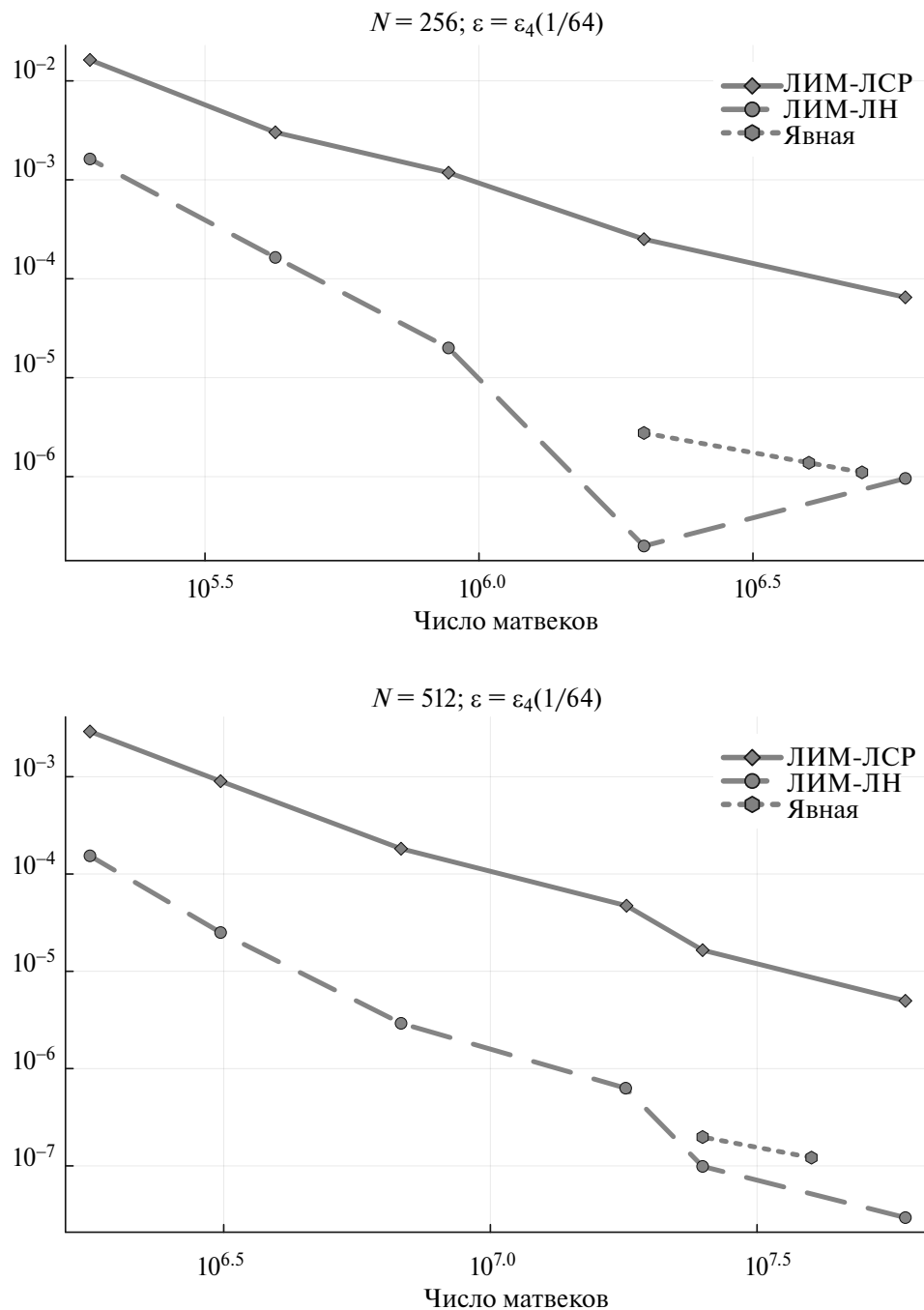
τ	Схема	Число матвеков / решений лин. систем $\times 10^6$	Ошибка (4.4)
8.0×10^{-7}	ЛИМ-ЛСР	5.75 / —	3.14×10^{-4}
	ЛСР	0.25 / 0.25	7.13×10^{-4}
	ЛИМ-ЛН	5.75 / —	3.83×10^{-6}
4.0×10^{-8}	ЛИМ-ЛСР	25 / —	1.65×10^{-5}
	ЛСР	5 / 5	3.56×10^{-5}
	ЛИМ-ЛН	25 / —	9.87×10^{-8}
8.0×10^{-9}	явная	25 / —	1.98×10^{-7}
	ЛИМ-ЛСР	75 / —	4.44×10^{-6}
	ЛСР	25 / 25	7.13×10^{-6}
	ЛИМ-ЛН	75 / —	5.14×10^{-8}

ЛИМ подтверждаются в тестах: схемы ЛИМ дают выигрыш по эффективности до 10 раз по сравнению с явной схемой. Выигрыш растет при сгущении сетки.

2. В рассмотренных численных тестах градиентно устойчивые схемы на основе расщепления Эйра оказываются менее точными, чем обычные линеаризованные схемы. В частности, обычная линеаризованная неявная схема ЛН (3.20) и основанная на ней схема локальных итераций ЛИМ-ЛН превосходят по точности схему линейно–стабилизированного расщепления ЛСР (3.13) и схему локальных итераций ЛИМ-ЛСР. В этих схемах градиентная устойчивость достигается за счет расщепления, приводящего к дополнительной ошибке.

3. В случае, если толщина диффузионной границы ϵ выбирается пропорционально размеру сетки, т.е. $\epsilon = \mathcal{O}(h)$, шаг по времени в явной схеме ограничен как $\tau = \mathcal{O}(h^2)$. Следовательно, можно ожидать, что потенциал схем локальных итераций для решения уравнений Кана–Хилларда сравним с потенциалом таких схем для параболических задач.

4. Для негладких начальных данных, например, если вектор начальных значений выбирается случайным образом, возникают дополнительные ограничения по точности на шаг по времени τ на временах $t \leq \mathcal{O}(\epsilon^2)$. По-



Фиг. 5. Зависимость достигаемой точности от числа матрично–векторных умножений (матвеков) для явной схемы и схем ЛИМ, сглаженное начальное решение $\mathbf{c}^0$, сетка $N = 256$ (верхний график) и $N = 512$ (нижний график), $\varepsilon = \varepsilon_4(1/64)$. Увеличение τ для явной схемы невозможно из-за ограничений устойчивости.

сколькx уменьшение шага в неявных схемах обычно означает снижение их эффективности, схемы локальных итераций оказываются особенно привлекательными (для них снижение шага приводит к уменьшению вычислительных затрат). В таких задачах разумно применение схем локальных итераций с адаптивным выбором шага по времени (см. [65]).

Авторы благодарят В.Т. Жукова (ИПМ им. М.В. Келдыша РАН) за полезные обсуждения и консультации по схемам локальных итераций.

СПИСОК ЛИТЕРАТУРЫ

1. *Cahn J.W., Hilliard J.E.* Free Energy of a Nonuniform System. I. Interfacial Free Energy // J. Chem. Phys. 1958. V. 28. № 2. P. 258–267. <https://doi.org/10.1063/1.1744102>
2. *Gurtin M.E.* Generalized Ginzburg-Landau and Cahn-Hilliard equations based on a microforce balance // Physica D: Nonlinear Phenomena. 1996. V. 92. Iss. 3–4. P. 178–192. [https://doi.org/10.1016/0167-2789\(95\)00173-5](https://doi.org/10.1016/0167-2789(95)00173-5)
3. *Provatas N., Elder K.* Phase-Field Methods in Materials Science and Engineering. First published: 7 October 2010 <https://doi.org/10.1002/9783527631520> 2010 Wiley-VCH Verlag GmbH & Co. KGaA
4. *Steinbach I., Salama H.* Lectures on Phase Field. Springer Cham, 2023. <https://doi.org/10.1007/978-3-031-21171-3>
5. *Скрянов В.П., Скрянов А.В.* Спинодальный распад (Фазовый переход с участием неустойчивых состояний) // УФН. 1979. Т. 128. Вып. 2. С. 193–231. <https://doi.org/10.3367/UFNr.0128.197906a.0193>
6. *Hohenberg P.C., Halperin B.I.* Theory of dynamic critical phenomena // Rev. Mod. Phys. 177. V. 49. Iss. 3. P. 435. <https://doi.org/10.1103/RevModPhys.49.435>
7. *Penrose O., Fife P.C.* Thermodynamically consistent models of phase-field type for the kinetic of phase transitions // Physica D: Nonlinear Phenomena. 1990. V. 43. Iss. 1. P. 44–62. [https://doi.org/10.1016/0167-2789\(90\)90015-H](https://doi.org/10.1016/0167-2789(90)90015-H)
8. *Bray A.J.* Theory of phase-ordering kinetics // Adv. Phys. 2002. V. 51. № 2. P. 481–587. <https://doi.org/10.1080/00018730110117433>
9. *Miranville A.* The Cahn–Hilliard Equation: Recent Advances and Applications // Soc. Indust. Appl. Math. 2019. <https://doi.org/10.1137/1.9781611975925>
10. *Pego R.L.* Front Migration in the Nonlinear Cahn-Hilliard Equation // Proceed. Royal Soc. London. Ser. A. Math. Phys. Sci. 1989. V. 422. № 863. P. 261–278. www.jstor.org/stable/2398477
11. *Bates P.W., Fife P.C.* The Dynamics of Nucleation for the Cahn-Hilliard Equation // SIAM J. Appl. Math. 1993. V. 53. № 4. P. 990–1008. www.jstor.org/stable/2102259
12. *de Mello E.V.L., Otton Teixeira da Silveira Filho* Numerical study of the Cahn–Hilliard equation in one, two and three dimensions // Physica A: Statistic. Mech. Appl. 2005. V. 347. P. 429–443. <https://doi.org/10.1016/j.physa.2004.08.076>
13. *Vollmayr-Lee B.P., Rutenberg A.D.* Fast and accurate coarsening simulation with an unconditionally stable time step // Phys. Rev. E. 2003. V. 68. Iss. 6. P. 066703. <https://doi.org/10.1103/PhysRevE.68.066703>
14. *Eyre D.J.* An unconditionally stable one-step scheme for gradient systems // Tech. Rep. Department of Mathematics, University of Utah. 1997. unpublished. <https://api.semanticscholar.org/CorpusID:117273508>
15. *Eyre D.J.* Unconditionally gradient stable time marching the Cahn-Hilliard equation // Comput. Math. Model. Microstructur. Evolut. Mater. Res. Soc. Symp. Proc. Ed. V. 529. Bullard J.W., Chen L.-Q., Kalia R.K., Stoneham A.M., 1998. P. 39–46.
16. *Tierra G., Guillen-Gonzalez F.* Numerical methods for solving the Cahn-Hilliard equation and its applicability to related Energy-based models // Necas Center for Math. Model. Preprint № 2013–035.
17. *Cueto-Felgueroso L., Peirairé J.* A time-adaptive finite volume method for the Cahn–Hilliard and Kuramoto–Sivashinsky equations // J. Comput. Phys. 2008. V. 227. Iss. 4. P. 9985–10017. <https://doi.org/10.1016/j.jcp.2008.07.024>
18. *Li Y., Choi Y., Kim J.* Computationally efficient adaptive time step method for the Cahn–Hilliard equation // Comput. Math. Appl. 2017. V. 73. Iss. 8. P. 1855–1864. <https://doi.org/10.1016/j.camwa.2017.02.021>
19. *Zhang Z., Qiao Z.* An Adaptive Time-Stepping Strategy for the Cahn-Hilliard Equation // Comm. Computat. Phys. 2012. V. 11. Iss. 4. P. 1261–1278. <https://doi.org/10.4208/cicp.300810.140411s>
20. *Minkoff S.E., Kridler N.M.* A comparison of adaptive time stepping methods for coupled flow and deformation modeling // Appl. Math. Model. 2006. V. 30. Iss. 9. P. 993–1009. <https://doi.org/10.1016/j.apm.2005.08.002>

21. Luo F., Tang T., Xie H. Parameter-Free Time Adaptivity Based on Energy Evolution for the Cahn-Hilliard Equation // Comm. Computa. Phys. 2016. V. 19. Iss. 5. P. 1542–1563. <https://doi.org/10.4208/cicp.scpscde14.45s>
22. Guillen-Gonzalez F., Tierra G. Second order schemes and time-step adaptivity for Allen-Cahn and Cahn-Hilliard models // Computers and Mathematics with Applications. 2014. V. 68. Iss. 8. P. 821–846. <https://doi.org/10.1016/j.camwa.2014.07.014>
23. Kassam A., Trefethen L., Fourth-order time-stepping for stiff PDEs // SIAM J. Sci. Comput. 2005. V. 26. Iss. 4. P. 1214–1233. <https://doi.org/10.1137/S1064827502410633>
24. He Y., Liu Y., Tang T. On large time-stepping methods for the Cahn–Hilliard equation // Appl. Numeric. Math. 2007. V. 57. Iss. 5–7. P. 616–628. <https://doi.org/10.1016/j.apnum.2006.07.026>
25. Song H. Energy stable and large time-stepping methods for the Cahn–Hilliard equation // Inter. J. Comput. Math. 2015. V. 92. Iss. 10. P. 2091–2108. <https://doi.org/10.1080/00207160.2014.964694>
26. Li D. Why large time-stepping methods for the Cahn-Hilliard equation is stable // Math. Comp. 2022. V. 91. № 238. P. 2501–2515. <https://doi.org/10.1090/mcom/3768>
27. Chen W., Wang C., Wang X., Wise S.M. Positivity-preserving, energy stable numerical schemes for the Cahn-Hilliard equation with logarithmic potential // J. Comput. Phys.: X. 2009. V. 3. P. 100031. <https://doi.org/10.1016/j.jcp.x.2019.100031>
28. Chen W., Wang X., Yan Y., Zhang Z. A Second Order BDF Numerical Scheme with Variable Steps for the Cahn-Hilliard Equation // SIAM J. Numer. Anal. 2019. V. 57. Iss. 1. P. 495–525. <https://doi.org/10.1137/18M1206084>
29. Zhang J., Jiang M., Gong Y., Zhao J. Energy-stable predictor-corrector schemes for the Cahn-Hilliard equation // J. Comput. Appl. Math. 2020. V. 376. P. 112832. <https://doi.org/10.1016/j.cam.2020.112832>
30. Zhou Q., Sun Y. Energy stability of exponential time differencing schemes for the nonlocal Cahn-Hilliard equation // Numer. Meth. Partial Differ. Eq. 2023. V. 39. Iss. 5. P. 4030–4058. <https://doi.org/10.1002/num.23035>
31. Lee S. Unconditionally strong energy stable scheme for Cahn–Hilliard equation with second-order temporal accuracy // Math. Meth. Appl. Sci. 2023. V. 46. Iss. 6. P. 6463–6469. <https://doi.org/10.1002/mma.8917>
32. Boyer F., Minjeaud S. Numerical schemes for a three component Cahn-Hilliard model // ESAIM: Math. Model. Numeric. Anal. 2011. V. 45. No. 4. P. 697–738. <https://doi.org/10.1051/m2an/2010072>
33. Brachet M., Chehab J.-P. Fast and Stable Schemes for Phase Fields Models // Comput. Math. Appl. 2020. V. 80. Iss. 6. P. 1683–1713. <https://doi.org/10.1016/j.camwa.2020.07.015>
34. Elliott C., French D.A. A nonconforming finite element method for the two- dimensional Cahn–Hilliard equation // SIAM J. Numer. Anal. 1989. V. 26. № 4. P. 884–903. www.jstor.org/stable/2157884
35. Barrett J.B. An error bound for the finite element approximation of the Cahn–Hilliard equation with logarithmic free energy // Numer. Math. 1995. V. 72. P. 1–20. <https://doi.org/10.1007/s002110050157>
36. Chen L.-Q., Shen J., Applications of semi-implicit fourier-spectral method to phase field equations // Comput. Phys. Commun. 1996. V. 108. Iss. 2–3. P. 147–158. [https://doi.org/10.1016/S0010-4655\(97\)00115-X](https://doi.org/10.1016/S0010-4655(97)00115-X)
37. Furihata D. A stable and conservative finite difference scheme for the Cahn– Hilliard equation // Numer. Math. 2001. V. 87. Iss. 4. P. 675–699. <https://doi.org/10.1007/PL00005429>
38. Feng X., Prohl A. Error analysis of a mixed finite element method for the Cahn– Hilliard equation // Numer. Math. 2004. V. 99. Iss. 1. P. 47–84. <https://doi.org/10.1007/s00211-004-0546-5>
39. Wells E., Kuhl K., Garikipati S. A discontinuous Galerkin method for the Cahn– Hilliard equation // J. Comput. Phys. 2006. V. 218. Iss. 2. P. 860–877. <https://doi.org/10.1016/j.jcp.2006.03.010>
40. Wise S.M., Wang C., Lowengrub J.S. An Energy-Stable and Convergent Finite-Difference Scheme for the Phase Field Crystal Equation // SIAM J. Numer. Anal. 2009. V. 47. Iss. 3. P. 2269–2288. <https://doi.org/10.1137/0807381>
41. Du Q., Ju L., Tian L. Finite element approximation of the Cahn–Hilliard equation on surfaces // Comput. Meth. Appl. Mech. Engineer. 2011. V. 200. Iss. 29–32. P. 458–2470. <https://doi.org/10.1016/j.cma.2011.04.018>

42. Xia Y., Xu Y., Shu C.-W. Local discontinuous Galerkin methods for the Cahn–Hilliard type equations // J. Comput. Phys. 2007. V. 227. Iss. 1. P. 472–491. <https://doi.org/10.1016/j.jcp.2007.08.001>
43. Brenner S.C., Diegel A.E., Sung L.-Y. A robust solver for a second order mixed finite element method for the Cahn–Hilliard equation // J. Comput. Appl. Math. 2020. V. 364. P. 112322. <https://doi.org/10.1016/j.cam.2019.06.038>
44. Gomez H., Calo V.M., Bazilevs Y., Hughes T.J.R. Isogeometric analysis of the Cahn–Hilliard phase-field model // Comput. Meth. Appl. Mech. Engineer. 2008. V. 197. Iss. 49–50. P. 4333–4352. <https://doi.org/10.1016/j.cma.2008.05.003>
45. Zhang R., Qian X. Triangulation-based isogeometric analysis of the Cahn–Hilliard phase-field model // Comput. Meth. Appl. Mech. Engineer. 2019. V. 357. P. 112569. <https://doi.org/10.1016/j.cma.2019.112569>
46. Kastner M., Metsch P., de Borst R. Isogeometric analysis of the Cahn–Hilliard equation — a convergence study // J. Comput. Phys. 2016. V. 305. P. 360–371. <https://doi.org/10.1016/j.jcp.2015.10.047>
47. Goudenège L., Martin D., Vial G. High Order Finite Element Calculations for the Cahn–Hilliard Equation // J. Sci. Comput. 2012. V. 52. P. 294–321. <https://doi.org/10.1007/s10915-011-9546-7>
48. Чжао-дин Ю. Об устойчивости разностных схем для решения дифференциальных уравнений параболического типа // Докл. АН. 1957. Т. 117. № 4. С. 578–581. www.mathnet.ru/rus/dan22546
49. Чжао-дин Ю. Некоторые разностные схемы численного решения дифференциального уравнения параболического типа // Матем. сб. 1960. Т. 50(92). № 4. С. 391–422. www.mathnet.ru/rus/sm4800
50. Гельфанд И.М., Локуцкий О.В. О разностных схемах для решения уравнения теплопроводности. В кн.: Годунов С.К., Рябенский В.С. Введение в теорию разностных схем. М.: Физматгиз. 1962. 340 С.
51. Бабенко К.И. Основы численного анализа. М.; Ижевск: Регулярная и хаотическая динамика. 2002. 848 С.
52. Локуцкий В.О., Локуцкий О.В. Применение чебышевских параметров для численного решения некоторых эволюционных задач // Препринты ИПМ им. М.В. Келдыша РАН, 1984. № 99. 30 С. <https://library.keldysh.ru/preprint.asp?id=1984-99>
53. Жуков В.Т. Численные эксперименты по решению уравнения теплопроводности методом локальных итераций // Препринты ИПМ им. М.В. Келдыша РАН, 1984. № 97. 22 С. <https://library.keldysh.ru/preprint.asp?id=1984-97>
54. Локуцкий В.О., Локуцкий О.В. О численном решении краевых задач для уравнений параболического типа // Докл. АН СССР. 1986. Т. 291. № 3. С. 540–544. www.mathnet.ru/rus/dan47741
55. Жуков В.Т. Разностные схемы локальных итераций для параболических уравнений // Препринты ИПМ им. М.В. Келдыша, 1986. № 173. 31 С. <https://library.keldysh.ru/preprint.asp?id=1986-173>
56. Жуков В.Т. Явно-итерационные схемы для параболических уравнений // Вопросы атомной науки и техники. Серия: Матем. моделирование физ. процессов, 1993. № 4. С. 40–46.
57. Shvedov A.S., Zhukov V.T. Explicit iterative difference schemes for parabolic equations // Russian J. Numer. Anal. Math. Modelling. 1998. V. 13. № 2. P. 133–148.
58. Жуков В.Т. О явных методах численного интегрирования для параболических уравнений // Матем. моделирование. 2010. Т. 22. № 10. С. 127–158; Math. Models Comput. Simul. 2011. V. 3. №3. P. 311–332.
59. Жуков В.Т., Новикова Н.Д., Феодоритова О.Б. О применении многосеточного и явно-итерационного методов к решению параболических уравнений с анизотропными разрывными коэффициентами // Препринты ИПМ им. М.В. Келдыша. 2014. № 085. 24 С.
60. Жуков В.Т., Феодоритова О.Б., Дубень А.П., Новикова Н.Д. Явное интегрирование по времени уравнений Навье–Стокса с помощью метода локальных итераций // Препринты ИПМ им. М.В. Келдыша. 2019. № 012. 32 С.

61. Жуков В.Т., Феодоритова О.Б. О развитии параллельных алгоритмов решения параболических и эллиптических уравнений // Матем. анализ, Итоги науки и техн. Сер. Современ. мат. и ее прил. Темат. обз., 155, ВИНТИ РАН, М., 2018, 20–37; J. Math. Sci. (N.Y.), 254:5 (2021), 606–624.
62. Жуков В.Т., Зайцев Н.А., Лысов В.Г., Рыков Ю.Г., Феодоритова О.Б. Численный анализ модели процессов кристаллизации металлов, двумерный случай // Матем. моделирование. 2012. Т. 24. № 1. С. 109–128; Math. Models Comput. Simul. 2012. V. 4. № 4. P. 440–453.
63. Lee D., Huh J.-Y., Jeong D., Shin J., Yun A., Kim J. Physical, mathematical, and numerical derivations of the Cahn–Hilliard equation // Comput. Materials Sci. 2014. V. 81. P. 216–225. <https://doi.org/10.1016/j.commatsci.2013.08.027>
64. Lee S., Lee C., Lee H., Kim J. Comparison of different numerical schemes for the Cahn–Hilliard equation // Journal of the Korea Society for Industrial and Applied Mathematics. 2013. V. 17. Iss. 3. P. 197–207. <https://doi.org/10.12941/jksiam.2013.17.197>
65. Botchev M.A., Zhukov V.T. Adaptive iterative explicit time integration for nonlinear heat conduction problems // Lobachevskii J. Math. 2024. V. 44. (To appear.)

EFFICIENT AND STABLE TIME INTEGRATION OF THE CAHN-HILLIARD EQUATIONS: EXPLICIT, IMPLICIT, AND EXPLICIT-ITERATIVE SCHEMES

M. A. Botchev^{a,*}, I. A. Fakhuridinov^{a,b,**}, E. B. Savenkov^{a,***}

^aKeldysh Institute of Applied Mathematics of the Russian Academy of Sciences, Miusskaya sq., 4, Moscow, 125047 Russia

^bNational Research Nuclear University “Moscow Engineering Physics Institute”, Kashirskoe shosse, 31, Moscow, 115409 Russia

*e-mail: botchev@kiam.ru

**e-mail: mv1451003@gmail.com

***e-mail: savenkov@keldysh.ru

Received 02 April, 2024

Revised 02 April, 2024

Accepted 02 May, 2024

Abstract. The article proposes a new algorithm for numerical integration over time of the Cahn–Hilliard equation, based on the combined application of the Eyre splitting method and the local iteration modified (LIM) scheme for solving a finite-dimensional problem at each time step. The proposed method is gradient-stable and allows calculations with large time steps and has an explicit nature of calculations. The results of numerical calculations are presented, demonstrating the capabilities of the proposed method and its comparison with common methods of time integration of the Cahn–Hilliard equation.

Keywords: Cahn–Hilliard equation, gradient-stable schemes, Ayre splitting, local iteration scheme, LIM.

ОЦЕНКИ p -НОРМ РЕШЕНИЙ И ОБРАТНЫХ МАТРИЦ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ С ЦИРКУЛЯНТНОЙ МАТРИЦЕЙ¹⁾

© 2024 г. Ю. С. Волков^{1,*}, В. В. Богданов^{1,**}¹ 630090 Новосибирск, пр-т Акад. Коптюга, 4, Институт математики им. С. Л. Соболева СО РАН, Россия

*e-mail: volkov@math.nsc.ru

**e-mail: bogdanov@math.nsc.ru

Поступила в редакцию 07.02.2024 г.

Переработанный вариант 07.02.2024 г.

Принята к публикации 02.05.2024 г.

Рассматривается задача оценки решений и обратных матриц систем линейных уравнений с циркулянтной матрицей в p -норме, $1 \leq p < \infty$. Получена оценка для циркулянтной матрицы, имеющей диагональное преобладание. На основе этого результата и идеи разложения матрицы в произведение матриц, связанных с разложением характеристического многочлена, предложена оценка для общей циркулянтной матрицы. Библ. 41.

Ключевые слова: разностное уравнение, циркулянтная матрица, диагональное преобладание, норма обратной матрицы, оценка решения.

DOI: 10.31857/S0044466924080042, EDN: YBBTNX

1. ВВЕДЕНИЕ

Во многих приложениях, например, при решении дифференциальных уравнений (см. [1]), восстановлении изображений (см. [2]), при разработке беспроводной связи (см. [3]) и многих других, мы сталкиваемся с необходимостью исследования и решения систем линейных уравнений

$$Ax = b, \quad (1)$$

с циркулянтной матрицей порядка $m \times m$

$$A = \text{circ}(a_0, a_1, \dots, a_{m-1}) = \begin{pmatrix} a_0 & a_1 & \dots & a_{m-1} \\ a_{m-1} & a_0 & \dots & a_{m-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \dots & a_0 \end{pmatrix},$$

называемой также *циркулянт*, где первая определяющая строка (см. [4], [5]) матрицы A имеет вид $(a_0, a_1, \dots, a_{m-1})$, где $b = (b_0, b_1, \dots, b_{m-1})^T = \{b_i\}_{i=0}^{m-1} = \{b_i\}$ — вектор правой части системы (1) и $x = (x_0, x_1, \dots, x_{m-1})^T = \{x_i\}_{i=0}^{m-1} = \{x_i\}$ — искомое решение. Например, в задачах сплайн-интерполяции система (1) с такой матрицей возникает при периодической интерполяции на равномерной сетке (см. [6], [7]).

Можно смотреть на эту задачу, как на задачу решения разностного уравнения

$$a_0 x_i + a_1 x_{i+1} + \dots + a_{m-1} x_{i+m-1} = b_i, \quad i = 0, \dots, m-1, \quad (2)$$

где $x_{i+m} = x_i$ при $i = 0, \dots, m-1$. Эта задача является периодическим случаем общей задачи решения разностного уравнения

$$\sum_{j \in \mathbb{Z}} a_{j-i} x_j = b_i, \quad i \in \mathbb{Z},$$

¹⁾ Работа выполнена в рамках государственного задания ИМ СО РАН (проект № FWNF-2022-0015).

для произвольной дважды бесконечной последовательности $\{b_i\}_{i \in \mathbb{Z}}$. Естественно, в общем случае решение $\{x_i\}_{i \in \mathbb{Z}}$ тоже не будет уже периодической последовательностью. Примером может выступать задача вычисления констант экстремальной функциональной интерполяции (см. [8], [9]).

С разностным уравнением (2) (или циркулянтном A) ассоциируется многочлен

$$Q(z) = a_0 + a_1 z + \dots + a_{m-1} z^{m-1},$$

называемый *ассоциированным* или *характеристическим*. Во многих случаях приходится рассматривать ленточные циркулянты, если для некоторого $n < m$ будет $a_n \neq 0$ и $a_{n+1} = \dots = a_{m-1} = 0$, то многочлен будет таким

$$Q(z) = a_0 + a_1 z + \dots + a_n z^n.$$

Мы будем рассматривать p -норму ($1 \leq p \leq \infty$) вектора $y = (y_0, y_1, \dots, y_{m-1})^T$, определяемую следующим образом:

$$\|y\|_\infty := \max\{|y_0|, |y_1|, \dots, |y_{m-1}|\}$$

и при $1 \leq p < \infty$,

$$\|y\|_p := (|y_0|^p + |y_1|^p + \dots + |y_{m-1}|^p)^{1/p}.$$

Операторная (индуцированная) p -норма произвольной $m \times m$ матрицы $B = (b_{ij})_{i,j=0}^{m-1}$ определяется стандартным образом:

$$\|B\|_p := \sup_{y \neq 0} \frac{\|By\|_p}{\|y\|_p}.$$

Для p -норм матриц явные выражения через элементы матрицы есть только для $p = 1$ и $p = \infty$:

$$\|B\|_1 = \max_{0 \leq j \leq m-1} \sum_{i=0}^{m-1} |b_{ij}|,$$

$$\|B\|_\infty = \max_{0 \leq i \leq m-1} \sum_{j=0}^{m-1} |b_{ij}|.$$

При $p = 2$ норма матрицы равна величине наибольшего сингулярного числа матрицы. Напомним, что *сингулярными числами* произвольной матрицы B называются квадратные корни из собственных чисел матрицы $B^* B$, где B^* — сопряженная матрица (для вещественной матрицы — транспонированная). Как известно, собственные числа симметрической матрицы $B^* B$ всегда неотрицательны. Приведенное свойство 2-нормы объясняет специальное название этой нормы — *спектральная*.

Для остальных p -норм ($p \neq 1, \infty$) в общем случае нет выражения через элементы матрицы, для спектральной нормы можно привести такую оценку

$$\|B\|_2 \leq \|B\|_E := \sqrt{\sum_{i=0}^{m-1} \sum_{j=0}^{m-1} b_{ij}^2},$$

здесь $\|\cdot\|_E$ — евклидова норма матрицы, называемая также нормой Фробениуса. Вообще спектральная норма является наименьшей из всех p -норм

$$\|B\|_2 \leq \|B\|_p$$

для всех $1 \leq p \leq \infty$.

Если диагональный элемент в каждой строке матрицы B больше по модулю суммы модулей остальных элементов строки, то такие матрицы принято называть *матрицами с диагональным преобладанием*. В работе [10] для таких матриц предложена очень простая и эффективная оценка ∞ -нормы обратной матрицы. Она является точной, однако ее постоянно уточняют для разных специальных случаев и переносят на другие классы матриц (см. [11]). В работах [12], [13] показано, что эта оценка остается справедливой и для бесконечных матриц.

Довольно часто изучаемые матрицы оказываются ленточными. Для ленточных матриц установлено важное свойство — экспоненциальное убывание элементов обратных матриц при удалении от главной диагонали (см. [14], [15]). Правда в случае циркулянтов матрицы являются циклическими ленточными, что вносит свою специфику, для элементов обратных матриц экспоненциальное убывание элементов сменяется на их экспоненциальный рост, т.е. наблюдается эффект цикличности, что также позволяет получить оценки элементов обратных матриц (см. [16], [17]).

Как уже отмечалось, только при $p = 1$ и $p = \infty$ значение нормы для матрицы общего вида выражается через ее элементы. Поэтому вычислить или получить оценки норм при других значениях p непросто: задача не только для обратных, но и самих матриц. Известно (см. [18], [19]) выражение для спектральной нормы циркулянтной матрицы, если все ее элементы одного знака или строго знакопеременны

$$\|A\|_2 = |a_0| + |a_1| + \dots + |a_n|. \quad (3)$$

Естественно, в этом случае $\|A\|_2 = \|A\|_1 = \|A\|_\infty$. Более того, формула (3) останется справедливой, если матрица $A^T A$ или ее степень, кратная степени 2, имеет элементы только одного знака (см. [20]).

Как отмечено выше, хоть и не через элементы матрицы, но для спектральной нормы все-таки есть явная формула для вычисления — через собственные и сингулярные числа. А для циркулянтных матриц известно (см., например, [21]) простое правило вычисления собственных значений, а именно, все собственные значения могут быть вычислены по формуле

$$\lambda_k = Q(\omega^k), \quad k = 0, 1, \dots, m-1,$$

где $\omega = \exp(2\pi i/m) = \cos(2\pi/m) + i \sin(2\pi/m)$, $i = \sqrt{-1}$. Это означает, циркулянтная матрица A невырождена, если числа ω^k , $k = 0, 1, \dots, m-1$, не являются корнями характеристического многочлена $Q(z)$. И есть ряд работ (см. [22]–[24]), в которых все-таки вычисляется спектральная норма циркулянта, но не в произвольном случае, а для матриц со специальными элементами.

Отметим еще работу [25], в которой устанавливаются оценки спектральной нормы и обратной матрицы для некоторых циркулянтов специального вида.

Каких-либо оценок обратных к циркулянтным матрицам для других p -норм кроме $p = 1, 2, \infty$ не известно. Есть только работы [26], [27] по вычислению p -норм самих циркулянтов, причем некоторого частного вида.

Ранее в работе [28] были установлены оценки решений систем линейных уравнений с циркулянтными матрицами в равномерной норме, наша цель распространить такие оценки на случай произвольных p -норм, $p < \infty$. В разделе 2 мы устанавливаем оценку для циркулянтной матрицы с диагональным преобладанием. В разделе 3 приводятся оценки для циркулянтных матриц общего вида через разложение матриц на основе идеи работы [29] о разложении матрицы в произведение матриц, ассоциированных с множителями при разложении характеристического многочлена. В последнем разделе для циркулянтных матриц, возникающих в задачах сплайн-интерполяции, известные оценки спектральной и бесконечной норм обратных матриц переносятся на все p -нормы.

2. ЦИРКУЛЯНТНЫЕ МАТРИЦЫ С ДИАГОНАЛЬНЫМ ПРЕОБЛАДАНИЕМ

Мы говорим, что циркулянтная матрица A является *матрицей с диагональным преобладанием*, если для некоторого $k \in \{0, 1, \dots, n\}$ выполняется неравенство

$$|a_k| - \sum_{j=0, j \neq k}^n |a_j| = r > 0. \quad (4)$$

Заметим, что мы допускаем, что не только центральная диагональ с элементом a_0 может быть доминирующей. Если в нашем определении $k \neq 0$, то считаем, что доминирующая диагональ с элементом a_k состоит из двух кусков, а именно, первый кусок составляет наддиагональ с $m-k$ элементами a_k и второй кусок — поддиагональ с k такими же элементами a_k .

Теорема 1. Если циркулянтная матрица A имеет диагональное преобладание, то система линейных уравнений (1) имеет единственное решение x и при $p < \infty$ справедлива оценка

$$\|x\|_p \leq \frac{\|b\|_p}{r}. \quad (5)$$

Доказательство. Предположим, что у циркулянта A доминирующей является диагональ с элементами a_k , $0 \leq k \leq n$. Пусть y — произвольный вектор $y = (y_0, \dots, y_{m-1})^T$ и доопределим $y_{i+m} = y_{i-m} = y_i$, $i = 0, \dots, m-1$. Имеем

$$\begin{aligned} \|Ay\|_p &= \|\{a_0 y_i + \dots + a_n y_{i+n}\}\|_p \geq \\ &\geq \|\{a_k y_{i+k}\}\|_p - \|\{a_0 y_i + \dots + a_{k-1} y_{i+k-1} + a_{k+1} y_{i+k+1} + \dots + a_n y_{i+n}\}\|_p \geq \\ &\geq |a_k| \|\{y_{i+k}\}\|_p - \|\{a_0 y_i\}\|_p - \dots - \|\{a_{k-1} y_{i+k-1}\}\|_p - \\ &\quad - \|\{a_{k+1} y_{i+k+1}\}\|_p - \dots - \|\{a_n y_{i+n}\}\|_p, \end{aligned}$$

заметим, что в правой части последнего неравенства стоят нормы векторов $\{a_j y_{i+j}\}$, которые по нашему определению по периодичности величин y_j с индексами за пределами множества $\{0, 1, \dots, m-1\}$ совпадают с нормами векторов $\{a_j y_i\}_{i=0}^{m-1}$. Тогда все выражение в правой части неравенства будет равно

$$\begin{aligned} & |a_k| \|y\|_p - |a_0| \|y\|_p - \dots - |a_{k-1}| \|y\|_p - |a_{k+1}| \|y\|_p - \dots - |a_n| \|y\|_p = \\ & = \left(|a_k| - \sum_{j=0, j \neq k}^n |a_j| \right) \|y\|_p = r \|y\|_p. \end{aligned}$$

Таким образом,

$$\|Ay\|_p \geq r \|y\|_p,$$

что означает, что существует обратная матрица A^{-1} и имеет место оценка

$$\|A^{-1}\|_p \leq \frac{1}{r}.$$

Следовательно, для решения x системы (1) получаем

$$\|x\|_p \leq \|A^{-1}y\|_p \leq \|A^{-1}\|_p \|y\|_p \leq \frac{1}{r} \|y\|_p.$$

Теорема доказана.

Следствие 1. Если циркулянтная матрица A имеет диагональное преобладание, то

$$\|A^{-1}\|_p \leq \frac{1}{r}. \quad (6)$$

Заметим, что теорема 1 справедлива и для случая когда среди чисел $a_0, a_1, \dots, a_n$ есть комплексные.

Следствие 2. Если циркулянтная матрица A с диагональным преобладанием имеет положительную доминирующую диагональ и неположительные остальные, то

$$\|A^{-1}\|_2 = \frac{1}{r}. \quad (7)$$

Справедливость следствия 2 вытекает из того, что в этом случае среди элементов обратной матрицы нет отрицательных элементов [11], поэтому норма обратной матрицы будет вычисляться по формуле (3) и, значит, значение спектральной нормы совпадает с известным значением ∞ -нормы. Утверждение останется справедливым и для матриц монотонного вида. Напомним, матрица B есть *матрица монотонного вида*, если все элементы обратной матрицы B^{-1} неотрицательны (см. [30]).

Следствие 2 говорит о том, что для спектральной нормы оценка (6) неулучшаемая как и для норм при $p = 1$ и $p = \infty$.

Пример 1. Пусть матрица A — трехдиагональная и $a_0 = 1, a_1 = 4, a_2 = 1$. Это хорошо известная матрица задачи интерполяции кубическими сплайнами на равномерной сетке. Диагональ, соответствующая $a_1 = 4$, является доминирующей, условия теоремы 1 выполнены, следовательно имеем оценку решения

$$\|x\|_p \leq \frac{1}{2} \|b\|_p.$$

В работе [31] явно вычислена обратная к такой матрице размерности 4. Применение равенства (3) демонстрирует точность полученной оценки для спектральной нормы.

Следствие 3. Пусть A — двухдиагональная циркулянтная матрица ($n = 1$) с условием $|a_0| \neq |a_1|$, тогда $r = ||a_0| - |a_1||$. Если числа a_0 и a_1 вещественны и имеют одинаковый знак, то значение диагонального преобладания равно $r = |a_0 - a_1| = |Q(-1)|$. Если числа a_0 и a_1 противоположного знака, то $r = |a_0 + a_1| = |Q(1)|$.

Таким образом, следствие 3 показывает, что любая двухдиагональная циркулянтная матрица с разными по модулю значениями на диагоналях является матрицей с диагональным преобладанием, для нее выполнено условие (4). Конечно, в общем случае это неверно.

3. ЦИРКУЛЯНТЫ ОБЩЕГО ВИДА

Пусть A произвольный невырожденный циркулянт. Используя идею работы [29] по разложению циркулянтной матрицы в произведение матриц, ассоциированных с множителями при разложении характеристического многочлена $Q(z)$, установим оценки решения системы (1) в p -норме.

Заметим, что множество коэффициентов $a_0, a_1, \dots, a_n$ однозначно определяет как циркулянтную матрицу A , так и ее ассоциированный многочлен $Q(z)$. Более того, множество циркулянтных матриц A размера $m \times m$ и множество многочленов $Q(z)$ изоморфно относительно умножения. Правда здесь нужно сделать оговорку, мы рассматриваем степень многочленов по модулю m , т.е. если при перемножении двух многочленов степень результирующего многочлена получается больше $m - 1$, то каждый моном $a_k z^k$ при $k \geq m$ заменяется на моном $a_k z^{k-m}$, где $0 \leq k - m \leq m - 1$.

Лемма 1. *Предположим, что многочлен $Q(z)$, соответствующий циркулянту A , разложен в произведение многочленов $Q_1(z)$ и $Q_2(z)$, соответствующих некоторым циркулянтным матрицам A_1 и A_2 , и эти матрицы определяются множествами чисел $a_0^1, a_1^1, \dots, a_{n-k}^1$ и $a_0^2, a_1^2, \dots, a_k^2$ соответственно. Тогда $A = A_1 A_2 = A_2 A_1$.*

Доказательство. Для многочленов первой и второй степени справедливость леммы устанавливается прямыми вычислениями. Для общего случая доказательство легко следует из того, что любой многочлен разлагается на линейные и квадратичные множители.

Теорема 2. *Если все корни характеристического многочлена $Q(z)$ циркулянтной матрицы A положительны и не равны 1, то для решения системы (1) справедлива оценка*

$$\|x\|_p \leq \frac{\|b\|_p}{|Q(1)|}. \quad (8)$$

Для нормы обратной матрицы имеем

$$\|A^{-1}\|_p \leq \frac{1}{|Q(1)|},$$

причем если матрица четного порядка, то

$$\|A^{-1}\|_2 = \|A^{-1}\|_1 = \|A^{-1}\|_\infty = \frac{1}{|Q(1)|}. \quad (9)$$

Теорема 3. *Если все корни характеристического многочлена $Q(z)$ циркулянтной матрицы A отрицательны и не равны -1 , то для решения системы (1) справедлива оценка*

$$\|x\|_p \leq \frac{\|b\|_p}{|Q(-1)|}. \quad (10)$$

Для нормы обратной матрицы имеем

$$\|A^{-1}\|_p \leq \frac{1}{|Q(-1)|},$$

причем если матрица четного порядка, то

$$\|A^{-1}\|_2 = \|A^{-1}\|_1 = \|A^{-1}\|_\infty = \frac{1}{|Q(-1)|}. \quad (11)$$

Доказательство. Справедливость неравенств (8) и (10) следует из следствия 3 и леммы 1, следовательно будут справедливы и оценки для p -норм обратной матрицы рассматриваемой системы уравнений.

Если все корни $Q(z)$ положительны, то в разложении характеристического многочлена на линейные множители каждому множителю соответствует двухдиагональная матрица с диагоналями разных знаков. Кроме того, каждая элементарная матрица в разложении будет иметь диагональное преобладание в силу того, что на одной из диагоналей элементы равны 1, а на другой равны соответствующему корню многочлена, отличному от 1. Это означает, что все элементы обратной матрицы одного знака (если корень меньше 1, то положительны и при больше 1 наоборот). Следовательно, обратная матрица A^{-1} , получаемая как произведение матриц с элементами одного знака, также будет содержать элементы только одного знака. Таким образом, для нормы обратной матрицы будет выполняться соотношение (3), означающее $\|A^{-1}\|_2 = \|A^{-1}\|_1 = \|A^{-1}\|_\infty$. А для матриц четного порядка

$$\|A^{-1}\|_\infty = \frac{1}{|Q(1)|},$$

что означает справедливость (9).

Если корни $Q(z)$ отрицательны, умножение матрицы A слева и справа на диагональную матрицу $D := \text{diag}\{-1, +1, \dots, (-1)^m\}$ сводит ситуацию к случаю когда все корни положительны, в характеристическом многочлене переменная z заменится на $-z$. Тогда будут доказаны равенства (11).

Теоремы 2 и 3 доказаны.

Условие того, что точное значение норм в теоремах 2 и 3 имеет место только для матриц четного порядка существенно. В работах [32], [33] для частного случая симметрической матрицы показано, что если порядок циркулянта нечетен, то ∞ -норма обратной матрицы строго меньше величины $1/|Q(1)|$ при положительных корнях характеристического многочлена и строго меньше $1/|Q(-1)|$ при отрицательных корнях.

Следствие 4. Если характеристический многочлен представлен в виде произведения $Q(z) = Q_1(z)Q_2(z)$, причем все корни многочлена $Q_1(z)$ положительны, а многочлена $Q_2(z)$ — отрицательны, и $Q(\pm 1) \neq 0$, тогда для решения системы (1) справедлива оценка

$$\|x\|_p \leq \frac{\|b\|_p}{|Q_1(1)Q_2(-1)|}.$$

Мы показали как получить оценки решения системы уравнений (1) с циркулянтной матрицей если корни характеристического многочлена $Q(z)$ произвольные вещественные. Но в общем случае корни могут быть комплексными.

Пусть характеристический многочлен имеет вид $Q(z) = (z - z_1)(z - z_2)$ с комплексными корнями, не лежащими на окружности $|z| = 1$, т.е. $z_1 = \bar{z}_2 \in \mathbb{C}$. Как отмечалось ранее, следствие 3 останется верным и для комплексных коэффициентов, следовательно

$$\|A^{-1}\|_p \leq \frac{1}{(|z_1| - 1)^2}.$$

Таким образом, имеет место

Теорема 4. Предположим, что все корни характеристического многочлена $Q(z) = a_n(z - z_1) \dots (z - z_n)$, соответствующего циркулянтной матрице A , комплексные и не лежат на окружности $|z| = 1$, тогда

$$\|A^{-1}\|_p \leq \frac{1}{|a_n(|z_1| - 1) \dots (|z_n| - 1)|}. \quad (12)$$

Следствие 5. Предположим, что характеристический многочлен $Q(z)$, соответствующий циркулянтной матрице A , не имеет корней на окружности $|z| = 1$ и представлен в виде $Q(z) = (z - z_1) \dots (z - z_k)Q_1(z)Q_2(z)$, причем все корни многочлена $Q_1(z)$ положительны, а многочлена $Q_2(z)$ — отрицательны, тогда

$$\|A^{-1}\|_p \leq \frac{1}{(|z_1| - 1) \dots (|z_k| - 1)Q_1(1)Q_2(-1)}. \quad (13)$$

Пример 2. Пусть $a_0 = 4$, $a_1 = 1$, $a_2 = 1$, $n = 2$. Многочлен $Q(z)$ не имеет вещественных корней. Применение теоремы 4 приводит к оценке $\|x\|_p \leq \|b\|_p$. но матрица A имеет диагональное преобладание с $r = 2$. Тогда теорема 1 дает более точную оценку $\|x\|_p \leq \|b\|_p/2$.

4. ЦИРКУЛЯНТЫ В ЗАДАЧАХ СПЛАЙН-ИНТЕРПОЛЯЦИИ

В настоящей работе мы предложили подход, который позволяет устанавливать оценки решения системы уравнений (1) с произвольной невырожденной циркулянтной матрицей, если корни соответствующего характеристического многочлена $Q(z)$ не лежат на окружности $|z| = 1$. Наш метод основан на концепции диагонального преобладания и на идее разложения матрицы в произведение матриц, связанных с разложением характеристического многочлена. Установленные оценки точны, указаны случаи с точным значением p -норм. Но если характеристический многочлен имеет комплексные корни, то оценки (12), (13), по-видимому, не точны.

Исторически первая работа по получению эффективной оценки нормы обратной матрицы была установлена в 1963 г. в связи с изучением задачи интерполяции кубическими сплайнами [10]. Пристальное внимание к циркулянтным матрицам также возникло при рассмотрении задач интерполяции сплайнами для случая равномерной сетки и периодических краевых условий. Здесь возникают именно ленточные циркулянты (см. [6]).

В 1973 г. при получении оценок погрешности интерполяции на равномерной сетке периодическим полиномиальным сплайном нечетной степени $2k - 1$ Альбасини и Хоскинс (см. [34]) нашли оценки спектральной и

бесконечной норм обратных матриц к циркулянтным матрицам C_{2k-1} , возникающим здесь при интерполяции, $n = 2k - 2$. Эти оценки имеют вид

$$\|C_{2k-1}^{-1}\|_2 = \|C_{2k-1}^{-1}\|_\infty \leq \frac{(2k)!}{2^{2k}(2^{2k}-1)|B_{2k}|}, \quad (14)$$

где B_{2k} — числа Бернулли. Позже [35] эти же неравенства были установлены совсем простым путем через разложение матрицы C_{2k-1} в произведение трехдиагональных циркулянтов, имеющих диагональное преобладание. Здесь же было показано, что выражение в правой части неравенства (14) равно $1/|Q_{2k-1}(-1)|$, где $Q_{2k-1}(z)$ — характеристический многочлен, ассоциированный с исходной матрицей C_{2k-1} . Тогда мы можем сказать, что применение теоремы 1 и следствия 1 устанавливает справедливость такой же оценки

$$\|C_{2k-1}^{-1}\|_p \leq \frac{(2k)!}{2^{2k}(2^{2k}-1)|B_{2k}|}, \quad (15)$$

для всех p -норм, $1 \leq p \leq \infty$.

Методы работы [34] были обобщены и расширены на случай любых симметрических циркулянтов с положительными элементами [32]. Это позволило рассмотреть задачу периодической сплайн интерполяции для случая сплайнов четной степени $2k$ и установить оценки бесконечной нормы обратной матрицы к циркулянтной матрице C_{2k} , возникающей в этой задаче, здесь $n = 2k$. Эта же оценка

$$\|C_{2k}^{-1}\|_\infty \leq \frac{(2k)!}{|E_{2k}|}, \quad (16)$$

где E_{2k} — числа Эйлера, с помощью более простой техники работы [35] выписана в работе [36]. Так как здесь также оценка выписывается через величину $Q_{2k}(-1)$ путем разложение циркулянта C_{2k} в произведение трехдиагональных матриц с диагональным преобладанием, то оценка (16) будет справедливой и для всех p -норм, $1 \leq p \leq \infty$. Следовательно,

$$\|C_{2k}^{-1}\|_p \leq \frac{(2k)!}{|E_{2k}|}, \quad 1 \leq p \leq \infty. \quad (17)$$

Упомянем еще работу [37], в которой вся рассматриваемая техника применена к задаче приближения гистосплайнами. Установленные там оценки бесконечной нормы соответствующих обратных матриц возникших там циркулянтов также распространяются на любые p -нормы при $p < \infty$.

Заметим, что числа Бернулли и числа Эйлера выражаются через константы Фавара [38], которые хорошо известны в теории приближения, их обычно определяют как сумму ряда

$$\mathcal{K}_n = \frac{4}{\pi} \sum_{v=0}^{\infty} \frac{(-1)^{v(n+1)}}{(2v+1)^{n+1}}, \quad n = 0, 1, \dots$$

Однако их лучше определять (см. [39], [40]) по рекуррентной формуле

$$\mathcal{K}_0 = 1, \quad \mathcal{K}_1 = \frac{\pi}{2}, \quad \mathcal{K}_n = \frac{\pi}{4n} \sum_{k=0}^{n-1} \mathcal{K}_k \mathcal{K}_{n-1-k}, \quad n = 2, 3, \dots$$

Тогда оценки (15) и (17) можно объединить в единое неравенство

$$\|C_n^{-1}\|_p \leq \frac{\pi^n}{2^n \mathcal{K}_n}, \quad 1 \leq p \leq \infty. \quad (18)$$

Необходимость получения оценок вида (18) возникает при получении оценок погрешности приближения в задачах интерполяции периодическими полиномиальными сплайнами на равномерных сетках. История задачи изучения точных оценок погрешности интерполяции, сходимости процессов интерполяции и связанных аспектов приведена в обзоре (см. [41]).

СПИСОК ЛИТЕРАТУРЫ

1. *Wilde A.C.* Differential equations involving circulant matrices // Rocky Mountain J. Math. 1983. V. 13. № 1. P. 1–14.
2. *Carrasquinha E., Amado C., Pires A.M., Oliveira L.* Image reconstruction based on circulant matrices // Signal Processing: Image Communication. 2018. V. 63. P. 72–80.

3. *Tse D., Viswanath P.* Fundamentals of wireless communication. Cambridge: Cambridge Univ. Press, 2005. 564 p.
4. *Marcus M., Minc H.* A Survey of Matrix Theory and Matrix Inequalities. Boston: Allyn and Bacon, Inc., 1964. 180 p.
Перевод: *Маркус М., Минк Х.* Обзор по теории матриц и матричных неравенств. М.: Наука, 1972. 232 с.
5. *Davis P.J.* Circulant matrices. New York: Wiley, 1979. 250 p.
6. *Ahlberg J.H., Nilson E.N., Walsh J.L.* The theory of splines and their applications. New York: Acad. Press, 1967. 284 p.
Перевод: *Алберг Дж., Нильсон Э., Уолш Дж.* Теория сплайнов и ее приложения. М.: Мир, 1972. 316 с.
7. *Стечкин С.Б., Субботин Ю.Н.* Сплайны в вычислительной математике. М.: Наука, 1976. 248 с.
8. *Субботин Ю.Н.* О связи между конечными разностями и соответствующими производными // Тр. МИАН СССР. 1965. Т. 78. С. 24–42.
Перевод: *Subbotin Yu.N.* On the relations between finite differences and the corresponding derivatives // Am. Math. Soc. Translations. 1967. P. 23–42.
9. *Субботин Ю.Н.* Функциональная интерполяция в среднем с наименьшей n -й производной // Тр. МИАН СССР. 1967. Т. 88. С. 30–60.
Перевод: *Subbotin Yu.N.* Interpolation by functions with n -th derivative of minimum norm // Am. Math. Soc. Translations. 1969. P. 31–63.
10. *Ahlberg J.H., Nilson E.N.* Convergence properties of the spline fit // J. Soc. Indust. Appl. Math. 1963. V. 11. № 1. P. 95–104.
11. *Волков Ю.С., Мирошниченко В.Л.* Оценки норм матриц, обратных к матрицам монотонного вида и вполне неотрицательным матрицам // Сиб. матем. журн. 2009. Т. 50. № 6. С. 1248–1254.
Перевод: *Volkov Yu.S., Miroshnichenko V.L.* Norm estimates for the inverses of matrices of monotone type and totally positive matrices // Siberian Math. J. 2009. V. 50, n. 6. P. 982–987.
12. *Volkov Yu.S., Novikov S.I.* Estimates for solutions of bi-infinite systems of linear equations // Eur. J. Math. 2022. V. 8, n. 2. P. 722–731.
13. *Волков Ю.С., Новиков С.И.* Оценки решений бесконечных систем линейных уравнений и задача интерполяции кубическими сплайнами на прямой // Сиб. матем. журн. 2022. Т. 63, № 4. С. 814–830.
Перевод: *Volkov Yu.S., Novikov S.I.* Estimates of solutions to infinite systems of linear equations and the problem of interpolation by cubic splines on the real line // Siberian Math. J. 2022. V. 63. № 4. P. 677–690.
14. *Demko S.* Inverses of band matrices and local convergence of spline projections // SIAM J. Numer. Anal. 1977. V. 14. № 4. P. 616–619.
15. *de Boor C.* Odd-degree spline interpolation at a biinfinite knot sequence // Approximation Theory: Proc. Int. Colloq., Bonn, 1976 / Ed. R. Schaback, K. Scherer. Lect. Notes Math., 556. Berlin–Heidelberg–New York: Springer, 1976. P. 30–53.
16. *Волков Ю.С.* Об оценке элементов матрицы, обратной к циклической ленточной матрице // Сиб. журн. вычисл. матем. 2003. Т. 6. № 3. С. 263–267.
17. *Волков Ю.С.* Обратные циклических ленточных матриц и сходимость процессов интерполяции для производных периодических интерполяционных сплайнов // Сиб. журн. вычисл. матем. 2010. Т. 13. № 3. С. 243–253.
Перевод: *Volkov Yu.S.* Inverses of cyclic band matrices and the convergence of interpolation processes for derivatives of periodic interpolation splines // Numer. Anal. Appl. 2010. V. 3. № 3. P. 199–207.
18. *Merikoski J.K., Haukkanen P., Mattila M., Tossavainen T.* The spectral norm of a circulant matrix // JP J. Algebra, Number Theory Appl. 2018. V. 40. № 4. P. 495–500.
19. *Jiang Z., Zhou J.* A note on spectral norms of even-order r -circulant matrices // Appl. Math. Comput. 2015. V. 250. P. 368–378.
20. *Lindner M.* Circulant matrices: norm, powers, and positivity // Opuscula Math. 2018. V. 38. № 6. P. 849–857.
21. *Searle S. R.* On inverting circulant matrices // Linear Algebra Appl. 1979. V. 25. P. 77–89.

22. *Solak S.* On the norms of circulant matrices with the Fibonacci and Lucas numbers // *Appl. Math. Comput.* 2005. V. 160. № 1. P. 125–132.
23. *Ipek A.* On the spectral norms of circulant matrices with classical Fibonacci and Lucas numbers entries // *Appl. Math. Comput.* 2011. V. 217. № 12. P. 6011–6012.
24. *Zhou J., Jiang Z.* Spectral norms of circulant-type matrices with binomial coefficients and harmonic numbers // *Int. J. Comput. Meth.* 2014. V. 11. № 5. Art. No. 1350076.
25. *Pan V.Y., Svadlenka J., Zhao L.* Estimating the norms of random circulant and Toeplitz matrices and their inverses // *Linear Algebra Appl.* 2015. V. 468. P. 197–210.
26. *Sahasranand K.R.* The p -norm of circulant matrices via Fourier analysis // *Concr. Oper.* 2022. V. 9. № 1. P. 71–78.
27. *Bouthat L., Khare A., Mashreghi J., Moreau-Guerin F.* The p -norm of circulant matrices // *Linear Multilinear Algebra.* 2022. V. 70. № 21. P. 7176–7188.
28. *Volkov Yu.S., Novikov S.I.* Estimates for solutions of systems of linear equations with circulant matrices // *J. Phys.: Conf. Ser.* 2021. V. 2099. Art. No. 012019.
29. *Волков Ю.С.* О неотрицательном решении системы уравнений с симметрической циркулянтной матрицей // *Матем. заметки.* 2001. Т. 70. № 2. С. 170–180.
Перевод: *Volkov Yu.S.* Nonnegative solutions to systems with symmetric circulant matrix // *Math. Notes.* 2001. V. 70. № 2. P. 154–162.
30. *Collatz L.* Functional analysis and computational mathematics. New York: Academic Press, 1966. 473 p.
Перевод: *Коллатц Л.* Функциональный анализ и вычислительная математика. М.: Мир, 1969. 447 с.
31. *Fuyong L.* The inverse of circulant matrix // *Appl. Math. Comput.* 2011. V. 217, n. 21. P. 8495–8503.
32. *Hoskins W.D., Meek D.S.* The infinity norm of a certain type of symmetric circulant matrix // *Math. Comput.* 1977. V. 31. № 139. P. 733–737.
33. *Киндалев Б.С.* Точная оценка нормы обратной матрицы для симметрического циркулянта // *Вычислительные системы.* Новосибирск: ИМ СО АН СССР, 1987. Вып. 121: Аппроксимация сплайнами. С. 37–45.
34. *Albasiny E.L., Hoskins W.D.* Explicit error bounds for periodic splines of odd order on a uniform mesh // *J. Inst. Math. Appl.* 1973. V. 12. № 3. P. 303–318.
35. *Kershaw D.* A bound on the inverse of a band matrix which occurs in interpolation by periodic odd order splines // *J. Inst. Math. Appl.* 1977. V. 20. № 2. P. 227–228.
36. *Dubeau F.* On band circulant matrices in the periodic spline interpolation theory // *Linear Algebra Appl.* 1985. V. 72. P. 177–182.
37. *Dubeau F., Savoie J.* On circulant matrices for certain periodic spline and histospline projections // *Bull. Australian Math. Soc.* 1987. V. 36. № 1. P. 49–59.
38. *Abramovitz M., Stegun I.A.* (eds.) Handbook of mathematical functions with formulas, graphs, and mathematical tables. National Bureau of Standards, Washington, 1972. 1046 p.
Перевод: *Абрамовиц М., Стиган И.* (ред.) Справочник по специальным функциям с формулами, графиками и математическими таблицами. М.: Наука, 1979. 832 с.
39. *Волков Ю.С.* Об одной задаче экстремальной функциональной интерполяции и константах Фавара // *Докл. АН. Математика, информатика, процессы управления.* 2020. Т. 495. С. 34–37.
Перевод: *Volkov Yu.S.* One problem of extremal functional interpolation and the Favard constants // *Dokl. Math.* 2020. V. 102. № 3. P. 474–477.
40. *Volkov Yu.S.* Efficient computation of Favard constants and their connection to Euler polynomials and numbers // *Сиб. электрон. матем. известия.* 2020. Т. 17. С. 1921–1942.
41. *Волков Ю.С., Субботин Ю.Н.* 50 лет задаче Шёнберга о сходимости сплайн-интерполяции // *Тр. Ин-та математики и механики УрО РАН.* 2014. Т. 20. № 1. С. 52–67.
Перевод: *Volkov Yu.S., Subbotin Yu.N.* Fifty years of Schoenberg's problem on the convergence of spline interpolation // *Proc. Steklov Inst. Math.* 2015. V. 288. Suppl. 1. P. S222–S237.

ESTIMATES OF THE p -NORMS OF SOLUTIONS AND INVERSE MATRICES OF SYSTEMS OF LINEAR EQUATIONS WITH A CIRCULANT MATRIX

Yu. S. Volkov^{a,*}, V. V. Bogdanov^{a,**}

^a*Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Acad. Koptug Ave., 4, Novosibirsk, 630090 Russia*

^{*}*e-mail: volkov@math.nsc.ru*

^{**}*e-mail: bogdanov@math.nsc.ru*

Received 07 February, 2024

Revised 07 February, 2024

Accepted 02 May, 2024

Abstract. The problem of estimating solutions and inverse matrices of systems of linear equations with a circulant matrix in the p -norm, $1 < p < \infty$, is considered. An estimate is obtained for a circulant matrix with diagonal dominance. Based on this result and the idea of decomposing a matrix into a product of matrices related to the decomposition of the characteristic polynomial, an estimate is proposed for a general circulant matrix.

Keywords: difference equation, circulant matrix, diagonal dominance, norm of inverse matrix, solution estimate.

ЧИСЛЕННОЕ РЕШЕНИЕ УРАВНЕНИЙ ИОНИЗАЦИИ МЕТОДОМ ДРОБНЫХ ШАГОВ

© 2024 г. М.Б. Гавриков^{1,*}, А.А. Таюрский^{1,**}

¹125047 Москва, Миусская пл., 4, ИПМ им. М.В. Келдыша РАН, Россия

*e-mail: mbgavrikov@yandex.ru

**e-mail: tayurskiy2001@mail.ru

Поступила в редакцию 11.03.2024 г.

Переработанный вариант 11.03.2024 г.

Принята к публикации 05.05.2024 г.

Метод дробных шагов применяется для решения уравнений ионизации электрон-ионной плазмы, содержащей нейтралы, применительно к процессам в стационарных плазменных двигателях (СПД). Использовано расщепление по физическим процессам: эволюция плазмы представлена как суперпозиция чистого переноса во внешнем поле в фазовом пространстве с последующей ионизацией. Для расчета переноса использован метод макрочастиц. Процесс ионизации в случае плоской симметрии рассчитывается по явным аналитическим формулам. Полученные формулы метода дробных шагов сравниваются с численным решением уравнения эволюции плазмы посредством известной физической модели ионизации. Библ. 23. Фиг. 5.

Ключевые слова: метод дробных шагов, ионизация, метод макрочастиц, физическая модель ионизации.

DOI: 10.31857/S0044466924080058, **EDN:** YAWELK

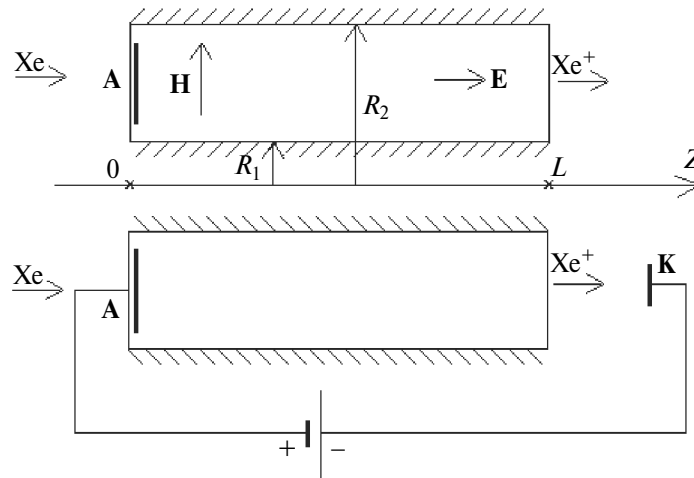
1. ВВЕДЕНИЕ

Рассмотрим процесс ионизации электрон-ионной плазмы, содержащей нейтралы. Пусть $f_i(t, \mathbf{x}, \mathbf{v})$, $f_a(t, \mathbf{x}, \mathbf{v})$ — функции распределения ионов и нейтральных атомов, соответственно. Тогда ионизация атомов задается системой уравнений [1, 2]:

$$\begin{aligned} \frac{\partial f_i}{\partial t} + \langle \mathbf{v}, \nabla_{\mathbf{x}} f_i \rangle + \langle \mathbf{F}, \nabla_{\mathbf{v}} f_i \rangle &= \beta n_i f_a, \\ \frac{\partial f_a}{\partial t} + \langle \mathbf{v}, \nabla_{\mathbf{x}} f_a \rangle &= -\beta n_i f_a, \quad n_i = \int_{\mathbb{R}^3} f_i d\mathbf{v}. \end{aligned} \quad (1)$$

где заданы $\beta = \beta(T_e) \geq 0$ — скалярная функция электронной температуры (коэффициент ионизации) и $\mathbf{F} = \mathbf{F}(t, \mathbf{x}, \mathbf{v})$ — векторная функция, подчиненная условию $\operatorname{div}_{\mathbf{v}} \mathbf{F} = 0$. Часто считается $\beta \equiv \beta_0 = \text{const}$, $\mathbf{F}(t, \mathbf{x}, \mathbf{v}) = \frac{e}{m_i} \left\{ \mathbf{E} + \frac{1}{c} [\mathbf{v}, \mathbf{H}] - \frac{\mathbf{j}}{\sigma} \right\}$, где $\mathbf{E}(t, \mathbf{x})$, $\mathbf{H}(t, \mathbf{x})$, $\mathbf{j}(t, \mathbf{x})$ — напряженности электрического и магнитного поля в плазме и плотность плазменного тока, σ — проводимость плазмы (в простейшем случае спитцеровская, $\sigma \sim T_e^{3/2}$).

Применительно к конкретным плазменным установкам система упрощается. Особый интерес представляет процесс ионизации плазмы в стационарных плазменных двигателях (СПД), которые используются для коррекции орбит космических летательных аппаратов. История вопроса отражена в [3–5]. Согласно упрощенной принципиальной схеме, приведенной на фиг. 1, СПД представляет собой коаксиальную камеру с керамическими боковыми стенками, в торцах которой расположены анод и катод с подведенной к ним разностью потенциалов, приводящей к появлению в камере продольного электрического поля $\mathbf{E}$. Кроме того, система обмоток с постоянными токами создает в камере почти радиальное магнитное поле $\mathbf{H}$. Нейтральный газ — ксенон (Xe) подается со стороны анода с постоянной скоростью $\mathbf{v}_a$, ионизируется посредством ударной ионизации электронами, и возникающие ионы Xe^+ ускоряются продольным электрическим полем до скоростей ~ 15 км/с и покидают камеру в сторону катода, нейтрализуясь на нем. В результате возникает реактивная сила тяги, передаваемая жестко связанному с камерой летательному аппарату, на получение которой и направлена работа СПД. Длина типичного [6] СПД-100М составляет $L = 2.5$ см, внутренний и внешний радиусы камеры установки равны



Фиг. 1. Принципиальная схема СПД.

$R_1 = 3.5$ см, $R_2 = 5$ см, время непрерывной работы двигателя после выхода на квазистационарный режим $\sim 10^4$ ч.

Практически важный класс решений системы (1) получим, считая $f_a(t, \mathbf{x}, \mathbf{v}) = n_a(t, \mathbf{x})\delta(\mathbf{v} - \mathbf{v}_a)$, где $\delta(\mathbf{v})$ есть δ -функция. Тогда, учитывая равенства $(v^i - v_a^i)\delta(\mathbf{v} - \mathbf{v}_a) = 0$, $1 \leq i \leq 3$, уравнение кинетики для нейтралов равносильно уравнению для концентрации нейтралов

$$\frac{\partial n_a}{\partial t} + \langle \mathbf{v}_a, \nabla n_a \rangle = -\beta n_i n_a,$$

и система (1) сводится к следующей равносильной системе:

$$\begin{aligned} \frac{\partial f_i}{\partial t} + \langle \mathbf{v}, \nabla_{\mathbf{x}} f_i \rangle + \langle \mathbf{F}, \nabla_{\mathbf{v}} f_i \rangle &= \beta n_i n_a \delta(\mathbf{v} - \mathbf{v}_a), \\ \frac{\partial n_a}{\partial t} + \langle \mathbf{v}_a, \nabla n_a \rangle &= -\beta n_i n_a, \quad n_i = \int_{\mathbb{R}^3} f_i d\mathbf{v}. \end{aligned} \quad (2)$$

В системе (2) n_a исключается из числа неизвестных (см. разд. 3). Для этого, интегрируя по характеристикам второе уравнение (2), выразим в явном виде n_a через n_i и, подставив найденное выражение в первое уравнение (2), получим линейное интегро-дифференциальное уравнение на функцию f_i , которое ниже решается методом дробных шагов.

Конкретные формулы метода дробных шагов получим ограничившись случаем плоской симметрии, когда $\mathbf{x} = (0, 0, z)$, $\mathbf{v} = (0, v_y, v_z)$, $\mathbf{F} = (0, F_y, F_z)$, $\mathbf{v}_a = (0, 0, v_a)$, $v_a > 0$, $f_i = f(t, z, v_y, v_z)$, $n_a = n_a(t, z)$, $\nabla_{\mathbf{x}} = \partial/\partial z$, $\nabla_{\mathbf{v}} = (0, \partial/\partial v_y, \partial/\partial v_z)$. При этом $\mathbf{H} = (H_x, 0, 0)$, $\mathbf{E} = (0, 0, E_z)$, $H_x = \text{const}$, $E_z = \text{const}$, $\mathbf{j} = 0$, $F_y = \frac{e}{m_i c} v_z H_x$, $F_z = \frac{e}{m_i} E_z - \frac{e}{m_i c} v_y H_x$, а система (2) преобразуется к виду

$$\begin{aligned} \frac{\partial f}{\partial t} + v_z \frac{\partial f}{\partial z} + F_y \frac{\partial f}{\partial v_y} + F_z \frac{\partial f}{\partial v_z} &= \beta n n_a \delta(v_y) \delta(v_z - v_a), \\ \frac{\partial n_a}{\partial t} + v_a \frac{\partial n_a}{\partial z} &= -\beta n n_a, \quad n = \int_{\mathbb{R}^2} f dv_y dv_z, \quad t \geq 0, \quad 0 \leq z \leq L, \end{aligned} \quad (3)$$

которая ниже и решается методом дробных шагов.

Приближение плоской симметрии часто применяется в работах по СПД [7, 8] и исходит из допущения, что радиальные процессы в СПД не вносят существенный вклад в ионизацию атомов Xe и ускорение возникающих ионов вдоль оси z . В частности, можно пренебречь радиальным движением электронов, их рассеянием на боковых стенках установки и пристеночной проводимостью [9]. Математический переход к случаю плоской симметрии получается, если выпрямить коаксиальную камеру установки посредством экспоненциального отображения. Тогда ось r перейдет в ось x , ось φ — в ось y , в частности, H_x — это радиальная компонента магнитного поля.

Наконец, подчеркнем, что в предыдущей постановке, как обычно в исследованиях по СПД, считается, что индукционное электромагнитное поле, порожденное переменными токами и зарядами, пренебрежимо мало, а $\mathbf{H}$, $\mathbf{E}$ это внешнее электромагнитное поле, создаваемое постоянными токами обмоток СПД и разностью потенциалов между анодом и катодом.

В разд. 2 рассмотрена общая схема метода дробных шагов применительно к системам (2), (3). В разд. 3 кратко изложен в необходимом объеме метод макрочастиц. В разд. 4 рассмотрена численная реализация метода дробных шагов для системы (3). В разд. 5 метод дробных шагов сравнивается с известным способом численного решения системы (2), основанном на физическом моделировании ионизации [10–12]. Показано, что 1) для достаточно сильных электрических полей E_z метод дробных шагов пригоден для практических вычислений, 2) из сравнения концентраций ионов и атомов Хе следует, что расчеты по двум указанным методам качественно совпадают. Учитывая, что физическая модель ионизации [10–12] не имеет математического обоснования, предпочтение при численном исследовании ионизации следует отдать методу дробных шагов. В то же время оба метода расчета дополняют друг друга. С одной стороны, принимая во внимание качественное совпадение результатов расчета, метод дробных шагов можно рассматривать как математическое обоснование алгоритма физического моделирования ионизации. С другой стороны, метод дробных шагов при численной реализации требует больших вычислительных ресурсов (см. разд. 3), поэтому при необходимости его можно заменить на алгоритм физического моделирования ионизации.

Полученные методы дробных шагов легко обобщаются для численного решения системы (2) в случае осевой симметрии и с учетом индукционных электромагнитных полей.

2. МЕТОД ДРОБНЫХ ШАГОВ

Представим кинетическое уравнение в виде

$$\frac{\partial f}{\partial t} = \mathcal{L}_n(f) + \mathcal{L}_u(f), \quad \mathcal{L}_n(f) \stackrel{\text{def}}{=} -\langle \mathbf{v}, \nabla_{\mathbf{x}} f \rangle - \langle \mathbf{F}, \nabla_{\mathbf{v}} f \rangle, \quad \mathcal{L}_u(f) \stackrel{\text{def}}{=} \beta n n_a \delta(\mathbf{v} - \mathbf{v}_a),$$

где $f = f_i$, $n = n_i$. Для численного решения кинетического уравнения используем расщепление по физическим процессам или метод дробных шагов [13, 14], суть которого в данном случае в следующем. Допустим, на отрезке $[t_0, t_0 + \Delta t]$ мы умеем численно (или аналитически) решать задачи Коши для уравнений

$$\frac{\partial f}{\partial t} = \mathcal{L}_n(f), \quad f|_{t=t_0} = f^0, \quad t_0 \leq t \leq t_0 + \Delta t, \quad (4)$$

$$\frac{\partial f}{\partial t} = \mathcal{L}_u(f), \quad f|_{t=t_0} = f^0, \quad t_0 \leq t \leq t_0 + \Delta t. \quad (5)$$

Тогда задача Коши

$$\frac{\partial f}{\partial t} = \mathcal{L}_n(f) + \mathcal{L}_u(f), \quad f|_{t=t_0} = f^0, \quad t_0 \leq t \leq t_0 + \Delta t, \quad (6)$$

решается следующим образом. Пусть $\bar{f}$ — значение решения задачи Коши в момент времени $t = t_0 + \Delta t$, тогда

$$\frac{\bar{f} - f^0}{\Delta t} = \mathcal{L}_n(\bar{f}). \quad (7)$$

Пусть f^1 — значение решения задачи Коши $\partial f / \partial t = \mathcal{L}_u(f)$, $f|_{t=t_0} = f^0$, $t_0 \leq t \leq t_0 + \Delta t$ в момент времени $t = t_0 + \Delta t$, тогда

$$\frac{f^1 - \bar{f}}{\Delta t} = \mathcal{L}_u(f^1). \quad (8)$$

Из (7) и (8) получим

$$(\bar{f} - \Delta t \mathcal{L}_n(\bar{f})) = f^0, \quad (f^1 - \Delta t \mathcal{L}_u(f^1)) = \bar{f}.$$

Умножая обе части второго равенства на $I - \Delta t \mathcal{L}_n$, где I — единичный оператор, имеем

$$(I - \Delta t \mathcal{L}_n)(I - \Delta t \mathcal{L}_u)f^1 = (I - \Delta t \mathcal{L}_n)\bar{f} = f^0,$$

откуда

$$f^1 - f^0 = \Delta t(\mathcal{L}_n + \mathcal{L}_u)(f^1) - \Delta t^2 \mathcal{L}_n \mathcal{L}_u(f^1)$$

или

$$\frac{f^1 - f^0}{\Delta t} = (\mathcal{L}_n + \mathcal{L}_u)(f^1) - \Delta t \mathcal{L}_n \mathcal{L}_u(f^1) = (\mathcal{L}_n + \mathcal{L}_u)(f^1) + O(\Delta t) \quad (9)$$

и, значит, f^1 — решение на верхнем слое ($t = t_0 + \Delta t$) задачи Коши (6) на отрезке $[t_0, t_0 + \Delta t]$. Заметим, что в правых частях (7), (8) могут стоять любые добавки порядка $O(\Delta t)$. Тогда все равно приходим к равенству (9). Обратимся теперь к решению задач Коши (4), (5). Выше, по умолчанию, предполагалось, что эти задачи решаются на отрезке $[0, L]$ с определенными граничными условиями. На левой границе $z = 0$ ставится условие зеркального отражения, а правая граница $z = L$ считается свободной: ион, достигающий правую границу с положительной z -й скоростью, покидает область $[0, L]$.

3. МЕТОД МАКРОЧАСТИЦ

Задача Коши (4) численно решается методом макрочастиц. Ключевым моментом при дискретизации уравнения (4) является выбор конечномерных агрегатов, аппроксимирующих неизвестное решение, совокупность которых образует аппроксимирующее пространство. В физической литературе [15] функция распределения f обычно аппроксимируется конечной суммой δ -функций:

$$f(t, \mathbf{x}, \mathbf{v}) \cong \sum_{j=1}^N q_j \delta(\mathbf{x} - \mathbf{x}_j(t)) (\mathbf{v} - \mathbf{v}_j(t)), \quad (10)$$

где $q_j > 0$ не зависят от t , $\sum_{j=1}^N q_j = \text{const}$ не зависит от N , $\mathbf{x}_j(t)$, $\mathbf{v}_j(t)$ — заданные функции. В методе макрочастиц эти функции ищутся из решения уравнения характеристик

$$\frac{d\mathbf{x}_j}{dt} = \mathbf{v}_j, \quad \frac{d\mathbf{v}_j}{dt} = F(t, \mathbf{x}_j, \mathbf{v}_j), \quad \mathbf{x}_j(0), \mathbf{v}_j(0) — \text{заданы} \quad (11)$$

и называются макрочастицами, а величины q_j — массами макрочастиц. Значения q_j , $\mathbf{x}_j(0)$, $\mathbf{v}_j(0)$, $1 \leq j \leq N$, определяются начальной функцией распределения $f_0(\mathbf{x}, \mathbf{v}) = f(0, \mathbf{x}, \mathbf{v})$, которая считается известной. Например, фиксируем разбиение фазового пространства в сумму N областей D_j , $1 \leq j \leq N$ (не обязательно компактных!), с кусочно-гладкими границами и непересекающимися и непустыми внутренностями. Тогда положим

$$q_j = \int_{D_j} f_0 d\mathbf{x} d\mathbf{v}, \quad \mathbf{x}_j(0) = \frac{1}{q_j} \int_{D_j} \mathbf{x} f_0 d\mathbf{x} d\mathbf{v}, \quad \mathbf{v}_j(0) = \frac{1}{q_j} \int_{D_j} \mathbf{v} f_0 d\mathbf{x} d\mathbf{v}, \quad 1 \leq j \leq N.$$

Иными словами, q_j — масса, $(\mathbf{x}_j(0), \mathbf{v}_j(0))$ — центр тяжести области D_j . Детали см. в [16]. При дискретизации (10) параметры плазменной среды в канале СПД задаются обобщенными функциями (для концентрации числа частиц это прямо следует из (10)), что серьезно затрудняет верификацию получаемых численных результатов и сравнение их с экспериментальными данными. В [17] указанный недостаток преодолевается коррекцией аппроксимации (10) путем применения к ее правой части линейного интегрального преобразования по $\mathbf{x}$ с симметричным или вырожденным ядром. Другая возможность — переход на язык конечных регулярных борелевских мер [18]. Каждая функция распределения в момент времени t порождает конечную регулярную борелевскую меру $\mu(t)$ на σ -алгебре S всех борелевских подмножеств фазового пространства:

$$\mu(t)(E) \stackrel{\text{def}}{=} \int_E f(t, \mathbf{x}, \mathbf{v}) d\mathbf{x} d\mathbf{v}, \quad E \in S.$$

При этом уравнение (4) переписывается в виде эволюционного уравнения относительно $\mu(t)$ [19]. Тогда дискретизация (10) может рассматриваться как аппроксимация $\mu(t) \cong \mu_N(t)$, где

$$\mu_N(t)(\Psi) \stackrel{\text{def}}{=} \sum_{j=1}^N q_j \Psi(\mathbf{x}_j(t), \mathbf{v}_j(t)),$$

и используется отождествление конечных регулярных борелевских мер с непрерывными положительными линейными функционалами на пространстве финитных ($\equiv$ с компактным носителем) непрерывных функций фазового пространства $\{\Psi\}$ с нормой максимум модуля [18]. В [19] дано обоснование метода макрочастиц (более точно, установлена слабая сходимость $\mu_N(t) \rightarrow \mu(t)$ при $N \rightarrow +\infty$ для всех t из конечного отрезка $[0, T]$) применительно к численному решению систем уравнений Максвелла–Власова и Власова–Пуассона при условии вычисления средних величин со сглаживанием.

Для практических вычислений предпочтительней использовать следующую дискретизацию:

$$f(t, \mathbf{x}, \mathbf{v}) \cong \sum_{j=1}^N q_j \delta_\varepsilon(\mathbf{x} - \mathbf{x}_j(t)) (\mathbf{v} - \mathbf{v}_j(t)), \quad \varepsilon > 0, t \geq 0,$$

где, как и в (10), $q_j > 0$ и не зависят от t , $\mathbf{x}_j(t)$, $\mathbf{v}_j(t)$ — заданные функции, которые в методе макрочастиц удовлетворяют уравнению характеристик (11), а q_j , $\mathbf{x}_j(0)$, $\mathbf{v}_j(0)$ однозначно определяются начальной функцией $f_0(\mathbf{x}, \mathbf{v})$. При этом функция δ_ε удовлетворяет следующим ограничениям:

- 1) $\delta_\varepsilon \geq 0$;
- 2) $\int_{\mathbb{R}^3} \delta_\varepsilon(\mathbf{x}) d\mathbf{x} = 1$;
- 3) δ_ε при $\varepsilon \rightarrow 0$ сходятся к δ -функции в нуле, что означает справедливость предельного равенства

$$U(0) = \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^3} U(\mathbf{x}) \delta_\varepsilon(\mathbf{x}) d\mathbf{x}$$

для любой непрерывной ограниченной функции $U(\mathbf{x})$;

- 4) $\text{supp} \delta_\varepsilon$ — компактен, $0 \in \text{supp} \delta_\varepsilon$ и $\lim_{\varepsilon \rightarrow 0} \text{diam} \text{supp} \delta_\varepsilon = 0$.

Выбор $\delta_\varepsilon(\mathbf{x})$ предопределяет вычислительный алгоритм метода макрочастиц. Конкретные примеры $\delta_\varepsilon(\mathbf{x})$ содержатся в [15, 20–22], в частности, как правило, выполнено условие $\text{supp} \delta_\varepsilon \subseteq [-\varepsilon, \varepsilon]^3$, что ниже предполагается.

В случае плоской симметрии координаты макрочастиц суть $z_j(t)$, $\mathbf{v}_j(t)$, $\mathbf{v} = (v_y, v_z)$, а задача Коши на отрезке $[t_0, t_0 + \Delta t]$ для уравнения характеристик (11) существенно упрощается:

$$\begin{aligned} \frac{dz}{dt} &= v_z, \quad \frac{dv_y}{dt} = F_y(t, z, \mathbf{v}), \quad \frac{dv_z}{dt} = F_z(t, z, \mathbf{v}), \quad t_0 \leq t \leq t_0 + \Delta t, \\ z(t_0) &= z_j(t_0), \quad \mathbf{V}(t_0) = \mathbf{V}_j(t_0), \\ F_y &= \frac{e}{m_i c} v_z H_x, \quad F_z = \frac{e}{m_i} E_z - \frac{e}{m_i c} v_y H_x, \quad E_z = \text{const}, \quad H_x = \text{const}. \end{aligned} \quad (12)$$

Система (12) — линейная с постоянными коэффициентами и может быть легко решена в явном виде. Однако с учетом граничных условий ее практическую интеграцию целесообразно проводить численно, например, по схеме leap-frog [15, 20–22]. Если $z_j(t)$, $\mathbf{v}_j(t)$, $t_0 \leq t \leq t_0 + \Delta t$ — решение задачи Коши (12), то имеем

$$\begin{aligned} f^0 &= \sum_j q_j \delta_\varepsilon(z - z_j(t_0)) \delta(\mathbf{v} - \mathbf{v}_j(t_0)), \\ \bar{f} &= \sum_j q_j \delta_\varepsilon(z - z_j(t_0 + \Delta t)) \delta(\mathbf{v} - \mathbf{v}_j(t_0 + \Delta t)). \end{aligned} \quad (13)$$

В общем случае при учете индукционных электромагнитных полей для вычисления поля $\mathbf{F}$ необходимо численно решить систему уравнений в частных производных для нахождения полей $\mathbf{E}$, $\mathbf{H}$. Если последняя система решается разностным методом, то метод макрочастиц необходимо дополнить двумя группами формул: формулами вычислений средних величин, для которых на разностной сетке сохранялись бы масса, импульс и т.д., и интерполяционными формулами, позволяющими по сеточным значениям полей $\mathbf{E}$, $\mathbf{H}$ вычислить значение $\mathbf{F}$ в точках — местоположениях макрочастиц. Подробности см. в [15, 20–22].

4. ЧИСЛЕННАЯ РЕАЛИЗАЦИЯ МЕТОДА ДРОБНЫХ ШАГОВ

Решение уравнения (5) получается из аналитического решения системы, дополненной естественными начальными и граничными условиями:

$$\begin{aligned} \frac{\partial f}{\partial t} &= \beta n n_a \delta(\mathbf{v} - \mathbf{v}_a), \quad f|_{t_0} = \bar{f}, \quad \mathbf{v}_a = (0, v_a), \quad v_a > 0, \quad t_0 \leq t \leq t_0 + \Delta t, \\ \frac{\partial n_a}{\partial t} + v_a \frac{\partial n_a}{\partial z} &= -\beta n n_a, \quad 0 \leq z \leq L, \quad t_0 \leq t \leq t_0 + \Delta t, \\ n_a|_{z=0} &= n_{a0}(t), \quad n_a|_{t=t_0} = n_a^0(z), \quad 0 \leq z \leq L, \end{aligned} \quad (14)$$

где $\bar{f}$ вычисляется по (13), $n_{a0}(t)$, $n_a^0(z)$ — известные функции, удовлетворяющие условию согласованности $n_{a0}(0) = n_a^0(0)$. Как уже отмечалось во введении, функция

$$n_a$$

выражается через n посредством явных формул. Действительно, характеристики уравнения переноса концентрации n_a имеют вид $z_c(t) = v_a t + C$, $C = \text{const}$. Пусть $X = \ln n_a$, тогда для любого C имеем

$$\frac{d}{dt} X(t, z_c(t)) = \frac{\partial X}{\partial t} + v_a \frac{\partial X}{\partial z} = -\beta n.$$

Теорема 1 (см. [23]). *Решение системы в полуполосе $0 \leq z \leq L$, $t \geq t_0$, удовлетворяющее указанным выше начальным и граничным условиям, имеет вид*

$$n(t, z) = n^0(z) \frac{\exp N(z)}{1 + \frac{\beta}{v_a} \int_0^z n^0(p) e^{N(p)} dp + \frac{\beta}{v_a} \int_0^{z-v_a(t-t_0)} n_a^0(p) e^{N(p)} dp},$$

$$n_a(t, z) = n_a^0(z - (t - t_0)v_a) \frac{\exp N(z - v_a(t - t_0))}{1 + \frac{\beta}{v_a} \int_0^z n^0(p) e^{N(p)} dp + \frac{\beta}{v_a} \int_0^{z-v_a(t-t_0)} n_a^0(p) e^{N(p)} dp},$$

$$(t - t_0)v_a \leq z \leq L, \quad t \geq t_0,$$

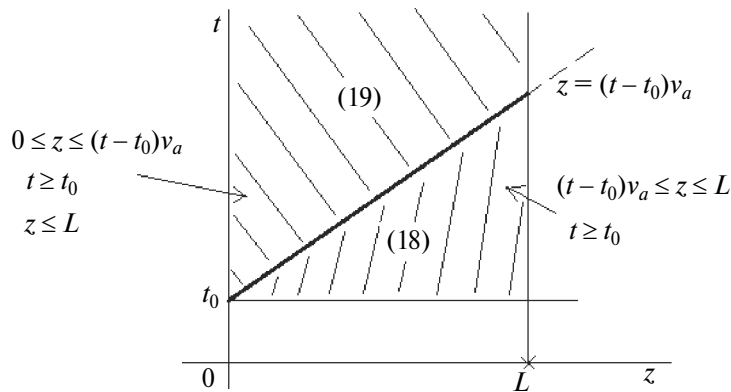
$$n(t, z) = n^0(z) \frac{\exp N(z)}{\frac{\beta}{v_a} \int_0^z n^0(p) e^{N(p)} dp + \exp \left[- \left(t - t_0 - \frac{z}{v_a} \right) \beta n_{a0} \right]},$$

$$n(t, z) = n_{a0}(z) \frac{\exp[-(t - t_0 - z/v_a)\beta n_{a0}]}{\frac{\beta}{v_a} \int_0^z n^0(p) e^{N(p)} dp + \exp \left[- \left(t - t_0 - \frac{z}{v_a} \right) \beta n_{a0} \right]},$$

$$0 \leq z \leq (t - t_0)v_a, \quad t \geq t_0, z \leq L,$$

где $N(z) = \frac{\beta}{v_a} \int_0^z [n^0(p) + n_a^0(p)] dp$.

Геометрическая иллюстрация решения (18), (19) представлена на фиг. 3.



Фиг. 3. Геометрическая иллюстрация формул (18), (19).

Очевидно, что на граничной прямой $z = (t - t_0)v_a$ оба комплекта формул (18), (19) совпадают. Вернемся к вычислению f^1 , считая $z_j = z_j(t_0 + \Delta t)$, $v_j = v_j(t_0 + \Delta t)$:

$$f^1 = \bar{f} + \int_{t_0}^{t_0 + \Delta t} \beta n(\tau, z) n_a(\tau, z) d\tau \cdot \delta(\mathbf{v} - \mathbf{v}_a) \cong \bar{f} + \Delta t \beta n(t_0, z) n_a(t_0, z) \delta(\mathbf{v} - \mathbf{v}_a) =$$

$$= \bar{f} + \Delta t \beta n^0(z) n_a^0(z) \delta(\mathbf{v} - \mathbf{v}_a) = \bar{f} + \Delta t \beta \sum_j q_j \delta_\epsilon(z - z_j) n_a^0(z) \delta(\mathbf{v} - \mathbf{v}_a) \cong$$

$$\cong \bar{f} + \Delta t \beta \sum_j q_j \delta_\epsilon(z - z_j) n_a^0(z_j) \delta(\mathbf{v} - \mathbf{v}_a) =$$

$$= \sum_j q_j \delta_\epsilon(z - z_j) \delta(\mathbf{v} - \mathbf{v}_j) + \sum_j \beta \Delta t q_j n_a^0(z_j) \delta_\epsilon(z - z_j) \delta(\mathbf{v} - \mathbf{v}_a).$$

Итак, согласно формуле (20), функция распределения f^1 получается из $\bar{f}$ вбрасыванием макрочастиц в точки с координатами $z_j = z_j(t_0 + \Delta t)$, скоростями $\mathbf{V}_a$ и массами $\Delta t \beta q_j n_a^0(z_j(t_0 + \Delta t))$. В цепочке (20) имеется два приближенных равенства. Первое — это формула прямоугольников для вычисления интеграла на отрезке $[t_0, t_0 + \Delta t]$. Второе — $\delta_\epsilon(z - z_j) n_a^0(z) \cong \delta_\epsilon(z - z_j) n_a^0(z_j)$ следует из обращения функции $\delta_\epsilon(z - z_j) n_a^0(z)$ в нуль вне ϵ -окрестности точки z_j (напомним, $\text{supp} \delta_\epsilon \subseteq [-\epsilon, \epsilon]$), поэтому $\delta_\epsilon(z - z_j) n_a^0(z) = \delta_\epsilon(z - z_j) (n_a^0(z_j) + O(\epsilon)) \cong$

$\cong \delta_\varepsilon(z - z_j)n_a^0(z_j)$. Приближенные равенства в цепочке (20) понимаются в слабом смысле, как приближенные равенства определяемых функциями распределениями линейных непрерывных функционалов на пространстве непрерывных финитных функций в фазовом пространстве (см. разд. 3). Согласно (20), начальное количество макрочастиц за n временных шагов увеличивается в 2^n раз, т.е. растет экспоненциально, что является недостатком метода дробных шагов. Однако в сильном продольном электрическом поле, типичном для СПД, количество покидаемых через правую границу макрочастиц, как показывают расчеты, тоже растет экспоненциально. В итоге в канале СПД остается ограниченное сверху и пригодное для расчета число макрочастиц (см. ниже). С другой стороны, в методе дробных шагов нет необходимости в использовании эйлеровой сетки в расчетной области для нахождения координат новых макрочастиц, что является преимуществом этого метода. Кроме того, можно увеличить порядок по Δt точности вычисления интеграла в цепочке преобразований (20), если посредством явных формул (18), (19) разложить $n(t, z)$, $n_a(t, z)$ по степеням $t - t_0 \leq \Delta t$ и ограничиться несколькими первыми членами разложений (при этом надо пользоваться формулами (18), справедливыми при $(t - t_0)v_a \leq z$). Для этой цели полезно преобразовать формулы (18), (19) с учетом тождеств

$$(e^{N(p)})' = \frac{\beta}{v_a} e^{N(p)} (n^0(p) + n_a^0(p)), \quad \frac{\beta}{v_a} \int_0^z n^0(p) e^{N(p)} dp = e^{N(z)} - 1 - \frac{\beta}{v_a} \int_0^z n_a^0(p) e^{N(p)} dp.$$

Тогда, например, формулы (18) преобразуются к виду

$$\begin{aligned} n(t, z) &= n^0(z) \left\{ 1 - \frac{\beta}{v_a} \int_{z-v_a(t-t_0)}^z n_a^0(p) \exp \left[-\frac{\beta}{v_a} \int_p^z (n_a^0(q) + n^0(q)) dq \right] dp \right\}^{-1}, \\ n_a(t, z) &= n_a^0(z - (t - t_0)v_a) \left\{ \exp \left[\frac{\beta}{v_a} \int_{z-v_a(t-t_0)}^z (n_a^0(p) + n^0(p)) dp \right] - \right. \\ &\quad \left. - \frac{\beta}{v_a} \int_{z-v_a(t-t_0)}^z n_a^0(p) \exp \left[-\frac{\beta}{v_a} \int_{z-v_a(t-t_0)}^p (n_a^0(q) + n^0(q)) dq \right] dp \right\}^{-1}, \\ &\quad (t - t_0)v_a \leq z \leq L. \end{aligned}$$

Однако в настоящей работе мы ограничимся аппроксимациями (20).

5. СРАВНЕНИЕ С ДРУГИМИ МЕТОДАМИ И РЕЗУЛЬТАТЫ ЧИСЛЕННЫХ РАСЧЕТОВ

Сравним метод дробных шагов с численным методом, основанном на физическом моделировании ионизации [10–12]. Его суть применительно к случаю плоской симметрии в следующем. Выберем эйлерову сетку $z_k = kh$, $0 \leq k \leq M$, $h = L/M$, на отрезке $[0, L]$. Тогда из физических соображений, согласно уравнению (3), следует, что в момент времени t в единицу времени в ячейке $[z_{k-1}, z_k]$, $1 \leq k \leq M$, вследствие ионизации, рождается ионов в количестве

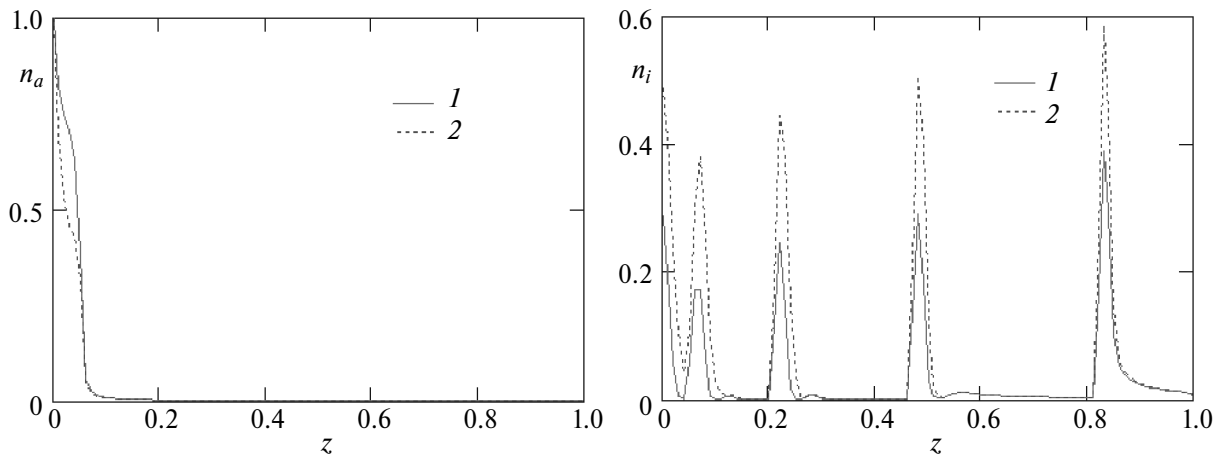
$$\int_{[z_{k-1}, z_k] \times \mathbb{R}^2} \beta n(t, z) n_a(t, z) \delta(\mathbf{v} - \mathbf{v}_a) dz d\mathbf{v} = \beta \int_{z_{k-1}}^{z_k} n(t, z) n_a(t, z) dz.$$

Поэтому на временном шаге $[t_0, t_0 + \Delta t]$ в ячейке $[z_{k-1}, z_k]$ число родившихся ионов приблизительно равно $\beta h \Delta t n(t_0, z_{k-1/2}) n_a(t_0, z_{k-1/2})$, $z_{k-1/2} = (k - 1/2)h$, а их скорость равна v_a . На этом основано физическое моделирование ионизации в методе макрочастиц: на шаге Δt вбрасывается M новых макрочастиц с координатами $z_{k-1/2}$, $1 \leq k \leq M$, скоростью v_a и массой $Q_k = \beta h \Delta t n(t_0, z_{k-1/2}) n_a(t_0, z_{k-1/2})$. Таким образом, число новых макрочастиц в методе физического моделирования ионизации за n временных шагов равно Mn , т.е. растет линейно, что является преимуществом указанного метода.

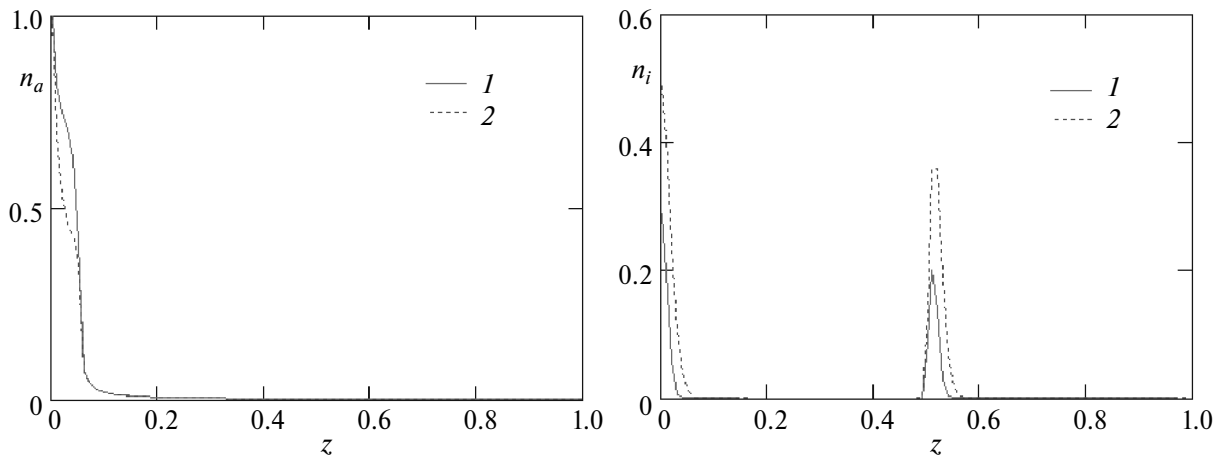
Движение макрочастиц рассчитывалось по системе уравнений:

$$\frac{dz}{dt} = v_z, \quad \frac{dv_y}{dt} = \varepsilon H_x v_z, \quad \frac{dv_z}{dt} = \varepsilon (E_z - H_x v_y), \quad \varepsilon = [t] \omega_{ci}, \quad \omega_{ci} = \frac{e[H]}{cm_i},$$

для следующих параметров $[t] = 10^{-4}$ с, $[H] = 200$ Гс, $m_i = 0.9018 \cdot 10^{-22}$ г, $e = 4.803 \cdot 10^{-10}$ ед. СГСЭ, $\varepsilon = 1$, $H_x = 1$. Здесь квадратные скобки обозначают характерный масштаб соответствующей физической величины.



Фиг. 4. Концентрация атомов $n_a(t, z)$ и ионов $n_i(t, z)$ при $t = 0.5$ и $E_z = 10$, посчитанная методами дробных шагов — 1 и физического моделирования ионизации — 2.



Фиг. 5. Концентрация атомов $n_a(t, z)$ и ионов $n_i(t, z)$ при $t = 0.5$ и $E_z = 100$, посчитанная методами дробных шагов — 1 и физического моделирования ионизации — 2.

Более подробно численный алгоритм, включающий полуявную разностную схему расчета уравнения переноса n_a , при котором значения n_i снимаются с нулевого слоя, изложен в [16]. Проведены расчеты для двух вариантов $E_z = 10$ и $E_z = 100$.

Как следует из приведенных результатов (фиг. 4, 5), качественная картина расчетов по двум представленным методам практически одинакова — различаются только максимальные значения концентраций ионов и атомов. Более того, для достаточно больших E_z количество макрочастиц в канале СПД с течением времени стабилизируется и их количество ограничено. Например, для $E_z = 2 \cdot 10^4$, $\epsilon = 1$, $H_x = 1$ число частиц в канале имеет порядок $\sim 2 \cdot 10^3$. Подчеркнем, что стабилизация количества макрочастиц в канале СПД со временем наступает для любых электрических полей, однако для практических вычислений важно, чтобы это количество было как можно меньше, тем более при расчете на персональном компьютере. Полученное совпадение концентраций n_i , n_a следует признать удовлетворительным, поскольку результаты расчетов по методу макрочастиц нельзя сравнивать в норме C , а самое большее в L_p , $p \geq 1$, так что в остальных точках могут быть существенные расхождения результатов, которые с течением времени, как правило, нарастают. В любом случае за эталонный следует брать результат, полученный по методу дробных шагов.

6. ЗАКЛЮЧЕНИЕ

В работе метод дробных шагов применяется для решения гибридной (кинетической по ионам, гидродинамической по нейтралам) системы уравнений ионизации электрон-ионной плазмы, содержащей нейтралы, применительно к процессам в стационарных плазменных двигателях (СПД), используемых для коррекции орбит летательных аппаратов. Основная идея метода основана на расщеплении по физическим процессам: эво-

люция плазмы представлена как суперпозиция чистого переноса во внешнем поле в фазовом пространстве и последующей ионизации. Процесс ионизации в случае плоской симметрии, которым в работе для простоты и ограничились, рассчитывается по явным аналитическим формулам. Полученные формулы метода дробных шагов сравниваются на примере расчета концентрации атомов и ионов с численным решением уравнений динамики плазмы посредством известной физической модели ионизации.

Метод дробных шагов расчета ионизации не требует использования эйлеровой сетки разбиения счетной области на ячейки, а количество вбрасываемых макрочастиц экспоненциально зависит от числа временных шагов. Результаты сравнения двух методов расчета ионизации, основанных на радикально различных идеях, демонстрирующие их качественную идентичность, следует считать удовлетворительными, поскольку расчеты по методу макрочастиц сравниваются не в равномерной метрике, а самое большее, в интегральной норме. Из сравнения двух методов следует, что метод дробных шагов можно рассматривать как математическое обоснование физической модели ионизации, а последнюю следует использовать, когда ужесточаются требования к объему вычислительных ресурсов при расчете методом дробных шагов. Метод дробных шагов показан к применению при расчете ионизации в сильных продольных электрических полях, типичных для СПД, когда недостаток метода, связанный с экспоненциальным ростом числа макрочастиц компенсируется экспоненциальным ростом числа макрочастиц, покидающих канал СПД через правую границу.

СПИСОК ЛИТЕРАТУРЫ

1. *Морозов А.И.* Физические основы космических электрореактивных двигателей. М.: Атомиздат, 1978.
2. *Морозов А.И.* Введение в плазмодинамику. М.: Физматлит, 2006. 576 с.
3. *Козубский К.Н., Мурашко В.М., Рылов В.П., Трифонов Ю.В., Ходенко В.П., Ким В.П., Попов Г.А., Обухов В.А.* СПД работает в космосе // Физика плазмы. 2003. Т. 29. № 3. С. 277–292.
4. *Kim V., Kozubsky K.N., Murashko V.M., Semenko A.V.* History of the Hall Thrusters Development in USSR // Paper IEPC-2007-142. 30th International Electric Propulsion Conference. September 17–20, 2007. Florence. Italy.
5. *Ким В.П., Семенкин А.В., Хартов С.А.* Конструктивные и физические особенности двигателей с замкнутым дрейфом электронов. М.: Изд-во МАИ, 2016. 160 с.
6. *Mitrofanova O.A., Gnizdor R.Yu., Murashko V.M., Koryakin A.I., Nesterenko A.N.* New Generation of SPT-100. 32nd International Electric Propulsion Conference, Wiesbaden, Germany. September 11–15, 2011, 7 p. IEPC-2011-041.
7. *Garrigues L., Heron A., Adam J.C., Boeuf J.P.* Hybrid and Particle-In-Cell Models of a Stationary Plasma Thruster // Plasma Source Sci. Technol. 9. 2000. P.219–226.
8. *Morozov A.I., Savelyev V.V.* Fundamentals of Stationary Plasma Thruster Theory // Reviews of Plasma Physics. P. 203–391.
9. *Морозов А.И.* Эффект пристеночной проводимости в хорошо замагниченной плазме // ПМТФ. 1968. № 3. С. 19–23.
10. *Волков Б.И., Якунин С.А.* Математические задачи плазмooптики. М.: Знание, 1982. 64 с. Новое в жизни, науке, технике. Сер. “Математика, кибернетика”; № 11.
11. *Быков А.А., Попов В.Ю., Свешников А.Г., Якунин С.А.* Внутренние переходные слои потенциала в сильно замагниченной плазме // Матем. моделирование. Т. 1. № 6. 1989. С. 33–47.
12. *Свешников А.Д., Якунин С.А.* Математическое моделирование нестационарных процессов в плазмooптических системах // Ж. вычисл. матем. и матем. физ. 1983. Т. 23. № 5. С. 1141–1157.
13. *Яненко Н.Н.* Метод дробных шагов решения многомерных задач математической физики. Новосибирск: Наука, 1967.
14. *Самарский А.А.* Теория разностных схем. М.: Наука, 1977.
15. *Климонтович Ю.Л.* Статистическая теория открытых схем. М.: ТОО Янус, 1995. 624 с.
16. *Гавриков М.Б., Таюрский А.А.* Гибридная модель стационарного плазменного двигателя // Препринты ИПМ им. М.В. Келдыша. 2021. № 35. 48 с.

17. Сигов Ю.С. Вычислительный эксперимент: мост между прошлым и будущим физики плазмы. Избранные труды / Сост. Г.И. Змиевская, В.Д. Левченко. М.: Физматлит, 2001. 288 с.
18. Халмош П. Теория меры. М.: Иностранлит., 1953. 291 с.
19. Арсеньев А.А. Лекции о кинетических уравнениях. М.: Наука, 1992. 216 с.
20. Березин Ю.А., Вишневский В.А. Метод частиц в динамике разреженной плазмы. Новосибирск: Наука, 1980. 95 с.
21. Хокни Р., Иствуд Дж. Численное моделирование методом частиц. М.: Мир, 1987. 640 с.
22. Бэдсел Ч., Ленгдон А. Физика плазмы и численное моделирование. М.: Энергоатомиздат, 1989. 452 с.
23. Гавриков М.Б., Таюрский А.А. Аналитическое решение смешанных задач для уравнений одномерной ионизации в случае постоянных скоростей атомов и ионов // Препринты ИПМ им. М.В. Келдыша. 2023. № 30. 36 с.

NUMERICAL SOLUTION OF THE IONIZATION EQUATIONS BY STEP FRACTIONS

М. Б. Гавриков^{a,*}, А. А. Таюрский^{a,**}

^a*Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences, Miusskaya sq., 4, Moscow, 125047 Russia*

^{*}*e-mail: mbgavrikov@yandex.ru*

^{**}*e-mail: tayurskiy2001@mail.ru*

Received 11 March, 2024

Revised 11 March, 2024

Accepted 05 May, 2024

Abstract. The fractional steps method is applied to solve electron-ion plasma ionization equations involving neutrals, with a focus on processes in stationary plasma thrusters (SPT). The approach is based on splitting by physical processes: plasma evolution is represented as a superposition of pure transport in an external field in phase space followed by ionization. The particle-in-cell method is used for calculating transport. In the case of planar symmetry, the ionization process is calculated using explicit analytical formulas. The derived formulas of the fractional steps method are compared with the numerical solution of the plasma evolution equation through a well-known physical ionization model.

Keywords: fractional steps method, ionization, particle-in-cell method, physical ionization model.

ОЦЕНКИ РЕШЕНИЙ ОДНОЙ БИОЛОГИЧЕСКОЙ МОДЕЛИ С БЕСКОНЕЧНЫМ РАСПРЕДЕЛЕННЫМ ЗАПАЗДЫВАНИЕМ¹⁾© 2024 г. Т. К. Искаков^{1,2,*}, М. А. Скворцова^{1,2,**}¹630090 Новосибирск, ул. Пирогова, 1, Новосибирский государственный университет, Россия²630090 Новосибирск, пр-т Акад. Коптюга, 4, Институт математики им. С. Л. Соболева СО РАН, Россия

*e-mail: istima92@mail.ru

**e-mail: sm-18-nsu@yandex.ru

Поступила в редакцию 02.04.2024 г.

Переработанный вариант 02.04.2024 г.

Принята к публикации 02.05.2024 г.

Рассматривается модель конкуренции нескольких видов микроорганизмов, описываемая системой нелинейных дифференциальных уравнений с бесконечным распределенным запаздыванием. Изучается случай асимптотической устойчивости положения равновесия, соответствующего выживанию только одного вида и вымиранию всех остальных. Указаны условия на начальные численности видов и начальную концентрацию питательного вещества, при которых система приходит в равновесное состояние, при этом установлены оценки скорости стабилизации. Результаты получены с использованием функционала Ляпунова—Красовского. Библ. 17.

Ключевые слова: модель конкуренции видов, хемостат, уравнения с запаздывающим аргументом, бесконечное распределенное запаздывание, положение равновесия, асимптотическая устойчивость, оценки решений, область притяжения, функционал Ляпунова—Красовского.

DOI: 10.31857/S0044466924080067, EDN: YATLNLK

1. ВВЕДЕНИЕ

Настоящая работа продолжает исследования [1] по изучению асимптотических свойств решений системы дифференциальных уравнений с бесконечным распределённым запаздыванием

$$\begin{aligned} \frac{d}{dt}S(t) &= D(S^0 - S(t)) - \sum_{i=1}^n x_i(t)p_i(S(t)), \\ \frac{d}{dt}x_i(t) &= -Dx_i(t) + \int_{-\infty}^t x_i(\theta)p_i(S(\theta))e^{-D(t-\theta)}K_i(t-\theta)d\theta, \quad i = 1, \dots, n, \end{aligned} \quad (1.1)$$

где

$$K_i(s) = \frac{\alpha_i^{r_i+1}s^{r_i}}{r_i!}e^{-\alpha_i s}, \quad \alpha_i > 0, \quad r_i \in \mathbb{N} \cup \{0\}, \quad s > 0, \quad i = 1, \dots, n. \quad (1.2)$$

Система была предложена в работе [2] для описания конкуренции нескольких видов микроорганизмов, находящихся в хемостате (хемостат — аппарат для непрерывного культивирования бактерий, обеспечивающий оптимальные температурные условия и постоянное поступление свежей питательной среды при одновременном удалении части бактериальной культуры). Функции и параметры, входящие в систему, имеют следующий биологический смысл: $S(t)$ — концентрация питательного вещества в момент времени t , $S^0 > 0$ — концентрация вводимого питательного вещества, $D > 0$ — скорость поступления питательного вещества, $x_i(t)$ — концентрация i -го вида микроорганизмов в момент времени t . Функции поглощения питательных веществ $p_i(\xi)$ предполагаются строго монотонно возрастающими, локально липшицевыми, при этом $p_i(0) = 0$. Показатели смертности, характерные для видов, считаются незначительными по сравнению со скоростью поступления питательного вещества и в данной модели не учитываются. Каждое ядро $K_i(s)$ связано с преобразованием питательного вещества в жизнеспособные клетки.

¹⁾ Работа выполнена при поддержке Математического Центра в Академгородке, соглашение № 075-15-2022-282 с Минобрнауки РФ.

Более подробное описание модели содержится в [2]. Некоторые другие модели, описывающие биологические процессы в хемостате, содержатся, например, в [3, 4].

Вместе с системой (1.1) рассмотрим начальные условия

$$\begin{aligned} S(t) &= \varphi_0(t), \quad t < 0, \quad S(+0) = \varphi_0(0), \\ x_i(t) &= \varphi_i(t), \quad t < 0, \quad x_i(+0) = \varphi_i(0), \quad i = 1, \dots, n, \end{aligned} \quad (1.3)$$

где $\varphi_i(t) \in C((-\infty, 0])$, $i = 0, 1, \dots, n$, — заданные непрерывные ограниченные функции. Учитывая биологический смысл задачи, будем предполагать, что начальные данные неотрицательны:

$$\varphi_i(t) \geq 0, \quad t \leq 0, \quad i = 0, 1, \dots, n. \quad (1.4)$$

Отметим, что начальная задача (1.1), (1.3) однозначно разрешима. Более того, в работе [2] было показано, что при условии (1.4) компоненты решения начальной задачи (1.1), (1.3) также неотрицательны при всех $t > 0$ и ограничены сверху.

Теперь рассмотрим положения равновесия системы (1.1) с неотрицательными компонентами, указанные в [2].

При любых значениях параметров у системы (1.1) существует положение равновесия $E_0 = (S^0, 0, \dots, 0)$, соответствующее полному вымиранию всех видов микроорганизмов. В работе [2] было показано, что при выполнении условий

$$p_i(S^0) \leq D \left(\frac{D + \alpha_i}{\alpha_i} \right)^{r_i+1}, \quad i = 1, \dots, n, \quad (1.5)$$

первая компонента $S(t)$ решения начальной задачи (1.1), (1.3) сходится к S^0 при $t \rightarrow +\infty$, а все остальные компоненты $x_i(t)$, $i = 1, \dots, n$, стремятся к нулю, т. е.

$$(S(t), x_1(t), \dots, x_n(t)) \rightarrow (S^0, 0, \dots, 0) \quad \text{при} \quad t \rightarrow +\infty.$$

Оценки скорости сходимости решений к положению равновесия $(S^0, 0, \dots, 0)$ были установлены в работе [1] в случае, когда в формуле (1.5) все неравенства — строгие. При получении оценок в [1] использовались функционалы Ляпунова—Красовского.

При определенных значениях параметров системы (1.1) также имеются положения равновесия $E_i = (S_i^*, 0, \dots, 0, x_i^*, 0, \dots, 0)$, соответствующие выживанию только i -го вида и вымиранию всех остальных. Сформулируем условия, при которых эти решения существуют и имеют неотрицательные компоненты. Вначале введем обозначения [2]. Пусть $\lambda_i \in (0, +\infty]$, $i = 1, \dots, n$, — такие величины, что выполнены неравенства:

$$\begin{aligned} p_i(S) &< D \left(\frac{D + \alpha_i}{\alpha_i} \right)^{r_i+1} \quad \text{при} \quad S < \lambda_i, \\ p_i(S) &> D \left(\frac{D + \alpha_i}{\alpha_i} \right)^{r_i+1} \quad \text{при} \quad S > \lambda_i. \end{aligned}$$

Заметим, что в силу строгой монотонности функций $p_i(S)$ величины λ_i всегда существуют, при этом если $\lambda_i < +\infty$, то

$$p_i(\lambda_i) = D \left(\frac{D + \alpha_i}{\alpha_i} \right)^{r_i+1}. \quad (1.6)$$

В этом случае положение равновесия E_i имеет вид

$$E_i = (S_i^*, 0, \dots, 0, x_i^*, 0, \dots, 0) = \left(\lambda_i, 0, \dots, 0, \left(\frac{\alpha_i}{D + \alpha_i} \right)^{r_i+1} (S^0 - \lambda_i), 0, \dots, 0 \right).$$

При условии $\lambda_i < S^0$ компонента решения $x_i(t) = x_i^*$ является положительной.

В работе [2] исследовался вопрос о сходимости решений системы (1.1) к положению равновесия E_i , однако вопрос об оценках скорости сходимости пока не рассматривался.

В настоящей работе мы укажем условия на функции и параметры системы, а также на начальные данные, при которых решение начальной задачи (1.1), (1.3) сходится к положению равновесия E_i , при этом мы получим оценки решений, характеризующие скорость сходимости. Так же, как и в работе [1], при получении оценок мы будем использовать функционалы Ляпунова—Красовского.

Отметим, что функционалы Ляпунова—Красовского широко используются для получения оценок решений различных классов систем с запаздыванием (см., например, работы [5–10] о системах с сосредоточенным запаздыванием, работы [11, 12] о системах с распределенным запаздыванием, обзор [13], посвященный применению второго метода Ляпунова для систем с запаздывающим аргументом, и имеющуюся там библиографию). Некоторые результаты об оценках решений биологических моделей, описываемых уравнениями с запаздыванием, содержатся, например, в работах [1, 14–17].

2. АСИМПТОТИЧЕСКАЯ УСТОЙЧИВОСТЬ ПОЛОЖЕНИЯ РАВНОВЕСИЯ E_i

В этом разделе мы укажем достаточные условия асимптотической устойчивости положения равновесия E_i , которые получаются с помощью теоремы об устойчивости по первому приближению. Без ограничения общности будем считать, что $i = 1$, т. е. положение равновесия имеет вид

$$E_1 = (S_1^*, x_1^*, 0, \dots, 0) = \left(\lambda_1, \left(\frac{\alpha_1}{D + \alpha_1} \right)^{r_1+1} (S^0 - \lambda_1), 0, \dots, 0 \right).$$

Предположим, что выполнено условие $\lambda_1 < S^0$, при котором $x_1^* > 0$, а также, что функция $p_1(S)$ является непрерывно дифференцируемой в некоторой окрестности точки $S = \lambda_1$.

Для начала сведем задачу об устойчивости положения равновесия E_1 к задаче об устойчивости нулевого решения. Сделаем замену переменных

$$S(t) = S_1^* + y_0(t), \quad x_1(t) = x_1^* + y_1(t), \quad x_i(t) = y_i(t), \quad i = 2, \dots, n.$$

Тогда начальная задача (1.1), (1.3) преобразуется к виду

$$\frac{d}{dt}y(t) = Ay(t) + F(y(t)) + \sum_{i=1}^n \int_{-\infty}^t \left(B_i y(\theta) + G_i(y(\theta)) \right) e^{-D(t-\theta)} K_i(t-\theta) d\theta, \quad (2.1)$$

$$y(t) = \psi(t), \quad t < 0, \quad y(+0) = \psi(0),$$

где

$$y(t) = \begin{pmatrix} y_0(t) \\ y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix}, \quad \psi(t) = \begin{pmatrix} \psi_0(t) \\ \psi_1(t) \\ \psi_2(t) \\ \vdots \\ \psi_n(t) \end{pmatrix} = \begin{pmatrix} \varphi_0(t) \\ \varphi_1(t) \\ \varphi_2(t) \\ \vdots \\ \varphi_n(t) \end{pmatrix} - \begin{pmatrix} S_1^* \\ x_1^* \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \varphi(t) - E_1,$$

$$A = \begin{pmatrix} -(D + x_1^* p_1'(S_1^*)) & -p_1(S_1^*) & -p_2(S_1^*) & \cdots & -p_n(S_1^*) \\ 0 & -D & 0 & \cdots & 0 \\ 0 & 0 & -D & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -D \end{pmatrix},$$

$$B_1 = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ x_1^* p_1'(S_1^*) & p_1(S_1^*) & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

$$B_2 = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & p_2(S_1^*) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \quad \dots, \quad B_n = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & p_n(S_1^*) \end{pmatrix},$$

$$F(y(t)) = - \begin{pmatrix} x_1^* f_0(y_0(t)) + \sum_{i=1}^n y_i(t) f_i(y_0(t)) \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad (2.2)$$

$$G_1(y(\theta)) = \begin{pmatrix} 0 \\ x_1^* f_0(y_0(\theta)) + y_1(\theta) f_1(y_0(\theta)) \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad (2.3)$$

$$G_2(y(\theta)) = \begin{pmatrix} 0 \\ 0 \\ y_2(\theta) f_2(y_0(\theta)) \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad G_n(y(\theta)) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ y_n(\theta) f_n(y_0(\theta)) \end{pmatrix},$$

$$f_0(y_0) = p_1(S_1^* + y_0) - p_1(S_1^*) - p_1'(S_1^*) y_0, \quad (2.4)$$

$$f_i(y_0) = p_i(S_1^* + y_0) - p_i(S_1^*), \quad i = 1, \dots, n. \quad (2.5)$$

Рассмотрим систему линейного приближения

$$\frac{d}{dt} y(t) = Ay(t) + \sum_{i=1}^n \int_{-\infty}^t B_i y(\theta) e^{-D(t-\theta)} K_i(t-\theta) d\theta. \quad (2.6)$$

Укажем достаточные условия асимптотической устойчивости нулевого решения данной системы. Для этого выпишем характеристическое уравнение:

$$P(\mu) = \det \left(\mu I - A - \sum_{i=1}^n B_i \int_{-\infty}^t e^{-(D+\mu)(t-\theta)} K_i(t-\theta) d\theta \right) = 0.$$

Учитывая явный вид (1.2) функций $K_i(s)$, $s > 0$, вычислим значение интеграла:

$$\omega_i(\mu) = \int_{-\infty}^t e^{-(D+\mu)(t-\theta)} K_i(t-\theta) d\theta = \left(\frac{\alpha_i}{D + \alpha_i + \mu} \right)^{r_i+1}. \quad (2.7)$$

Тем самым, характеристическое уравнение будет иметь вид

$$P(\mu) = \det \left(\mu I - A - \sum_{i=1}^n \omega_i(\mu) B_i \right) = 0.$$

Используя явный вид матриц A , B_i , $i = 1, \dots, n$, получим

$$P(\mu) = (\mu + D) \left(\mu + D - \omega_1(\mu) p_1(S_1^*) + x_1^* p_1'(S_1^*) \right) \prod_{i=2}^n \left(\mu + D - \omega_i(\mu) p_i(S_1^*) \right) = 0.$$

Наконец, в силу формул (1.6) и (2.7) функцию $P(\mu)$ можно переписать в виде

$$P(\mu) = (\mu + D) \left(\mu + D - \frac{\omega_1(\mu)}{\omega_1(0)} D + x_1^* p_1'(\lambda_1) \right) \prod_{i=2}^n \left(\mu + D - \frac{\omega_i(\mu)}{\omega_i(0)} \frac{p_i(\lambda_1)}{p_i(\lambda_i)} D \right) = 0.$$

Утверждается, что при выполнении условий

$$x_1^* > 0, \quad p_1'(\lambda_1) > 0, \quad \lambda_1 < \lambda_i, \quad i = 2, \dots, n, \quad (2.8)$$

все корни характеристического уравнения $P(\mu) = 0$ содержатся в левой полуплоскости $\mathbb{C}_- = \{\mu \in \mathbb{C} : \operatorname{Re} \mu < 0\}$. Действительно, пусть число $\mu \in \mathbb{C}$ такое, что $\operatorname{Re} \mu \geq 0$. В этом случае из определения (2.7) следует, что $|\omega_i(\mu)| \leq |\omega_i(0)|$, а в силу монотонности функций $p_i(S)$ имеем $p_i(\lambda_1) < p_i(\lambda_i)$, $i = 2, \dots, n$. Тогда

$$\left| \mu + D - \frac{\omega_1(\mu)}{\omega_1(0)} D + x_1^* p_1'(\lambda_1) \right| \geq |\mu + D + x_1^* p_1'(\lambda_1)| - \left| \frac{\omega_1(\mu)}{\omega_1(0)} D \right| \geq x_1^* p_1'(\lambda_1) > 0,$$

$$\left| \mu + D - \frac{\omega_i(\mu)}{\omega_i(0)} \frac{p_i(\lambda_1)}{p_i(\lambda_i)} D \right| \geq |\mu + D| - \left| \frac{\omega_i(\mu)}{\omega_i(0)} \frac{p_i(\lambda_1)}{p_i(\lambda_i)} D \right| > 0,$$

тем самым, число $\mu \in \mathbb{C}$, $\operatorname{Re} \mu \geq 0$, не может быть корнем характеристического уравнения.

Итак, при выполнении условий (2.8) нулевое решение линейной системы (2.6) асимптотически устойчиво, а следовательно, асимптотически устойчиво положение равновесия E_1 системы (1.1). Всюду далее будем предполагать, что условия (2.8) выполняются.

3. ФУНКЦИОНАЛ ЛЯПУНОВА–КРАСОВСКОГО

В этом параграфе мы рассмотрим линейную систему (2.6) в предположении, что выполнены условия (2.8), гарантирующие асимптотическую устойчивость нулевого решения. Мы построим функционал Ляпунова–Красовского, с помощью которого установим оценки решений линейной системы (2.6), характеризующие экспоненциальное убывание на бесконечности. В дальнейшем на основе построенного функционала будут получены условия на начальные данные и оценки решений исходной системы (1.1), характеризующие скорость сходимости к положению равновесия E_1 .

Принимая во внимание результаты работ [1, 12], рассмотрим функционал Ляпунова–Красовского следующего вида:

$$V(t, y) = \langle Hy(t), y(t) \rangle + \sum_{i=1}^n \int_0^\infty \int_{t-\theta}^t e^{-m(t-\eta)} \langle \Upsilon_i(\theta) y(\eta), y(\eta) \rangle d\eta d\theta, \quad (3.1)$$

где матрицы $H = H^* > 0$, $\Upsilon_i(\theta) = \Upsilon_i^*(\theta) \geq 0$, $i = 1, \dots, n$, и величина $m > 0$ будут определены ниже.

Вычислим производную от функционала $V(t, y)$ вдоль решения системы (2.6):

$$\begin{aligned} \frac{d}{dt} V(t, y) &= \left\langle H \frac{d}{dt} y(t), y(t) \right\rangle + \left\langle Hy(t), \frac{d}{dt} y(t) \right\rangle + \\ &+ \sum_{i=1}^n \int_0^\infty \langle \Upsilon_i(\theta) y(t), y(t) \rangle d\theta - \sum_{i=1}^n \int_0^\infty e^{-m\theta} \langle \Upsilon_i(\theta) y(t-\theta), y(t-\theta) \rangle d\theta - \\ &- m \sum_{i=1}^n \int_0^\infty \int_{t-\theta}^t e^{-m(t-\eta)} \langle \Upsilon_i(\theta) y(\eta), y(\eta) \rangle d\eta d\theta. \end{aligned}$$

Поскольку $y(t)$ — решение (2.6), то производная переписывается в виде

$$\begin{aligned} \frac{d}{dt} V(t, y) &= \left\langle H \left(Ay(t) + \sum_{i=1}^n \int_0^\infty B_i y(t-\xi) e^{-D\xi} K_i(\xi) d\xi \right), y(t) \right\rangle + \\ &+ \left\langle Hy(t), \left(Ay(t) + \sum_{i=1}^n \int_0^\infty B_i y(t-\xi) e^{-D\xi} K_i(\xi) d\xi \right) \right\rangle + \\ &+ \sum_{i=1}^n \int_0^\infty \langle \Upsilon_i(\xi) y(t), y(t) \rangle d\xi - \sum_{i=1}^n \int_0^\infty e^{-m\xi} \langle \Upsilon_i(\xi) y(t-\xi), y(t-\xi) \rangle d\xi - \\ &- m \sum_{i=1}^n \int_0^\infty \int_{t-\theta}^t e^{-m(t-\eta)} \langle \Upsilon_i(\theta) y(\eta), y(\eta) \rangle d\eta d\theta. \end{aligned}$$

Положим

$$\Upsilon_i(\xi) = \delta_i e^{-D\xi} K_i(\xi) e^{m\xi/2} B_i^* B_i, \quad \delta_i > 0, \quad i = 1, \dots, n, \quad (3.2)$$

тогда

$$\begin{aligned} \frac{d}{dt} V(t, y) &= \langle (HA + A^* H) y(t), y(t) \rangle + \\ &+ \sum_{i=1}^n \int_0^\infty e^{-D\xi} K_i(\xi) \left(\delta_i e^{m\xi/2} \langle B_i^* B_i y(t), y(t) \rangle + \langle H B_i y(t - \xi), y(t) \rangle + \right. \\ &\quad \left. + \langle B_i^* H y(t), y(t - \xi) \rangle - \delta_i e^{-m\xi/2} \langle B_i^* B_i y(t - \xi), y(t - \xi) \rangle \right) d\xi - \\ &- m \sum_{i=1}^n \int_0^\infty \int_{t-\theta}^t e^{-m(t-\eta)} \langle \Upsilon_i(\theta) y(\eta), y(\eta) \rangle d\eta d\theta. \end{aligned}$$

Введем обозначения:

$$H = \begin{pmatrix} h_{00} & h_{01} & 0 & \dots & 0 \\ h_{01} & h_{11} & 0 & \dots & 0 \\ 0 & 0 & h_{22} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & h_{nn} \end{pmatrix}, \quad \tilde{H}_1 = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ h_{01} & h_{11} & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \quad (3.3)$$

$$\tilde{H}_2 = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & h_{22} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \quad \dots, \quad \tilde{H}_n = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & h_{nn} \end{pmatrix}. \quad (3.4)$$

Учитывая явный вид матриц $B_i, i = 1, \dots, n$, получим

$$B_i^* H = B_i^* \tilde{H}_i, \quad H B_i = \tilde{H}_i^* B_i, \quad i = 1, \dots, n.$$

В этом случае

$$\begin{aligned} &\langle H B_i y(t - \xi), y(t) \rangle + \langle B_i^* H y(t), y(t - \xi) \rangle - \delta_i e^{-m\xi/2} \langle B_i^* B_i y(t - \xi), y(t - \xi) \rangle = \\ &= \langle \tilde{H}_i^* B_i y(t - \xi), y(t) \rangle + \langle B_i^* \tilde{H}_i y(t), y(t - \xi) \rangle - \delta_i e^{-m\xi/2} \langle B_i^* B_i y(t - \xi), y(t - \xi) \rangle \leq \\ &\leq \frac{e^{m\xi/2}}{\delta_i} \langle \tilde{H}_i^* \tilde{H}_i y(t), y(t) \rangle. \end{aligned}$$

Отсюда следует неравенство

$$\begin{aligned} \frac{d}{dt} V(t, y) &\leq \langle (HA + A^* H) y(t), y(t) \rangle + \\ &+ \sum_{i=1}^n \int_0^\infty e^{-(D-m/2)\xi} K_i(\xi) \left\langle \left(\delta_i B_i^* B_i + \frac{1}{\delta_i} \tilde{H}_i^* \tilde{H}_i \right) y(t), y(t) \right\rangle d\xi - \\ &- m \sum_{i=1}^n \int_0^\infty \int_{t-\theta}^t e^{-m(t-\eta)} \langle \Upsilon_i(\theta) y(\eta), y(\eta) \rangle d\eta d\theta. \end{aligned}$$

В силу обозначения (2.7) данную оценку можно переписать в виде

$$\frac{d}{dt} V(t, y) \leq -\langle U y(t), y(t) \rangle - m \sum_{i=1}^n \int_0^\infty \int_{t-\theta}^t e^{-m(t-\eta)} \langle \Upsilon_i(\theta) y(\eta), y(\eta) \rangle d\eta d\theta, \quad (3.5)$$

где

$$U = -HA - A^*H - \sum_{i=1}^n \omega_i \left(-\frac{m}{2} \right) \left(\delta_i B_i^* B_i + \frac{1}{\delta_i} \tilde{H}_i^* \tilde{H}_i \right). \quad (3.6)$$

Наша цель — подобрать элементы матрицы H , величину $m > 0$ и числа $\delta_i > 0$, $i = 1, \dots, n$, так, чтобы было выполнено матричное неравенство $U > 0$. Имеем

$$U = \begin{pmatrix} u_{00} & u_{01} & u_{02} & \cdots & u_{0n} \\ u_{01} & u_{11} & u_{12} & \cdots & u_{1n} \\ u_{02} & u_{12} & u_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u_{0n} & u_{1n} & 0 & \cdots & u_{nn} \end{pmatrix},$$

где

$$\begin{aligned} u_{00} &= 2h_{00}(D + x_1^* p_1'(S_1^*)) - \omega_1 \left(-\frac{m}{2} \right) \left(\delta_1 (x_1^* p_1'(S_1^*))^2 + \frac{h_{01}^2}{\delta_1} \right), \\ u_{01} &= h_{00} p_1(S_1^*) + h_{01} (2D + x_1^* p_1'(S_1^*)) - \omega_1 \left(-\frac{m}{2} \right) \left(\delta_1 x_1^* p_1'(S_1^*) p_1(S_1^*) + \frac{h_{01} h_{11}}{\delta_1} \right), \\ u_{11} &= 2h_{01} p_1(S_1^*) + 2h_{11} D - \omega_1 \left(-\frac{m}{2} \right) \left(\delta_1 p_1^2(S_1^*) + \frac{h_{11}^2}{\delta_1} \right), \\ u_{0i} &= h_{00} p_i(S_1^*), \quad i = 2, \dots, n, \\ u_{1i} &= h_{01} p_i(S_1^*), \quad i = 2, \dots, n, \\ u_{ii} &= 2h_{ii} D - \omega_i \left(-\frac{m}{2} \right) \left(\delta_i p_i^2(S_1^*) + \frac{h_{ii}^2}{\delta_i} \right), \quad i = 2, \dots, n. \end{aligned}$$

Выберем числа h_{ii} , $i = 1, \dots, n$, так, чтобы величины u_{ii} , $i = 1, \dots, n$, принимали максимальное значение:

$$h_{11} = \delta_1 D \left(\omega_1 \left(-\frac{m}{2} \right) \right)^{-1}, \quad (3.7)$$

$$h_{ii} = \delta_i D \left(\omega_i \left(-\frac{m}{2} \right) \right)^{-1}, \quad i = 2, \dots, n. \quad (3.8)$$

В этом случае

$$\begin{aligned} u_{11} &= 2h_{01} p_1(S_1^*) - \delta_1 \left(\omega_1 \left(-\frac{m}{2} \right) \right)^{-1} \left[\left(\omega_1 \left(-\frac{m}{2} \right) \right)^2 p_1^2(S_1^*) - D^2 \right], \\ u_{ii} &= \delta_i \left(\omega_i \left(-\frac{m}{2} \right) \right)^{-1} \left[D^2 - \left(\omega_i \left(-\frac{m}{2} \right) \right)^2 p_i^2(S_1^*) \right], \quad i = 2, \dots, n. \end{aligned} \quad (3.9)$$

Числа h_{00} , h_{01} выберем так, чтобы выполнялись равенства $u_{01} = 0$, $u_{00} = u_{11}$, т. е.

$$h_{01} = \frac{\delta_1}{p_1(S_1^*)} \left(\omega_1 \left(-\frac{m}{2} \right) \right)^{-1} \left[\sqrt{W} - \left[(D + \gamma_1)^2 + p_1^2(S_1^*) \right] \right], \quad (3.10)$$

$$\begin{aligned} h_{00} &= \frac{\delta_1}{p_1^2(S_1^*)} (D + \gamma_1) \left(\omega_1 \left(-\frac{m}{2} \right) \right)^{-1} \\ &\times \left[\left(\omega_1 \left(-\frac{m}{2} \right) \right)^2 \frac{\gamma_1 p_1^2(S_1^*)}{(D + \gamma_1)} + \left[(D + \gamma_1)^2 + p_1^2(S_1^*) \right] - \sqrt{W} \right], \end{aligned} \quad (3.11)$$

где

$$\begin{aligned} W &= \left[(D + \gamma_1)^2 + p_1^2(S_1^*) \right]^2 + p_1^2(S_1^*) \left[\left(\omega_1 \left(-\frac{m}{2} \right) \right)^2 (2D\gamma_1 + \gamma_1^2 + p_1^2(S_1^*)) - D^2 \right], \\ \gamma_1 &= x_1^* p_1'(S_1^*) > 0. \end{aligned}$$

Тогда

$$u_{00} = u_{11} = \delta_1 \left(\omega_1 \left(-\frac{m}{2} \right) \right)^{-1} \left\{ 2 \left[\sqrt{W} - \left[(D + \gamma_1)^2 + p_1^2(S_1^*) \right] \right] - \left[\left(\omega_1 \left(-\frac{m}{2} \right) \right)^2 p_1^2(S_1^*) - D^2 \right] \right\}, \quad u_{01} = 0.$$

Величину $m > 0$ найдем из условий $u_{ii} > 0, i = 0, 1, \dots, n$. Условие $u_{00} = u_{11} > 0$ дает

$$\left(\omega_1 \left(-\frac{m}{2} \right) \right)^2 p_1^2(S_1^*) < D^2 + 2D\gamma_1. \quad (3.12)$$

Условия $u_{ii} > 0, i = 2, \dots, n$, дают

$$\left(\omega_i \left(-\frac{m}{2} \right) \right)^2 p_i^2(S_1^*) < D^2, \quad i = 2, \dots, n. \quad (3.13)$$

Заметим, что в силу обозначений (1.6), (2.7) при $m = 0$ условие (3.12) эквивалентно неравенству $\gamma_1 = x_1^* p_1'(\lambda_1) > 0$, а условия (3.13) — неравенствам $\lambda_1 < \lambda_i, i = 2, \dots, n$, т. е. при $m = 0$ неравенства (3.12), (3.13) совпадают с (2.8). Следовательно, в силу непрерывности функций $\omega_i(\mu), i = 1, \dots, n$, условия (3.12), (3.13) выполнены и при некотором $m > 0$.

Осталось подобрать числа $\delta_i > 0, i = 1, \dots, n$. Учитывая структуру матрицы U , для произвольного вектора $z = (z_0, z_1, \dots, z_n)^T \in \mathbb{R}^{n+1}$ будем иметь

$$\begin{aligned} \langle Uz, z \rangle &= u_{00}z_0^2 + u_{00}z_1^2 + \sum_{i=2}^n u_{ii}z_i^2 + \sum_{i=2}^n 2u_{0i}z_0z_i + \sum_{i=2}^n 2u_{1i}z_1z_i \\ &\geq \left(u_{00} - \sum_{i=2}^n \beta_{0i}u_{0i} \right) z_0^2 + \left(u_{00} - \sum_{i=2}^n \beta_{1i}u_{1i} \right) z_1^2 + \sum_{i=2}^n \left(u_{ii} - \frac{u_{0i}}{\beta_{0i}} - \frac{u_{1i}}{\beta_{1i}} \right) z_i^2, \end{aligned}$$

где $\beta_{0i} > 0, \beta_{1i} > 0, i = 2, \dots, n$. Положим

$$\beta_{0i} = \frac{u_{00}}{2(n-1)u_{0i}}, \quad \beta_{1i} = \frac{u_{00}}{2(n-1)u_{1i}}, \quad i = 2, \dots, n,$$

тогда

$$\langle Uz, z \rangle \geq \frac{u_{00}}{2} z_0^2 + \frac{u_{00}}{2} z_1^2 + \sum_{i=2}^n \left(u_{ii} - \frac{2(n-1)(u_{0i}^2 + u_{1i}^2)}{u_{00}} \right) z_i^2.$$

Учитывая явный вид (3.9) величин $u_{ii}, i = 2, \dots, n$, выберем числа $\delta_i, i = 2, \dots, n$, следующим образом:

$$\begin{aligned} \delta_i &= \left(\omega_i \left(-\frac{m}{2} \right) \right) \left[D^2 - \left(\omega_i \left(-\frac{m}{2} \right) \right)^2 p_i^2(S_1^*) \right]^{-1} \\ &\quad \times \frac{4(n-1)(u_{0i}^2 + u_{1i}^2)}{u_{00}}, \quad i = 2, \dots, n. \end{aligned} \quad (3.14)$$

Величину δ_1 можно выбрать произвольно, например,

$$\delta_1 = 1. \quad (3.15)$$

Тем самым, функционал $V(t, y)$ полностью определен, при этом справедлива оценка

$$\langle Uz, z \rangle \geq \frac{u_{00}}{2} z_0^2 + \frac{u_{00}}{2} z_1^2 + \sum_{i=2}^n \frac{u_{ii}}{2} z_i^2,$$

откуда следует, что матрица U является положительно определенной.

Заметим, что при выполнении всех вышеперечисленных условий матрица H также положительно определена. Действительно, из формулы (3.6) вытекает, что $HA + A^*H \leq -U$, т. е. матрица H является решением матричного уравнения Ляпунова

$$HA + A^*H = -C, \quad C = C^* \geq U > 0.$$

Так как все собственные значения матрицы A содержатся в левой полуплоскости $\mathbb{C}_-$, то $H = H^* > 0$.
Теперь найдем такую величину $\sigma > 0$, чтобы было выполнено неравенство

$$U \geq \sigma H. \quad (3.16)$$

Поскольку

$$\begin{pmatrix} \frac{u_{00}}{2} & 0 \\ 0 & \frac{u_{00}}{2} \end{pmatrix} \geq \sigma_1 \begin{pmatrix} h_{00} & h_{01} \\ h_{01} & h_{11} \end{pmatrix}, \quad \frac{u_{ii}}{2} \geq \sigma_i h_{ii}, \quad i = 2, \dots, n,$$

где

$$\sigma_1 = \frac{u_{00}}{4(h_{00}h_{11} - h_{01}^2)} \left[(h_{00} + h_{11}) - \sqrt{(h_{00} + h_{11})^2 - 4(h_{00}h_{11} - h_{01}^2)} \right],$$

$$\sigma_i = \frac{u_{ii}}{2h_{ii}}, \quad i = 2, \dots, n,$$

то положим

$$\sigma = \min \{ \sigma_1, \sigma_2, \dots, \sigma_n \}, \quad (3.17)$$

откуда следует, что

$$\langle Uz, z \rangle \geq \frac{u_{00}}{2} z_0^2 + \frac{u_{00}}{2} z_1^2 + \sum_{i=2}^n \frac{u_{ii}}{2} z_i^2 \geq \sigma \langle Hz, z \rangle.$$

Тогда, учитывая неравенство (3.5), получим оценку на производную функционала $V(t, y)$:

$$\frac{d}{dt} V(t, y) \leq -\sigma \langle Hy(t), y(t) \rangle - m \sum_{i=1}^n \int_0^\infty \int_{t-\theta}^t e^{-m(t-\eta)} \langle \Upsilon_i(\theta) y(\eta), y(\eta) \rangle d\eta d\theta.$$

Полагая

$$\varepsilon = \min \{ \sigma, m \},$$

будем иметь

$$\frac{d}{dt} V(t, y) \leq -\varepsilon V(t, y),$$

откуда

$$V(t, y) \leq V(0, y) e^{-\varepsilon t}, \quad t > 0.$$

Отсюда следует оценка для решения $y(t)$ линейной системы (2.6):

$$\|y(t)\| \leq \sqrt{\|H^{-1}\|} V^{1/2}(0, y) e^{-\varepsilon t/2}, \quad t > 0.$$

4. ОСНОВНОЙ РЕЗУЛЬТАТ

В предыдущем разделе мы построили функционал Ляпунова–Красовского для линеаризованной системы (2.6) вида (3.1), указав значения матриц $H = H^* > 0$, $\Upsilon_i(\xi) = \Upsilon_i^*(\xi) \geq 0$, $i = 1, \dots, n$, и числа $m > 0$, при которых матрица U из (3.6) будет положительно определенной. Это влечет асимптотическую устойчивость нулевого решения системы (2.6). Обозначим через $u_{\min}^H > 0$ минимальное собственное значение положительно определенной матрицы $U_H = H^{-1/2} U H^{-1/2}$.

Определим матрицы H_i , $i = 1, \dots, n$:

$$H_1 = \begin{pmatrix} h_{00} & h_{01} & 0 & \cdots & 0 \\ h_{01} & h_{11} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \quad H_i = \tilde{H}_i, \quad i = 2, \dots, n, \quad (4.1)$$

где $\tilde{H}_i$, $i = 2, \dots, n$, заданы в (3.4). Отметим, что

$$H = \sum_{i=1}^n H_i. \quad (4.2)$$

Для получения оценок решений начальной задачи (2.1) мы будем использовать следующий функционал Ляпунова–Красовского:

$$\widehat{V}(t, y) = \langle Hy(t), y(t) \rangle + \sum_{i=1}^n \int_0^\infty \int_{t-\theta}^t e^{-m(t-\eta)} \langle (\Upsilon_i(\theta) + Q_i(\theta)) y(\eta), y(\eta) \rangle d\eta d\theta, \quad (4.3)$$

где матрица H определена в (3.3), (3.7), (3.8), (3.10), (3.11), матрицы $\Upsilon_i(\xi)$ — в (3.2), число $m > 0$ — в (3.12), (3.13), числа $\delta_i > 0$ — в (3.14), (3.15),

$$Q_i(\xi) = \frac{u_{\min}^H}{4\omega_i(-\frac{m}{2})} e^{-D\xi} K_i(\xi) e^{m\xi/2} H_i, \quad i = 1, \dots, n. \quad (4.4)$$

Введем неубывающие функции $c_i(\kappa)$, $\kappa > 0$, $i = 0, 1, \dots, n$:

$$c_0(\kappa) = \sup C_0(\kappa), \quad C_0(\kappa) = \{\tau > 0 : |f_0(y_0)| < \kappa |y_0| \quad \forall y_0 \quad |y_0| < \tau\}, \quad (4.5)$$

$$c_i(\kappa) = \sup C_i(\kappa), \quad C_i(\kappa) = \{\tau > 0 : |f_i(y_0)| < \kappa \quad \forall y_0 \quad |y_0| < \tau\}, \quad i = 1, \dots, n, \quad (4.6)$$

где функции $f_i(y_0)$, $i = 0, 1, \dots, n$, определены в (2.4), (2.5). Отметим, что множество $C_0(\kappa)$ не пусто, поскольку в силу формулы Тейлора $f_0(y_0) = o(y_0)$. Множества $C_i(\kappa)$, $i = 1, \dots, n$, также непустые в силу непрерывности в нуле функций $f_i(y_0)$, $i = 1, \dots, n$.

Введем обозначения:

$$v(H) = \|H\| \|H^{-1}\|, \quad (4.7)$$

$$\kappa_0 = \min \left\{ \frac{u_{\min}^H}{8x_1^* \sqrt{(n+1)v(H)}}, \frac{u_{\min}^H}{4\sqrt{2}x_1^* \omega_1(-\frac{m}{2}) \sqrt{v(H)}} \right\}, \quad (4.8)$$

$$\kappa_1 = \min \left\{ \frac{u_{\min}^H}{8\sqrt{(n+1)v(H)}}, \frac{u_{\min}^H}{4\sqrt{2}\omega_1(-\frac{m}{2}) \sqrt{v(H)}} \right\}, \quad (4.9)$$

$$\kappa_i = \min \left\{ \frac{u_{\min}^H}{8\sqrt{(n+1)v(H)}}, \frac{u_{\min}^H}{4\omega_i(-\frac{m}{2}) \sqrt{v(H)}} \right\}, \quad i = 2, \dots, n, \quad (4.10)$$

$$\rho = \min \{c_0(\kappa_0), c_1(\kappa_1), c_2(\kappa_2), \dots, c_n(\kappa_n)\}, \quad (4.11)$$

$$\widehat{V}(0, \psi) = \langle H\psi(0), \psi(0) \rangle + \sum_{i=1}^n \int_0^\infty \int_{-\theta}^0 e^{m\eta} \langle (\Upsilon_i(\theta) + Q_i(\theta)) \psi(\eta), \psi(\eta) \rangle d\eta d\theta. \quad (4.12)$$

Теорема. Пусть выполнены условия (2.8), тогда для решения начальной задачи (2.1) с начальными данными из множества

$$\mathcal{E} = \left\{ \psi(t) \in C((-\infty, 0]) : \sup_{t \leq 0} |\psi_0(t)| < \rho, \quad \sqrt{\|H^{-1}\|} \widehat{V}^{1/2}(0, \psi) < \rho \right\}, \quad (4.13)$$

где ρ и $\widehat{V}(0, \psi)$ определены в (4.11) и (4.12) соответственно, справедлива оценка

$$\|y(t)\| \leq \sqrt{\|H^{-1}\|} \widehat{V}^{1/2}(0, \psi) e^{-\varepsilon_0 t/2}, \quad (4.14)$$

где

$$\varepsilon_0 = \min \left\{ \frac{u_{\min}^H}{4}, m \right\}. \quad (4.15)$$

Доказательство. Разобьем доказательство на несколько частей.

1) Продифференцируем функционал Ляпунова–Красовского (4.3) вдоль решения начальной задачи (2.1), проведем рассуждения, аналогичные рассуждениям в предыдущем разделе, и получим аналог (3.5):

$$\frac{d}{dt} \widehat{V}(t, y) \leq - \left\langle \left(U - \sum_{i=1}^n \int_0^\infty Q_i(\theta) d\theta \right) y(t), y(t) \right\rangle +$$

$$\begin{aligned}
& + 2 \left\langle Hy(t), \left(F(y(t)) + \sum_{i=1}^n \int_{-\infty}^t G_i(y(\theta)) e^{-D(t-\theta)} K_i(t-\theta) d\theta \right) \right\rangle - \\
& - \sum_{i=1}^n \int_{-\infty}^t e^{-m(t-\theta)} \langle Q_i(t-\theta)y(\theta), y(\theta) \rangle d\theta - \\
& - m \sum_{i=1}^n \int_0^\infty \int_{t-\theta}^t e^{-m(t-\eta)} \langle (\Upsilon_i(\theta) + Q_i(\theta))y(\eta), y(\eta) \rangle d\eta d\theta.
\end{aligned}$$

Используя (2.7), (4.2) и (4.4), несложно убедиться, что

$$\sum_{i=1}^n \int_0^\infty Q_i(\theta) d\theta = \frac{u_{\min}^H}{4} \sum_{i=1}^n H_i = \frac{u_{\min}^H}{4} H.$$

Следовательно, в силу обозначения $u_{\min}^H$ и формулы (4.2) имеем

$$\begin{aligned}
\frac{d}{dt} \widehat{V}(t, y) & \leq -\frac{u_{\min}^H}{4} \langle Hy(t), y(t) \rangle + \sum_{i=0}^n I_i(t) - \\
& - m \sum_{i=1}^n \int_0^\infty \int_{t-\theta}^t e^{-m(t-\eta)} \langle (\Upsilon_i(\theta) + Q_i(\theta))y(\eta), y(\eta) \rangle d\eta d\theta,
\end{aligned} \tag{4.16}$$

где

$$I_i(t) = 2 \langle Hy(t), F(y(t)) \rangle - \frac{1}{4} u_{\min}^H \langle Hy(t), y(t) \rangle, \tag{4.17}$$

$$\begin{aligned}
I_i(t) & = 2 \left\langle Hy(t), \int_{-\infty}^t G_i(y(\theta)) e^{-D(t-\theta)} K_i(t-\theta) d\theta \right\rangle - \frac{1}{4} u_{\min}^H \langle Hy(t), y(t) \rangle - \\
& - \int_{-\infty}^t e^{-m(t-\theta)} \langle Q_i(t-\theta)y(\theta), y(\theta) \rangle d\theta, \quad i = 1, \dots, n,
\end{aligned} \tag{4.18}$$

$H_i, i = 1, \dots, n$, определены в (4.1).

2) Покажем, что существует такое число $t^* > 0$, что при всех $t \in [0, t^*)$ выполнены неравенства $I_i(t) \leq 0$, $i = 0, 1, \dots, n$.

2а) Оценим $I_0(t)$ из (4.17):

$$I_0(t) \leq 2\sqrt{\|H\|} \|F(y(t))\| \langle Hy(t), y(t) \rangle^{1/2} - \frac{1}{4} u_{\min}^H \langle Hy(t), y(t) \rangle.$$

В силу определения $F(y(t))$ в (2.2) и неравенства $\sum_{i=0}^n |y_i| \leq \sqrt{n+1} \left(\sum_{i=0}^n y_i^2 \right)^{1/2}$ имеем

$$\begin{aligned}
I_0(t) & \leq \langle Hy(t), y(t) \rangle^{1/2} \left(\left[2\sqrt{\|H\|} x_1^* |f_0(y_0(t))| - \frac{u_{\min}^H}{4\sqrt{(n+1)\|H^{-1}\|}} |y_0(t)| \right] + \right. \\
& \left. + \sum_{i=1}^n |y_i(t)| \left[2\sqrt{\|H\|} |f_i(y_0(t))| - \frac{u_{\min}^H}{4\sqrt{(n+1)\|H^{-1}\|}} \right] \right).
\end{aligned} \tag{4.19}$$

Покажем, что выражения в квадратных скобках отрицательны при $t = 0$. Тогда в силу непрерывности функций $f_i(y_0)$, $i = 0, 1, \dots, n$, и компоненты решения $y_0(t)$ будет существовать такое число $t_0 > 0$, что данные выражения будут отрицательными при всех $t \in [0, t_0)$. В этом случае будет выполнено $I_0(t) \leq 0$, $t \in [0, t_0)$.

Отрицательность выражения в первой квадратной скобке в (4.19) при $t = 0$ эквивалентна следующему неравенству:

$$|f_0(\psi_0(0))| < \frac{u_{\min}^H}{8x_1^* \sqrt{(n+1)v(H)}} |\psi_0(0)|, \quad (4.20)$$

где $v(H)$ определено в (4.7). Из первого неравенства в определении множества $\mathcal{E}$ (см. (4.13)) и из (4.11) вытекает, что $|\psi_0(0)| < c_0(\kappa_0)$. Следовательно, в силу (4.5) имеем

$$|f_0(\psi_0(0))| < \kappa_0 |\psi_0(0)|.$$

Учитывая определение κ_0 в (4.8), получаем выполнение неравенства (4.20).

Отрицательность i -го выражения под суммой в (4.19), $i = 1, \dots, n$, при $t = 0$ эквивалентна неравенству

$$|f_i(\psi_0(0))| < \frac{u_{\min}^H}{8\sqrt{(n+1)v(H)}}.$$

Это неравенство выполняется в силу (4.13), (4.11), (4.9), (4.10) и (4.6).

2б) В данной части доказательства оценим $I_1(t)$ из (4.18). В силу (4.1) и того, что вектор-функция $G_1(y(\theta))$, определенная в (2.3), имеет только одну ненулевую компоненту, верно равенство

$$\langle Hz, G_1(y(\theta)) \rangle = \langle H_1 z, G_1(y(\theta)) \rangle, \quad z \in \mathbb{R}^{n+1}. \quad (4.21)$$

Выпишем $I_1(t)$, учитывая (4.4) и (4.21):

$$\begin{aligned} I_1(t) = & \int_{-\infty}^t 2 \langle H_1 y(t), G_1(y(\theta)) \rangle e^{-D(t-\theta)} K_1(t-\theta) d\theta - \frac{1}{4} u_{\min}^H \langle H_1 y(t), y(t) \rangle - \\ & - \frac{u_{\min}^H}{4\omega_1(-\frac{m}{2})} \int_{-\infty}^t \langle H_1 y(\theta), y(\theta) \rangle e^{-D(t-\theta)} K_1(t-\theta) e^{-m(t-\theta)/2} d\theta. \end{aligned}$$

Так как H_1 является эрмитовой неотрицательно определенной матрицей, то существует $(H_1)^{1/2}$. Тогда, используя неравенство

$$\begin{aligned} 2 \langle H_1 y(t), G_1(y(\theta)) \rangle &= 2 \langle (H_1)^{1/2} y(t), (H_1)^{1/2} G_1(y(\theta)) \rangle \leq \\ &\leq \frac{u_{\min}^H e^{m(t-\theta)/2}}{4\omega_1(-\frac{m}{2})} \|(H_1)^{1/2} y(t)\|^2 + \frac{4\omega_1(-\frac{m}{2}) e^{-m(t-\theta)/2}}{u_{\min}^H} \|(H_1)^{1/2} G_1(y(\theta))\|^2 \end{aligned}$$

и (2.7), получим

$$\begin{aligned} I_1(t) \leq & \int_{-\infty}^t \left(\frac{4\omega_1(-\frac{m}{2})}{u_{\min}^H} \langle H_1 G_1(y(\theta)), G_1(y(\theta)) \rangle - \frac{u_{\min}^H}{4\omega_1(-\frac{m}{2})} \langle H_1 y(\theta), y(\theta) \rangle \right) \times \\ & \times e^{-D(t-\theta)} K_1(t-\theta) e^{-m(t-\theta)/2} d\theta. \end{aligned} \quad (4.22)$$

В силу определений матриц H и H_1 имеем

$$\langle H_1 y(\theta), y(\theta) \rangle \geq \frac{1}{\|H^{-1}\|} (y_0^2(\theta) + y_1^2(\theta)) \geq \frac{1}{2\|H^{-1}\|} (|y_0(\theta)| + |y_1(\theta)|)^2.$$

Используя данное неравенство и формулу (4.21), из (4.22) получим оценку

$$\begin{aligned} I_1(t) \leq & \frac{4\omega_1(-\frac{m}{2}) \|H\|}{u_{\min}^H} \int_{-\infty}^t \left[\|G_1(y(\theta))\|^2 - \left(\frac{u_{\min}^H}{4\sqrt{2}\omega_1(-\frac{m}{2}) \sqrt{v(H)}} \right)^2 (|y_0(\theta)| + |y_1(\theta)|)^2 \right] \times \\ & \times e^{-D(t-\theta)} K_1(t-\theta) e^{-m(t-\theta)/2} d\theta. \end{aligned} \quad (4.23)$$

Если мы покажем, что выражение в квадратных скобках отрицательно при всех $\theta \leq 0$, то в силу непрерывности будет существовать такое число $t_1 > 0$, что данное выражение будет отрицательным при всех $\theta < t_1$,

следовательно, $I_1(t) \leq 0$ при $t \in [0, t_1]$. Отрицательность выражения в квадратных скобках при $\theta \leq 0$ в (4.23) эквивалентна неравенству

$$\|G_1(\psi(\theta))\| < \frac{u_{\min}^H}{4\sqrt{2}\omega_1(-\frac{m}{2})\sqrt{v(H)}}(|\psi_0(\theta)| + |\psi_1(\theta)|).$$

В силу (2.3) справедливость данного неравенства вытекает из оценки

$$x_1^*|f_0(\psi_0(\theta))| + |\psi_1(\theta)||f_1(\psi_0(\theta))| < \frac{u_{\min}^H}{4\sqrt{2}\omega_1(-\frac{m}{2})\sqrt{v(H)}}(|\psi_0(\theta)| + |\psi_1(\theta)|). \quad (4.24)$$

Проведя рассуждения, подобные рассуждениям после (4.20), и используя первое неравенство в (4.13), а также (4.11), (4.8), (4.9), (4.5), (4.6), несложно показать выполнение неравенства (4.24).

2в) Оценивая $I_i(t)$ из (4.18), $i = 2, \dots, n$, так же, как $I_1(t)$ в пункте 2б), мы получим существование положительных t_i , таких, что $I_i(t) \leq 0$ при всех $t \in [0, t_i]$.

3) Положим $t^* = \min\{t_0, t_1, \dots, t_n\} > 0$, тогда

$$\sum_{i=0}^n I_i(t) \leq 0, \quad t \in [0, t^*].$$

Следовательно, при $t \in (0, t^*)$ из (4.16) имеем

$$\begin{aligned} \frac{d}{dt} \widehat{V}(t, y) &\leq -\frac{u_{\min}^H}{4} \langle Hy(t), y(t) \rangle - \\ &- m \sum_{i=1}^n \int_0^t \int_{t-\theta}^{\infty} e^{-m(t-\eta)} \langle (\Upsilon_i(\theta) + Q_i(\theta))y(\eta), y(\eta) \rangle d\eta d\theta. \end{aligned} \quad (4.25)$$

Используя (4.3) и (4.15), получим

$$\frac{d}{dt} \widehat{V}(t, y) + \varepsilon_0 \widehat{V}(t, y) \leq 0.$$

Из данного неравенства вытекает справедливость оценки (4.14) при $t \in (0, t^*)$.

4) Покажем, что данная оценка выполнена при всех $t > 0$. Докажем от противного: пусть $t^* > 0$ — это такое максимальное конечное число, что оценка (4.14) справедлива при $t \in (0, t^*]$. Следовательно, из (4.14) и второго неравенства в определении $\mathcal{E}$ из (4.13) при $t \in (0, t^*]$ имеем

$$|y_0(t)| \leq \|y(t)\| \leq \sqrt{\|H^{-1}\|} \widehat{V}^{1/2}(0, \psi) < \rho.$$

Повторяя рассуждения, аналогичные рассуждениям в пункте 2), получим существование такого числа $t^{**} > t^* > 0$, что $I_i(t) \leq 0$, $i = 0, 1, \dots, n$, при $t \in [t^*, t^* + t^{**}]$, где $I_i(t)$ определены в (4.17) и (4.18). Тогда из (4.16) вытекает (4.25) при $t \in [t^*, t^* + t^{**}]$, следовательно, оценка (4.14) справедлива при $t \in (0, t^* + t^{**})$, откуда получаем противоречие. Это значит, что оценка (4.14) выполнена при всех $t > 0$.

Теорема доказана.

Замечание. Если в формулировке теоремы величину $u_{\min}^H > 0$ заменить на $\sigma > 0$ из (3.17), то результат останется справедливым, поскольку в силу (3.16) выполнено неравенство $u_{\min}^H \geq \sigma$.

Авторы выражают благодарность Г. В. Демиденко и И. И. Матвеевой за внимание к работе.

СПИСОК ЛИТЕРАТУРЫ

1. Сковороцова М.А., Ыскак Т. (Искаков Т.К.) Оценки решений дифференциальных уравнений с распределенным запаздыванием, описывающих конкуренцию нескольких видов микроорганизмов // Сиб. журн. инду. триал. матем. 2022. Т. 25. № 4. С. 193–205.
2. Wolkowicz G.S.K., Xia H., Wu J. Global dynamics of a chemostat competition model with distributed delay // J. Math. Biology. 1999. V. 38. P. 285–316.
3. Wolkowicz G.S.K., Xia H. Global asymptotic behavior of a chemostat model with discrete delays // SIAM J. Appl. Math. 1997. V. 57. № 4. P. 1019–1043.

4. *MacDonald N.* Time delays in chemostat models // in: *Microbial Population Dynamics*, M.J. Bazin, ed., CRC Press, Florida. 1982. P. 33–53.
5. *Демиденко Г.В., Матвеева И.И.* Асимптотические свойства решений дифференциальных уравнений с запаздывающим аргументом // *Вестник НГУ. Сер.: матем., мех., информатика*. 2005. Т. 5. № 3. С. 20–28.
6. *Демиденко Г.В., Матвеева И.И.* Устойчивость решений дифференциальных уравнений с запаздывающим аргументом и периодическими коэффициентами в линейных членах // *Сиб. матем. журн.* 2007. Т. 48. № 5. С. 1025–1040.
7. *Демиденко Г.В., Матвеева И.И.* Об оценках решений систем дифференциальных уравнений нейтрального типа с периодическими коэффициентами // *Сиб. матем. журн.* 2014. Т. 55. № 5. С. 1059–1077.
8. *Матвеева И.И.* Об экспоненциальной устойчивости решений периодических систем нейтрального типа // *Сиб. матем. журн.* 2017. Т. 58. № 2. С. 344–352.
9. *Демиденко Г.В., Матвеева И.И., Скворцова М.А.* Оценки решений дифференциальных уравнений нейтрального типа с периодическими коэффициентами в линейных членах // *Сиб. матем. журн.* 2019. Т. 60. № 5. С. 1063–1079.
10. *Матвеева И.И.* Оценки экспоненциального убывания решений одного класса нелинейных систем нейтрального типа с периодическими коэффициентами // *Ж. вычисл. матем. и матем. физ.* 2020. Т. 60. № 4. С. 612–620.
11. *Ыскак Т. (Искаков Т.К.)* Оценки решений одного класса систем нелинейных дифференциальных уравнений с распределенным запаздыванием // *Сиб. электрон. матем. изв.* 2020. Т. 17. С. 2204–2215.
12. *Ыскак Т. (Искаков Т.К.)* Устойчивость решений систем нелинейных дифференциальных уравнений с бесконечным распределенным запаздыванием // *Челябинский физ.-матем. журн.* 2023. Т. 8. № 4. С. 542–552.
13. *Demidenko G.V., Matveeva I.I.* The second Lyapunov method for time-delay systems // *Functional Differential Equations and Applications* (Editors: Domoshnitsky A., Rasin A., Padhi S.). Ser.: Springer Proceed. in Math. Statist. Singapore: Springer Nature, 2021. V. 379. P. 145–167.
14. *Skvortsova M.A.* Asymptotic properties of solutions to a system describing the spread of avian influenza // *Сиб. электрон. матем. известия*. 2016. Т. 13. С. 782–798.
15. *Скворцова М.А.* Оценки решений в модели взаимодействия популяций с несколькими запаздываниями // *Итоги науки и техники. Серия «Современная математика и ее приложения. Тематические обзоры»*. 2020. Т. 188. С. 84–105.
16. *Скворцова М.А., Ыскак Т. (Искаков Т.К.)* Асимптотическое поведение решений в одной модели «хищник-жертва» с запаздыванием // *Сиб. матем. журн.* 2021. Т. 62. № 2. С. 402–416.
17. *Скворцова М.А.* Оценки решений для одной биологической модели // *Матем. труды*. 2022. Т. 25. № 1. С. 152–176.

ESTIMATES FOR SOLUTIONS OF A BIOLOGICAL MODEL WITH INFINITE DISTRIBUTED DELAY

T. K. Iskakov^{a,b,*}, M. A. Skvortsova^{a,b,**}

^aNovosibirsk State University, Pirogov St., 1, Novosibirsk, 630090 Russia

^bSobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Acad. Koptug Ave., 4,
Novosibirsk, 630090 Russia

*e-mail: istima92@mail.ru

**e-mail: sm-18-nsu@yandex.ru

Received 02 April, 2024

Revised 02 April, 2024

Accepted 02 May, 2024

Abstract. A model of competition between several microorganism species described by a system of nonlinear differential equations with infinite distributed delay is considered. The case of asymptotic stability of the equilibrium point corresponding to the survival of only one species and the extinction of all others is studied. The conditions for the initial numbers of species and the initial concentration of the nutrient at which the system reaches an equilibrium state are specified, and estimates of the stabilization rate are established. The results are obtained using the Lyapunov–Krasovskii functional.

Keywords: model of species competition, chemostat, delay differential equations, infinite distributed delay, equilibrium point, asymptotic stability, estimates for solutions, attraction domain, Lyapunov–Krasovskii functional.

ОЦЕНКИ СХОДИМОСТИ ИТЕРАЦИОННЫХ МЕТОДОВ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ ТРЕХМЕРНЫХ ПРОЦЕССОВ В МАГНИТНОЙ ГИДРОДИНАМИКЕ

© 2024 г. А.Ю. Круковский^{1,*}, И.В. Попов^{1,**}, Ю.А. Повещенко^{1,***}

¹ 125047 Москва, Миусская пл., 4, Институт прикладной математики им. М.В. Келдыша РАН, Россия

*e-mail: alexander-krukovskiy@yandex.ru

**e-mail: piv2964@mail.ru

***e-mail: hecon@mail.ru

Поступила в редакцию 15.04.2024 г.

Переработанный вариант 15.04.2024 г.

Принята к публикации 05.05.2024 г.

В настоящей работе рассматривается вопрос сходимости итерационных процессов, применяемых к неявным полностью консервативным разностным схемам трехмерной магнитной гидродинамики методов раздельного и комбинированного решения групп разностных уравнений, которые разбиваются по физическим процессам. Получены оценки сходимости итерационных процессов для рассматриваемых в данной работе численных методов. Исследован вопрос области применимости как метода комбинированных, так и раздельных способов решения разностных трехмерных уравнений магнитной гидродинамики. Учитывая, что проведенные исследования представленных алгоритмов носят в основном качественный характер, справедливость полученных оценок была подтверждена численными экспериментами, причем рассматривались как модельные, так и реальные задачи. Отметим, что полученные оценки сходимости итерационных процессов дают возможность при решении трехмерных задач магнитной гидродинамики на любом шаге по времени выбрать оптимальный численный метод решения разностных уравнений. Библ. 12. Фиг. 7.

Ключевые слова: уравнения трехмерной магнитной гидродинамики, неявная полностью консервативная разностная схема, сходимость итерационного процесса.

DOI: 10.31857/S0044466924080076, EDN: YARBOF

1. ВВЕДЕНИЕ

При моделировании задач на первом этапе определяется ее специфика и, исходя из этого, выбирается метод или методы ее оптимального численного решения. Созданный на основе такого численного метода (методов) комплекс программ (КП) может решать лишь узкий класс задач. Задачи, которые в данный класс не входят, либо решаются не эффективно, либо не решаются вообще. Создание достаточно универсального алгоритма для решения широкого класса задач математической физики, как правило, не приводит к положительным результатам из-за чрезвычайной громоздкости подобного алгоритма.

В работе уделяется большое внимание исследованию сходимости метода, примененного к неявной полностью консервативной разностной схеме (ПКРС) трехмерной магнитной гидродинамики (МГД), раздельного и комбинированного решения групп трехмерных разностных уравнений. Эти уравнения разбиты по физическим процессам [1–3], а именно по группам: динамическая, тепловая и электромагнитная.

Метод раздельного решения разностных уравнений заключается в том, что эти три группы решаются раздельно.

Комбинированный метод решения разностных уравнений заключается в том, что три группы объединяются следующим образом: в первую группу включаются уравнения динамики и электромагнитного поля, а во вторую группу — уравнения энергии.

Уравнение состояния является уравнением — связкой между тремя (двумя) группами. Уравнения трех (обеих) групп решаются некоторым итерационным методом, например, методом Ньютона, относительно определенных переменных, а не входящие в эту группу переменные считаются “замороженными” и берутся с предыдущей итерации. Последовательное решение групп уравнений связывается внешним итерационным процессом. Сходимость этого процесса контролируется по точности выполнения общего баланса энергии и точности распределения полной энергии по ее отдельным видам.

Полученные в настоящей работе теоретические оценки сходимости итерационных процессов для системы разностных уравнений трехмерных МГД подтверждается численными расчетами как модельных, так и реальных физических задач.

2. СИСТЕМА ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ МАГНИТНОЙ ГИДРОДИНАМИКИ

В декартовой системе координат система уравнений МГД имеет следующий вид (см. [4]):

$$\frac{dx}{dt} = u; \quad \frac{dy}{dt} = w; \quad \frac{dz}{dt} = v; \quad (1)$$

$$\frac{d\rho}{dt} + \rho \frac{\partial u}{\partial x} + \rho \frac{\partial w}{\partial y} + \rho \frac{\partial v}{\partial z} = 0; \quad (2)$$

$$\rho \frac{du}{dt} = -\frac{\partial}{\partial x} \left(P + \frac{B_x^2}{8\pi} + \frac{B_y^2}{8\pi} + \frac{B_z^2}{8\pi} \right) + \frac{1}{4\pi} \frac{\partial(B_x^2)}{\partial x} + \frac{1}{4\pi} \frac{\partial(B_x B_y)}{\partial y} + \frac{1}{4\pi} \frac{\partial(B_x B_z)}{\partial z}; \quad (3)$$

$$\rho \frac{dw}{dt} = -\frac{\partial}{\partial y} \left(P + \frac{B_x^2}{8\pi} + \frac{B_y^2}{8\pi} + \frac{B_z^2}{8\pi} \right) + \frac{1}{4\pi} \frac{\partial(B_x B_y)}{\partial x} + \frac{1}{4\pi} \frac{\partial(B_y^2)}{\partial y} + \frac{1}{4\pi} \frac{\partial(B_y B_z)}{\partial z}; \quad (4)$$

$$\rho \frac{dv}{dt} = -\frac{\partial}{\partial z} \left(P + \frac{B_x^2}{8\pi} + \frac{B_y^2}{8\pi} + \frac{B_z^2}{8\pi} \right) + \frac{1}{4\pi} \frac{\partial(B_x B_z)}{\partial x} + \frac{1}{4\pi} \frac{\partial(B_y B_z)}{\partial y} + \frac{1}{4\pi} \frac{\partial(B_z^2)}{\partial z}; \quad (5)$$

$$P \frac{d\varepsilon}{dt} = -\rho P \frac{\partial(1/\rho)}{\partial t} - \frac{\partial W_x}{\partial x} - \frac{\partial W_y}{\partial y} - \frac{\partial W_z}{\partial z} + G_{\text{Joul}}; \quad (6)$$

$$G_{\text{Joul}} = \frac{1}{4\pi} (E_x J_x + E_y J_y + E_z J_z); \quad (7)$$

$$J_x = \frac{\partial B_z}{\partial y} - \frac{\partial B_y}{\partial z}; \quad J_y = \frac{\partial B_x}{\partial z} - \frac{\partial B_z}{\partial x}; \quad J_z = \frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}; \quad (8)$$

$$W_x = -\kappa_{xx} \frac{\partial T}{\partial x} - \kappa_{xy} \frac{\partial T}{\partial y} - \kappa_{xz} \frac{\partial T}{\partial z}; \quad (9)$$

$$W_y = -\kappa_{yx} \frac{\partial T}{\partial x} - \kappa_{yy} \frac{\partial T}{\partial y} - \kappa_{yz} \frac{\partial T}{\partial z}; \quad (10)$$

$$W_z = -\kappa_{zx} \frac{\partial T}{\partial x} - \kappa_{zy} \frac{\partial T}{\partial y} - \kappa_{zz} \frac{\partial T}{\partial z}; \quad (11)$$

$$\kappa_{\alpha\alpha} = \kappa_{\parallel} \left(\frac{B_{\alpha}}{B} \right)^2 - \kappa_{\perp} \left(1 - \left(\frac{B_{\alpha}}{B} \right)^2 \right), \quad \alpha = x, y, z; \quad (12)$$

$$\kappa_{\alpha\beta} = \kappa_{\beta\alpha} = (\kappa_{\parallel} - \kappa_{\perp}) \frac{B_{\alpha} B_{\beta}}{B^2}, \quad \alpha = x, y, z, \quad \beta = x, y, z, \quad \alpha \neq \beta; \quad (13)$$

$$\rho \frac{d}{dt} \left(\frac{B_x}{\rho} \right) = \frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y} + \left(B_x \frac{\partial u}{\partial x} + B_y \frac{\partial u}{\partial y} + B_z \frac{\partial u}{\partial z} \right); \quad (14)$$

$$\rho \frac{d}{dt} \left(\frac{B_y}{\rho} \right) = \frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z} + \left(B_x \frac{\partial w}{\partial x} + B_y \frac{\partial w}{\partial y} + B_z \frac{\partial w}{\partial z} \right); \quad (15)$$

$$\rho \frac{d}{dt} \left(\frac{B_z}{\rho} \right) = \frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} + \left(B_x \frac{\partial v}{\partial x} + B_y \frac{\partial v}{\partial y} + B_z \frac{\partial v}{\partial z} \right); \quad (16)$$

$$E_{\alpha} = \frac{1}{4\pi\sigma_{\perp}} J_{\alpha} + \frac{\sigma_{\perp} - \sigma_{\parallel}}{4\pi\sigma_{\perp}\sigma_{\parallel}} \frac{B_{\alpha}}{B^2} (\mathbf{J}, \mathbf{B}), \quad \alpha = x, y, z; \quad (17)$$

$$(\mathbf{J}, \mathbf{B}) = (J_x B_x + J_y B_y + J_z B_z). \quad (18)$$

Здесь t — время, $\mathbf{U} = (u, w, v)$ — вектор скорости вещества, ρ — плотность, T — температура, P — гидродинамическое давление, ε — удельная внутренняя энергия, отнесенная к единице массы, $\mathbf{W} = (W_x, W_y, W_z)$ — вектор потока тепла, $\mathbf{B} = (B_x, B_y, B_z)$ — вектор магнитного поля, $\mathbf{E} = (E_x, E_y, E_z)$ — вектор электрического поля, $\mathbf{J} = (J_x, J_y, J_z)$ — вектор плотности тока, G_{Joul} — массовая плотность энергии джоулева нагрева, κ — коэффициенты теплопроводности (тензор), σ — коэффициенты электропроводности (тензор).

3. СИСТЕМА РАЗНОСТНЫХ УРАВНЕНИЙ ТРЕХМЕРНОЙ МАГНИТНОЙ ГИДРОДИНАМИКИ

Анализ сходимости итерационного процесса комбинированного решения системы разностных уравнений МГД проведен в одномерном плоском случае в [5] и в двумерном случае в декартовых координатах в [3]. Предполагалось, что $\kappa = 0$ и $\sigma = \infty$. Было показано, что для неявных схем процесс безусловно устойчивый лишь при выполнении следующего условия:

$$\frac{B^2}{8\pi} \geq (\gamma - 2)P. \quad (19)$$

При $\frac{B^2}{8\pi} < (\gamma - 2)P$ условие устойчивости в одномерном случае имеет вид

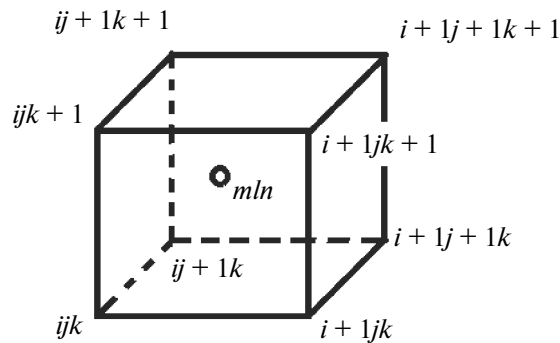
$$\Delta t < \frac{\Delta m}{2\rho \sqrt{(\gamma - 2)\frac{P}{\rho} - \frac{B^2}{8\pi\rho}}}. \quad (20)$$

В двумерном случае условие устойчивости записывается в виде

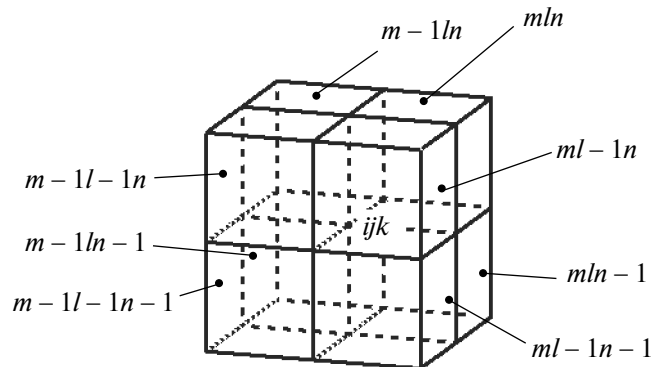
$$\Delta t < \frac{\Delta m}{2\rho(h_x + h_y) \sqrt{(\gamma - 2)\frac{P}{\rho} - \frac{B^2}{8\pi\rho}}}. \quad (21)$$

При учете диссипативных процессов [6] ограничения на временной шаг становятся менее обременительными. В дальнейшем при исследовании устойчивости процесса решения системы разностных уравнений методами раздельного и комбинированного решения предполагаем отсутствие диссипативных процессов. Ограничимся также случаем течения в магнитном поле, ориентированным перпендикулярно плоскости течения ($B_x = B_y = 0$).

Введем пространства сеточных функций определенных в узлах и ячейках разностной сетки (см. фиг. 1 и фиг. 2). Обозначив их H_Ω и H_w , соответственно. Для записи сеточных функций $f \in H_\Omega$ используем индексы (i, j, k) : $f_{ijk} = f \in H_\Omega$, а сеточные функции $g \in H_w$ будем отмечать индексами (m, l, n) : $g_{mln} = g \in H_w$. Полагаем, что массы всех ячеек $\Delta m_{mln} = \Delta m = \text{const}$.



Фиг. 1. Π_1 — шаблон ячейки.



Фиг. 2. Π_2 — шаблон узла.

В описанных выше предположениях система (1)–(18) аппроксимируется в неявных ПКРС в лагранжевых переменных следующим образом:

$$\Delta mu_t = \sum_{r \in \mathcal{M}_2} \left(\frac{\partial V_r}{\partial x} \right) \left(R \hat{\rho}_r \hat{T}_r + \frac{\hat{B}_r B_r}{8\pi} \right); \quad (22)$$

$$\Delta mw_t = \sum_{r \in \mathcal{M}_2} \left(\frac{\partial V_r}{\partial y} \right) \left(R \hat{\rho}_r \hat{T}_r + \frac{\hat{B}_r B_r}{8\pi} \right); \quad (23)$$

$$\Delta mv_t = \sum_{r \in \mathcal{M}_2} \left(\frac{\partial V_r}{\partial z} \right) \left(R \hat{\rho}_r \hat{T}_r + \frac{\hat{B}_r B_r}{8\pi} \right); \quad (24)$$

$$\hat{B} \hat{V} = BV; \quad (25)$$

$$\Delta m \left(\frac{1}{\rho} \right)_t = V_t = \sum_{q \in \mathcal{M}_1} \left(\left(\frac{\partial V}{\partial x_q} \right) u_q^{(0,5)} + \left(\frac{\partial V}{\partial y_q} \right) w_q^{(0,5)} + \left(\frac{\partial V}{\partial z_q} \right) v_q^{(0,5)} \right); \quad (26)$$

$$\frac{1}{\gamma - 1} (\hat{T} - T) = - \frac{\hat{\rho} \hat{T}}{\Delta m} (\hat{V} - V), \quad (27)$$

где объем ячейки

$$V_{\min} = \frac{1}{6} \left[(x_{10} - x_{13})(S_1^x + S_3^x) + (y_{10} - y_{13})(S_1^y + S_3^y) + (z_{10} - z_{13})(S_1^z + S_3^z) + \right. \\ \left. + (x_{12} - x_{11})(S_2^x + S_4^x) + (y_{12} - y_{11})(S_2^y + S_4^y) + (z_{12} - z_{11})(S_2^z + S_4^z) + \right. \\ \left. + (x_{14} - x_9)(S_5^x + S_6^x) + (y_{14} - y_9)(S_5^y + S_6^y) + (z_{14} - z_9)(S_5^z + S_6^z) \right].$$

Здесь

$$r_9 = \frac{1}{4}(r_1 + r_2 + r_3 + r_4); \quad r_{10} = \frac{1}{4}(r_1 + r_2 + r_5 + r_6); \quad r_{11} = \frac{1}{4}(r_1 + r_4 + r_5 + r_8); \\ r_{12} = \frac{1}{4}(r_2 + r_3 + r_6 + r_7); \quad r_{13} = \frac{1}{4}(r_3 + r_4 + r_7 + r_8); \quad r_{14} = \frac{1}{4}(r_5 + r_6 + r_7 + r_8); \\ r_1 = r_{i+1jk}; \quad r_2 = r_{i+1j+1k}; \quad r_3 = r_{ij+1k}; \quad r_4 = r_{ijk}; \quad r_5 = r_{i+1jk+1}; \quad r_6 = r_{i+1j+1k+1}; \\ r_7 = r_{ij+1k+1}; \quad r_8 = r_{ijk+1}; \quad r = x, y, z; \\ S_1^r = \frac{1}{2}[(\psi_2 - \psi_5)(\theta_6 - \theta_1) + (\psi_6 - \psi_1)(\theta_5 - \theta_2)]; \\ S_2^r = \frac{1}{2}[(\psi_7 - \psi_2)(\theta_6 - \theta_3) + (\psi_6 - \psi_3)(\theta_2 - \theta_7)]; \\ S_3^r = \frac{1}{2}[(\psi_3 - \psi_8)(\theta_7 - \theta_4) + (\psi_7 - \psi_4)(\theta_8 - \theta_3)]; \\ S_4^r = \frac{1}{2}[(\psi_8 - \psi_1)(\theta_5 - \theta_4) + (\psi_5 - \psi_4)(\theta_1 - \theta_8)]; \\ S_5^r = \frac{1}{2}[(\psi_1 - \psi_3)(\theta_2 - \theta_4) + (\psi_2 - \psi_4)(\theta_3 - \theta_1)]; \\ S_6^r = \frac{1}{2}[(\psi_5 - \psi_7)(\theta_6 - \theta_8) + (\psi_6 - \psi_8)(\theta_7 - \theta_5)];$$

где $r = x, y, z$.

Если $r = x$, то $\psi = y, \theta = z$.

Если $r = y$, то $\psi = z, \theta = x$.

Если $r = z$, то $\psi = x, \theta = y$.

$\Delta m_{mln} = \hat{\rho}_{mln} \hat{V}_{mln} = \rho_{mln} V_{mln}$.

Сеточные функции x, y, z, u, w и v определены в узлах разностной сетки, а остальные сеточные функции определены в ячейках.

При записи (22)–(27) использовалось безындексное представление сеточных функций [1]. В (22)–(24) и (27) учитывается, что $P = R \rho T$, $\varepsilon = \frac{RT}{\gamma - 1}$.

Разностные уравнения (22)–(27) записаны на разностных шаблонах $\mathcal{M}_1(m, l, n)$ и $\mathcal{M}_2(i, j, k)$.

Производные объемов по координатам, входящие в (22)–(24), представим в следующем виде:

$$\begin{aligned}\frac{\partial V_{mln}}{\partial r_1} &= \frac{1}{12}(S_{18}^r - S_{13}^r - S_{16}^r); & \frac{\partial V_{mln}}{\partial r_2} &= \frac{1}{12}(S_{27}^r + S_{24}^r - S_{25}^r); & \frac{\partial V_{mln}}{\partial r_3} &= \frac{1}{12}(-S_{36}^r + S_{13}^r - S_{38}^r); \\ \frac{\partial V_{mln}}{\partial r_4} &= \frac{1}{12}(-S_{45}^r - S_{24}^r - S_{47}^r); & \frac{\partial V_{mln}}{\partial r_5} &= \frac{1}{12}(S_{45}^r - S_{57}^r + S_{25}^r); & \frac{\partial V_{mln}}{\partial r_6} &= \frac{1}{12}(S_{36}^r + S_{68}^r + S_{16}^r); \\ \frac{\partial V_{mln}}{\partial r_7} &= \frac{1}{12}(-S_{27}^r + S_{57}^r + S_{47}^r); & \frac{\partial V_{mln}}{\partial r_8} &= \frac{1}{12}(-S_{18}^r - S_{68}^r + S_{38}^r).\end{aligned}\quad (28)$$

Здесь

$$\begin{aligned}S_{18}^r &= \frac{1}{2}(S_1^r + S_3^r) + (\psi_{12} - \psi_{11})(\theta_5 - \theta_4) + (\theta_{12} - \theta_{11})(\psi_4 - \psi_5); \\ S_{13}^r &= \frac{1}{2}(S_2^r + S_4^r) + (\psi_{14} - \psi_9)(\theta_2 - \theta_4) + (\theta_{14} - \theta_9)(\psi_4 - \psi_2); \\ S_{16}^r &= \frac{1}{2}(S_5^r + S_6^r) + (\psi_{10} - \psi_{13})(\theta_2 - \theta_5) + (\theta_{10} - \theta_{13})(\psi_5 - \psi_2); \\ S_{27}^r &= \frac{1}{2}(S_1^r + S_3^r) + (\psi_{12} - \psi_{11})(\theta_6 - \theta_3) + (\theta_{12} - \theta_{11})(\psi_3 - \psi_6); \\ S_{24}^r &= \frac{1}{2}(S_2^r + S_4^r) + (\psi_{14} - \psi_9)(\theta_1 - \theta_3) + (\theta_{14} - \theta_9)(\psi_3 - \psi_1); \\ S_{25}^r &= \frac{1}{2}(S_5^r + S_6^r) + (\psi_{10} - \psi_{13})(\theta_6 - \theta_1) + (\theta_{10} - \theta_{13})(\psi_1 - \psi_6); \\ S_{36}^r &= \frac{1}{2}(S_1^r + S_3^r) + (\psi_{12} - \psi_{11})(\theta_7 - \theta_2) + (\theta_{12} - \theta_{11})(\psi_2 - \psi_7); \\ S_{38}^r &= \frac{1}{2}(S_5^r + S_6^r) + (\psi_{10} - \psi_{13})(\theta_7 - \theta_4) + (\theta_{10} - \theta_{13})(\psi_4 - \psi_7); \\ S_{45}^r &= \frac{1}{2}(S_1^r + S_3^r) + (\psi_{12} - \psi_{11})(\theta_8 - \theta_1) + (\theta_{12} - \theta_{11})(\psi_1 - \psi_8); \\ S_{47}^r &= \frac{1}{2}(S_5^r + S_6^r) + (\psi_{10} - \psi_{13})(\theta_3 - \theta_8) + (\theta_{10} - \theta_{13})(\psi_8 - \psi_3); \\ S_{57}^r &= \frac{1}{2}(S_2^r + S_4^r) + (\psi_{14} - \psi_9)(\theta_6 - \theta_8) + (\theta_{14} - \theta_9)(\psi_8 - \psi_6); \\ S_{68}^r &= \frac{1}{2}(S_2^r + S_4^r) + (\psi_{14} - \psi_9)(\theta_5 - \theta_7) + (\theta_{14} - \theta_9)(\psi_7 - \psi_5).\end{aligned}$$

4. ОЦЕНКИ СХОДИМОСТИ ЧИСЛЕННЫХ АЛГОРИТМОВ

Для расчетов течений излучающей плазмы исходные модели типа (1)–(18) существенно усложняются – поскольку могут добавляться уравнения кинетики ионизации, радиационного энергообмена, учитываться двухтемпературность, использоваться табличные данные по состоянию вещества, его транспортным коэффициентам и оптическим свойствам и т.д.

Полный анализ сходимости методов решения нелинейных трехмерных уравнений МГД провести зачастую или не удастся, или количество упрощающих предположений делает результаты анализ не пригодными для практического исследования.

В настоящей работе сходимость алгоритмов исследуется для линеаризированной системы (22)–(27). Данный подход к изучению методов решения нелинейной трехмерной системы МГД, несмотря на теоретическую неполноту, позволяет сформулировать физически содержательные критерии сходимости итерационного процесса. Заметим, что этими критериями удобно пользоваться при решении практических задач. Полученные при анализе линеаризированной системы уравнений оценки можно также использовать для изучения сходимости численного метода в нелинейном случае, поскольку, линеаризация является центральным пунктом многочисленных методик [1, 7, 8]. Выводы из линейного анализа в работе проверены в методических расчетах и при расчете реальной задачи.

При построении численных методов решения системы нелинейных уравнений используется принцип локального расщепления физических процессов. Реализующийся осуществляется через разделение системы разностных уравнений по видам физических процессов: уравнения динамики плазмы (D), уравнения электромагнитного поля (ЕМ) и уравнения теплового баланса (Т), описывающие энергообмен между компонентами плазмы в двухтемпературном случае. В простейшем однотемпературном случае уравнение теплового баланса

(Т) одно — (27). Для построения конкретного алгоритма уравнения различных видов объединяются в группы. Далее рассматривается как алгоритм, в котором разбиение на группы происходит следующим образом:

- к первой группе относят динамические (D) и электромагнитные (ЕМ) уравнения;
- ко второй — тепловые (Т) уравнения (метод комбинированного решения системы разностных уравнений) и алгоритм, в котором все три группы решаются отдельно (метод раздельного решения системы разностных уравнений).

Уравнения в каждой из этих двух (трех) групп решаются относительно определенных переменных, переменные же, не входящие в данную группу, считаются “замороженными” и берутся с предыдущей итерации. Последовательное решение групп разностных уравнений связано внешним итерационным процессом, в котором контролируется как точность решения разностных уравнений, так и точность выполнения энергетического баланса.

Линеаризованные уравнения системы (22)–(27), соответствующие принятым выше ограничениям и характеру итерационного процесса, который соответствует методу комбинированного решения, запишем в виде

$$\Delta m \frac{u_{ijk}^{s+1} - u_{ijk}}{\Delta t} = \sum_{r \in \mathbb{M}_2} \left(\frac{\partial V_r}{\partial x_{ijk}} \right) \left(R(\rho_r T_r + T_r \delta \rho_r^{s+1} + \rho_r \delta T_r^s) + \frac{B_r^{s+1} B_r}{8\pi} \right); \quad (29)$$

$$\Delta m \frac{w_{ijk}^{s+1} - w_{ijk}}{\Delta t} = \sum_{r \in \mathbb{M}_2} \left(\frac{\partial V_r}{\partial y_{ijk}} \right) \left(R(\rho_r T_r + T_r \delta \rho_r^{s+1} + \rho_r \delta T_r^s) + \frac{B_r^{s+1} B_r}{8\pi} \right); \quad (30)$$

$$\Delta m \frac{v_{ijk}^{s+1} - v_{ijk}}{\Delta t} = \sum_{r \in \mathbb{M}_2} \left(\frac{\partial V_r}{\partial z_{ijk}} \right) \left(R(\rho_r T_r + T_r \delta \rho_r^{s+1} + \rho_r \delta T_r^s) + \frac{B_r^{s+1} B_r}{8\pi} \right); \quad (31)$$

$$\delta \rho_{mln}^{s+1} = -\frac{\rho_{mln}^2}{\Delta m} (V_{mln}^{s+1} - V_{mln}); \quad (32)$$

$$B_{mln}^{s+1} = B_{mln} \left[1 - \frac{\rho_{mln}}{\Delta m} (V_{mln}^{s+1} - V_{mln}) \right]; \quad (33)$$

$$\delta T_{mln}^s = -\frac{\gamma - 1}{\Delta m} \rho_{mln} T_{mln} (V_{mln}^s - V_{mln}); \quad (34)$$

где

$$V_{mln}^{s+1} - V_{mln} = \frac{\Delta t}{12} \sum_{q \in \mathbb{M}_1} \left[\left(\frac{\partial V_{mln}}{\partial x_q} \right) (u_q + u_q^{s+1}) + \left(\frac{\partial V_{mln}}{\partial y_q} \right) (w_q + w_q^{s+1}) + \left(\frac{\partial V_{mln}}{\partial z_q} \right) (v_q + v_q^{s+1}) \right].$$

Здесь искомые величины обозначаются индексом s — соответственно номеру итерации. Исключив из (29)–(31) $\delta \rho_{mln}^{s+1}$, B_{mln}^{s+1} и δT_{mln}^s с помощью (32)–(34), получим систему трехмерных уравнений, которая связывает U_{ijk}^{s+1} и U_{ijk}^s :

$$(\Lambda_1 U^{s+1})_{ijk} = -(\Lambda_2 U^s)_{ijk} + \mathbf{F}_{ijk}, \quad (35)$$

где $\mathbf{F}_{ijk} = (F_{ijk}^{(1)}, F_{ijk}^{(2)}, F_{ijk}^{(3)}), F_{ijk}^{(1)}, F_{ijk}^{(2)}$ и $F_{ijk}^{(3)}$ — свободные члены в преобразованных уравнениях (29), (30) и (31) соответственно, в которые входят величины с предыдущего слоя по времени:

$$(\Lambda_1 U)_{ijk} = (\Lambda_1(u)_{ijk}, \Lambda_1(w)_{ijk}, \Lambda_1(v)_{ijk}), \quad (\Lambda_2 U)_{ijk} = (\Lambda_2(u)_{ijk}, \Lambda_2(w)_{ijk}, \Lambda_2(v)_{ijk}), \quad (36)$$

$$\Lambda_1(u)_{ijk} = \kappa u_{ijk} + \sum_{r \in \mathbb{M}_2} c1_r \left(\frac{\partial V_r}{\partial x_{ijk}} \right) H_r, \quad \Lambda_1(w)_{ijk} = \kappa w_{ijk} + \sum_{r \in \mathbb{M}_2} c1_r \left(\frac{\partial V_r}{\partial y_{ijk}} \right) H_r,$$

$$\Lambda_1(v)_{ijk} = \kappa v_{ijk} + \sum_{r \in \mathbb{M}_2} c1_r \left(\frac{\partial V_r}{\partial z_{ijk}} \right) H_r, \quad \Lambda_2(u)_{ijk} = \kappa u_{ijk} + \sum_{r \in \mathbb{M}_2} c2_r \left(\frac{\partial V_r}{\partial x_{ijk}} \right) H_r,$$

$$\Lambda_2(w)_{ijk} = \kappa w_{ijk} + \sum_{r \in \mathbb{M}_2} c2_r \left(\frac{\partial V_r}{\partial y_{ijk}} \right) H_r, \quad \Lambda_2(v)_{ijk} = \kappa v_{ijk} + \sum_{r \in \mathbb{M}_2} c2_r \left(\frac{\partial V_r}{\partial z_{ijk}} \right) H_r, \quad \text{где } \kappa = 2 \left(\frac{\Delta m}{\Delta t} \right)^2,$$

$$c1_{mln} = RT_{mln} \rho_{mln}^2 + \frac{B_{mln}^2 \rho_{mln}}{8\pi} = P_{mln} \rho_{mln} + \frac{B_{mln}^2 \rho_{mln}}{8\pi}, \quad c2_{mln} = (\gamma - 1) RT_{mln} \rho_{mln}^2 = (\gamma - 1) P_{mln} \rho_{mln}, \quad (37)$$

$$H_{mln} = \sum_{q \in \mathbb{M}_1} \left(\frac{\partial V_{mln}}{\partial x_q} u_q^{s+1} + \frac{\partial V_{mln}}{\partial y_q} w_q^{s+1} + \frac{\partial V_{mln}}{\partial z_q} v_q^{s+1} \right).$$

Рассмотрим итерационный процесс (35) на множестве сеточных функций, принимающих нулевые значения на границе:

$$U \in \Omega_h^\circ, \quad U|_{\Gamma_h} = 0.$$

Покажем, что разностный оператор Λ_1 обладает следующим свойством: $\Lambda_1 = \Lambda_1^* > 0$.

Вычислим скалярное произведение $(\Lambda_1 U, U)$. Учитывая форму записи производной объема по координатам имеем

$$\begin{aligned} (\Lambda_1 U, U) &= \sum_{ijk} (\Lambda_1(u)_{ijk} u_{ijk} \Delta m + \Lambda_1(w)_{ijk} w_{ijk} \Delta m + \Lambda_1(v)_{ijk} v_{ijk} \Delta m) = \\ &= \Delta m \kappa \sum_{ijk} (u_{ijk}^2 + w_{ijk}^2 + v_{ijk}^2) + \Delta m \sum_{ijk} \left[\sum_{r \in \mathcal{H}_2} c_{1r} H_r \left(\frac{\partial V_r}{\partial x_{ijk}} u_{ijk} + \frac{\partial V_r}{\partial y_{ijk}} w_{ijk} + \frac{\partial V_r}{\partial z_{ijk}} v_{ijk} \right) \right] = \\ &= \Delta m \kappa \sum_{ijk} (u_{ijk}^2 + w_{ijk}^2 + v_{ijk}^2) + \Delta m \sum_{ijk} \left[\sum_{r \in \mathcal{H}_2} c_{1r} \left(\frac{\partial V_r}{\partial x_{ijk}} u_{ijk} + \frac{\partial V_r}{\partial y_{ijk}} w_{ijk} + \frac{\partial V_r}{\partial z_{ijk}} v_{ijk} \right) \times \right. \\ &\quad \left. \times \sum_{q \in \mathcal{H}_1} \left(\frac{\partial V_r}{\partial x_q} u_q + \frac{\partial V_r}{\partial y_q} w_q + \frac{\partial V_r}{\partial z_q} v_q \right) \right] = \\ &= \Delta m \kappa \sum_{ijk} (u_{ijk}^2 + w_{ijk}^2 + v_{ijk}^2) + \frac{\Delta m}{144} \sum_{ijk} c_{1mln} I_{ijk}^2 = (U, \Lambda_1 U) > 0. \end{aligned}$$

Таким образом, $\Lambda_1 = \Lambda_1^* > 0$.

Суммирование идет по всем узлам, кроме граничных; $m = i, l = j, n = k$. Здесь

$$\begin{aligned} I &= \left[(u_1 - u_8) S_{18}^x + (u_3 - u_1) S_{13}^x + (u_6 - u_1) S_{16}^x + (u_2 - u_7) S_{27}^x + (u_2 - u_4) S_{24}^x + (u_5 - u_2) S_{25}^x + \right. \\ &\quad + (u_6 - u_3) S_{36}^x + (u_8 - u_3) S_{38}^x + (u_5 - u_4) S_{45}^x + (u_7 - u_4) S_{47}^x + (u_7 - u_5) S_{57}^x + (u_6 - u_8) S_{68}^x + \\ &\quad + (w_1 - w_8) S_{18}^x + (w_3 - w_1) S_{13}^x + (w_6 - w_1) S_{16}^x + (w_2 - w_7) S_{27}^x + (w_2 - w_4) S_{24}^x + (w_5 - w_2) S_{25}^x + \\ &\quad + (w_6 - w_3) S_{36}^x + (w_8 - w_3) S_{38}^x + (w_5 - w_4) S_{45}^x + (w_7 - w_4) S_{47}^x + (w_7 - w_5) S_{57}^x + (w_6 - w_8) S_{68}^x + \\ &\quad + (v_1 - v_8) S_{18}^x + (v_3 - v_1) S_{13}^x + (v_6 - v_1) S_{16}^x + (v_2 - v_7) S_{27}^x + (v_2 - v_4) S_{24}^x + (v_5 - v_2) S_{25}^x + \\ &\quad \left. + (v_6 - v_3) S_{36}^x + (v_8 - v_3) S_{38}^x + (v_5 - v_4) S_{45}^x + (v_7 - v_4) S_{47}^x + (v_7 - v_5) S_{57}^x + (v_6 - v_8) S_{68}^x \right]. \end{aligned}$$

Покажем, что разностный оператор Λ_2 обладает следующим свойством: $\Lambda_2 = \Lambda_2^* > 0$.

Вычислим скалярное произведение $(\Lambda_2 U, U)$. Учитывая (28), получим

$$\begin{aligned} (\Lambda_2 U, U) &= \sum_{ijk} (\Lambda_2(u)_{ijk} u_{ijk} \Delta m + \Lambda_2(w)_{ijk} w_{ijk} \Delta m + \Lambda_2(v)_{ijk} v_{ijk} \Delta m) = \\ &= \Delta m \sum_{ijk} \left[\sum_{r \in \mathcal{H}_2} c_{2r} H_r \left(\frac{\partial V_r}{\partial x_{ijk}} u_{ijk} + \frac{\partial V_r}{\partial y_{ijk}} w_{ijk} + \frac{\partial V_r}{\partial z_{ijk}} v_{ijk} \right) \right] = \frac{\Delta m}{144} \sum_{ijk} c_{2mln} I_{ijk}^2 = (U, \Lambda_2 U) > 0, \end{aligned}$$

из этого следует, что $\Lambda_2 = \Lambda_2^* > 0$.

Запишем итерационный процесс (35) в каноническом виде:

$$BU_t + AU = F, \quad U_t = \frac{U^{s+1} - U^s}{\tau_s}. \quad (38)$$

Здесь $B = \Lambda_1$, $A = \Lambda_1 + \Lambda_2$, $\tau_s \equiv 1$. Заметим, что $B = B^* > 0$, $A = A^* > 0$.

Поскольку операторы A и B не зависят от номера итерации, то процесс является стационарным. Достаточным условием для сходимости этого итерационного процесса является выполнение неравенства [8]:

$$B \geq 0.5 \tau_s A \quad (39)$$

или $\Lambda_1 \geq \Lambda_2$.

Вычислим скалярное произведение

$$\begin{aligned} RR = ((\Lambda_1 - \Lambda_2)U, U) &= \sum_{ijk} ((\Lambda_1 - \Lambda_2)(u)_{ijk} u_{ijk} \Delta m + (\Lambda_1 - \Lambda_2)(w)_{ijk} w_{ijk} \Delta m + (\Lambda_1 - \Lambda_2)(v)_{ijk} v_{ijk} \Delta m) = \\ &= \Delta m \kappa \sum_{ijk} (u_{ijk}^2 + w_{ijk}^2 + v_{ijk}^2) + \frac{\Delta m}{144} \sum_{ijk} (c_{1mln} - c_{2mln}) I_{ijk}^2. \end{aligned}$$

Для всех (i, j, k) справедливо неравенство

$$\begin{aligned} \max & \left(|u_{i+1jk} - u_{ijk+1}|, |u_{ij+1k} - u_{i+1jk}|, |u_{i+1j+1k+1} - u_{i+1jk}|, |u_{i+1j+1k} - u_{ij+1k+1}|, \right. \\ & |u_{i+1j+1k} - u_{ijk}|, |u_{i+1jk+1} - u_{i+1j+1k}|, |u_{i+1j+1k+1} - u_{ij+1k}|, |u_{ijk+1} - u_{ij+1k}|, \\ & |u_{i+1jk+1} - u_{ijk}|, |u_{ij+1k+1} - u_{ijk}|, |u_{ij+1k+1} - u_{i+1jk+1}|, |u_{i+1j+1k+1} - u_{i+1jk+1}|, \\ & |w_{i+1jk} - w_{ijk+1}|, |w_{ij+1k} - w_{i+1jk}|, |w_{i+1j+1k+1} - w_{i+1jk}|, |w_{i+1j+1k} - w_{ij+1k+1}|, \\ & |w_{i+1j+1k} - w_{ijk}|, |w_{i+1jk+1} - w_{i+1j+1k}|, |w_{i+1j+1k+1} - w_{ij+1k}|, |w_{ijk+1} - w_{ij+1k}|, \\ & |w_{i+1jk+1} - w_{ijk}|, |w_{ij+1k+1} - w_{ijk}|, |w_{ij+1k+1} - w_{i+1jk+1}|, |w_{i+1j+1k+1} - w_{i+1jk+1}|, \\ & |v_{i+1jk} - v_{ijk+1}|, |v_{ij+1k} - v_{i+1jk}|, |v_{i+1j+1k+1} - v_{i+1jk}|, |v_{i+1j+1k} - v_{ij+1k+1}|, \\ & |v_{i+1j+1k} - v_{ijk}|, |v_{i+1jk+1} - v_{i+1j+1k}|, |v_{i+1j+1k+1} - v_{ij+1k}|, |v_{ijk+1} - v_{ij+1k}|, \\ & \left. |v_{i+1jk+1} - v_{ijk}|, |v_{ij+1k+1} - v_{ijk}|, |v_{ij+1k+1} - v_{i+1jk+1}|, |v_{i+1j+1k+1} - v_{i+1jk+1}| \right) \leq \\ & \leq \sqrt{2 \sum_{k_1=0}^1 \sum_{k_2=0}^1 \sum_{k_3=0}^1 (u_{i+k_1j+k_2k+k_3}^2 + w_{i+k_1j+k_2k+k_3}^2 + v_{i+k_1j+k_2k+k_3}^2)}. \end{aligned}$$

Поэтому, для любого U будет справедливо неравенство:

$$\sum_{ijk} I_{ijk}^2 \leq 2 \sum_{ijk} \left(\sum_{k_1=0}^1 \sum_{k_2=0}^1 \sum_{k_3=0}^1 (u_{i+k_1j+k_2k+k_3}^2 + w_{i+k_1j+k_2k+k_3}^2 + v_{i+k_1j+k_2k+k_3}^2) \right).$$

$$\begin{aligned} & \left(S_{18}^x + S_{13}^x + S_{16}^x + S_{27}^x + S_{24}^x + S_{25}^x + S_{36}^x + S_{38}^x + S_{45}^x + S_{47}^x + S_{57}^x + S_{68}^x + S_{18}^y + S_{13}^y + S_{16}^y + S_{27}^y + S_{24}^y + S_{25}^y + S_{36}^y + S_{38}^y + \right. \\ & S_{45}^y + S_{47}^y + S_{57}^y + S_{68}^y + S_{18}^z + S_{13}^z + S_{16}^z + S_{27}^z + S_{24}^z + S_{25}^z + S_{36}^z + S_{38}^z + S_{45}^z + S_{47}^z + S_{57}^z + S_{68}^z \Big)^2 \leq \\ & \leq 2 \times 9 \times 144 \times (S^x + S^y + S^z)^2 \cdot \sum_{ijk} \left(\sum_{k_1=0}^1 \sum_{k_2=0}^1 \sum_{k_3=0}^1 (u_{i+k_1j+k_2k+k_3}^2 + w_{i+k_1j+k_2k+k_3}^2 + v_{i+k_1j+k_2k+k_3}^2) \right). \end{aligned}$$

Здесь $S^x = h_y h_z$; $S^y = h_x h_z$; $S^z = h_x h_y$;

$$\begin{aligned} h_r = \max_{ijk} & \left\{ |r_{i+1jk} - r_{ijk+1}|, |r_{ij+1k} - r_{i+1jk}|, |r_{i+1j+1k+1} - r_{i+1jk}|, |r_{i+1j+1k} - r_{ij+1k+1}|, \right. \\ & |r_{i+1j+1k} - r_{ijk}|, |r_{i+1jk+1} - r_{i+1j+1k}|, |r_{i+1j+1k+1} - r_{ij+1k}|, |r_{ijk+1} - r_{ij+1k}|, \\ & \left. |r_{i+1jk+1} - r_{ijk}|, |r_{ij+1k+1} - r_{ijk}|, |r_{ij+1k+1} - r_{i+1jk+1}|, |r_{i+1j+1k+1} - r_{ij+1k+1}| \right\}. \end{aligned}$$

Здесь учитываем, что $S_{k_1 k_2}^r \leq 3S^r$, $r = x, y, z$. Кроме того, справедливо тождество:

$$\begin{aligned} \sum_{ijk} (u_{ijk}^2 + w_{ijk}^2 + v_{ijk}^2) &= \frac{1}{8} \sum_{ijk} (u_{ijk}^2 + u_{i+1jk}^2 + u_{ij+1k}^2 + u_{i+1j+1k}^2 + u_{ijk+1}^2 + u_{i+1jk+1}^2 + u_{ij+1k+1}^2 + u_{i+1j+1k+1}^2 + \\ & + w_{ijk}^2 + w_{i+1jk}^2 + w_{ij+1k}^2 + w_{i+1j+1k}^2 + w_{ijk+1}^2 + w_{i+1jk+1}^2 + w_{ij+1k+1}^2 + w_{i+1j+1k+1}^2 + \\ & + v_{ijk}^2 + v_{i+1jk}^2 + v_{ij+1k}^2 + v_{i+1j+1k}^2 + v_{ijk+1}^2 + v_{i+1jk+1}^2 + v_{ij+1k+1}^2 + v_{i+1j+1k+1}^2). \end{aligned}$$

Следовательно, для любого U выполняется неравенство

$$2 \times 9 \times 144 \times (S^x + S^y + S^z)^2 \sum_{ijk} (u_{ijk}^2 + w_{ijk}^2 + v_{ijk}^2) \geq \sum_{ijk} I_{ijk}^2.$$

Учитывая последнее неравенство, получим

$$\begin{aligned} RR &\geq \frac{\Delta m}{144} \sum_{ijk} \left(\frac{\kappa}{9 \times 16 \times (S^x + S^y + S^z)} + c1_{mln} - c2_{mln} \right) I_{ijk}^2 \geq \\ &\geq \frac{\Delta m}{144} N \cdot \min_{ijk} \left(\frac{\kappa}{9 \times 16 \times (S^x + S^y + S^z)} + c1_{mln} - c2_{mln} \right) I_{ijk}^2. \end{aligned}$$

Здесь N — число узлов разностной сетки, кроме граничных.

Сходимость итерационного процесса (35) имеет место, если

$$\frac{\kappa}{9 \times 16 \times (S^x + S^y + S^z)} + c1_{mln} - c2_{mln} \geq 0$$

выполняется во всех ячейках разностной сетки, т.е. при

$$\frac{B^2}{8\pi} \geq (\gamma - 2)P. \quad (40)$$

В противном случае условие сходимости имеет следующий вид:

$$\Delta t < \frac{\sqrt{2} \cdot \Delta m}{12\rho(S^x + S^y + S^z) \sqrt{(\gamma - 2) \frac{P}{\rho} - \frac{B^2}{8\pi}}}. \quad (41)$$

Таким образом, критерий устойчивости для трехмерного метода комбинированного решения (40) совпадает с критериями для одномерного и двумерного аналогов метода (19). При этом ограничения на шаг по времени при нарушении условия (40), несколько строже ограничения на шаг по времени как для одномерного (20), так и для двумерного (21) аналогов.

Поскольку для большинства практических задач $\gamma < 2$, в этом случае условие сходимости (40) выполняется автоматически.

Рассмотрим теперь сходимость метода раздельного решения основных разностных уравнений (22)–(27). Тогда система разностных уравнений разделяется на три группы: к первой группе относятся динамические уравнения, ко второй — уравнения энергии, к третьей — уравнения электромагнитного поля. Далее решение проводится аналогично предыдущему по группам и также связывается общим итерационным циклом.

Анализ устойчивости данного алгоритма решения разностных уравнений проводится тем же способом, как и показано выше. Итерационный процесс, соответствующий этому алгоритму, сводится к (36) и (37) при

$$c1_{mln} = RT_{mln}\rho_{mln}^2; \quad c2_{mln} = (\gamma - 1)RT_{mln}\rho_{mln}^2 + \frac{B_{mln}^2\rho_{mln}}{8\pi}.$$

Таким образом, метод раздельного решения трехмерных МГД-уравнений будет сходиться, при справедливости условия

$$\frac{B^2}{8\pi} \leq (2 - \gamma)P \quad (42)$$

в каждой ячейке разностной сетки. В противном случае этот метод сходится при условии:

$$\Delta t < \frac{\sqrt{2} \cdot \Delta m}{12\rho(S^x + S^y + S^z) \sqrt{(\gamma - 2) \frac{P}{\rho} + \frac{B^2}{8\pi}}}. \quad (43)$$

5. ПРИМЕРЫ ЧИСЛЕННЫХ РАСЧЕТОВ

Учитывая, что проведенные выше исследования алгоритмов раздельного и комбинированного решения групп разностных уравнений носят в основном качественный характер, справедливость полученных выше оценок необходимо проверить численными расчетами.

Рассмотрим модельную задачу. В начальный момент времени $t = 0$ покоящееся вещество занимает область: $-3 \leq x \leq 3$, $-3 \leq y \leq 3$, $-3 \leq z \leq 3$. Начальная плотность и температура равны соответственно $\rho_0 = 0.2$, $T_0 = 0.9$. Вещество является идеальным газом $P = \rho T$, внутренняя энергия $\epsilon = \frac{T}{\gamma - 1}$. Газ нетеплопроводный,

т.е. $\kappa = 0$. Он помещен в однородное магнитное поле $B = (B_x, B_y, B_z)$, и является практически “вмороженным”: $\sigma = 10^{12}$. На границах $z = z_{\text{лев}}(t)$ и $z = z_{\text{прав}}(t)$ задана малая скорость $U_{\text{лев}}(t) = U_{\text{прав}}(t) = \pm 0.01$, на границах $x = x_{\text{лев}}(t)$, $x = x_{\text{прав}}(t)$, $y = y_{\text{лев}}(t)$ и $y = y_{\text{прав}}(t)$ задается нулевая скорость.

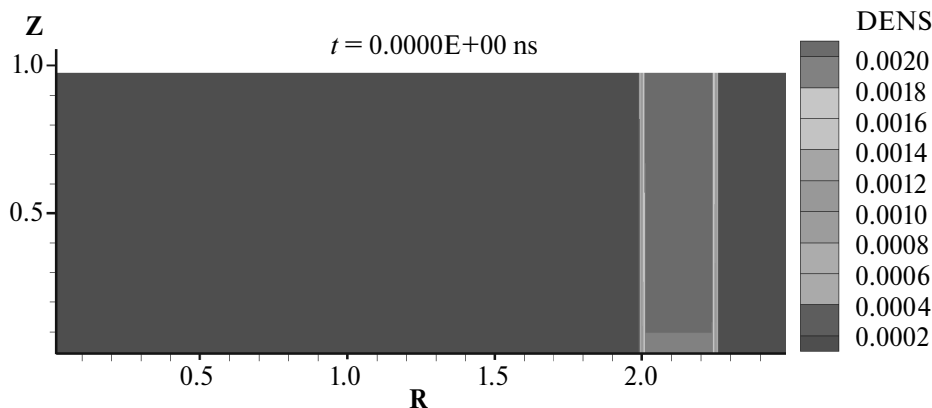
В данной постановке акустические колебания в среде распространяются со скоростью, много большей скорости выдвигания (вдвижения) поршня. Следовательно, в любой момент времени распределение плотности газа и магнитного поля практически однородно.

Задача решалась для различных значений величин $B = \sqrt{B_x^2 + B_y^2 + B_z^2}$ и γ : $0 \leq B \leq 2000$, $1 < \gamma \leq 15$. Данный класс задач решался как в лагранжевых, так и в смешанных эйлерово-лагранжевых переменных (СЭЛ) [12].

Расчеты проводились методами отдельного и комбинированного решения групп разностных уравнений

Ограничения на временной шаг в численных расчетах хорошо согласуются с оценками, полученными в настоящей работе. При $\gamma \leq 2$ расчеты с любыми магнитными полями комбинированным методом проходили практически без ограничения на шаг по времени (в расчетах $\Delta t = 1$). Расчеты отдельным методом абсолютно устойчивы только при $B < 4.6$, при $B = 100$ шаг интегрирования $t \approx 0.0013$, а при $B = 1000$ — $\Delta t \approx 0.00012$, что неприемлемо при решении практических задач. Оценки полученные в работе, справедливы как для лагранжевых, так и для СЭЛ переменных.

Рассмотрим расчет, моделирует сжатие плазмы лайнером (газоплазменной оболочкой), ускоряемым электродинамически импульсом тока [9, 10]). В начальный момент времени неподвижная плазма занимает область $0 \leq R \leq 2.5$, $0 \leq Z \leq 1$ и $0 < \varphi \leq 2\pi$ (рассматривается (R, φ, Z) геометрия). Расчет проводился в СЭЛ переменных. Начальная температура $T_e = T_i = 1$ эВ (рассматривается двухтемпературная модель). В области $2 \leq R \leq 2.25$ моделируется лайнер с плотностью, соответствующей $\rho_l = 1.97 \cdot 10^{-3} \cdot (1 + 0.1 \cdot \sin(\frac{\pi z}{2}))$ г/см³, в остальной области моделируется плазма с плотностью $\rho_{\text{пл}} = 2 \cdot 10^{-6}$ г/см³ (см. фиг. 3).



Фиг. 3. Начальное распределение плотности, $\varphi = \pi$.

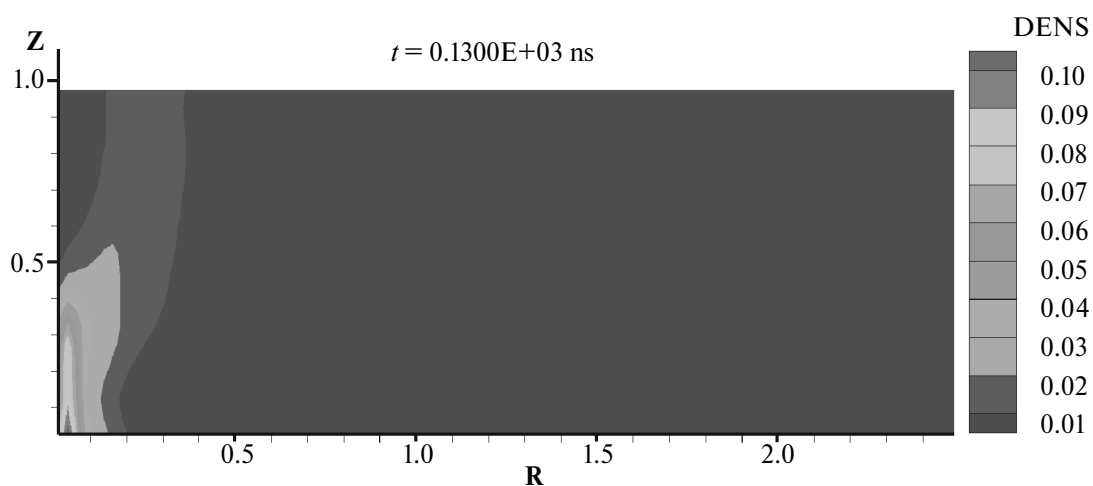
Вещество-вольфрам. Уравнение состояния приводится в работе [11]. Газ помещен в однородное магнитное поле $B_z = 50$ кГс. На внешней границе $B_\varphi(r_{\text{прав}}, \varphi, z) = B_0 \cdot \sin(\frac{\pi t}{0.16})$. Здесь t время, которое измеряется в мкс., $B_0 = 2$ МГс.

На фиг. 4–7 представлены результаты расчета динамики лайнера на момент времени 0.13 мкс в виде распределенной плотности в г/см³; радиальной компоненты скорости U_R в 10^6 см/с; $B_\varphi \cdot R$ — компоненты магнитной индукции B_φ в МГс, умноженной на радиус R ; компоненты магнитной индукции B_z в МГс. Около этого момента времени лайнер достигает осевой линии и продолжает схлопываться.

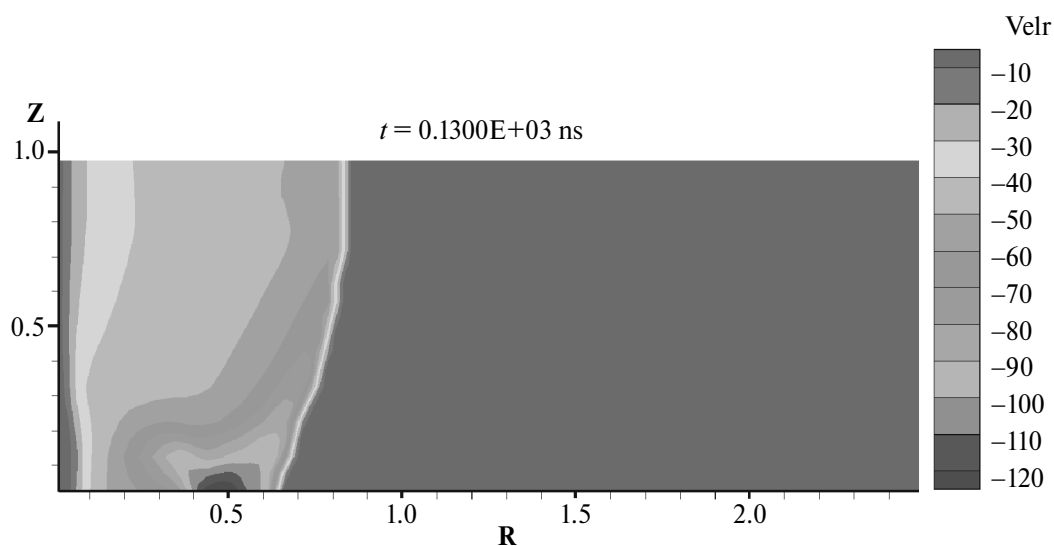
Провести расчет удалось только методом комбинированного решения групп разностных уравнений. При решении задачи методом отдельного решения, при росте $B_\varphi(t)$ шаг по времени резко падает, что подтверждено оценками, полученными в данной работе (см. формулы (42), (43)).

6. ЗАКЛЮЧЕНИЕ

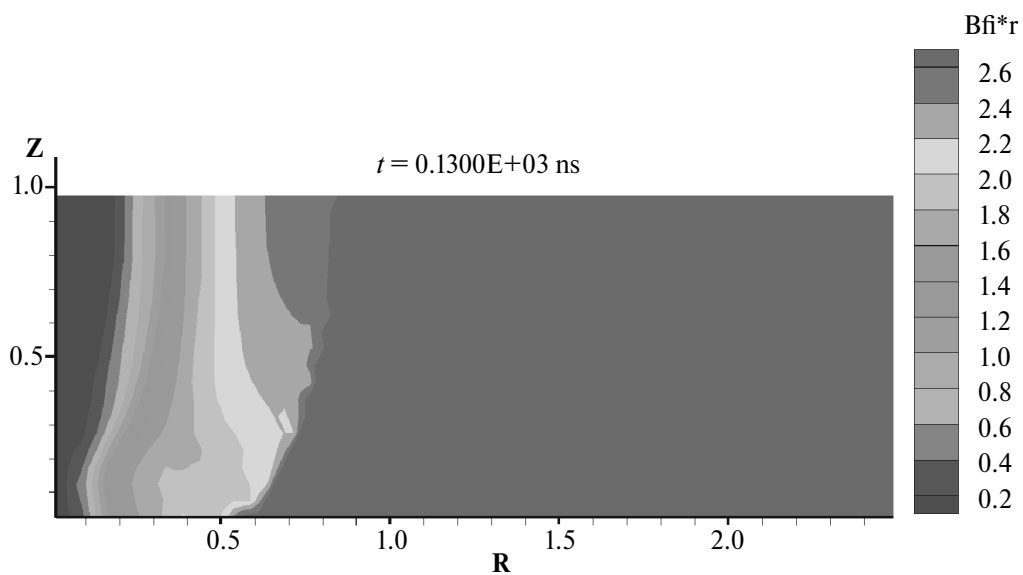
Полученные в работе оценки сходимости алгоритмов отдельного и комбинированного решения системы трехмерных разностных уравнений подкреплены численными модельными расчетами и расчетами практических задач. При решении конкретной задачи представляется целесообразным задействовать оба описанных в настоящей работе численных метода. Вначале каждого временного шага анализируем физическое соотношение



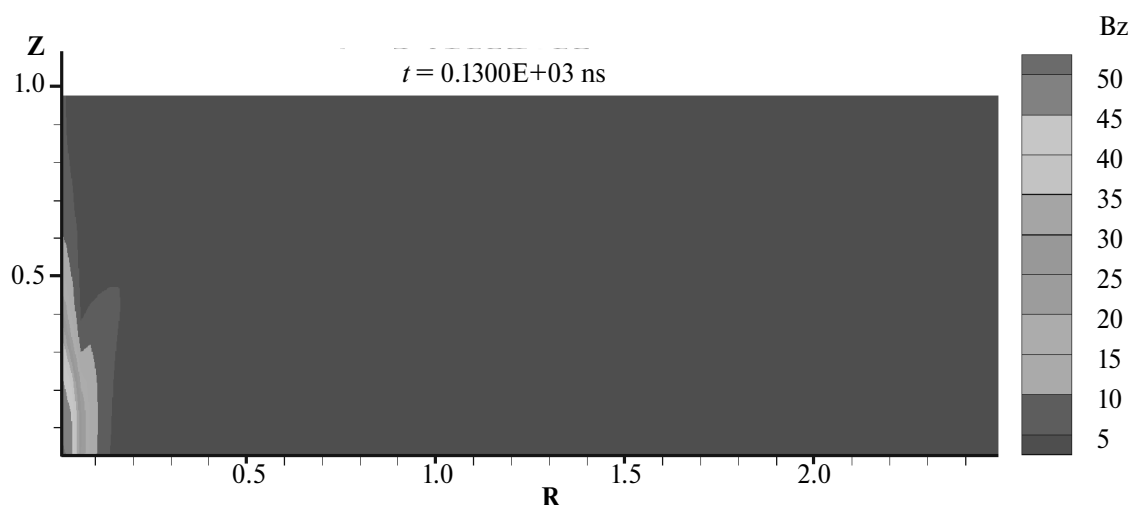
Фиг. 4. Распределение плотности на момент времени 0.13 мкс, $\varphi = \pi$.



Фиг. 5. Распределение радиальной компоненты скорости U_R на момент времени 0.13 мкс, $\varphi = \pi$.



Фиг. 6. Распределение компоненты магнитной индукции B_φ , умноженной на радиус R на момент времени 0.13 мкс, $\varphi = \pi$.



Фиг. 7. Распределение компоненты магнитной индукции B_z на момент времени 0.13 мкс, $\varphi = \pi$.

параметров вещества и магнитного поля. В случае выполнения на данном временном шаге условия сходимости для метода раздельного решения системы разностных уравнений целесообразно использовать именно этот метод. В противном случае необходимо применить более универсальный метод комбинированного решения системы разностных уравнений.

СПИСОК ЛИТЕРАТУРЫ

1. Самарский А.А., Попов Ю.П. Разностные методы решения задач газовой динамики. М.: Наука, 1992. 424 с.
2. Гасилов В.А., Круковский А.Ю., Новикова Т.П., Оточин А.А. Оценки сходимости некоторых итерационных алгоритмов численного решения задач двумерной магнитной гидродинамики. М.: 1995. Препринт ИММ РАН, № 6, 20 с.
3. Круковский А.Ю., Гасилов В.А., Повещенко Ю.А., Шарова Ю.С., Ключкова Л.В. Визуализация полностью консервативной лагранжево-эйлеровой схемы для двухмерных задач магнитной гидродинамики // Матем. моделирование. 2020. Т. 32. № 1. С. 50–70, <https://doi.org/10/20948/mm-2020-01-04>
4. Куликовский А.Г., Любимов В.А. Магнитная гидродинамика. М.: Логос. 2005. 328 с.
5. Круковский А.Ю. Сходимость метода комбинированных прогонок для разностной схемы одномерной магнитной гидродинамики. М.: 1988. Препринт ИПМ им. М.В. Келдыша № 113, 12 с.
6. Самарский А.А., Волосевич П.П., Волчинская М.И., Курдюмов С.П. Метод конечных разностей для решения одномерных нестационарных задач магнитной гидродинамики // Ж. вычисл. матем. и матем. физ. 1968. Т. 8. № 5. С. 1025–1038.
7. Федоренко Р.П. Введение в вычислительную физику. М.: Изд-во физико-технического института, 1994. 504 с.
8. Самарский А.А., Гулин А.В. Численные методы. М.: Наука, 1989. 432 с.
9. Дюдерштадт Дж., Мозес Г. Инерциальный термоядерный синтез. М.: Энергоатомиздат, 1984. 300 с.
10. Zakharov S.V., Smirnov V.P., Grabovskii E.V., Nedoseev S.L., Oleinik G.M., Zaitsev V.I. Imploding liner as driver for indirect driven target physics studies // Proc. of the IAEA Technical Committee Meeting on Drivers for Inertial Confinement Fusion. Paris, 1994 (International Atomic Energy Agency Vienna, 1995), p. 395.
11. Никифоров А.Ф., Новиков В.Г., Уваров В.Б. Квантово-статистические модели высокотемпературной плазмы и методы расчета росселандовых пробегов и уравнений состояния. М.: Физматлит, 2000, 400 с.; Nikiforov A.F., Novikov V.G., Uvarov V.B. Quantum-statistical models of hot dense matter. Methods for computation opacity and equation of state. Switzerland: Birkhauser, 2005. 447 p.

12. Головизин В.М., Рязанов В.А., Сороковицова О.С. Об одном классе полностью консервативных схем МГД в смешанных эйлерово-лагранжевых переменных // Ж. вычисл. матем. и матем. физ. 1984.Т. 24. №4. С. 520–533.

CONVERGENCE ESTIMATES OF ITERATIVE METHODS FOR NUMERICAL MODELING OF THREE-DIMENSIONAL PROCESSES IN MAGNETOHYDRODYNAMICS

A.Yu. Krukovsky^{a,*}, I.V. Popov^{a,**}, Yu.A. Poveshchenko^{a,***}

^a Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences, Miusskaya sq., 4, Moscow, 125047 Russia

*e-mail:alexander-krukovskiy@yandex.ru,

**e-mail:piv2964@mail.ru

***e-mail:hecon@mail.ru

Received 15 April, 2024

Revised 15 April, 2024

Accepted 05 May, 2024

Abstract. This paper addresses the convergence of iterative processes applied to implicit, fully conservative difference schemes in three-dimensional magnetohydrodynamics, using both separate and combined solution methods for groups of difference equations split by physical processes. Convergence estimates for the iterative processes of the numerical methods considered in this study are obtained. The applicability of both combined and separate methods for solving three-dimensional difference equations in magnetohydrodynamics is examined. Given that the presented algorithm analysis is mainly qualitative, the validity of the obtained estimates was confirmed through numerical experiments on both model and real-world problems. Notably, the convergence estimates of the iterative processes allow for the selection of an optimal numerical method for solving difference equations in three-dimensional magnetohydrodynamic problems at any time step.

Keywords: three-dimensional magnetohydrodynamics equations, implicit fully conservative difference scheme, convergence of iterative process.

ОБ ОДНОМ ПОДХОДЕ К ПОВЫШЕНИЮ ТОЧНОСТИ И ЭФФЕКТИВНОСТИ МЕТОДА РОМБ ДЛЯ РЕШЕНИЯ НЕСТАЦИОНАРНОГО УРАВНЕНИЯ ТЕПЛОПРОВОДНОСТИ

© 2024 г. Н.Я. Моисеев^{1,*}

¹456770 Снежинск, Челябинская обл., ул. Победы, 17-87, Россия

*e-mail: nik.moiseev.43@mail.ru

Поступила в редакцию 02.04.2024 г.

Переработанный вариант 02.04.2024 г.

Принята к публикации 05.05.2024 г.

Представлена модификация метода РОМБ для численного решения нестационарных нелинейных уравнений теплопроводности. Модифицированные разностные схемы являются консервативными и на гладких решениях аппроксимируют модельные дифференциальные уравнения с постоянными коэффициентами со вторым порядком по времени и по пространству. Для модифицированных разностных схем приведены первые дифференциальные приближения, которые позволяют оценить погрешности аппроксимации в схемах. Библ. 14. Табл. 1.

Ключевые слова: нестационарное нелинейное уравнение теплопроводности, разностные схемы предиктор-корректор, метод РОМБ.

DOI: 10.31857/S0044466924080084, EDN: YAPRJA

1. ВВЕДЕНИЕ

Метод РОМБ является одним из методов численного решения нестационарных нелинейных одномерных [1], двумерных [2–3] и трехмерных [4] уравнений теплопроводности. Разностные схемы в методе РОМБ являются схемами типа предиктор-корректор [5–6], позволяют получать монотонные численные решения широкого класса прикладных задач и аппроксимируют исходные дифференциальные уравнения теплопроводности с первым порядком по времени и вторым по пространству [3]. Метод предназначен для его применения при решении уравнений газовой динамики, поэтому уравнение теплопроводности — это квазилинейное уравнений вида $\partial E / \partial t = \text{div}(k \text{ grad } T)$, где внутренняя энергия E и температура T вещества связаны уравнением состояния, в общем случае нелинейным.

Численные решения многомерных (2D или 3D) уравнений теплопроводности в схеме РОМБ находятся итерациями методом переменных направлений со стабилизирующей поправкой [6–7].

Если внутренняя энергия зависит от температуры нелинейно, то уравнения теплопроводности путем линеаризации энергии по температуре преобразуются к уравнению для вычисления температуры. Это уравнение в дальнейшем будем называть линеаризованным уравнением теплопроводности. Разностные схемы РОМБ (1D–3D) относятся к схемам типа предиктор-корректор и конструируются на основе линеаризованного уравнения теплопроводности.

В результате линеаризации энергии по температуре в линеаризованных уравнениях теплопроводности появляются недивергентные члены, которые можно рассматривать как источник погрешностей аппроксимации. Погрешности аппроксимации проявляются в нарушении консервативности разностных схем: дисбаланс по энергии отличен от нуля. Поэтому вопрос устранения источника этих погрешностей в методе РОМБ является актуальным.

Здесь представлена модификация метода РОМБ, которая позволяет повысить точность метода, обеспечить консервативность и устранить отмеченный выше источник погрешностей аппроксимации. Суть модификации состоит в том, что на этапе предиктор температуры и тепловые потоки на гранях ячеек вычисляются на промежуточных дробных временных шагах по методу РОМБ; а на этапе корректор вместо расчета температуры из линеаризованного уравнения вычисляется удельная внутренняя энергия вещества из дивергентного нелинейного уравнения теплопроводности без линеаризации энергии по температуре.

Особенности модификации метода РОМБ излагаются на примере решения двумерного уравнения теплопроводности в случае плоской симметрии.

2. МЕТОД РОМБ ДЛЯ ЧИСЛЕННОГО РЕШЕНИЯ УРАВНЕНИЯ ТЕПЛОПРОВОДНОСТИ

2.1. Постановка задачи

Двумерное уравнение теплопроводности в случае плоской симметрии запишем в виде

$$\frac{\partial E}{\partial t} = \frac{1}{\rho} \left[\frac{\partial}{\partial x} \left(\kappa \frac{\partial T}{\partial x} \right) + \frac{\partial}{\partial y} \left(\kappa \frac{\partial T}{\partial y} \right) \right] + Q. \quad (2.1)$$

Здесь x, y и t — независимые переменные по пространству и времени соответственно, E, T — неизвестные функции, E — удельная внутренняя энергия вещества, T — температура вещества, $\kappa(\rho, T)$ — коэффициент теплопроводности, ρ, Q — известные функции, ρ — удельная плотность вещества, Q — функция источника. Уравнение (2.1) дополняется уравнением состояния вещества в форме $E = E(\rho, T)$. Начальные условия: $T(0, x, y) = T_0(x, y)$ заданы на плоскости x, y в области D , ограниченной контуром L . Граничные условия: $\alpha T + \beta \mathbf{S} * \mathbf{n} = \varphi(t)$ для $(x, y) \in L$, $\mathbf{n}$ — внешняя нормаль к L , $\mathbf{S}(S_x, S_y) = -\kappa \nabla T$ — вектор плотности теплового потока, S_x, S_y — компоненты вектора плотности теплового потока $\mathbf{S}$ вдоль координатных осей x, y соответственно, $\nabla = (\partial/\partial x, \partial/\partial y)^T$, α, β, φ — параметры, с помощью которых можно задать различные граничные условия.

Требуется найти решение начально-краевой задачи для уравнения (2.1) при $t > 0$.

2.2. Основы подхода к решению нестационарного нелинейного уравнения теплопроводности методом РОМБ

Поскольку метод РОМБ подробно описан во многих публикациях, в частности, в работах [1–4], то здесь приводится краткое описание метода РОМБ для решения нестационарного нелинейного уравнения теплопроводности в двумерном случае.

Уравнение теплопроводности (2.1) в методе РОМБ записывается в потоковой форме в виде

$$\begin{aligned} \frac{\partial E}{\partial t} + \frac{1}{\rho} \left(\frac{\partial S_x}{\partial x} + \frac{\partial S_y}{\partial y} \right) &= Q, \\ S_x + \kappa \frac{\partial T}{\partial x} &= 0, \\ S_y + \kappa \frac{\partial T}{\partial y} &= 0. \end{aligned} \quad (2.2)$$

Численное решение уравнения теплопроводности (2.2) находится итерациями методом переменных направлений со стабилизирующей поправкой [6–7] расщеплением по направлениям (называемых в схеме РОМБ каналами). Эти итерации в дальнейшем будем называть *внешними*. На каждой внешней итерации решения уравнения теплопроводности находятся в два этапа. На первом этапе решение находится прогонками вдоль одного каналов (отвечающего координате x), на втором — вдоль второго (отвечающего координате y). Переход от внешней итерации μ к итерации $\mu + 1$ имеет вид

$$\frac{\partial E^{\mu+1/2}}{\partial t} + \frac{1}{\rho} \left(\left(\frac{\partial S_x}{\partial x} \right)^{\mu+1/2} + \left(\frac{\partial S_y}{\partial y} \right)^{\mu} \right) = Q, \quad (2.3)$$

$$\begin{aligned} (S_x)^{\mu+1/2} + \kappa^{\mu} \frac{\partial T^{\mu+1/2}}{\partial x} &= 0, \\ \frac{\partial E^{\mu+1}}{\partial t} + \frac{1}{\rho} \left(\left(\frac{\partial S_x}{\partial x} \right)^{\mu+1/2} + \left(\frac{\partial S_y}{\partial y} \right)^{\mu+1} \right) &= Q, \\ (S_y)^{\mu+1} + \kappa^{\mu} \frac{\partial T^{\mu+1}}{\partial y} &= 0. \end{aligned} \quad (2.4)$$

Если коэффициент теплопроводности зависит от температуры, то при его вычислении температура берется с предыдущей μ -й итерации.

Если внутренняя энергия E зависит от температуры нелинейно, то уравнения (2.3), (2.4) преобразуются к уравнениям для вычисления температуры с помощью линеаризации энергии по температуре:

$$E(\rho, T^{v+1}) \equiv E^{v+1} \cong E^v + E_T^v(\rho, T^v)(T^{v+1} - T^v). \quad (2.5)$$

Из этого соотношения по найденной энергии E^{v+1} находится температура T^{v+1} . Здесь v — это номер итерации по нелинейности энергии, $E_T^v = (\partial E / \partial T)^v$. Эти итерации будем называть *внутренними*. Из (2.5) следует

$$\frac{\partial E^{v+1}}{\partial t} \cong \frac{\partial E^v}{\partial t} + E_T^v(\rho, T^v) \frac{\partial T^{v+1}}{\partial t}.$$

Точность выполнения внешних и внутренних итераций задается некоторыми постоянными, называемыми *константами сходимости*.

Систему уравнений (2.3), (2.4) с линеаризованным уравнением теплопроводности для вычисления температуры запишем в виде

$$\frac{\partial T^{v+1}}{\partial t} + \frac{1}{\rho E_T(\rho, T^v)} \left(\left(\frac{\partial S_x}{\partial x} \right)^{v+1} + \left(\frac{\partial S_y}{\partial y} \right)^\mu \right) = \frac{1}{E_T(\rho, T^v)} \left(Q + \frac{\partial E^v}{\partial t} \right), \quad (2.6)$$

$$(S_x)^{v+1} + \varkappa^\mu \frac{\partial T^{v+1}}{\partial x} = 0,$$

$$\frac{\partial T^{v+1}}{\partial t} + \frac{1}{\rho E_T(\rho, T^v)} \left(\left(\frac{\partial S_x}{\partial x} \right)^{\mu+1/2} + \left(\frac{\partial S_y}{\partial y} \right)^{v+1} \right) = \frac{1}{E_T(\rho, T^v)} \left(Q + \frac{\partial E^v}{\partial t} \right), \quad (2.7)$$

$$(S_y)^{v+1} + \varkappa^\mu \frac{\partial T^{v+1}}{\partial y} = 0.$$

Решения уравнений (2.6), (2.7) находятся внутренними итерациями по v . Для вычисления промежуточных температур и тепловых потоков на гранях ячеек система уравнений (2.6), (2.7) дополняется соотношениями связи вдоль каналов между температурой и тепловыми потоками, заданными в центрах и на соответствующих гранях ячеек [3]. Промежуточные температуры и тепловые потоки на гранях ячеек вычисляются из решения расширенной системы уравнений по неявным разностным схемам в момент времени $t^{n+1} = t^n + \tau$ (это этап предиктор).

Температура в центрах ячеек вычисляется из разностных уравнений (2.6), (2.7) (это этап корректор).

Если решается уравнение (2.6), то $T^{v=0} = T^n$, т.е. температура берется с нижнего слоя по времени. При сходимости внутренних v -итераций с заданной точностью температура $T^{\mu+1/2}$, отвечающая внешней μ -итерации, полагается равной температуре T^{v+1} на последней внутренней итерации. Если решается уравнение (2.7), то $T^{v=0} = T^{\mu+1/2}$. Затем, при сходимости внешних итераций с заданной точностью температура $T^{\mu+1}$ на внешней итерации с номером $\mu + 1$ полагается равной температуре T^{v+1} . Контроль сходимости внутренних итераций проверяется в каждой точке i, j по условию

$$|T_{i,j}^{v+1} - T_{i,j}^v| < \varepsilon_1 (1 + T_{i,j}^v).$$

Здесь ε_1 — константа сходимости внутренних итераций, $T_{i,j}^{v+1}, T_{i,j}^v$ — температура вещества в центре ячейки i, j на $v + 1$ и v итерациях соответственно. Если условие выполнено, то итерации заканчиваются.

Контроль сходимости внешних итераций проверяется по условию

$$|T_{i,j}^{\mu+1} - T_{i,j}^\mu| < \varepsilon_2 (1 + T_{i,j}^\mu).$$

Здесь ε_2 — константа сходимости внешних итераций. Температура $T_{i,j}^{n+1}$ в центре ячейки при сходимости внешних μ -итераций полагается равной температуре на последней итерации

$$T_{i,j}^{n+1} = T_{i,j}^{\mu+1}.$$

Удельная внутренняя энергия вещества в центрах ячеек вычисляется из уравнения состояния.

Поскольку уравнения теплопроводности (2.6), (2.7) нелинейные и решения находятся итерациями, то консервативность и дисбаланс схем зависят от констант сходимости $\varepsilon_1, \varepsilon_2$, что можно отнести к недостаткам метода РОМБ.

3. МОДИФИЦИРОВАННЫЙ МЕТОД РОМБ ДЛЯ РЕШЕНИЯ НЕСТАЦИОНАРНОГО НЕЛИНЕЙНОГО УРАВНЕНИЯ ТЕПЛОПРОВОДНОСТИ

Основная цель модификации метода РОМБ направлена на повышение точности численных решений за счет устранения отмеченных выше недостатков в методе РОМБ, а именно, первый порядок аппроксимации по времени и нарушение консервативности в случае нелинейной зависимости внутренней энергии от температуры. В предлагаемой модифицированной схеме повышение порядка аппроксимации достигается за счет перехода к вычислению промежуточных величин температуры и тепловых потоков на гранях ячеек в другой момент времени $t^{n+1/2} = t^n + \tau^*$, $0 < \tau^* < \tau$. Параметр τ^* подбирается так, чтобы решения были монотонными, и на гладких решениях разностная схема аппроксимировала бы дифференциальные уравнения с повышенным вторым порядком по времени [8–9].

Обеспечение консервативности схем в случае нелинейной зависимости внутренней энергии от температуры достигается за счет перехода на этапе корректор от вычисления температуры к вычислению удельной внутренней энергии из дивергентных уравнений теплопроводности (2.3), (2.4) без линеаризации энергии.

В результате модифицированная разностная схема является консервативной, т.е. дисбаланс внутренней энергии равен нулю и не зависит от итераций по нелинейности энергии. На верхнем слое по времени температура вещества $T_{i,j}^{n+1}$ вычисляется в центрах ячеек из уравнения состояния, например, итерационным методом Ньютона, но эта процедура уже не нарушает баланс внутренней энергии.

Рассмотрим аппроксимацию модифицированной разностной схемы. Для простоты изложения предположим, что плотность вещества $\rho = 1$, источник энергии $Q = 0$, внутренняя энергия зависит от температуры линейно $E(\rho, T) = T$ и коэффициент теплопроводности $\kappa(\rho, T) = k = \text{const}$. Для исключения влияния погрешности внешних μ -итерации, доводим их до сходимости с точностью, при которой погрешность расщепления по направлениям можно считать равной нулю. Область интегрирования — прямоугольник со сторонами параллельными координатным осям. Разностное уравнение для вычисления удельной внутренней энергии (температуры) в этом случае записывается в виде

$$\frac{T_{i,j}^{n+1} - T_{i,j}^n}{\tau} + \frac{S_{i,j+1/2}^{n+1/2} - S_{i,j-1/2}^{n+1/2}}{\Delta x} + \frac{S_{i+1/2,j}^{n+1/2} - S_{i-1/2,j}^{n+1/2}}{\Delta y} = 0. \quad (3.1)$$

Здесь $\tau, \Delta x, \Delta y$ — шаги интегрирования вдоль осей t, x, y соответственно, $S_{i,j+1/2}^{n+1/2}$ — тепловые потоки, которые вычислены на гранях ячеек по неявной разностной схеме методом РОМБ [3] в момент времени $t^{n+1/2} = t^n + \tau^*$, $\tau^* > 0$. Разностные сетки являются равномерными вдоль каждого направления. Первое дифференциальное приближение (ПДП) [10] схемы (3.1) выпишем в виде

$$\frac{\partial T}{\partial t} + \frac{\partial S_x}{\partial x} + \frac{\partial S_y}{\partial y} = w. \quad (3.2)$$

Здесь $w \equiv w(\tau^*, \tau, \Delta x, \Delta y)$,

$$w = k^2(2\tau^* - 0.5\tau) \left(\frac{\partial^4 T}{\partial x^4} + 2 \frac{\partial^4 T}{\partial x^2 \partial y^2} + \frac{\partial^4 T}{\partial y^4} \right) + k \left(\frac{(\Delta x)^2}{12} \frac{\partial^4 T}{\partial x^4} + \frac{(\Delta y)^2}{12} \frac{\partial^4 T}{\partial y^4} \right).$$

Правую часть w ПДП (3.2) можно рассматривать как основную погрешность аппроксимации дифференциального уравнения теплопроводности

$$\frac{\partial T}{\partial t} - k \frac{\partial^2 T}{\partial x^2} - k \frac{\partial^2 T}{\partial y^2} = 0.$$

Замечание. ПДП упрощенной разностной схемы в [4] для решения нестационарного уравнения теплопроводности с постоянными коэффициентами и с тремя пространственными переменными запишется в виде

$$\frac{\partial T}{\partial t} - k \frac{\partial^2 T}{\partial x^2} - k \frac{\partial^2 T}{\partial y^2} - k \frac{\partial^2 T}{\partial z^2} = w, \quad (3.3)$$

$$w = (2\tau^* - 0.5\tau)k^2 \left(\frac{\partial^4 T}{\partial x^4} + \frac{\partial^4 T}{\partial y^4} + \frac{\partial^4 T}{\partial z^4} + 2 \frac{\partial^4 T}{\partial x^2 \partial y^2} + 2 \frac{\partial^4 T}{\partial x^2 \partial z^2} + 2 \frac{\partial^4 T}{\partial y^2 \partial z^2} \right) + \\ + \frac{1}{12}(\Delta x)^2 k \frac{\partial^4 T}{\partial x^4} + \frac{1}{12}(\Delta y)^2 k \frac{\partial^4 T}{\partial y^4} + \frac{1}{12}(\Delta z)^2 k \frac{\partial^4 T}{\partial z^4}.$$

Из ПДП (3.2), (3.3) следует, что разностные схемы с двумя и тремя пространственными переменными аппроксимируют соответствующие дифференциальные уравнения теплопроводности с постоянными коэффициентами с первым порядком по времени и вторым по пространству. Погрешность аппроксимации $O(\tau, (\Delta x)^2, (\Delta y)^2)$. Погрешности аппроксимации при $\tau^* = 0$ и $\tau^* = 0.5\tau$ в явной и неявной разностных схемах соответственно для решения одномерного уравнения теплопроводности совпадают с главными членами ошибок аппроксимации, которые приведены в работе [11].

Если $\tau^* = 0.25\tau$, то разностные схемы аппроксимируют упрощенные дифференциальные уравнения теплопроводности на гладких решениях со вторым порядком по времени и по пространству. Погрешность аппроксимации $O(\tau^2, (\Delta x)^2, (\Delta y)^2)$. Если $\tau^* > 0.25\tau$, то коэффициенты в ПДП положительные. Известно, что в схемах второго порядка точности могут появляться осцилляции. Поэтому конструируется гибридная разностная схема, в которой на гладких решениях применяются схемы повышенной точности, а в зоне больших градиентов —

схемы первого порядка точности [8, 9]. Это достигается выбором параметра τ^* , который подбирается так, чтобы решения были монотонными [9].

Поскольку погрешность аппроксимации w линейно зависит от параметра τ^* , то минимум и максимум погрешности достигаются на концах интервала $0.25\tau \leq \tau^* \leq \tau$, на котором коэффициенты ПДП положительные. Поэтому для погрешности выполняются условия

$$w(\tau^* = 0.25\tau) \leq w(\tau^*) \leq w(\tau^* = \tau).$$

Если $\tau^* = \tau$, то $w(\tau^* = \tau)$ — это погрешность аппроксимации метода РОМБ. Следовательно, погрешность аппроксимации модифицированного метода РОМБ для $0.25\tau \leq \tau^* < \tau$ не превосходит погрешности метода РОМБ.

Из ПДП (3.2) следует оценка погрешности аппроксимации, которая на гладких решениях с ограниченными четвертыми производными

$$|\partial^4 T / \partial x^4| \leq M_1, \quad |\partial^4 T / \partial x^4| \leq M_2, \quad |\partial^4 T / \partial x^2 \partial y^2| \leq M_3$$

записывается в виде

$$w(\tau^*) \leq \{(\Delta x)^2/12 + (\Delta y)^2/12 + 4k(2\tau^* - 0.5\tau)\}kM, \quad M = \max\{M_1, M_2, M_3\}.$$

Константы M_1, M_2, M_3 не зависят от шагов интегрирования по времени и по пространству. Следовательно, с уменьшением шагов интегрирования по времени и по пространству погрешность аппроксимации стремится к нулю, а решения ПДП (3.2) сходятся к решению дифференциального уравнения. Это дает основание надеяться, что модифицированные разностные схемы с переменными коэффициентами обладают этими же свойствами.

Эффективность модифицированной схемы можно повысить за счет ослабления требований к сходимости внешних и внутренних итераций (по v и по μ соответственно) выбором более грубых констант сходимости. Погрешность аппроксимации с грубой константой сходимости компенсируется за счет вычисления удельной внутренней энергии по консервативным схемам, как отмечено в [12]. В результате число итераций уменьшается и время счета шага сокращается.

4. РЕЗУЛЬТАТЫ РАСЧЕТОВ МОДЕЛЬНЫХ ЗАДАЧ

Для проверки эффективности и работоспособности модифицированного метода РОМБ проведены расчеты модельных одномерных и двумерных задач. Результаты этих расчетов сравнивались с результатами расчетов по методике РОМБ и с аналитическими решениями. Эффективность модифицированного метода по сравнению с методом РОМБ демонстрируется по результатам расчетов двух модельных задач: одномерной и двумерной.

Одномерная начально-краевая задача с плоской симметрией. В области $\{0 \leq x \leq 10\}$ задана начальная температура $T(0, x) = 0.0001$. Плотность вещества $\rho = 1$. На левой границе задана температура $T(t, 0) = at^b$, $a = 1.54, b = 1/8$. На правой границе задан поток $S(t, 10) = ct^b$, $c = -49.35$. Скорость тепловой волны $D = 32$, коэффициент теплопроводности — $k(\rho, T) = 256T^8$. Уравнение состояния линейное — $E(\rho, T) = T$. Требуется рассчитать положение тепловых волн в момент времени $t = 0.15$. Задача имеет точное решение [13].

Расчеты проводились на сгущающихся равномерных разностных сетках с шагами интегрирования по пространству h_k и по времени $\tau_k = h_k/D$, с константами сходимости $\varepsilon_{1,1} = \varepsilon_1 = \varepsilon_2 = 0.001$ и $\varepsilon_{2,1} = \varepsilon_1 = \varepsilon_2 = 0.00001$. Тепловая волна с такими шагами интегрирования проходит за один шаг одну ячейку. Промежуточные величины на этапе предиктор в расчетах по модифицированной методике рассчитывались с параметром $\tau^* = 0.75\tau$.

Погрешность аппроксимации $w(\tau^*)$ в расчетах одномерных задач вычислялась в сеточном аналоге нормы L_1 по формуле

$$w_k = h_k \sum_{i=1}^{N_k} |T_i - T_i^{\text{exact}}|, \quad k = 1, 2, \dots, 8, \quad N_k = 25, 50, \dots, 3200.$$

Здесь T_i^{exact} — это точное решение, T_i — численное решение, h_k — шаг интегрирования по пространству, N_k — число интервалов по пространству.

В табл. 1 приведены в первом столбце — число интервалов в области, во втором и третьем — погрешности $w(\varepsilon_{1,1})$, $w(\varepsilon_{2,1})$, полученные в расчетах по методу РОМБ с константами сходимости $\varepsilon_{1,1}$, $\varepsilon_{2,1}$ соответственно, в четвертом — погрешности $f(\varepsilon_{1,1})$, полученные в расчетах по предложенному в данной статье модифицированному методу, в столбцах $k1, k2, k3, k4, k5, k6$ — коэффициенты эффективности модифицированного метода относительно методики РОМБ: $k1 = f(\varepsilon_{1,1})/w(\varepsilon_{1,1})$, $k4 = f(\varepsilon_{1,1})/w(\varepsilon_{2,1})$, $k2 = tf(\varepsilon_{1,1})/tw(\varepsilon_{1,1})$, $k5 = tf(\varepsilon_{1,1})/tw(\varepsilon_{2,1})$ — отношения времен счета, $k3 = vf(\varepsilon_{1,1})/vw(\varepsilon_{1,1})$, $k6 = vf(\varepsilon_{1,1})/vw(\varepsilon_{2,1})$ — отношения

Таблица 1.

N_k	$w(\varepsilon_{1,1})$	$w(\varepsilon_{2,1})$	$f(\varepsilon_{1,1})$	k_1	k_2	k_3	k_4	k_5	k_6
25	.2763	.2746	.2208	0.8	0.45	0.53	0.8	0.47	0.36
50	.1511	.1491	.1142	0.76	0.66	0.54	0.76	0.66	0.33
100	.0745	.0736	.0615	0.83	0.5	1.82	0.83	0.29	0.31
200	.0371	.0369	.0328	0.88	0.61	0.6	0.88	0.31	0.32
400	.0187	.0187	.0174	0.93	0.64	0.63	0.93	0.33	0.34
800	.0096	.0096	.0092	0.96	0.67	0.66	0.96	0.35	0.35
1600	.0051	.0051	.0049	0.96	0.72	0.71	0.96	0.37	0.37
3200	.0027	.0027	.0026	0.96	0.75	0.73	0.96	0.41	0.4

среднего числа итераций на одном шаге. Здесь tf, vf — времена счета, число итераций, соответственно, по модифицированной методике, tw, vw — по методу РОМБ.

Из анализа данных в табл. 1 следует, что погрешности с уменьшением шагов интегрирования по времени и по пространству стремятся к нулю с показателем скорости сходимости близким к единице (см. столбцы 2–4). Погрешности расчетов температуры, времена расчетов и число итераций с различными константами сходимости по модифицированному методу меньше, чем по методу РОМБ: все коэффициенты эффективности меньше единицы. Положения тепловых волн, рассчитанные по обоим методикам, хорошо согласуются с точными решениями.

Двумерная задача с плоской симметрией. В двумерной задаче рассчитывается движение тепловой волны в прямоугольной области $\{0 \leq x \leq 18, 0 \leq y \leq 20\}$ с нетеплопроводной вставкой в области $\{10 \leq x \leq 12, 4 \leq y \leq 20\}$, где $k(\rho, T) = 0$. Геометрия задачи взята из работы [2]. Плотность вещества $\rho = 0.1$. Начальная температура $T(0, x, y) = 0.001$. Коэффициент теплопроводности $k(\rho, T) = 6T^6$. Уравнение состояния является нелинейным и записывается в виде [14]

$$E(\rho, T) = aT + bT^4,$$

где $a = 10, b = 13.7$. На границах области тепловые потоки равны нулю $S = 0$. В области $\{0 \leq x \leq 1, 19 \leq y \leq 20\}$ задано энерговыделение $dE/dt = e, e = 15000$. Требуется рассчитать состояние в момент времени $t = 0.26$.

Точного решения задача не имеет. Задача решена в двух вариантах: с нетеплопроводной вставкой и без вставки. Задача существенно двумерная: тепловая волна движется от левого верхнего угла в направлении правого нижнего угла, обтекает на своем пути нетеплопроводную вставку и распространяется между вставкой и правой границей от нижней границы к верхней. На начальной стадии движения имеется выраженный фронт тепловой волны, который исчезает после достижения нижней границы.

Расчеты проводились на равномерных разностных сетках с шагами интегрирования по пространству $\Delta x = 0.1, \Delta y = 0.1$ и по времени $\tau = 1$ и $\tau = 0.01$ с константами сходимости $\varepsilon_{1,1} = 5e - 4, \varepsilon_{2,1} = 5e - 3$. Дополнительно проводились расчеты по модифицированной методике РОМБ с параметром $\tau^* = 0.56\tau$.

Анализ результатов расчетов показал, что температурные поля, рассчитанные по обоим методикам, хорошо согласуются между собой. Дисбаланс по энергии в расчетах без вставки с константой сходимости $\varepsilon_{1,1}, \tau^* = 0.75\tau$ и шагом интегрирования по времени $\tau = 1$ по модифицированной методике равен $\delta \sim 1e - 14$, число внешних итераций — 991, по методике РОМБ — $\delta \sim 1e - 6$, число внешних итераций — 1155. Время счета по модифицированной методике составляет 0.9 от времени счета по исходной методике. Число внешних итераций в расчетах по модифицированной методике с константой сходимости $\varepsilon_{2,1}$ равно 726, время счета составляет 0.7 относительно времени счета с константой сходимости ε_1 .

Сравнение температур вдоль координатной линии $y = 13.4$ в расчетах задачи со вставкой показывает, что распределения температуры, полученные по методике РОМБ с шагами интегрирования по времени $\tau = 1$ и $\tau = 0.01$ различаются для $y > 10$.

Распределения, полученные по модифицированной методике с шагами интегрирования по времени $\tau = 1$ и $\tau = 0.01$, хорошо согласуются между собой и с распределением по методике РОМБ в расчете с $\tau = 0.01$.

Дисбаланс по энергии в расчетах задачи со вставкой с константой сходимости $\varepsilon_{1,1}$ и шагом интегрирования по времени $\tau = 1$ по модифицированной методике порядка $\delta \sim 10^{-12}$, по методике РОМБ — $\delta \sim 10^{-5}$, среднее число итераций на одном шаге равно 106 по методике РОМБ и равно 73 по модифицированной методике. Время счета по модифицированной методике составляет 0.7 от времени счета по исходной методике.

Результаты решения задачи 1 с линейным уравнением состояния и задачи 2 со сложным нелинейным уравнением состояния показывают, что предложенная модификация методики РОМБ повышает точность численных решений с преимуществом по времени счета.

5. ЗАКЛЮЧЕНИЕ

Представлена модификация метода РОМБ, применяемого для решения нестационарных нелинейных уравнений теплопроводности в рамках интегрирования уравнений газодинамики. Построенная разностная схема является консервативной и на гладких решениях аппроксимирует модельные дифференциальные уравнения с постоянными коэффициентами со вторым порядком по времени и по пространству. Приведенные для модельных разностных схем первые дифференциальные приближения позволяют оценить погрешности в схемах. Полученные численные результаты подтверждают теоретические выводы — модифицированный метод РОМБ обеспечивает повышение точности, вычислительной эффективности и может быть рекомендован для решения многомерных уравнений теплопроводности.

СПИСОК ЛИТЕРАТУРЫ

1. Гаджиев А.Д., Писарев В.Н. Неявный конечно-разностный метод для численного решения уравнений газовой динамики с теплопроводностью // Ж. вычисл. матем. и матем. физ. 1979. Т. 19. № 5. С. 1288.
2. Гаджиев А.Д., Писарев В.Н., Шестаков А.А. Метод расчета задач теплопроводности на неортогональных сетках // Ж. вычисл. матем. и матем. физ. 1982. Т. 22. № 2. С. 339.
3. Гаджиев А.Д., Писарев В.Н., Рыкованова В.В., Шестаков А.А. Методика и программа ТОМ1 для решения двумерного уравнения теплопроводности // ВАНТ. Сер. Методики и программы для численного решения задач математической физики. 1985. Вып.1. С. 53.
4. Писарев В.Н., Чернова С.В. Численная методика РОМБ для решения трехмерного уравнения теплопроводности в криволинейной системе координат // ВАНТ. Сер. Методики и программы для численного решения задач математической физики. 2007. Вып. 3(4). С. 3.
5. Брайн П.Л.И. Конечно-разностный метод высокого порядка точности для решения трехмерных задач теплопроводности // A.I. Ch. E. J. 1961. № 7. С. 367.
6. Яненко Н.Н. Метод дробных шагов решения многомерных задач математической физики. Новосибирск: Наука, 1967. С. 26—43.
7. Марчук Г.И. Методы расщепления. М.: Наука, 1988. С. 129—136.
8. Федоренко Р.П. Применение разностных схем высокой точности для решения гиперболических уравнений // Ж. вычисл. матем. и матем. физ. 1962. Т. 2. № 6. С.1122.
9. Моисеев Н.Я. Неявные разностные схемы бегущего счета повышенной точности // Ж. вычисл. матем. и матем. физ. 2011. Т. 51. № 5. С. 920.
10. Шокин Ю.И. Метод дифференциального приближения. Новосибирск: Наука, 1979.
11. Флетчер К. Вычислительные методы в динамике жидкостей 1. М.: Мир, 1991. С. 110—112.
12. Забродин А.В., Пикарчук С.Б. Об одном методе численного решения нелинейного уравнения теплопроводности на параллелограммной сетке точек // ВАНТ. Сер. Методики и программы для численного решения задач математической физики. 1982. Вып. 2(10). С. 14.
13. Самарский А.А., Попов Ю.П. Разностные методы решения задач газовой динамики. М.: Наука, 1992. С. 133—135.
14. Крайко А.Н. Теоретическая газовая динамика. М.: ТОРУС ПРЕСС, 2010. С. 43—48.

ON AN APPROACH TO IMPROVING THE ACCURACY AND EFFICIENCY OF THE ROMB METHOD FOR SOLVING THE UNSTEADY HEAT CONDUCTION EQUATION

N. Ya. Moiseev*

Chelyabinsk region, Snezhinsk, Pobedy st., 17-87, 456770 Russia

**e-mail: nik.moiseev.43@mail.ru*

Received 02 April, 2024

Revised 02 April, 2024

Accepted 05 May, 2024

Abstract. A modification of the ROMB method is presented for numerically solving unsteady nonlinear heat conduction equations. The modified difference schemes are conservative and, for smooth solutions, approximate the model differential equations with constant coefficients with second-order accuracy in both time and space. First differential approximations are provided for the modified difference schemes, enabling the assessment of approximation errors in the schemes.

Keywords: unsteady nonlinear heat conduction equation, predictor-corrector difference schemes, ROMB method.

УДК 517.922

ОБ УСЛОВИЯХ РАЗРЕШИМОСТИ ЗАДАЧИ КОШИ ДЛЯ
ОДНОЙ ПСЕВДОГИПЕРБОЛИЧЕСКОЙ СИСТЕМЫ¹⁾© 2024 г. Л. Н. Бондарь^{1,*}, С. Б. Мингнарлов^{1,**}¹630090 Новосибирск, ул. Пирогова, 2, Новосибирский государственный университет, Россия

*e-mail: l.bondar@g.nsu.ru

**e-mail: s.mingnarov@g.nsu.ru

Поступила в редакцию 05.05.2024 г.

Переработанный вариант 05.05.2024 г.

Принята к публикации 05.05.2024 г.

Рассматривается задача Коши для одной системы, не разрешенной относительно старшей производной по времени. Исследуемая система относится к классу псевдогиперболических. Система описывает поперечные изгибно-крутильные колебания упругого стержня. Получены условия разрешимости задачи Коши в соболевских пространствах и оценки решения. Библ. 12.

Ключевые слова: псевдогиперболическая система, задача Коши, изгибно-крутильные колебания, условия разрешимости.

DOI: 10.31857/S0044466924080092, EDN: YANCYM

1. ВВЕДЕНИЕ

В работе рассматривается следующая система

$$\begin{pmatrix} I - \alpha D_x^2 & 0 & a_1 \\ 0 & I - \alpha D_x^2 & -a_2 \\ \sigma a_1 & -\sigma a_2 & I - \alpha D_x^2 \end{pmatrix} D_t^2 \begin{pmatrix} u \\ v \\ \theta \end{pmatrix} + \beta D_x^4 \begin{pmatrix} u \\ v \\ \theta \end{pmatrix} = \begin{pmatrix} f^1(t, x) \\ f^2(t, x) \\ f^3(t, x) \end{pmatrix}, \quad (1.1)$$

где $\alpha, \beta > 0$, $\sigma(a_1^2 + a_2^2) = 1$. Система (1.1) описывает поперечные изгибно-крутильные колебания упругого стержня (см. [1]). Эта система является не разрешенной относительно старшей производной по времени и входит в класс псевдогиперболических систем [2].

В монографии [2] был введен класс псевдогиперболических уравнений и изучена задача Коши для него. Дальнейшие исследования разрешимости задачи Коши для псевдогиперболических уравнений проводились в [3]–[5] и др. Для псевдогиперболических систем общей теории разрешимости задачи Коши нет, есть лишь единичные результаты для конкретных систем (см., например, [6]–[9] и др.).

Отметим, что при $a_1^2 + a_2^2 = 0$, система (1.1) распадается на три псевдогиперболических уравнения и разрешимость задачи Коши для таких уравнений следует из работы [4]. Случай $0 < \sigma(a_1^2 + a_2^2) < 1$ изучен в работе [9]. В данной работе будет исследован случай, когда $\sigma(a_1^2 + a_2^2) = 1$.

Сделаем замену:

$$\begin{aligned} \tilde{x} &= \frac{x}{\sqrt{\alpha}}, & \tilde{t} &= \frac{\sigma}{\alpha\sqrt{\beta}}t, & \tilde{\theta}(\tilde{t}, \tilde{x}) &= \frac{\theta(t, x)}{\sqrt{\sigma}}, & \tilde{u}(\tilde{t}, \tilde{x}) &= u(t, x), \\ \tilde{v}(\tilde{t}, \tilde{x}) &= v(t, x), & \varepsilon_1 &= \sqrt{\sigma}a_1, & \varepsilon_2 &= \sqrt{\sigma}a_2, \\ \tilde{f}^1(\tilde{t}, \tilde{x}) &= \frac{\alpha^2\beta}{\sigma^2}f^1(t, x), & \tilde{f}^2(\tilde{t}, \tilde{x}) &= \frac{\alpha^2\beta}{\sigma^2}f^2(t, x), & \tilde{f}^3(\tilde{t}, \tilde{x}) &= \frac{\alpha^2\beta}{\sigma^{\frac{3}{2}}}f^3(t, x), & \tilde{\sigma}^2 &= \frac{\beta^2}{\sigma^2}, \end{aligned}$$

систему (1.1) перепишем в виде:

$$\begin{pmatrix} I - D_{\tilde{x}}^2 & 0 & \varepsilon_1 \\ 0 & I - D_{\tilde{x}}^2 & -\varepsilon_2 \\ \varepsilon_1 & -\varepsilon_2 & I - D_{\tilde{x}}^2 \end{pmatrix} D_{\tilde{t}}^2 \begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{\theta} \end{pmatrix} + \tilde{\sigma}^2 D_{\tilde{x}}^4 \begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{\theta} \end{pmatrix} = \begin{pmatrix} \tilde{f}^1(\tilde{t}, \tilde{x}) \\ \tilde{f}^2(\tilde{t}, \tilde{x}) \\ \tilde{f}^3(\tilde{t}, \tilde{x}) \end{pmatrix}.$$

¹⁾ Работа выполнена при поддержке Математического Центра в Академгородке, соглашение № 075-15-2022-282 с Минобрнауки РФ.

Для сокращения записи, в системе оставим прежние обозначения, т.е. далее будем рассматривать следующую систему:

$$\begin{pmatrix} I - D_x^2 & 0 & \varepsilon_1 \\ 0 & I - D_x^2 & -\varepsilon_2 \\ \varepsilon_1 & -\varepsilon_2 & I - D_x^2 \end{pmatrix} D_t^2 \begin{pmatrix} u \\ v \\ \theta \end{pmatrix} + \sigma^2 D_x^4 \begin{pmatrix} u \\ v \\ \theta \end{pmatrix} = \begin{pmatrix} f^1(t, x) \\ f^2(t, x) \\ f^3(t, x) \end{pmatrix}, \quad (1.2)$$

$$\varepsilon_1^2 + \varepsilon_2^2 = 1.$$

Наша цель — доказательство разрешимости задачи Коши для псевдогиперболической системы (1.2) в соболевских пространствах, получение оценок на решение.

2. ФОРМУЛИРОВКА РЕЗУЛЬТАТОВ

Рассмотрим задачу Коши для псевдогиперболической системы:

$$\begin{pmatrix} I - D_x^2 & 0 & \varepsilon_1 \\ 0 & I - D_x^2 & -\varepsilon_2 \\ \varepsilon_1 & -\varepsilon_2 & I - D_x^2 \end{pmatrix} D_t^2 U + \sigma^2 D_x^4 U = F(t, x), \quad t > 0, x \in R, \quad (2.1)$$

$$U|_{t=0} = \Phi(x), \quad D_t U|_{t=0} = \Psi(x),$$

где

$$U(t, x) = (u(t, x), v(t, x), \theta(t, x))^T, \quad F(t, x) = (f^1(t, x), f^2(t, x), f^3(t, x))^T, \\ \Phi(x) = (\varphi^1(x), \varphi^2(x), \varphi^3(x))^T, \quad \Psi(x) = (\psi^1(x), \psi^2(x), \psi^3(x))^T,$$

$\sigma > 0$, при этом $\varepsilon_1^2 + \varepsilon_2^2 = 1$.

Дадим определения анизотропных соболевских пространств (см., например, [2]), которые понадобятся нам при доказательстве разрешимости задачи (2.1).

Определение 1. Функция $u(t, x) \in L_2(G)$ принадлежит анизотропному соболевскому пространству $W_2^{l_1, l_2}(G)$, $G \subseteq R^2$, $l_1, l_2 \in N$, если существуют обобщенные производные

$$D_t^{\alpha_1} D_x^{\alpha_2} u(t, x), \quad \alpha_1/l_1 + \alpha_2/l_2 \leq 1,$$

в области G , при этом $D_t^{\alpha_1} D_x^{\alpha_2} u(t, x) \in L_2(G)$. Введем норму

$$\|u, W_2^{l_1, l_2}(G)\| = \sum_{(\alpha_1, \alpha_2): \alpha_1/l_1 + \alpha_2/l_2 \leq 1} \|D_t^{\alpha_1} D_x^{\alpha_2} u, L_2(G)\|.$$

Определение 2. Функция $u(t, x)$ принадлежит анизотропному соболевскому пространству с экспоненциальным весом $W_{2, \gamma}^{l_1, l_2}(G)$, $\gamma > 0$, если $e^{-\gamma t} u(t, x) \in W_2^{l_1, l_2}(G)$. Полагаем

$$\|u(t, x), W_{2, \gamma}^{l_1, l_2}(G)\| = \|e^{-\gamma t} u(t, x), W_2^{l_1, l_2}(G)\|.$$

Определение 3. Функция $f(t, x)$ принадлежит анизотропному соболевскому пространству $W_{2, \gamma}^{0, 1}(G)$, $\gamma > 0$, если $e^{-\gamma t} f(t, x) \in L_2(G)$, существует обобщенная производная $D_x f(t, x)$ в G , при этом $e^{-\gamma t} D_x f(t, x) \in L_2(G)$. Полагаем

$$\|f(t, x), W_{2, \gamma}^{0, 1}(G)\| = \|e^{-\gamma t} f(t, x), L_2(G)\| + \|e^{-\gamma t} D_x f(t, x), L_2(G)\|.$$

Будем говорить, что $V(t, x) = (v^1(t, x), v^2(t, x), v^3(t, x))^T \in W_{2, \gamma}^{l_1, l_2}(G)$, если $v^j(t, x) \in W_{2, \gamma}^{l_1, l_2}(G)$, $j = 1, 2, 3$. Полагаем

$$\|V(t, x), W_{2, \gamma}^{l_1, l_2}(G)\| = \sum_{j=1}^3 \|v^j(t, x), W_{2, \gamma}^{l_1, l_2}(G)\|.$$

В работе доказана следующая теорема.

Теорема. Пусть $\Phi(x) \in W_2^4(R)$, $\Psi(x) \in W_2^3(R)$ и вектор-функция $F(t, x) = (f^1(t, x), f^2(t, x), f^3(t, x))^T \in W_{2, \gamma}^{0, 1}(R_+^2)$, $\gamma > 0$, такая, что

$$(1 + x^2)F(t, x) \in L_{2, \gamma}(R_+; L_1(R))$$

и выполнены условия:

$$\varepsilon_1 \int_R f^1(t, x) dx - \varepsilon_2 \int_R f^2(t, x) dx - \int_R f^3(t, x) dx = 0, \quad (2.2)$$

$$\varepsilon_1 \int_R x f^1(t, x) dx - \varepsilon_2 \int_R x f^2(t, x) dx - \int_R x f^3(t, x) dx = 0, \quad t > 0. \quad (2.3)$$

Тогда задача Коши (2.1) имеет единственное решение $U(t, x)$ в пространстве вектор-функций $W_{2,\gamma}^{2,4}(R_+^2)$, $\gamma > 0$, таких, что $D_t^2 D_x^2 U \in L_{2,\gamma}(R_+^2)$, при этом справедлива оценка:

$$\begin{aligned} & \|U(t, x), W_{2,\gamma}^{2,4}(R_+^2)\| + \|D_t^2 D_x^2 U(t, x), L_{2,\gamma}(R_+^2)\| \leq \\ & \leq c(\gamma) \left(\|\Phi(x), W_2^4(R)\| + \|\Psi(x), W_2^3(R)\| + \|F(t, x), W_{2,\gamma}^{0,1}(R_+^2)\| + \right. \\ & \left. + \| (1+x^2)F(t, x), L_1(R) \|, L_{2,\gamma}(R_+) \| \right), \end{aligned} \quad (2.4)$$

где $c(\gamma)$ — константа, зависящая от коэффициентов системы и γ .

Условия, подобные (2.2), (2.3) возникали и раньше при исследовании разрешимости задачи Коши для строго псевдогиперболических уравнений в соболевских пространствах (см. [2, 4]). В [2] были приведены примеры, показывающие, что подобные условия являются необходимыми для разрешимости задачи Коши для строго псевдогиперболических уравнений в соответствующих классах.

Из условий теоремы следует, что для разрешимости задачи Коши (2.1) помимо условий ортогональности (2.2), (2.3) существенными требованиями являются условия гладкости на правую часть системы $F(t, x)$ и начальные данные $\Phi(x)$, $\Psi(x)$. Отметим, что требования дополнительной гладкости правых частей и начальных данных является характерной особенностью для задачи Коши для гиперболических уравнений и систем (см., например, [10], [11]).

3. РАЗРЕШИМОСТЬ ЗАДАЧИ КОШИ

Для построения решения задачи (2.1) рассмотрим вспомогательную задачу Коши для системы обыкновенных дифференциальных уравнений с параметром $\xi \in R$, которая получается при формальном применении оператора Фурье по x к задаче (2.1):

$$\begin{pmatrix} 1 + \xi^2 & 0 & \varepsilon_1 \\ 0 & 1 + \xi^2 & -\varepsilon_2 \\ \varepsilon_1 & -\varepsilon_2 & 1 + \xi^2 \end{pmatrix} D_t^2 V + \sigma^2 \xi^4 V = \widehat{F}(t, \xi), \quad t > 0, \xi \in R, \quad (3.1)$$

$$V|_{t=0} = \widehat{\Phi}(\xi), \quad D_t V|_{t=0} = \widehat{\Psi}(\xi),$$

где $\widehat{F}(t, \xi)$ — преобразование Фурье вектор-функции $F(t, x)$ по x .

Поскольку матрица при производной $D_t^2 V$

$$L^0(i\xi) = \begin{pmatrix} 1 + \xi^2 & 0 & \varepsilon_1 \\ 0 & 1 + \xi^2 & -\varepsilon_2 \\ \varepsilon_1 & -\varepsilon_2 & 1 + \xi^2 \end{pmatrix}$$

вырождается при $\xi = 0$, $\varepsilon_1^2 + \varepsilon_2^2 = 1$, то будем рассматривать систему (3.1) при $\xi \in R \setminus \{0\}$. Нетрудно получить формулы решения задачи Коши (3.1):

$$V(t, \xi) = V_1(t, \xi) + V_2(t, \xi) + V_3(t, \xi), \quad (3.2)$$

где

$$V_1(t, \xi) = (B_1 \cos(b_1(\xi)t) + B_2 \cos(b_2(\xi)t) + B_3 \cos(b_3(\xi)t)) \widehat{\Phi}(\xi), \quad (3.3)$$

$$V_2(t, \xi) = \left(B_1 \frac{\sin(b_1(\xi)t)}{b_1(\xi)} + B_2 \frac{\sin(b_2(\xi)t)}{b_2(\xi)} + B_3 \frac{\sin(b_3(\xi)t)}{b_3(\xi)} \right) \widehat{\Psi}(\xi), \quad (3.4)$$

$$V_3(t, \xi) = \int_0^t \left(B_1 \frac{\sin(b_1(\xi)(t-s))}{b_1(\xi)} + B_2 \frac{\sin(b_2(\xi)(t-s))}{b_2(\xi)} + B_3 \frac{\sin(b_3(\xi)(t-s))}{b_3(\xi)} \right) G(s, \xi) ds. \quad (3.5)$$

Здесь

$$B_1 = \begin{pmatrix} \varepsilon_2^2 & \varepsilon_1^2 & 0 \\ \varepsilon_1 \varepsilon_2 & 1 - \varepsilon_1 \varepsilon_2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad 2B_2 = \begin{pmatrix} \varepsilon_1^2 & -\varepsilon_1^2 & -\varepsilon_1 \\ -\varepsilon_1 \varepsilon_2 & \varepsilon_1 \varepsilon_2 & \varepsilon_2 \\ -\varepsilon_1 & \varepsilon_1 & 1 \end{pmatrix},$$

$$2B_3 = \begin{pmatrix} \varepsilon_1^2 & -\varepsilon_1^2 & \varepsilon_1 \\ -\varepsilon_1 \varepsilon_2 & \varepsilon_1 \varepsilon_2 & -\varepsilon_2 \\ \varepsilon_1 & -\varepsilon_1 & 1 \end{pmatrix}, \quad b_1(\xi) = \frac{\sigma \xi^2}{\sqrt{1 + \xi^2}}, \quad b_2(\xi) = \sigma |\xi|, \quad b_3(\xi) = \frac{\sigma \xi^2}{\sqrt{2 + \xi^2}}, \quad (3.6)$$

$$G(t, \xi) = (L^0(i\xi))^{-1} \hat{F}(t, \xi) = \begin{pmatrix} (1 + \xi^2)^2 - \varepsilon_2^2 & -\varepsilon_1 \varepsilon_2 & -\varepsilon_1(1 + \xi^2) \\ -\varepsilon_1 \varepsilon_2 & (1 + \xi^2)^2 - \varepsilon_1^2 & \varepsilon_2(1 + \xi^2) \\ -\varepsilon_1(1 + \xi^2) & \varepsilon_2(1 + \xi^2) & (1 + \xi^2)^2 \end{pmatrix} \frac{\hat{F}(t, \xi)}{\xi^2(1 + \xi^2)(2 + \xi^2)}.$$

Представим вектор-функцию $G(t, \xi)$ в виде:

$$G(t, \xi) = G_1(t, \xi) + G_2(t, \xi), \quad (3.7)$$

$$G_1(t, \xi) = \begin{pmatrix} 2 + \xi^2 & 0 & -\varepsilon_1 \\ 0 & 2 + \xi^2 & \varepsilon_2 \\ -\varepsilon_1 & \varepsilon_2 & 2 + \xi^2 \end{pmatrix} \frac{\hat{F}(t, \xi)}{(1 + \xi^2)(2 + \xi^2)},$$

$$G_2(t, \xi) = \begin{pmatrix} \varepsilon_1^2 & -\varepsilon_1 \varepsilon_2 & -\varepsilon_1 \\ -\varepsilon_1 \varepsilon_2 & \varepsilon_2^2 & \varepsilon_2 \\ -\varepsilon_1 & \varepsilon_2 & 1 \end{pmatrix} \frac{\hat{F}(t, \xi)}{\xi^2(1 + \xi^2)(2 + \xi^2)}.$$

Введем

$$\chi^m(\xi) = \begin{cases} 1, & |\xi| > 1/m, \\ 0, & |\xi| < 1/m, \end{cases} \quad V^m(t, \xi) = \chi^m(\xi) V(t, \xi).$$

Тогда для всех $t > 0$ вектор-функции $V^m(t, \xi) \in L_2(R)$ и к ним можно применить обратный оператор Фурье по ξ , и задать, как и в работе [4], в качестве последовательности приближенных решений задачи Коши (2.1) $\{U^m(t, x)\}$, где $U^m(t, x) = F^{-1}[V^m](t, x)$. Далее мы покажем, что последовательность $\{U^m(t, x)\}$, действительно, является последовательностью приближенных решений задачи (2.1).

Учитывая представление $V(t, \xi)$ из (3.2), перепишем $U^m(t, x)$ в виде:

$$U^m(t, x) = \sum_{j=1}^3 U_j^m(t, x), \quad U_j^m(t, x) = F^{-1}[\chi^m(\xi) V_j](t, x), \quad (3.8)$$

где $V_1(t, \xi)$, $V_2(t, \xi)$, $V_3(t, \xi)$ из (3.3)–(3.5), соответственно.

В следующих леммах проведем оценки последовательности $\{U_j^m(t, x)\}$, $j = 1, 2, 3$, в норме пространства $W_{2,\gamma}^{2,4}(R_+^2)$.

Лемма 1. Справедлива следующая оценка:

$$\sum_{\frac{\beta_1}{2} + \frac{\beta_2}{4} < 1} \left(\|D_t^{\beta_1} D_x^{\beta_2} U_1^m(t, x), L_{2,\gamma}(R_+^2)\| + \|D_t^{\beta_1} D_x^{\beta_2} U_2^m(t, x), L_{2,\gamma}(R_+^2)\| \right) \leq$$

$$\leq c \left(\|\Phi(x), W_2^3(R)\| + \|\Psi(x), W_2^2(R)\| \right),$$

где константа $c > 0$ не зависит от m , $\Phi(x)$, $\Psi(x)$, более того, для любого $k \geq 1$ имеет место сходимость

$$\sum_{\frac{\beta_1}{2} + \frac{\beta_2}{4} < 1} \|D_t^{\beta_1} D_x^{\beta_2} [U_j^{m+k}(t, x) - U_j^m(t, x)], L_{2,\gamma}(R_+^2)\| \rightarrow 0, \quad m \rightarrow \infty, \quad j = 1, 2. \quad (3.9)$$

Доказательство леммы 1. Учитывая (3.3), (3.6), (3.8), равенство Парсеваля, оценку

$$|b_1(\xi)| + |b_2(\xi)| + |b_3(\xi)| \leq c_1 |\xi|, \quad (3.10)$$

имеем

$$\begin{aligned} & \sum_{\frac{\beta_1}{2} + \frac{\beta_2}{4} < 1} \|D_t^{\beta_1} D_x^{\beta_2} U_1^m(t, x), L_{2,\gamma}(R_+^2)\| \leq \\ & \leq \frac{c_2}{\sqrt{\gamma}} \sum_{\frac{\beta_1}{2} + \frac{\beta_2}{4} < 1} \sum_{i=1}^3 \| |\xi|^{\beta_1 + \beta_2} \widehat{\varphi}^i(\xi), L_2(R) \| \leq \frac{c_3}{\sqrt{\gamma}} \|\Phi(x), W_2^3(R)\|. \end{aligned} \quad (3.11)$$

Учитывая оценку $b_i(\xi) \geq c|\xi|$, $|\xi| \geq 1$, будем иметь

$$\left| \frac{\sin(b_i(\xi)t)}{b_i(\xi)} \right| \leq \frac{c_4}{|\xi|}, \quad |\xi| \geq 1, \quad i = 1, 2, 3. \quad (3.12)$$

а из формулы

$$\sin(b_i(\xi)t) = b_i(\xi)t \int_0^1 \cos(sb_i(\xi)t) ds, \quad i = 1, 2, 3,$$

следует

$$\left| \frac{\sin(b_i(\xi)t)}{b_i(\xi)} \right| \leq t, \quad |\xi| < 1, \quad i = 1, 2, 3. \quad (3.13)$$

Оценки (3.12), (3.13) перепишем в одно неравенство:

$$\left| \frac{\sin(b_i(\xi)t)}{b_i(\xi)} \right| \leq \frac{c_3(t+1)}{1+|\xi|}, \quad \xi \in R, \quad i = 1, 2, 3. \quad (3.14)$$

В силу равенства Парсеваля, учитывая (3.4), (3.6), (3.10), (3.14), получим

$$\begin{aligned} & \sum_{\frac{\beta_1}{2} + \frac{\beta_2}{4} < 1} \|D_t^{\beta_1} D_x^{\beta_2} U_2^m(t, x), L_{2,\gamma}(R_+^2)\| \leq \\ & \leq c_4(\gamma) \left(\sum_{i=1}^3 \|\widehat{\psi}^i(\xi), L_2(R)\| + \sum_{0 < \frac{\beta_1}{2} + \frac{\beta_2}{4} < 1} \sum_{i=1}^3 \| |\xi|^{\beta_1 + \beta_2 - 1} \widehat{\psi}^i(\xi), L_2(R) \| \right) \leq \\ & \leq c_5(\gamma) \|\Psi(x), W_2^2(R)\|. \end{aligned} \quad (3.15)$$

Из оценок (3.11), (3.15) вытекает нужная оценка и, рассуждая аналогично, получим сходимость (3.9).

Лемма доказана.

Лемма 2. Справедлива следующая оценка:

$$\begin{aligned} & \sum_{j=1}^2 \left(\|D_t^2 U_j^m(t, x), L_{2,\gamma}(R_+^2)\| + \|D_t^2 D_x^2 U_j^m(t, x), L_{2,\gamma}(R_+^2)\| + \right. \\ & \left. + \|D_x^4 U_j^m(t, x), L_{2,\gamma}(R_+^2)\| \right) \leq c(\gamma) (\|\Phi(x), W_2^4(R)\| + \|\Psi(x), W_2^3(R)\|), \end{aligned}$$

где константа $c(\gamma) > 0$ зависит от γ и не зависит от $\Phi(x)$, $\Psi(x)$, m , более того, для любого $k \geq 1$ имеют место сходимости

$$\begin{aligned} & \|D_t^2 [U_j^{m+k}(t, x) - U_j^m(t, x)], L_{2,\gamma}(R_+^2)\| \rightarrow 0, \quad j = 1, 2, \\ & \|D_t^2 D_x^2 [U_j^{m+k}(t, x) - U_j^m(t, x)], L_{2,\gamma}(R_+^2)\| \rightarrow 0, \\ & \|D_x^4 [U_j^{m+k}(t, x) - U_j^m(t, x)], L_{2,\gamma}(R_+^2)\| \rightarrow 0 \quad \text{при } m \rightarrow \infty. \end{aligned}$$

Доказательство леммы 2 повторяет рассуждения из предыдущей леммы.

Проведем оценку третьего слагаемого из (3.8).

Лемма 3. При выполнении условий теоремы имеет место оценка

$$\begin{aligned} \sum_{l=0}^4 \|D_x^l U_3^m(t, x), L_{2,\gamma}(R_+^2)\| &\leq c_1(\gamma) \|F(t, x), W_{2,\gamma}^{0,1}(R_+^2)\| + \\ &+ c_2(\gamma) \| (1+x^2)F(t, x), L_1(R) \|, L_{2,\gamma}(R_+) \|, \end{aligned} \quad (3.16)$$

где константы $c_1(\gamma)$, $c_2(\gamma) > 0$ зависят от γ и не зависят от m , $F(t, x)$, более того, для любого $k \geq 1$ имеет место сходимость

$$\sum_{l=0}^4 \|D_x^l [U_3^{m+k}(t, x) - U_3^m(t, x)], L_{2,\gamma}(R_+^2)\| \rightarrow 0, \quad m \rightarrow \infty. \quad (3.17)$$

Доказательство леммы 3. Обозначим через

$$G_\gamma(t, \xi) = e^{-\gamma t} \theta(t) G(t, \xi),$$

где $\theta(t)$ — функция Хевисайда, вектор-функция $G(t, \xi)$ определена в (3.7).

Учитывая (3.5), преобразование Фурье для свертки, получим

$$\begin{aligned} \sum_{l=0}^4 \|D_x^l U_3^m(t, x), L_{2,\gamma}(R_+^2)\| &= \sum_{l=0}^4 \| |\xi|^l V_3^m(t, \xi), L_{2,\gamma}(R_+^2) \| = \\ &= \sum_{l=0}^4 \left\| \chi^m(\xi) |\xi|^l \int_{-\infty}^{+\infty} \theta(t-s) e^{-\gamma(t-s)} \left(B_1 \frac{\sin(b_1(\xi)(t-s))}{b_1(\xi)} + \right. \right. \\ &+ B_2 \frac{\sin(b_2(\xi)(t-s))}{b_2(\xi)} + B_3 \frac{\sin(b_3(\xi)(t-s))}{b_3(\xi)} \Big) G(s, \xi) \theta(s) e^{-\gamma s} ds, L_2(R^2) \Big\| = \\ &= \sum_{l=0}^4 \left\| \chi^m(\xi) |\xi|^l \int_{-\infty}^{\infty} \theta(t) e^{-(i\eta+\gamma)t} \left(B_1 \frac{\sin(b_1(\xi)t)}{b_1(\xi)} + \right. \right. \\ &+ B_2 \frac{\sin(b_2(\xi)t)}{b_2(\xi)} + B_3 \frac{\sin(b_3(\xi)t)}{b_3(\xi)} \Big) dt \widehat{G}_\gamma(\eta, \xi), L_2(R^2) \Big\|. \end{aligned}$$

Нетрудно показать, что выполнены соотношения

$$\int_0^{+\infty} e^{-(i\eta+\gamma)t} \frac{\sin(b_j(\xi)t)}{b_j(\xi)} dt = \frac{1}{(i\eta+\gamma)^2 + (b_j(\xi))^2}, \quad j = 1, 2, 3,$$

и имеют место оценки:

$$|(i\eta+\gamma)^2 + (b_j(\xi))^2| \geq c\gamma \sqrt{|\eta|^2 + \gamma^2 + (b_j(\xi))^2}, \quad \gamma > 0, (\eta, \xi) \in R^2,$$

где $c > 0$ — константа. Тогда будем иметь

$$\begin{aligned} \sum_{l=0}^4 \|D_x^l U_3^m(t, x), L_{2,\gamma}(R_+^2)\| &\leq \sum_{j=1}^3 \sum_{l=0}^4 \|\chi^m(\xi) \frac{|\xi|^l B_j \widehat{G}_\gamma(\eta, \xi)}{(i\eta+\gamma)^2 + (b_j(\xi))^2}, L_2(R^2)\| \leq \\ &\leq \frac{c}{\gamma} \sum_{j=1}^3 \sum_{l=0}^4 \|\chi^m(\xi) \frac{|\xi|^l B_j \widehat{G}_\gamma(\eta, \xi)}{\sqrt{|\eta|^2 + \gamma^2 + (b_j(\xi))^2}}, L_2(R^2)\|. \end{aligned} \quad (3.18)$$

Учитывая (3.18), равенство Парсеваля, имеем

$$\begin{aligned} & \sum_{l=2}^4 \|D_x^l U_3^m(t, x), L_{2,\gamma}(R_+^2)\| \leq \\ & \leq \frac{c}{\gamma} \sum_{j=1}^3 \sum_{l=2}^4 \left(\|\chi^m(\xi) \frac{|\xi|^l B_j G(t, \xi)}{b_j(\xi)}, L_2(|\xi| > 1)\|, L_{2,\gamma}(R_+)\| + \right. \\ & \quad \left. + \|\chi^m(\xi) \frac{|\xi|^l B_j G(t, \xi)}{\gamma}, L_2(|\xi| \leq 1)\|, L_{2,\gamma}(R_+)\| \right). \end{aligned}$$

Пользуясь

$$|b_j(\xi)| \geq c|\xi|, \quad |\xi| > 1,$$

и (3.7), получим

$$\|D_x^2 U_3^m(t, x), L_{2,\gamma}(R_+^2)\| + \|D_x^3 U_3^m(t, \xi), L_{2,\gamma}(R_+^2)\| \leq \frac{c_4}{\gamma} \|F(t, x), L_{2,\gamma}(R_+^2)\|, \quad (3.19)$$

$$\|D_x^4 U_3^m(t, x), L_{2,\gamma}(R_+^2)\| \leq \frac{c_5}{\gamma} \|F(t, x), W_{2,\gamma}^{0,1}(R_+^2)\|. \quad (3.20)$$

Из (3.18) также будем иметь

$$\sum_{l=0}^1 \|D_x^l U_3^m(t, x), L_{2,\gamma}(R_+^2)\| \leq \frac{c}{\gamma^2} \sum_{j=1}^3 \sum_{l=0}^1 \left\| \chi^m(\xi) |\xi|^l B_j G(t, \xi), L_{2,\gamma}(R_+^2) \right\|.$$

Воспользуемся представлением (3.7) и неравенством Минковского:

$$\begin{aligned} & \sum_{l=0}^1 \|D_x^l U_3^m(t, x), L_{2,\gamma}(R_+^2)\| \leq \\ & \leq \frac{c}{\gamma^2} \sum_{j=1}^3 \sum_{l=0}^1 \left(\|\chi^m(\xi) |\xi|^l B_j \begin{pmatrix} 2 + \xi^2 & 0 & -\varepsilon_1 \\ 0 & 2 + \xi^2 & \varepsilon_2 \\ -\varepsilon_1 & \varepsilon_2 & 2 + \xi^2 \end{pmatrix} \frac{\widehat{F}(t, \xi)}{(1 + \xi^2)(2 + \xi^2)}, L_{2,\gamma}(R_+^2)\| + \right. \\ & \quad \left. + \|\chi^m(\xi) |\xi|^l B_j \begin{pmatrix} \varepsilon_1^2 & -\varepsilon_1 \varepsilon_2 & -\varepsilon_1 \\ -\varepsilon_1 \varepsilon_2 & \varepsilon_2^2 & \varepsilon_2 \\ -\varepsilon_1 & \varepsilon_2 & 1 \end{pmatrix} \frac{\widehat{F}(t, \xi)}{\xi^2(1 + \xi^2)(2 + \xi^2)}, L_{2,\gamma}(R_+^2)\| \right) \leq \\ & \leq \frac{c_5}{\gamma^2} \|F(t, x), L_{2,\gamma}(R_+^2)\| + \frac{c_6}{\gamma^2} \sum_{j=1}^3 \sum_{l=0}^1 \|\chi^m(\xi) |\xi|^l B_j \times \\ & \quad \times \frac{\widehat{g}(t, \xi)}{\xi^2(1 + \xi^2)(2 + \xi^2)}, L_{2,\gamma}(R_+ \times \{|\xi| \leq 1\})\| \leq \\ & \leq \frac{c_5}{\gamma^2} \|F(t, x), L_{2,\gamma}(R_+^2)\| + \frac{c_6}{\gamma^2} \sum_{j=1}^3 \sum_{l=0}^1 \|\chi^m(\xi) B_j \frac{\widehat{g}(t, \xi)}{\xi^{2-l}}, L_{2,\gamma}(R_+ \times \{|\xi| \leq 1\})\|, \quad (3.21) \end{aligned}$$

где

$$\widehat{g}(t, \xi) = \begin{pmatrix} \varepsilon_1^2 & -\varepsilon_1 \varepsilon_2 & -\varepsilon_1 \\ -\varepsilon_1 \varepsilon_2 & \varepsilon_2^2 & \varepsilon_2 \\ -\varepsilon_1 & \varepsilon_2 & 1 \end{pmatrix} \widehat{F}(t, \xi).$$

Поскольку выполнено условие (2.2), то

$$\int_R g(t, x) dx = 0,$$

следовательно, учитывая, что $(1+x)F(t, x) \in L_{2,\gamma}(R_+; L_1(R))$ имеет место представление

$$\begin{aligned}\widehat{g}(t, \xi) &= \frac{1}{\sqrt{2\pi}} \int_R e^{-ix\xi} g(t, x) dx = \frac{1}{\sqrt{2\pi}} \int_R \int_0^1 D_{\lambda_1}(e^{-ix\xi\lambda_1}) g(t, x) d\lambda_1 dx + \\ &+ \frac{1}{\sqrt{2\pi}} \int_R g(t, x) dx = \frac{1}{\sqrt{2\pi}} \int_R \int_0^1 e^{-ix\xi\lambda_1} (-ix\xi) g(t, x) d\lambda_1 dx.\end{aligned}\quad (3.22)$$

В силу (2.3) имеем

$$\int_R xg(t, x) dx = 0$$

и, учитывая, представление (3.22), а также $(1+x^2)F(t, x) \in L_{2,\gamma}(R_+; L_1(R))$, получим

$$\begin{aligned}\widehat{g}(t, \xi) &= \frac{1}{\sqrt{2\pi}} \int_R \int_0^1 \left[\int_0^1 D_{\lambda_2}(e^{-ix\xi\lambda_1\lambda_2}) (-ix\xi) g(t, x) d\lambda_2 + (-ix\xi) g(t, x) \right] d\lambda_1 dx = \\ &= \frac{1}{\sqrt{2\pi}} \int_R \int_0^1 \int_0^1 e^{-ix\xi\lambda_1\lambda_2} (-ix\xi)^2 \lambda_1 g(t, x) d\lambda_1 d\lambda_2 dx - \frac{i\xi}{\sqrt{2\pi}} \int_R xg(t, x) dx = \\ &= -\frac{\xi^2}{\sqrt{2\pi}} \int_0^1 \int_0^1 \int_R e^{-ix\xi\lambda_1\lambda_2} x^2 \lambda_1 g(t, x) dx d\lambda_1 d\lambda_2.\end{aligned}\quad (3.23)$$

В силу оценки (3.21), представления (3.22), имеем

$$\begin{aligned}\|D_x U_3^m(t, x), L_{2,\gamma}(R_+^2)\| &\leq \frac{c_5}{\gamma^2} \|F(t, x), L_{2,\gamma}(R_+^2)\| + \\ &+ \frac{c_7}{\gamma^2} \|\chi^m(\xi) \int_R \int_0^1 e^{-ix\xi\lambda_1} xg(t, x) d\lambda_1 dx, L_{2,\gamma}(R_+ \times \{|\xi| \leq 1\})\| \leq \\ &\leq \frac{c_5}{\gamma^2} \|F(t, x), L_{2,\gamma}(R_+^2)\| + \frac{c_8}{\gamma^2} \| \|xF(t, x), L_1(R)\|, L_{2,\gamma}(R_+)\|.\end{aligned}\quad (3.24).$$

Аналогично, учитывая (3.21) и представление (3.23), будем иметь

$$\begin{aligned}\|U_3^m(t, x), L_{2,\gamma}(R_+^2)\| &\leq \frac{c_5}{\gamma^2} \|F(t, x), L_{2,\gamma}(R_+^2)\| + \\ &+ \frac{c_9}{\gamma^2} \| \|x^2 F(t, x), L_1(R)\|, L_{2,\gamma}(R_+)\|.\end{aligned}\quad (3.25)$$

Пользуясь (3.19), (3.20), (3.24), (3.25), будем иметь требуемую оценку (3.16). Аналогично доказывается сходимость (3.17).

Лемма доказана.

Лемма 4. При выполнении условий теоремы имеет место оценка

$$\begin{aligned}\|D_t^2 U_3^m(t, x), L_{2,\gamma}(R_+^2)\| &+ \|D_t^2 D_x^2 U_3^m(t, x), L_{2,\gamma}(R_+^2)\| \leq c_1(\gamma) \|F(t, x), W_{2,\gamma}^{0,1}(R_+^2)\| + \\ &+ c_2(\gamma) \| \|(1+x^2)F(t, x), L_1(R)\|, L_{2,\gamma}(R_+)\|,\end{aligned}\quad (3.26)$$

где константы $c_1, c_2 > 0$ не зависят от m , $F(t, x)$, более того, для любого $k \geq 1$ имеет место сходимость

$$\|D_t^2 [U_3^{m+k}(t, x) - U_3^m(t, x)], L_{2,\gamma}(R_+^2)\| + \|D_t^2 D_x^2 [U_3^{m+k}(t, x) - U_3^m(t, x)], L_{2,\gamma}(R_+^2)\| \rightarrow 0 \quad (3.27)$$

при $m \rightarrow \infty$.

Доказательство леммы 4. Учитывая (3.5), представим $D_t^2 V_3(t, \xi)$ в виде

$$D_t^2 V_3(t, \xi) = G(t, \xi) + \int_0^t \left(B_1 b_1(\xi) \sin(b_1(\xi)(s-t)) + \right. \\ \left. + B_2 b_2(\xi) \sin(b_2(\xi)(s-t)) + B_3 b_3(\xi) \sin(b_3(\xi)(s-t)) \right) G(s, \xi) ds.$$

Тогда

$$\begin{aligned} & \|D_t^2 U_3^m(t, x), L_{2,\gamma}(R_+^2)\| + \|D_t^2 D_x^2 U_3^m(t, x), L_{2,\gamma}(R_+^2)\| \leq \\ & \leq \|\chi^m(\xi) D_t^2 V_3(t, \xi), L_{2,\gamma}(R_+^2)\| + \|\xi^2 \chi^m(\xi) D_t^2 V_3(t, \xi), L_{2,\gamma}(R_+^2)\| \leq \\ & \leq \|\chi^m(\xi) G(t, \xi), L_{2,\gamma}(R_+^2)\| + \|\chi^m(\xi) \xi^2 G(t, \xi), L_{2,\gamma}(R_+^2)\| + \\ & + \sum_{j=1}^3 \left\| \chi^m(\xi) \int_{-\infty}^{+\infty} \theta(t-s) e^{-\gamma(t-s)} b_j(\xi) B_j \sin(b_j(\xi)(t-s)) G(s, \xi) \theta(s) e^{-\gamma s} ds, L_2(R^2) \right\| + \\ & + \sum_{j=1}^3 \left\| \chi^m(\xi) \xi^2 \int_{-\infty}^{+\infty} \theta(t-s) e^{-\gamma(t-s)} b_j(\xi) B_j \sin(b_j(\xi)(t-s)) G(s, \xi) \theta(s) e^{-\gamma s} ds, L_2(R^2) \right\|. \end{aligned}$$

Повторяя рассуждения, как и при доказательстве леммы 3, получим требуемую оценку (3.26) и сходимость (3.27).

Лемма доказана.

Доказательство теоремы. В силу лемм 1–4 последовательность $\{U^m(t, x)\}$, где $U^m(t, x) = \sum_{j=1}^3 U_j^m(t, x)$, фундаментальна в $W_{2,\gamma}^{2,4}(R_+^2)$, более того, $\{D_t^2 D_x^2 U^m(t, x)\}$ фундаментальна в $L_{2,\gamma}(R_+^2)$. В силу полноты $W_{2,\gamma}^{2,4}(R_+^2)$, $L_{2,\gamma}(R_+^2)$, теоремы о слабой замкнутости оператора обобщенного дифференцирования, существует вектор-функция $U(t, x) \in W_{2,\gamma}^{2,4}(R_+^2)$ такая, что существует обобщенная производная $D_t^2 D_x^2 U(t, x)$ в R_+^2 , $D_t^2 D_x^2 U(t, x) \in L_{2,\gamma}(R_+^2)$, более того, справедлива оценка (2.4). Покажем, что вектор-функция $U(t, x)$ — решение задачи Коши (2.1). Введем обозначение

$$L(D_t, D_x) = \begin{pmatrix} I - D_x^2 & 0 & \varepsilon_1 \\ 0 & I - D_x^2 & -\varepsilon_2 \\ \varepsilon_1 & -\varepsilon_2 & I - D_x^2 \end{pmatrix} D_t^2 + \sigma^2 D_x^4 I,$$

тогда задача Коши (2.1) перепишется в виде

$$\begin{aligned} L(D_t, D_x)U &= F(t, x), \quad t > 0, \quad x \in R, \\ U|_{t=0} &= \Phi(x), \\ D_t U|_{t=0} &= \Psi(x). \end{aligned}$$

С одной стороны, учитывая равенство Парсеваля и определение $U^m(t, x)$, имеем

$$\begin{aligned} \|L(D_t, D_x)U^m(t, x) - F(t, x), L_{2,\gamma}(R_+^2)\| &= \|\chi^m(\xi) L(D_t, i\xi) V(t, \xi) - \widehat{F}(t, \xi), L_{2,\gamma}(R_+^2)\| = \\ &= \|\widehat{F}(t, \xi), L_2(|\xi| \leq 1/m)\|, L_{2,\gamma}(R_+^2) \rightarrow 0, \quad m \rightarrow \infty. \end{aligned}$$

С другой стороны,

$$\begin{aligned} & \|L(D_t, D_x)U^m(t, x) - L(D_t, D_x)U(t, x), L_{2,\gamma}(R_+^2)\| \leq \\ & \leq c_1 \|D_t^2 D_x^2 [U^m(t, x) - U(t, x)], L_{2,\gamma}(R_+^2)\| + c_2 \|D_t^2 [U^m(t, x) - U(t, x)], L_{2,\gamma}(R_+^2)\| + \\ & + c_3 \|D_x^4 [U^m(t, x) - U(t, x)], L_{2,\gamma}(R_+^2)\| \rightarrow 0, \quad m \rightarrow \infty. \end{aligned}$$

В силу единственности предела получим

$$L(D_t, D_x)U(t, x) = F(t, x).$$

Проверим, что для вектор-функции $U(t, x)$ выполнены начальные условия задачи Коши (2.1). В силу теоремы о следах для анизотропных соболевских пространств (см., например, [2], [12]) имеем

$$\begin{aligned} \|U(t, x)|_{t=0} - \Phi(x), L_2(R)\| &= \|U(t, x)|_{t=0} - U^m(t, x)|_{t=0}, L_2(R)\| \leq \\ &\leq c\|U(t, x) - U^m(t, x), W_{2,\gamma}^{2,4}(R_+^2)\| \rightarrow 0, \quad m \rightarrow \infty, \\ \|D_t(e^{-\gamma t}U(t, x))|_{t=0} - (\Psi(x) - \gamma\Phi(x)), L_2(R)\| &= \\ = \|D_t(e^{-\gamma t}U(t, x))|_{t=0} - D_t(e^{-\gamma t}U^m(t, x))|_{t=0}, L_2(R)\| &\leq \\ \leq c\|U(t, x) - U^m(t, x), W_{2,\gamma}^{2,4}(R_+^2)\| &\rightarrow 0, \quad m \rightarrow \infty. \end{aligned}$$

Следовательно,

$$U(t, x)|_{t=0} = \Phi(x)$$

и

$$D_t U(t, x)|_{t=0} = \Psi(x).$$

Доказательство единственности повторяет рассуждения из [2].
Теорема доказана.

Авторы выражают благодарность Г.В. Демиденко за постановку задачи и внимание к работе.

СПИСОК ЛИТЕРАТУРЫ

1. Власов В.З. Тонкостенные упругие стержни. 2-е изд., перераб. и доп. М.: Физматгиз, 1959.
2. Демиденко Г. В., Успенский С. В. Уравнения и системы, не разрешенные относительно старшей производной. Новосибирск: Научная книга, 1998.
3. Fedotov I., Volevich L. V. The Cauchy problem for hyperbolic equations not resolved with respect to the highest time derivative // Russ. J. Math. Phys. 2006. V. 13. No. 3. P. 278–292.
4. Демиденко Г. В. Условия разрешимости задачи Коши для псевдогиперболических уравнений // Сиб. мат. журн. 2015. Т. 56. No. 6. С. 1289–1303.
5. Бондарь Л.Н., Демиденко Г.В. О корректности задачи Коши для псевдогиперболических уравнений в весовых соболевских пространствах // Сиб. матем. журнал. 2023. Т. 64. No. 5. С. 895–911.
6. Герасимов С. И., Ерофеев В. И. Задачи волновой динамики элементов конструкций. Саров: ФГУП “РФЯЦ-ВНИИЭФ”, 2014.
7. Бондарь Л. Н., Демиденко Г. В., Пинтус Г. М. Задача Коши для одной псевдогиперболической системы // Ж. вычисл. матем. и матем. физ. 2020. Т. 60. No. 4. С. 626–638.
8. Bondar L. N., Demidenko G. V. Solvability of the Cauchy problem for a pseudohyperbolic system // Complex Variables and Elliptic Equations. 2021. 66:6-7, 1084–1099.
9. Бондарь Л. Н., Мингнарлов С. Б. О задаче Коши для одной системы псевдогиперболического типа // Математические заметки СВФУ. 2023. Т.30. No.4. С. 3–11.
10. Лере Ж. Гиперболические дифференциальные уравнения. М.: Наука, 1984.
11. Годунов С.К. Уравнения математической физики. М.: Наука, 1979.
12. Бесов О.В., Ильин В.П., Никольский С.М. Интегральные представления функций и теоремы вложения. Изд. 2, доп. и перераб. Изд-во: М.: Наука, Физматлит, 1996.

CONDITIONS FOR SOLVABILITY OF THE CAUCHY PROBLEM FOR ONE PSEUDOHYPERBOLIC SYSTEM

L. N. Bondar^{a,*}, S. B. Mingnarov^{a,**}

^a *Novosibirsk State University, Pirogov St., 1, Novosibirsk, 630090 Russia*

*e-mail: l.bondar@g.nsu.ru

**e-mail: s.mingnarovQg.nsu.ru

Received: 05 May, 2024

Revised: 05 May, 2024

Accepted: 05 May, 2024

Abstract. The Cauchy problem for one system, not resolved with respect to the highest derivative with respect to time, is considered. The system under study belongs to the class of pseudohyperbolic systems. The system describes transverse bending-torsional vibrations of an elastic rod. Conditions for the solvability of the Cauchy problem in Sobolev spaces and estimates of the solution are obtained.

Keywords: pseudohyperbolic system, Cauchy problem, bending-torsional oscillations, solvability conditions.

СФЕРИЧЕСКИЕ СПЛАЙН-РЕШЕНИЯ НЕОДНОРОДНОГО БИГАРМОНИЧЕСКОГО УРАВНЕНИЯ¹⁾

© 2024 г. В. Л. Васкевич^{1,2,*}¹630090 Новосибирск, пр-т Акад. Коптюга, 4, Институт математики им. С.Л. Соболева СО РАН, Россия²630090 Новосибирск, ул. Пирогова, 1, Новосибирский государственный университет, Россия

*e-mail: v.vaskevich@ngsu.ru

Поступила в редакцию 05.03.2024 г.

Переработанный вариант 05.03.2024 г.

Принята к публикации 05.05.2024 г.

На единичной сфере трехмерного пространства рассматривается неоднородное бигармоническое уравнение. Принадлежащее сферическому пространству Соболева решение этого уравнения аппроксимируется последовательностью решений этого же уравнения, но со специальными правыми частями, представляющими собой линейные комбинации сдвигов дельта-функции Дирака. Доказано, что при заданных на сфере узлах, определяющих сдвиги, специальные решения уравнения — сферические бигармонические сплайны — существуют, а соответствующие каждому из них веса являются решениями сопутствующей невырожденной системы линейных алгебраических уравнений. Установлена связь качества аппроксимации решения дифференциальной задачи сферическими бигармоническими сплайнами с задачей о скорости сходимости оптимальных весовых сферических кубатурных формул. Библ. 10.

Ключевые слова: бигармоническое уравнение, сферические пространства Соболева, экстремальные функции, сплайны.

DOI: 10.31857/S0044466924080107, EDN: YAMREM

1. ВВЕДЕНИЕ

В трехмерном пространстве с декартовыми координатами (x, y, z) сферические координаты задаются равенствами

$$\begin{aligned} z &= r \cos \theta, \\ x &= r \sin \theta \cos \varphi, \\ y &= r \sin \theta \sin \varphi. \end{aligned}$$

Угловые переменные θ и φ здесь изменяются в промежутках $0 \leq \theta \leq \pi$ и $0 \leq \varphi < 2\pi$ соответственно. Пусть на единичной сфере $S \subset \mathbb{R}^3$ определена непрерывная функция $p(\theta, \varphi)$, подчиненная следующему условию ортогональности:

$$\int p(\theta, \varphi) dS = 0. \quad (1)$$

В сферических координатах оператор Лапласа Δ , с помощью которого задается бигармоническое уравнение, представляется в виде суммы радиальной и угловой части. Угловая его часть имеет при этом вид

$$\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2}.$$

Далее эту угловую часть условимся обозначать символом Δ_S . Для любой функции $u(\theta, \varphi)$ из $C^{(4)}(S)$ определено дифференциальное выражение $\Delta_S^2 u(\theta, \varphi)$, задающее на сфере S непрерывную функцию.

Неоднородное сферическое бигармоническое уравнение [1] записывается как следующее поточечное равенство:

$$\Delta_S^2 u(\theta, \varphi) = p(\theta, \varphi). \quad (2)$$

¹⁾Работа выполнена в рамках государственного задания Института математики им. С.Л. Соболева СО РАН (проект № FWNF-2022-0008).

Дифференциальный оператор Δ_S^2 является самосопряженным в том смысле, что для любых двух функций f и g из класса $C^{(4)}(S)$ справедливо интегральное соотношение

$$\int (\Delta_S^2 f) g(\theta, \varphi) dS = \int f(\theta, \varphi) (\Delta_S^2 g) dS.$$

Взяв здесь в качестве f решение $u(\theta, \varphi)$ уравнения (2), а в качестве g тождественно единичную функцию, заключаем, что условие (1) необходимо для разрешимости сферического бигармонического уравнения (2).

В исследовании дифференциальных уравнений на сфере S важную роль играют сферические гармоники различных порядков.

Напомним, что функцию $Y = Y(\theta, \varphi)$ называют сферической гармоникой порядка k , если эта функция получается сужением на единичную сферу S некоторого гармонического и однородного по переменным (x, y, z) полинома степени k , [2]. При этом выполняется следующее дифференциальное уравнение сферических гармоник:

$$\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 Y}{\partial \varphi^2} + k(k+1)Y = 0.$$

Это означает, что любая сферическая гармоника порядка k представляет собой собственную функцию оператора $-\Delta_S$ с соответствующим собственным числом $\lambda_k = k(k+1)$, где $k = 0, 1, 2, \dots$.

Вместе с пространством $C^4(S)$ функций, имеющих на сфере S непрерывные производные до четвертого порядка включительно, будем рассматривать также сферические пространства Соболева $H^2 = H^2(S)$.

Пространство H^2 на единичной сфере S получается как пополнение по некоторой норме множества сферических полиномов. По определению, сферический полином $f = f(\theta, \varphi)$ — это линейная комбинация сферических гармоник различных порядков:

$$f(\theta, \varphi) = \sum_{k=0}^{\infty} \sum_{l=-k}^k a_{k,l} Y_{k,l}(\theta, \varphi), \quad a_{k,l} = a_{k,l}(f). \quad (3)$$

В этом разложении лишь конечное число коэффициентов $a_{k,l}$ не равны нулю, а сферические гармоники $Y_{k,l}(\theta, \varphi)$ при $k \geq 1$ и $l = -k, \dots, 0, \dots, k$ определяются соотношениями

$$\begin{aligned} Y_{k,0}(\theta, \varphi) &= \sqrt{\frac{2k+1}{2\pi}} P_k(\cos \theta); \\ Y_{k,l}(\theta, \varphi) &= \sqrt{\frac{2k+1}{2\pi}} P_k^l(\cos \theta) \cos l\varphi \quad \text{при } l = 1, \dots, k; \\ Y_{k,l}(\theta, \varphi) &= \sqrt{\frac{2k+1}{2\pi}} P_k^{|l|}(\cos \theta) \sin |l|\varphi \quad \text{при } l = -k, -k+1, \dots, -1. \end{aligned}$$

Здесь $P_k(t)$ обозначает полином Лежандра степени k , а через $P_k^l(t)$, где $l = 1, 2, \dots, k$, обозначена присоединенная функция Лежандра степени k .

Сферическая гармоника $Y_{0,0}(\theta, \varphi)$ нулевого порядка тождественно постоянна и равна $1/\sqrt{4\pi}$.

Множество $\{Y_{k,l}(\theta, \varphi) \mid l = -k, -k+1, \dots, 0, \dots, k-1, k\}$ представляет собой ортонормированный базис в пространстве сферических гармоник порядка k :

$$\int Y_{k,l}(\theta, \varphi) Y_{k,p}(\theta, \varphi) dS = \delta_l^p.$$

Любой сферический полином f имеет конечную норму, определяемую соотношением

$$\|f \mid H^2(S)\| = \left\{ \left| \frac{1}{4\pi} \int f(\theta, \varphi) dS \right|^2 + \sum_{k=1}^{\infty} k^2(k+1)^2 \sum_{l=-k}^k |a_{k,l}(f)|^2 \right\}^{\frac{1}{2}}. \quad (4)$$

Пополнение по этой норме множества всех сферических полиномов представляет собой линейное функциональное пространство, называемое далее *сферическим пространством* Соболева и обозначаемое через $H^2 = H^2(S)$.

Отметим, что норма (4) тождественно единичной функции из пространства H^2 равна единице.

Любая функция f из пространства $H^2(S)$ раскладывается в ряд (3) по базисным сферическим гармоникам, сходящийся по норме (4). Коэффициенты $a_{k,l} = a_{k,l}(f)$ в разложении (3) представимы в виде следующих поверхностных интегралов:

$$a_{0,0}(f) = \frac{1}{\sqrt{4\pi}} \int f(\theta, \varphi) dS; \quad a_{k,l}(f) = \int f(\theta, \varphi) Y_{k,l}(\theta, \varphi) dS, \quad k \geq 1.$$

Заметим, что $H^2(S)$ — это гильбертово пространство со скалярным произведением

$$(f, g)_2 = \frac{1}{4\pi} a_{0,0}(f) a_{0,0}(g) + \sum_{k=1}^{\infty} k^2 (k+1)^2 \sum_{l=-k}^k a_{k,l}(f) a_{k,l}(g).$$

Объединение сферических гармоник $Y_{k,l}(\theta, \varphi)$ по всем $k = 0, 1, \dots$ и $l = -k, -k+1, \dots, k-1, k$ задает в гильбертовом пространстве H^2 ортогональный базис.

Классы функций, близкие к сферическим пространствам Соболева, рассматривались ранее во многих работах, например, в [3,4].

Разрешимость неоднородного сферического полигармонического уравнения в пространстве H^2 подробно исследована в статье [1], где было доказано, в частности, следующее утверждение.

Теорема 1. Для любого функционала $l = l(\theta, \varphi)$ из пространства $C^*(S)$, удовлетворяющего условию $(l, 1) = 0$, задача

$$\Delta_S^2 u = l(\theta, \varphi), \quad \int u(\theta, \varphi) dS = 0, \quad (5)$$

имеет решение $u(\theta, \varphi)$, принадлежащее пространству $H^2(S)$. Других решений в пространстве $H^2(S)$ у задачи (5) нет. Это единственное в $H^2(S)$ решение задачи (5) принадлежит также пространству $C^{(1)}(S)$.

Принадлежность решения задачи пространству $H^2(S)$ автоматически влечет за собой непрерывность этого решения на сфере S . Более того, как известно [5,6], каждая функция $f(\theta, \varphi)$ из пространства $H^2(S)$, а значит, и искомое решение задачи, удовлетворяет оценке

$$\sup_{(\theta, \varphi)} |f(\theta, \varphi)| \leq A_2 \|f\|_{H^2(S)}. \quad (6)$$

Здесь A_2 — это константа вложения, определяемая соотношением

$$A_2 = \left\{ 1 + \frac{1}{4\pi} \sum_{k=1}^{\infty} \frac{2k+1}{k^2(k+1)^2} \right\}^{1/2}.$$

Постоянная A_2 представляет собой норму оператора вложения сферического пространства Соболева $H^2(S)$ в пространство $C(S)$ непрерывных на единичной сфере функций.

Известно также [2], что в $H^2(S)$ существует функция вложения, для которой в (6) достигается точное равенство. Функцией вложения $G(\theta, \varphi)$ является, например, сумма следующего абсолютно сходящегося ряда:

$$G(\theta, \varphi) = 1 + \frac{1}{4\pi} \sum_{k=1}^{\infty} \frac{2k+1}{k^2(k+1)^2} P_k(\cos \theta). \quad (7)$$

Здесь $0 \leq \theta \leq \pi$, а полином Лежандра $P_k(t)$ степени k удовлетворяет равенствам

$$\max_{-1 \leq t \leq 1} |P_k(t)| = P_k(+1) = 1.$$

Учитывая это, заключаем, что ряд (7) сходится абсолютно и равномерно по θ , где $0 \leq \theta \leq \pi$.

Функция вложения пространства H^2 в пространство $C(S)$ неединственна: для произвольного $0 \leq \theta_0 \leq \pi$ любая функция вида $G(\theta - \theta_0, \varphi)$ также удовлетворяет всем требуемым условиям.

2. СФЕРИЧЕСКИЕ СПЛАЙНЫ И СПЛАЙН-РЕШЕНИЯ

Вместе с неоднородным сферическим бигармоническим уравнением (2) будем рассматривать уравнения с более общей нежели просто непрерывная функция правой частью. Точнее, будем иметь дело с линейными дифференциальными уравнениями вида

$$\Delta_S^2 u = l(\theta, \varphi), \quad (8)$$

где правая часть $l = l(\theta, \varphi)$ представляет собой линейный и ограниченный на пространстве $C(S)$ функционал, значение которого на тождественно постоянной функции равно нулю:

$$(l(\theta, \varphi), 1) = 0. \quad (9)$$

Таким образом, правая часть l неоднородного уравнения (8) принадлежит пространству $C^*(S)$, сопряженному к $C(S)$.

Из-за принадлежности линейного функционала $l(\theta, \varphi)$ пространству $C^*(S)$ и в силу оценки (6) этот же функционал $l(\theta, \varphi)$ ограничен и на пространстве $H^2(S)$. При этом справедливо неравенство

$$\|l \mid H^{2*}(S)\| \leq A_2 \|l \mid C^*(S)\|.$$

Принадлежность функционала $l(\theta, \varphi)$ сопряженному пространству $H^{2*}(S)$ означает сходимость к нему частичных сумм ряда по сферическим гармоникам

$$\sum_{k=1}^{\infty} \sum_{l=-k}^k (l, Y_{k,l}) Y_{k,l}(\theta, \varphi)$$

по норме

$$\|l \mid H^{2*}(S)\| = \left\{ \sum_{k=1}^{\infty} \frac{1}{k^2(k+1)^2} \sum_{l=-k}^k |(l(\theta, \varphi), Y_{k,l}(\theta, \varphi))|^2 \right\}^{1/2} < \infty.$$

Решение $u(\theta, \varphi)$ уравнения (8), принадлежащее классу H^2 , представимо при этом в следующем виде:

$$u(\theta, \varphi) = \sum_{k=1}^{\infty} \frac{1}{k^2(k+1)^2} \sum_{l=-k}^k (l, Y_{k,l}) Y_{k,l}(\theta, \varphi).$$

В качестве важного примера неоднородного сферического бигармонического уравнения рассмотрим уравнение

$$\Delta_S^2 u = l(\mathbf{x} \mid \Theta^{(0)}), \quad \mathbf{x} = (x, y, z), \quad |\Theta^{(0)}| = 1,$$

со специальной правой частью $l(\mathbf{x} \mid \Theta^{(0)})$, определив действие этого линейного функционала следующим образом.

Для вектора $\mathbf{x} = (x, y, z)$ из пространства $\mathbb{R}^3$ и лежащей на сфере S точки $\Theta^{(0)}$ полагаем, что линейный функционал $l(\mathbf{x} \mid \Theta^{(0)})$, действует на пробную функцию $f(\mathbf{x})$ по формуле

$$(l(\mathbf{x} \mid \Theta^{(0)}), f(\mathbf{x})) = f(\Theta^{(0)}) - \frac{1}{4\pi} \int f(\Theta') dS_{\Theta'}. \quad (10)$$

Таким образом, действие $l(\mathbf{x} \mid \Theta^{(0)})$ на непрерывную функцию $f(x, y, z)$ представляет собой разность между значением функции $f(x, y, z)$ в точке $\Theta^{(0)}$ сферы и ее же средним значением по сфере S . В частности, на тождественно постоянной функции функционал $l(\mathbf{x} \mid \Theta^{(0)})$ обращается в нуль, а его носитель сосредоточен на сфере S .

Полагая $\Theta = \frac{1}{r}\mathbf{x}$, где $r = |\mathbf{x}|$, получаем следующее разложение функционала $l(\Theta \mid \Theta^{(0)})$ в ряд по сферическим гармоникам:

$$l(\Theta \mid \Theta^{(0)}) = \sum_{k=1}^{\infty} \sum_{l=-k}^k Y_{k,l}(\Theta^{(0)}) Y_{k,l}(\Theta) = \sum_{k=1}^{\infty} \frac{2k+1}{4\pi} P_k(\cos \omega^{(0)}). \quad (11)$$

Последнее равенство здесь справедливо по теореме сложения для сферических гармоник [8]. В равенстве (11) $\cos \omega^{(0)}$ — это скалярное произведение векторов Θ и $\Theta^{(0)}$, то есть $\omega^{(0)}$ представляет собой угол между векторами Θ и $\Theta^{(0)}$.

Ряд (11) сходится по норме пространства $H^{2*}(S)$ и, следовательно, функционал $l(\Theta | \Theta^{(0)})$ принадлежит пространству $H^{2*}(S)$.

Принадлежащее классу $H^2(S)$ решение неоднородного сферического бигармонического уравнения с функционалом $l(\Theta | \Theta^{(0)})$ в правой части представимо как следующий абсолютно сходящийся ряд:

$$\begin{aligned} G(\Theta \cdot \Theta^{(0)}) &= \sum_{k=1}^{\infty} \frac{1}{k^2(k+1)^2} \sum_{l=-k}^k Y_{k,l}(\Theta^{(0)}) Y_{k,l}(\Theta) = \\ &= \sum_{k=1}^{\infty} \frac{2k+1}{4\pi k^2(k+1)^2} P_k(\Theta \cdot \Theta^{(0)}). \end{aligned}$$

Сумма $G(\Theta \cdot \Theta^{(0)})$ этого ряда известна как функция Грина сферического бигармонического оператора Δ_S^2 . По переменной Θ она принадлежит пространству $H^2(S)$ и при этом ортогональна на S тождественно постоянной функции. Согласно теореме 1 по переменной Θ функция $G(\Theta \cdot \Theta^{(0)})$ непрерывна на S и более того принадлежит классу $C^{(1)}(S)$.

Еще один важный класс неоднородных сферических бигармонических уравнений образуют уравнения вида

$$\Delta_S^2 u = \sum_{j=1}^N c_j \delta(\Theta - \Theta^{(j)}), \quad |\Theta| = |\Theta^{(j)}| = 1. \quad (12)$$

Правая часть в (12) представляет собой линейную комбинацию сдвигов дельта-функции Дирака на заданные и принадлежащие сфере S векторы $\Theta^{(j)}$, $j = 1, \dots, N$. Необходимым условием существования у уравнения (12) решения из пространства $H^2(S)$ является обращение в нуль функционала в его правой части на любой тождественно постоянной функции, то есть равенство

$$\sum_{j=1}^N c_j = 0. \quad (13)$$

Это же равенство (13) является достаточным для существования у уравнения (12) решения из пространства $H^2(S)$, ортогонального к тому же тождественно постоянной функции на сфере S .

Если условие (13) выполнено, то справедливо равенство

$$\sum_{j=1}^N c_j \delta(\Theta - \Theta^{(j)}) = \sum_{j=1}^N c_j l(\Theta | \Theta^{(j)}),$$

в котором каждый из функционалов $l(\Theta | \Theta^{(j)})$, $j = 1, \dots, N$, задается формулой вида (10). Решение уравнения (12) из пространства $H^2(S)$ в этом случае имеет вид

$$u_N(\theta, \varphi) = u_N(\Theta) = \sum_{j=1}^N c_j G(\Theta \cdot \Theta^{(j)}), \quad \sum_{j=1}^N c_j = 0, \quad (14)$$

где $G(\Theta \cdot \Theta^{(j)})$ — это введенная выше функция Грина задачи (8), (9).

Определяемая равенством (14) функция $u_N(\Theta)$ непрерывна на сфере S и более того принадлежит классу $C^{(1)}(S)$.

Произвольную функцию $u_N(\Theta)$ вида (14) условимся называть *сферическим $H^2(S)$ -сплайном*. Входящие в представление (14) точки $\Theta^{(j)}$, $j = 1, \dots, N$, единичной сферы называем узлами сплайна $u_N(\Theta)$, а коэффициенты c_j — его весами. Нетривиальные $H^2(S)$ -сплайны существуют лишь при условии, что $N \geq 2$.

Далее нас интересует решение следующей экстремальной задачи: для известной непрерывной на единичной сфере функции $p(\theta, \varphi)$ с условием ортогональности (1) найти сферический $H^2(S)$ -сплайн $u_N(\Theta)$ с заданными узлами $\Theta^{(j)}$, $j = 1, \dots, N$, наименее уклоняющийся от решения $u(\Theta)$ дифференциальной задачи

$$\Delta_S^2 u(\theta, \varphi) = p(\theta, \varphi), \quad \int u(\theta, \varphi) dS = 0,$$

в норме пространства $H^2(S)$.

Решение этой экстремальной задачи, как будет доказано в следующем разделе, существует и единственно в классе функций, ортогональных на S тождественной единице. Обозначив это решение как $u_N^{(0)}(\theta, \varphi) = u_N^{(0)}(\Theta)$, назовем его *сплайн-решением* исходной дифференциальной задачи.

3. СУЩЕСТВОВАНИЕ И ЕДИНСТВЕННОСТЬ СФЕРИЧЕСКОГО СПЛАЙН-РЕШЕНИЯ

Установим разрешимость сформулированной в предыдущем разделе экстремальной задачи и попутно докажем для нее теорему единственности.

Теорема 2. Среди всех $H^2(S)$ -сплайнов с заданным на единичной сфере множеством узлов $\Theta^{(j)}$, $j = 1, \dots, N$, т.е. среди функций вида

$$u_N(\theta, \varphi) = u_N(\Theta) = \sum_{j=1}^N c_j G(\Theta \cdot \Theta^{(j)}), \quad \sum_{j=1}^N c_j = 0,$$

существует в точности один, отклоняющийся в $H^2(S)$ -норме от решения $u(\Theta)$ исходной краевой задачи (1), (2) на минимально возможное значение.

Доказательство. Функция $u(\Theta)$, как установлено, принадлежит пространству $H^2(S)$. Зафиксировав на единичной сфере множество узлов $\Theta^{(j)}$, $j = 1, \dots, N$, где $N \geq 2$, рассмотрим следующий числовой параметр

$$S_N = \inf \left\{ \|u(\Theta) - \sum_{j=1}^N c_j G(\Theta \cdot \Theta^{(j)})\|_{H^2(S)} \mid \sum_{j=1}^N c_j = 0 \right\}. \quad (15)$$

Ясно, что число S_N неотрицательно и конечно.

Сферический $H^2(S)$ -сплайн $u_N^{(0)}(\Theta)$ с весовым вектором $c^{(0)} = (c_1^0, \dots, c_N^0)$ отклоняется от функции $u(\Theta)$ на минимально возможное значение $H^2(S)$ -нормы тогда и только тогда, когда выполняется равенство

$$\|u(\Theta) - \sum_{j=1}^N c_j^0 G(\Theta \cdot \Theta^{(j)})\|_{H^2(S)} = S_N. \quad (16)$$

Это означает, что весовой вектор $c^{(0)} = (c_1^0, \dots, c_N^0)$ доставляет на гиперплоскости $\sum_{j=1}^N c_j = 0$ пространства $\mathbb{R}^n$ минимум квадратичной функции $\psi(c)$ переменных $c = (c_1, \dots, c_N)$, задаваемой равенством

$$\psi(c) = \|u(\Theta) - \sum_{j=1}^N c_j G(\Theta \cdot \Theta^{(j)})\|_{H^2(S)}^2.$$

Запишем $\psi(c)$ в общепринятом для квадратичных функций виде, т.е. следующим образом:

$$\psi(c) = \sum_{i,j=1}^N g_{ij} c_i c_j + 2 \sum_{j=1}^N b_j c_j + C. \quad (17)$$

С этой целью заметим, что разность $v(\Theta) = u(\Theta) - \sum_{j=1}^N c_j G(\Theta \cdot \Theta^{(j)})$ является решением следующей дифференциальной задачи:

$$\Delta_S^2 v(\Theta) = p(\Theta) - \sum_{j=1}^N c_j \delta(\Theta - \Theta^{(j)}), \quad \int v(\Theta) dS = 0.$$

Как известно [1], в этом случае справедливы равенства

$$\begin{aligned} \|v(\Theta)\|_{H^2(S)}^2 &= (p(\Theta) - \sum_{j=1}^N c_j \delta(\Theta - \Theta^{(j)}), v(\Theta)) = \\ &= \int p(\Theta) v(\Theta) dS - \sum_{j=1}^N c_j v(\Theta^{(j)}). \end{aligned}$$

Запишем это равенство в эквивалентной форме, подставив в правую часть разложение $v(\Theta) = u(\Theta) - \sum_{j=1}^N c_j G(\Theta \cdot \Theta^{(j)})$. В результате получим

$$\psi(c) = \sum_{i,j=1}^N G(\Theta^{(i)} \cdot \Theta^{(j)}) c_i c_j - 2 \sum_{j=1}^N c_j u(\Theta^{(j)}) + \int p(\Theta) u(\Theta) dS.$$

Это означает, что в стандартном представлении (17) достаточно взять

$$g_{ij} = G(\Theta^{(i)} \cdot \Theta^{(j)}), \quad b_j = -u(\Theta^{(j)}), \quad C = \int p(\Theta) u(\Theta) dS.$$

Исходная экстремальная задача эквивалентна, таким образом, стандартной задаче на поиск условного экстремума неотрицательной квадратичной функции $\psi(c)$, решаемой методом неопределенных множителей Лагранжа.

Приравняв нулю частные производные от вспомогательной функции $\psi_1(c, \lambda) = \psi(c) + 2\lambda \sum_{j=1}^N c_j$, получим сопутствующую задаче на условный экстремум систему линейных уравнений. В покомпонентном виде эта система записывается следующим образом:

$$\begin{aligned} \sum_{j=1}^N G(\Theta^{(i)} \cdot \Theta^{(j)}) c_j + \lambda &= u(\Theta^{(i)}), \\ i &= 1, \dots, N; \\ \sum_{j=1}^N c_j &= 0. \end{aligned} \quad (18)$$

Пусть G — это квадратная $N \times N$ —матрица с элементами g_{ij} , $\mathbf{u}$ — это вектор-столбец значений функции $u(\Theta)$ в узлах сплайна, а P — это вектор-строка длины N с единичными элементами. С учетом этих обозначений запишем систему (17) в матричном виде:

$$\begin{pmatrix} G & P^* \\ P & 0 \end{pmatrix} \begin{pmatrix} c \\ \lambda \end{pmatrix} = \begin{pmatrix} \mathbf{u} \\ 0 \end{pmatrix}.$$

Матрица Q этой сопутствующей системы линейных уравнений задается равенством

$$Q = \begin{pmatrix} G & P^* \\ P & 0 \end{pmatrix}. \quad (19)$$

Это симметричная и, как доказано, к примеру, в [7], положительно определенная матрица. В частности, система (18) не вырождена.

Таким образом, решение $(c^{(0)}, \lambda^{(0)})$ системы (18), сопутствующей задаче на условный экстремум, существует и единственно.

По определению, точка $(c^{(0)}, \lambda^{(0)})$ является стационарной для функции $\psi_1(c, \lambda) = \psi(c) + 2\lambda \sum_{j=1}^N c_j$. Из теории условного экстремума известно, что соответствующая компонента $c^{(0)}$ доставляет квадратичной функции $\psi(c)$ минимум на гиперплоскости $\sum_{j=1}^N c_j = 0$, если на этой гиперплоскости положительно определена квадратичная форма

$$\Phi(c) = \sum_{i,j=1}^N \frac{\partial^2 \psi_1}{\partial c_i \partial c_j} c_i c_j = 2 \sum_{i,j=1}^N g_{ij} c_i c_j.$$

Установим, что это условие действительно выполняется.

Заметим, что при любом ненулевом $c = (c_1, \dots, c_N)$ с условием $\sum_{j=1}^N c_j = 0$ дифференциальная задача

$$\Delta_S^2 u_N = \sum_{j=1}^N c_j \delta(\Theta - \Theta^{(j)}), \quad \int u_N(\Theta) dS = 0$$

имеет в $H^2(S)$ единственное решение. Это решение — нетривиальный сферический сплайн, задаваемый равенством

$$u_N(\Theta) = \sum_{j=1}^N c_j G(\Theta \cdot \Theta^{(j)}).$$

При этом справедливы соотношения

$$\sum_{i,j=1}^N g_{ij} c_i c_j = \sum_{i,j=1}^N G(\Theta^{(i)} \cdot \Theta^{(j)}) c_i c_j =$$

$$= \left(\sum_{i=1}^N c_i \delta(\Theta - \Theta^{(i)}), u_N(\Theta) \right) = \|u_N\|_{H^{2*}(S)}^2 > 0.$$

Таким образом, квадратичная форма

$$\Phi(c) = \sum_{i,j=1}^N \frac{\partial^2 \Psi_1}{\partial c_i \partial c_j} c_i c_j = 2 \sum_{i,j=1}^N g_{ij} c_i c_j$$

положительно определена.

Следовательно, сферический сплайн со свойством минимальности (16) существует. Это означает, что сплайн-решение рассматриваемой дифференциальной задачи также существует. Установим его единственность.

Пусть существуют два разных сферических сплайна $u_N^{(1)}(\Theta)$ и $u_N^{(2)}(\Theta)$ с одинаковыми узлами $\Theta^{(j)}$, $j = 1, \dots, N$, и свойством минимальности (16). Тогда разности $v_N^{(1)}(\Theta) = u(\Theta) - u_N^{(1)}(\Theta)$ и $v_N^{(2)}(\Theta) = u(\Theta) - u_N^{(2)}(\Theta)$ имеют в $H^2(S)$ одну и ту же минимальную норму:

$$\|v_N^{(1)}\|_{H^2(S)} = \|v_N^{(2)}\|_{H^2(S)} = S_N.$$

В гильбертовом пространстве $H^2(S)$ справедливо известное тождество параллелограмма

$$\left\| \frac{v_N^{(1)} + v_N^{(2)}}{2} \right\|_{H^2}^2 + \left\| \frac{v_N^{(1)} - v_N^{(2)}}{2} \right\|_{H^2}^2 = \frac{\|v_N^{(1)}\|_{H^2}^2}{2} + \frac{\|v_N^{(2)}\|_{H^2}^2}{2}.$$

Воспользовавшись им, получим строгое неравенство

$$\left\| \frac{v_N^{(1)} + v_N^{(2)}}{2} \right\|_{H^2(S)}^2 = S_N^2 - \left\| \frac{u_N^{(1)} - u_N^{(2)}}{2} \right\|_{H^2(S)}^2 < S_N^2. \quad (19)$$

Последнее строгое неравенство справедливо в силу предположения, что функции $u_N^{(1)}(\Theta)$ и $u_N^{(2)}(\Theta)$ различны.

Полусумма $\frac{u_N^{(1)} + u_N^{(2)}}{2}$ представляет собой некоторый сферический сплайн с тем же самым набором узлов, что и у сплайнов $u_N^{(1)}(\Theta)$ и $u_N^{(2)}(\Theta)$, т.е.

$$\frac{u_N^{(1)} + u_N^{(2)}}{2} = \sum_{j=1}^N c_j^{(3)} G(\Theta \cdot \Theta^{(j)}) = u_N^{(3)}(\Theta).$$

Учитывая это, получаем равенство

$$\left\| \frac{v_N^{(1)} + v_N^{(2)}}{2} \right\|_{H^2(S)}^2 = \|u(\Theta) - u_N^{(3)}(\Theta)\|_{H^2(S)}^2.$$

Таким образом и в силу (19), норма $\|u(\Theta) - u_N^{(3)}(\Theta)\|_{H^2(S)}$ должна быть строго меньше величины S_N . Но это противоречит определению S_N как точной нижней грани равенством (15). Следовательно, функции $u_N^{(1)}(\Theta)$ и $u_N^{(2)}(\Theta)$ обязаны совпадать.

Перепишем первые N уравнений системы (18) в виде равенств

$$\lambda = u(\Theta^{(j)}) - \sum_{i=1}^N G(\Theta^{(j)} \cdot \Theta^{(i)}) c_i, \quad j = 1, \dots, N;$$

или, что в силу определения сплайна $u_N^{(0)}(\Theta)$ то же самое:

$$\lambda = u(\Theta^{(j)}) - u_N^{(0)}(\Theta^{(j)}), \quad j = 1, \dots, N.$$

Таким образом, получаем

Следствие 1. Сферический $H^2(S)$ -сплайн $u_N^{(0)}(\Theta)$ с весовым вектором $c^{(0)} = (c_1^0, \dots, c_N^0)$ в каждом из своих узлов $\Theta^{(j)}$ отклоняется от значения $u(\Theta^{(j)})$ на одно и то же значение. Это значение равно нулю в том и только том случае, если сплайн $u_N^{(0)}(\Theta)$ является интерполяционным.

Задавшись на сфере S множеством узлов $\Theta^{(j)}$, $j = 1, \dots, N$, рассмотрим соответствующее им множество сферических весовых кубатурных формул вида

$$\int p(\theta, \varphi) f(\theta, \varphi) dS \approx \sum_{j=1}^N c_j f(\Theta^{(j)}). \quad (20)$$

При этом предположим, что выполнены равенства

$$\sum_{j=1}^N c_j = \int p(\theta, \varphi) dS = 0. \quad (21)$$

Функционал погрешности $l_N(\Theta)$ весовой формулы (20) задается следующим равенством [2,9,10]:

$$(l_N, f) = \int p(\theta, \varphi) f(\theta, \varphi) dS - \sum_{j=1}^N c_j f(\Theta^{(j)}) \quad \forall f \in C(S).$$

Этот функционал линеен и ограничен на пространстве $C(S)$ и в силу условия (21) точен на любой тождественно постоянной функции: $(l_N, 1) = 0$.

В частности, в пространстве $H^2(S)$ дифференциальная задача

$$\Delta_S^2 v = l_N(\Theta), \quad \int v(\Theta) dS = 0,$$

как уже установлено, имеет единственное решение $v_N(\Theta)$. Эта функция $v_N(\Theta)$ представляет собой разность между решением $u(\Theta)$ исходной задачи (1), (2) и сферическим $H^2(S)$ -сплайном $u_N(\Theta)$, имеющим в качестве весов компоненты вектора $c = (c_1, \dots, c_N)$, т.е. веса рассматриваемой кубатурной формулы. При этом справедливы равенства

$$\|u(\Theta) - u_N(\Theta) | H^2(S)\|^2 = (l_N(\Theta), v_N(\Theta)) = \|l_N(\Theta) | H^{2*}(S)\|^2. \quad (22)$$

Функция $v_N(\Theta) = u(\Theta) - u_N(\Theta)$ со свойством (22) является экстремальной для функционала $l_N(\Theta)$, [9, 10].

Таким образом, сферический $H^2(S)$ -сплайн $u_N^{(0)}(\Theta)$ с весовым вектором $c^{(0)} = (c_1^0, \dots, c_N^0)$ отклоняется от функции $u(\Theta)$ на минимально возможное значение $H^2(S)$ -нормы тогда и только тогда, когда весовой функционал погрешности

$$(l_N^{(0)}, f) = \int p(\theta, \varphi) f(\theta, \varphi) dS - \sum_{j=1}^N c_j^0 f(\Theta^{(j)}), \quad \forall f \in C(S),$$

имеет в сопряженном пространстве $H^{2*}(S)$ минимально возможную норму. Иными словами, тогда и только тогда, когда сферическая кубатурная формула

$$\int p(\theta, \varphi) f(\theta, \varphi) dS \approx \sum_{j=1}^N c_j^0 f(\Theta^{(j)})$$

с весами $c_1^0, \dots, c_N^0$ оптимальна в пространстве $H^2(S)$.

Таким образом, исследование сходимости сплайн-решений к решению дифференциальной задачи при неограниченном увеличении узлов сплайна тесно связано с установлением сходимости к нулю последовательности норм функционалов погрешности оптимальных кубатурных формул, [7].

СПИСОК ЛИТЕРАТУРЫ

1. Васкевич В.Л. Полигармоническое уравнение на сфере трехмерного пространства // Матем. заметки СВФУ. 2022. Т. 29. № 3. С. 22–30.
2. Соболев С.Л. Введение в теорию кубатурных формул. М.: Наука, 1974. 808 с.
3. Михлин С.Г. Многомерные сингулярные интегралы и интегральные уравнения. М.: ГИФМЛ, 1962. 256 с.
4. Игнатов М.И., Певный А.Б. Натуральные сплайны многих переменных. Л.: Наука, 1991. 125 с.

5. Васкевич В.Л. Константы и функции вложения пространств соболевского типа на единичной сфере // Докл. АН. 2010. Т. 433. № 4. С. 441–446.
6. Васкевич В.Л. Константы вложения периодических пространств Соболева дробного порядка // Сиб. матем. ж. 2008. Т. 49. № 5. С. 1019–1027.
7. Васкевич В.Л. Сферические кубатурные формулы в пространствах Соболева // Сиб. матем. ж. 2017. Т. 58. № 3. С. 530–542.
8. Muller C. Spherical harmonics. Berlin: Springer-Verlag. 1966. 46 p.
9. Соболев С.Л., Васкевич В.Л. Кубатурные формулы. Новосибирск: Изд-во Ин-та математики, 1996. 484 с.
10. Sobolev S. L., Vaskevich V. L. *The Theory of Cubature Formulas*. Dordrecht: Kluwer Academic Publishers. 1997. XXII+416 pp.

SPHERICAL SPLINE SOLUTIONS OF THE INHOMOGENEOUS BIHARMONIC EQUATION

V. L. Vaskevich^{a,b,*}

^a Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Acad. Koptyug Ave., 4, Novosibirsk, 630090 Russia

^b Novosibirsk State University, Pirogov St., 1, Novosibirsk, 630090 Russia

*e-mail: v.vaskevich@g.nsu.ru

Received 05 March, 2024

Revised 05 March, 2024

Accepted 05 May, 2024

Abstract. An inhomogeneous biharmonic equation is considered on the unit sphere in three-dimensional space. The solution of this equation, belonging to the Sobolev space on the sphere, is approximated by a sequence of solutions of the same equation but with specific right-hand sides, represented as linear combinations of shifts of the Dirac delta function. It is proven that, given specified nodes on the sphere determining the shifts, special solutions of the equation — spherical biharmonic splines — exist, and the weights corresponding to each are solutions of an associated non-degenerate system of linear algebraic equations. The connection between the approximation quality of the differential problem solution by spherical biharmonic splines and the problem of the convergence rate of optimal weighted spherical cubature formulas is established.

Keywords: biharmonic equation, Sobolev spaces on spheres, extremal functions, splines.

ЭНЕРГЕТИЧЕСКИЕ ОЦЕНКИ ДЛЯ ОДНОГО КЛАССА ПСЕВДОГИПЕРБОЛИЧЕСКИХ ОПЕРАТОРОВ С ПЕРЕМЕННЫМИ КОЭФФИЦИЕНТАМИ¹⁾

© 2024 г. Г. В. Демиденко^{1,*}¹ 630090 Новосибирск, ул. Пирогова, 2, Новосибирский государственный университет, Россия

*e-mail: demidenk@math.nsc.ru

Поступила в редакцию 10.03.2024 г.

Переработанный вариант 10.03.2024 г.

Принята к публикации 05.05.2024 г.

Рассмотрен класс строго псевдогиперболических операторов четвертого порядка с переменными коэффициентами. При некоторых условиях на коэффициенты установлены энергетические оценки. Из этих оценок вытекает единственность решения задачи Коши, а также априорные оценки. Библ. 14.

Ключевые слова: уравнения, не разрешенные относительно старшей производной, псевдогиперболические операторы, весовые соболевские пространства, энергетические оценки.

DOI: 10.31857/S0044466924080111, EDN: YAGPPA

1. ВВЕДЕНИЕ

В монографии [1] была введена некоторая классификация линейных уравнений с частными производными следующего вида:

$$L(D_t, D_x)u \equiv L_0(D_x)D_t^l u + \sum_{k=0}^{l-1} L_{l-k}(D_x)D_t^k u = f(t, x), \quad t > 0, x \in \mathbb{R}^n, \quad (1.1)$$

где $L_0(D_x)$ — квазиэллиптический оператор, и для них изучен широкий класс краевых задач. В литературе такие уравнения часто называют *уравнениями соболевского типа*, поскольку исследования С. Л. Соболева [2] были первыми глубокими исследованиями дифференциальных уравнений, не разрешенных относительно старшей производной. В настоящее время имеется большое число работ, посвященных изучению различных задач для уравнений вида (1.1) (см., например, монографии [1], [3] и имеющуюся там библиографию). Однако для класса *псевдогиперболических уравнений*, введенного в [1], теория краевых задач является мало изученной, в частности, по задаче Коши для таких уравнений с переменными коэффициентами в литературе пока нет результатов.

Напомним определение псевдогиперболических операторов [1] в случае постоянных коэффициентов

$$L(D_t, D_x) = L_0(D_x)D_t^l + \sum_{k=0}^{l-1} L_{l-k}(D_x)D_t^k.$$

Условие 1. Будем предполагать, что символ $L(i\eta, i\xi)$ оператора $L(D_t, D_x)$ однороден относительно некоторого вектора $(\alpha_0, \alpha_1, \dots, \alpha_n)$, $\alpha_0 > 0$, $1/\alpha_j \in \mathbb{N}$, $j = 1, \dots, n$, т. е.

$$L(c^{\alpha_0} i\eta, c^{\alpha_1} i\xi_1, \dots, c^{\alpha_n} i\xi_n) = cL(i\eta, i\xi_1, \dots, i\xi_n), \quad c > 0.$$

Условие 2. Оператор $L_0(D_x)$ является квазиэллиптическим.

Условие 3. Уравнение

$$(i\eta)^l + \sum_{k=0}^{l-1} \frac{L_{l-k}(i\xi)}{L_0(i\xi)} (i\eta)^k = 0, \quad \xi \in \mathbb{R}^n \setminus \{0\}, \quad (1.2)$$

имеет только вещественные корни $\eta_1(\xi), \dots, \eta_l(\xi)$.

¹⁾ Работа выполнена при финансовой поддержке РФФИ (код проекта 24-21-00370), <https://rscf.ru/project/24-21-00370/>.

Определение. Дифференциальный оператор $L(D_t, D_x)$ называется *псевдогиперболическим*, если для его символа $L(i\eta, i\xi)$ выполнены условия 1–3. Если корни уравнения (1.2) являются различными, то оператор $L(D_t, D_x)$ будем называть *строго псевдогиперболическим* и соответствующее дифференциальное уравнение (1.1) будем называть *строго псевдогиперболическим уравнением*.

Первые теоремы о разрешимости задачи Коши для строго псевдогиперболических уравнений с постоянными коэффициентами были доказаны в [1, гл. 2], [4]. В работе [5] была изучена задача Коши для этого класса уравнений, содержащих младшие члены. Следует отметить, что для задачи Коши для псевдогиперболических уравнений имеется ряд особенностей. А именно, для ее разрешимости в соболевских пространствах $W_2^{l,r}((0, T) \times \mathbb{R}^n)$ правая часть $f(t, x)$ уравнения должна иметь дополнительную гладкость и в зависимости от порядков дифференциальных операторов и размерности n должна быть ортогональна некоторым мономам x^β (см. также [6]–[8]).

В данной работе мы начинаем изучение строго псевдогиперболических операторов с переменными коэффициентами. Мы будем рассматривать класс операторов четвертого порядка

$$\mathcal{L}(x; D_t, D_x) = \mathcal{L}^1(D_t, D_x) + \mathcal{L}_a^2(x; D_t, D_x), \quad t \in \mathbb{R}, \quad x \in \mathbb{R}^n, \quad (1.3)$$

где $\mathcal{L}^1(D_t, D_x)$ — однородный строго псевдогиперболический оператор с постоянными вещественными коэффициентами имеет вид

$$\begin{aligned} \mathcal{L}^1(D_t, D_x) &= L_0^1(D_x)D_t^2 + L_1^1(D_x)D_t + L_2^1(D_x), \\ L_0^1(D_x) &= \sum_{|\beta|=2} a_\beta^0 D_x^\beta, \quad L_1^1(D_x) = \sum_{|\beta|=3} a_\beta^1 D_x^\beta, \quad L_2^1(D_x) = \sum_{|\beta|=4} a_\beta^2 D_x^\beta, \end{aligned} \quad (1.4)$$

при этом $L_0^1(D_x)$ — эллиптический оператор. Будем предполагать, что выполнена оценка

$$q_2 |\xi|^2 \geq L_0^1(i\xi) \equiv - \sum_{|\beta|=2} a_\beta^0 \xi^\beta \geq q_1 |\xi|^2, \quad \xi \in \mathbb{R}^n, \quad (1.5)$$

где $q_2 \geq q_1 > 0$ — постоянные.

Второй дифференциальный оператор в (1.3) с вещественнозначными переменными коэффициентами $\mathcal{L}_a^2(x; D_t, D_x)$ имеет вид

$$\mathcal{L}_a^2(x; D_t, D_x) = L_{a,0}^2(x; D_x)D_t^2 + L_1^2(x; D_x)D_t + L_2^2(x; D_x), \quad (1.6)$$

где

$$L_{a,0}^2(x; D_x) = \sum_{|\beta|=2} \alpha_\beta^0(x) D_x^\beta + aI, \quad L_1^2(x; D_x) = \sum_{|\beta|=3} \alpha_\beta^1(x) D_x^\beta, \quad L_2^2(x; D_x) = \sum_{|\beta|=4} \alpha_\beta^2(x) D_x^\beta,$$

при этом $\alpha_\beta^k(x) \in C_0^\infty(\mathbb{R}^n)$, $k = 0, 1, 2$, $a > 0$ — константа.

Оператор $\mathcal{L}_a^2(x; D_t, D_x)$ можно рассматривать, как возмущение псевдогиперболического оператора $\mathcal{L}^1(D_t, D_x)$.

В качестве примера строго псевдогиперболического уравнения с младшими членами можно рассмотреть многомерные аналоги уравнения Власова [9, 10] и уравнения Рэлея–Бишопа [11], [12], [13]

$$(a_0 I - a_1 \Delta) D_t^2 u + a_2 \Delta^2 u + \sum_{|\beta| \leq 3} a_\beta(x) D_x^\beta u = f(t, x),$$

где $a_0, a_1, a_2 > 0$.

Наша цель — получение энергетических оценок для оператора $\mathcal{L}(x; D_t, D_x)$.

В дальнейшем через $W_{2,\gamma}^{2,4}(\mathbb{R}^{n+1})$, $\gamma > 0$, будем обозначать соболевское пространство с экспоненциальным весом $e^{-\gamma t}$, т. е. функция $u(t, x) \in W_{2,\gamma}^{2,4}(\mathbb{R}^{n+1})$, если функция $u_\gamma(t, x) = e^{-\gamma t} u(t, x) \in W_2^{2,4}(\mathbb{R}^{n+1})$. Символом $\hat{u}_\gamma(\eta, \xi)$ будем обозначать преобразование Фурье функции $u_\gamma(t, x) \in L_2(\mathbb{R}^{n+1})$.

В работе будет доказана следующая

Теорема. Существует $\gamma_0 > 0$ такое, что если коэффициенты $\alpha_\beta^k(x)$ оператора (1.6) вместе со своими производными до третьего порядка включительно достаточно малы, то для любой функции $u(t, x) \in W_{2,\gamma}^{2,4}(\mathbb{R}^{n+1})$ при $\gamma > \gamma_0$ такой, что

$$D_t^2 D_x^\beta u(t, x) \in L_{2,\gamma}(\mathbb{R}^{n+1}), \quad |\beta| = 2,$$

имеет место оценка

$$\gamma \|(|\xi|^2 + a)(|\eta| + \gamma + |\xi|) \widehat{u}_\gamma(\eta, \xi), L_2(\mathbb{R}^{n+1})\| \leq c \|\mathcal{L}(x; D_t, D_x)u(t, x), L_{2,\gamma}(\mathbb{R}^{n+1})\| \quad (1.7)$$

с константой $c > 0$, не зависящей от $u(t, x)$.

Оценка (1.7) является аналогом энергетического неравенства для строго гиперболических операторов [14], [15].

Отметим, что полученные энергетические оценки можно использовать для изучения корректности задачи Коши для строго псевдогиперболических уравнений

$$\mathcal{L}(x; D_t, D_x)u = f(t, x), \quad t > 0, x \in \mathbb{R}^n,$$

$$u|_{t=0} = \varphi_1(x), \quad D_t u|_{t=0} = \varphi_2(x),$$

в весовом соболевском пространстве $W_{2,\gamma}^{2,4}(\mathbb{R}_+^{n+1})$.

2. ЭНЕРГЕТИЧЕСКИЕ ОЦЕНКИ ДЛЯ ОПЕРАТОРОВ С ПОСТОЯННЫМИ КОЭФФИЦИЕНТАМИ

Из общего определения псевдогиперболических операторов [1] вытекает, что оператор $\mathcal{L}^1(D_t, D_x)$ является псевдогиперболическим, если уравнение

$$L_0^1(i\xi)(i\eta)^2 + L_1^1(i\xi)i\eta + L_2^1(i\xi) = 0, \quad \xi \in \mathbb{R}^n \setminus \{0\},$$

имеет только вещественные корни $\eta_1(\xi), \eta_2(\xi)$, и если корни различные, то этот оператор — строго псевдогиперболический. Отметим, что для оператора (1.4) $\alpha_0 = \alpha_1 = \dots = \alpha_n = 1/4$.

В силу определения операторов $L_k^1(D_x)$, $k = 0, 1, 2$, дифференциальный оператор $\mathcal{L}^1(D_t, D_x)$ является строго псевдогиперболическим тогда и только тогда, когда выполняется неравенство

$$d(\xi) = \left(\sum_{|\beta|=3} a_\beta^1 \xi^\beta \right)^2 - 4 \left(\sum_{|\beta|=2} a_\beta^0 \xi^\beta \right) \left(\sum_{|\beta|=4} a_\beta^2 \xi^\beta \right) > 0, \quad \xi \in \mathbb{R}^n \setminus \{0\}. \quad (2.1)$$

Многочлен $d(\xi)$ является однородным шестой степени, поэтому существуют постоянные $p_2 \geq p_1 > 0$ такие, что

$$p_2 |\xi|^6 \geq d(\xi) \geq p_1 |\xi|^6, \quad \xi \in \mathbb{R}^n. \quad (2.2)$$

В монографии [1] при получении энергетических оценок для строго псевдогиперболических операторов (1.1) с постоянными коэффициентами и квазиоднородными символами использовалась схема Лере [15], предложенная для изучения корректности задачи Коши для строго гиперболических уравнений. В частности, в [1, гл.2] была получена оценка

$$M(i\eta + \gamma, i\xi) \equiv -\operatorname{Im} \left(L(i\eta + \gamma, i\xi) \overline{D_\eta L(i\eta + \gamma, i\xi)} \right) \geq q \gamma \langle \xi \rangle^{2(1-l_{\alpha_0})} \left(|i\eta + \gamma| + \langle \xi \rangle^{\alpha_0} \right)^{2l-2},$$

$$\langle \xi \rangle^2 = \sum_{j=1}^n \xi_j^{2/\alpha_j}, \quad \gamma > 0, \quad (\eta, \xi) \in \mathbb{R}^{n+1},$$

где $q = \operatorname{const} > 0$, $L(i\eta, i\xi)$ — символ строго псевдогиперболического оператора $L(D_t, D_x)$, который однороден относительно некоторого вектора $(\alpha_0, \alpha_1, \dots, \alpha_n)$, $\alpha_0 > 0$, $1/\alpha_j \in \mathbb{N}$, $j = 1, \dots, n$.

Выпишем аналогичный многочлен $M(i\eta + \gamma, i\xi)$ для оператора $\mathcal{L}^1(D_t, D_x)$. Поскольку

$$\mathcal{L}^1(i\eta + \gamma, i\xi) = L_0^1(i\xi)(i\eta + \gamma)^2 + L_1^1(i\xi)(i\eta + \gamma) + L_2^1(i\xi),$$

$$D_\eta \mathcal{L}^1(i\eta + \gamma, i\xi) = 2iL_0^1(i\xi)(i\eta + \gamma) + iL_1^1(i\xi),$$

то

$$M(i\eta + \gamma, i\xi) = -\operatorname{Im} \left(\left(L_0^1(i\xi)(-\eta^2 + 2i\eta\gamma + \gamma^2) + L_1^1(i\xi)(i\eta + \gamma) + L_2^1(i\xi) \right) \times \right. \\ \left. \times \overline{\left(2iL_0^1(i\xi)(i\eta + \gamma) + iL_1^1(i\xi) \right)} \right), \quad \gamma \geq 0, \quad (\eta, \xi) \in \mathbb{R}^{n+1}. \quad (2.3)$$

Проводя вычисления, получим

$$\begin{aligned} M(i\eta + \gamma, i\xi) &= \gamma \left(\left(\sum_{|\beta|=3} a_{\beta}^1 \xi^{\beta} \right)^2 + 2 \left(\sum_{|\beta|=2} a_{\beta}^0 \xi^{\beta} \right)^2 (\eta^2 + \gamma^2) - \right. \\ &\quad \left. - 2 \left(\sum_{|\beta|=2} a_{\beta}^0 \xi^{\beta} \right) \left(\sum_{|\beta|=4} a_{\beta}^2 \xi^{\beta} \right) + 2 \left(\sum_{|\beta|=2} a_{\beta}^0 \xi^{\beta} \right) \left(\sum_{|\beta|=3} a_{\beta}^1 \xi^{\beta} \right) \eta \right). \end{aligned}$$

Учитывая обозначение (2.1), это можно переписать в следующем виде:

$$\begin{aligned} M(i\eta + \gamma, i\xi) &= 2\gamma \left(\left(\eta \sum_{|\beta|=2} a_{\beta}^0 \xi^{\beta} + \frac{1}{2} \sum_{|\beta|=3} a_{\beta}^1 \xi^{\beta} \right)^2 + \right. \\ &\quad \left. + \left[\frac{1}{4} \left(\sum_{|\beta|=3} a_{\beta}^1 \xi^{\beta} \right)^2 - \left(\sum_{|\beta|=2} a_{\beta}^0 \xi^{\beta} \right) \left(\sum_{|\beta|=4} a_{\beta}^2 \xi^{\beta} \right) \right] + \gamma^2 \left(\sum_{|\beta|=2} a_{\beta}^0 \xi^{\beta} \right)^2 \right) = \\ &= 2\gamma \left(\left(\eta L_0^1(\xi) + \frac{1}{2} \sum_{|\beta|=3} a_{\beta}^1 \xi^{\beta} \right)^2 + \frac{1}{4} d(\xi) + \gamma^2 \left(L_0^1(\xi) \right)^2 \right) \end{aligned}$$

или

$$M(i\eta + \gamma, i\xi) = 2\gamma (L_0^1(\xi))^2 N(\eta, \gamma, \xi),$$

где

$$N(\eta, \gamma, \xi) = \left(\eta + \frac{1}{2} (L_0^1(\xi))^{-1} \left(\sum_{|\beta|=3} a_{\beta}^1 \xi^{\beta} \right) \right)^2 + \frac{1}{4} (L_0^1(\xi))^{-2} d(\xi) + \gamma^2 \quad (2.4)$$

есть неотрицательная однородная функция второй степени относительно переменных $\gamma \geq 0$, $(\eta, \xi) \in \mathbb{R}^{n+1}$. В силу неравенств (1.5), (2.2), очевидно, что $N(\eta, \gamma, \xi) = 0$ тогда и только тогда, когда $\gamma = \eta = |\xi| = 0$. Следовательно, существуют постоянные $r_2 \geq r_1 > 0$ такие, что выполняется оценка

$$r_2(\eta^2 + |\xi|^2 + \gamma^2) \geq N(\eta, \gamma, \xi) \geq r_1(\eta^2 + |\xi|^2 + \gamma^2). \quad (2.5)$$

Поэтому, учитывая (1.5), получим неравенство

$$2q_2^2 r_2 \gamma |\xi|^4 (\eta^2 + |\xi|^2 + \gamma^2) \geq M(i\eta + \gamma, i\xi) \geq 2q_1^2 r_1 \gamma |\xi|^4 (\eta^2 + |\xi|^2 + \gamma^2), \quad (2.6)$$

$$\gamma > 0, \quad (\eta, \xi) \in \mathbb{R}^{n+1}.$$

Из проведенных рассуждений вытекает следующее утверждение.

Лемма 1. Оператор $\mathcal{L}^1(D_t, D_x)$ является строго псевдодогиперболическим тогда и только тогда, когда выполняется неравенство (2.6).

Определим дифференциальный оператор

$$\mathcal{L}_1^1(D_t, D_x) = 2iL_0^1(D_x)D_t + iL_1^1(D_x) \quad (2.7)$$

и рассмотрим форму

$$\mathcal{M}^1 u = -\operatorname{Im} \int_{\mathbb{R}^{n+1}} e^{-\gamma t} \mathcal{L}^1(D_t, D_x) u(t, x) \overline{\left(e^{-\gamma t} \mathcal{L}_1^1(D_t, D_x) u(t, x) \right)} dz, \quad z = (t, x). \quad (2.8)$$

Учитывая (2.3) и (2.4), для любой функции $u(t, x) \in C_0^\infty(\mathbb{R}^{n+1})$ нетрудно получить двухстороннюю оценку

$$\begin{aligned} &2q_2^2 r_2 \gamma \int_{\mathbb{R}^{n+1}} |\xi|^4 (\eta^2 + |\xi|^2 + \gamma^2) |\widehat{u}_\gamma(\eta, \xi)|^2 d\xi \geq \mathcal{M}^1 u \geq \\ &\geq 2q_1^2 r_1 \gamma \int_{\mathbb{R}^{n+1}} |\xi|^4 (\eta^2 + |\xi|^2 + \gamma^2) |\widehat{u}_\gamma(\eta, \xi)|^2 d\xi, \quad \xi = (\eta, \xi). \end{aligned} \quad (2.9)$$

Действительно, используя свойства преобразования Фурье, имеем

$$\begin{aligned}\mathcal{M}^1 u &= -\operatorname{Im} \int_{\mathbb{R}^{n+1}} \mathcal{L}^1(D_t + \gamma I, D_x) u_\gamma(t, x) \overline{\left(\mathcal{L}_1^1(D_t + \gamma I, D_x) u_\gamma(t, x) \right)} dz = \\ &= -\operatorname{Im} \int_{\mathbb{R}^{n+1}} \mathcal{L}^1(i\eta + \gamma, i\xi) \hat{u}_\gamma(\eta, \xi) \overline{\left(\mathcal{L}_1^1(i\eta + \gamma, i\xi) \hat{u}_\gamma(\eta, \xi) \right)} d\xi = \int_{\mathbb{R}^{n+1}} M(i\eta + \gamma, i\xi) |\hat{u}_\gamma(\eta, \xi)|^2 d\xi.\end{aligned}$$

Отсюда в силу (2.6) вытекает (2.9).

Из неравенства (2.9), очевидно, следует энергетическая оценка для псевдогоиперболического оператора $\mathcal{L}^1(D_t, D_x)$.

Лемма 2. Для любой функции $u(t, x) \in W_{2,\gamma}^{2,4}(\mathbb{R}^{n+1})$, $\gamma > 0$, такой, что $D_t^2 D_x^\beta u(t, x) \in L_{2,\gamma}(\mathbb{R}^{n+1})$, $|\beta| = 2$, имеет место оценка

$$\gamma \| |\xi|^2 (|\eta| + |\xi| + \gamma) \hat{u}_\gamma(\eta, \xi), L_2(\mathbb{R}^{n+1}) \| \leq c \| \mathcal{L}^1(D_t, D_x) u(t, x), L_{2,\gamma}(\mathbb{R}^{n+1}) \|, \quad (2.10)$$

где $c > 0$ — константа, не зависящая от $u(t, x)$.

Из этой леммы вытекает единственность решения задачи Коши для псевдогоиперболического уравнения с постоянными коэффициентами без младших членов в соболевском пространстве $W_{2,\gamma}^{2,4}(\mathbb{R}^{n+1})$, $\gamma > 0$.

Прежде чем переходить к получению энергетической оценки (1.7) для псевдогоиперболического оператора с переменными вещественными коэффициентами, покажем как можно получить такую оценку для псевдогоиперболического оператора с младшим членом

$$\mathcal{L}_a^1(D_t, D_x) = \mathcal{L}^1(D_t, D_x) + a D_t^2 = (L_0^1(D_x) + a I) D_t^2 + L_1^1(D_x) D_t + L_2^1(D_x), \quad a > 0. \quad (2.11)$$

Вначале выпишем для этого оператора аналог многочлена (2.3)

$$\begin{aligned}M_a(i\eta + \gamma, i\xi) &= -\operatorname{Im} \left(\left((L_0^1(i\xi) + a)(-\eta^2 + 2i\eta\gamma + \gamma^2) + L_1^1(i\xi)(i\eta + \gamma) + L_2^1(i\xi) \right) \times \right. \\ &\quad \left. \times \overline{\left(2i(L_0^1(i\xi) + a)(i\eta + \gamma) + iL_1^1(i\xi) \right)} \right), \quad \gamma \geq 0, \quad (\eta, \xi) \in \mathbb{R}^{n+1}.\end{aligned} \quad (2.12)$$

Проводя вычисления, получим

$$M_a(i\eta + \gamma, i\xi) = \gamma \left(\left(L_1^1(\xi) \right)^2 + 2 \left(L_0^1(\xi) - a \right)^2 (\eta^2 + \gamma^2) - 2 \left(L_0^1(\xi) - a \right) L_2^1(\xi) + 2 \left(L_0^1(\xi) - a \right) L_1^1(\xi) \eta \right).$$

Перепишем это в следующем виде:

$$M_a(i\eta + \gamma, i\xi) = 2\gamma \left(-L_0^1(\xi) + a \right)^2 \left(\left(\eta - \frac{L_1^1(\xi)}{2(-L_0^1(\xi) + a)} \right)^2 + \frac{(L_1^1(\xi))^2 + 4(-L_0^1(\xi) + a)L_2^1(\xi)}{4(-L_0^1(\xi) + a)^2} + \gamma^2 \right).$$

Введем обозначение

$$Q(\eta, \gamma, \xi, a) = \left(\eta - \frac{L_1^1(\xi)}{2(-L_0^1(\xi) + a)} \right)^2 + \frac{(L_1^1(\xi))^2 + 4(-L_0^1(\xi) + a)L_2^1(\xi)}{4(-L_0^1(\xi) + a)^2} + \gamma^2. \quad (2.13)$$

Тогда полином (2.12) можно представить в виде

$$M_a(i\eta + \gamma, i\xi) = 2\gamma \left(-L_0^1(\xi) + a \right)^2 Q(\eta, \gamma, \xi, a). \quad (2.14)$$

В силу определений (2.4), (2.13), очевидно, имеем

$$Q(\eta, \gamma, \xi, 0) = N(\eta, \gamma, \xi).$$

Напомним, что функция $N(\eta, \gamma, \xi)$ является однородной второй степени, и в силу строгой псевдогоиперболическости оператора $\mathcal{L}^1(D_t, D_x)$ для нее выполнена оценка (2.5), которую можно переписать в виде

$$\begin{aligned}r_2 \geq Q(\eta', \gamma', \xi', 0) = N(\eta', \gamma', \xi') \geq r_1 > 0, \quad s = (\eta', \gamma', \xi'), \\ \eta' = \eta/\Delta, \quad \gamma' = \gamma/\Delta, \quad \xi' = \xi/\Delta, \quad \Delta = \sqrt{\eta^2 + |\xi|^2 + \gamma^2}.\end{aligned}$$

Запишем теперь функцию (2.13) следующим образом

$$Q(\eta, \gamma, \xi, a) = \Delta^2 Q(\eta', \gamma', \xi', \frac{a}{\Delta^2}).$$

Тогда в силу равномерной непрерывности функции

$$Q(\eta', \gamma', \xi', \alpha), \quad (\eta')^2 + (\gamma')^2 + |\xi'|^2 = 1, \quad \alpha \in [0, \alpha_0],$$

из этой оценки следует, что существует $\gamma_1 > 0$ такое, что при всех $(\eta, \xi) \in \mathbb{R}^{n+1}$, $\gamma \geq \gamma_1$ будет выполняться неравенство

$$2r_2(\eta^2 + |\xi|^2 + \gamma^2) \geq Q(\eta, \gamma, \xi, a) \geq \frac{r_1}{2}(\eta^2 + |\xi|^2 + \gamma^2).$$

Отсюда в силу (2.14) получаем

$$4r_2\gamma \left(-L_0^1(\xi) + a \right)^2 (\eta^2 + |\xi|^2 + \gamma^2) \geq M_a(i\eta + \gamma, i\xi) \geq r_1\gamma \left(-L_0^1(\xi) + a \right)^2 (\eta^2 + |\xi|^2 + \gamma^2).$$

Следовательно, учитывая (1.5), имеем аналог неравенства (2.6)

$$4r_2\gamma(q_2|\xi|^2 + a)^2(\eta^2 + |\xi|^2 + \gamma^2) \geq M_a(i\eta + \gamma, i\xi) \geq r_1\gamma(q_1|\xi|^2 + a)^2(\eta^2 + |\xi|^2 + \gamma^2), \quad \gamma \geq \gamma_1, \quad (\eta, \xi) \in \mathbb{R}^{n+1}. \quad (2.15)$$

Используя (2.15), нетрудно получить энергетическую оценку для оператора (2.11). Для этого, определим дифференциальный оператор

$$\mathcal{L}_{a,1}^1(D_t, D_x) = 2i(L_0^1(D_x) + aI)D_t + iL_1^1(D_x) \quad (2.16)$$

и рассмотрим аналог формы (2.8)

$$\mathcal{M}_a^1 u = -\text{Im} \int_{\mathbb{R}^{n+1}} e^{-\gamma t} \mathcal{L}_a^1(D_t, D_x) u(t, x) \overline{\left(e^{-\gamma t} \mathcal{L}_{a,1}^1(D_t, D_x) u(t, x) \right)} dz, \quad z = (t, x). \quad (2.17)$$

$$u(t, x) \in C_0^\infty(\mathbb{R}^{n+1}).$$

Используя свойства преобразования Фурье, это можно записать в виде

$$\mathcal{M}_a^1 u = -\text{Im} \int_{\mathbb{R}^{n+1}} \mathcal{L}_a^1(i\eta + \gamma, i\xi) \hat{u}_\gamma(\eta, \xi) \overline{\left(\mathcal{L}_{a,1}^1(i\eta + \gamma, i\xi) \hat{u}_\gamma(\eta, \xi) \right)} d\xi = \int_{\mathbb{R}^{n+1}} M_a(i\eta + \gamma, i\xi) |\hat{u}_\gamma(\eta, \xi)|^2 d\xi.$$

Отсюда в силу (2.15) получаем двухстороннее неравенство

$$\begin{aligned} 4r_2\gamma \int_{\mathbb{R}^{n+1}} (q_2|\xi|^2 + a)^2(\eta^2 + |\xi|^2 + \gamma^2) |\hat{u}_\gamma(\eta, \xi)|^2 d\xi &\geq \mathcal{M}_a^1 u \geq \\ &\geq r_1\gamma \int_{\mathbb{R}^{n+1}} (q_1|\xi|^2 + a)^2(\eta^2 + |\xi|^2 + \gamma^2) |\hat{u}_\gamma(\eta, \xi)|^2 d\xi, \quad \gamma \geq \gamma_1. \end{aligned} \quad (2.18)$$

Из этого неравенства, очевидно, вытекает энергетическая оценка для оператора (2.11).

Лемма 3. Существует $\gamma_1 > 0$ такое, что для любой функции $u(t, x) \in W_{2,\gamma}^{2,4}(\mathbb{R}^{n+1})$ при $\gamma \geq \gamma_1$ такой, что $D_t^2 D_x^\beta u(t, x) \in L_{2,\gamma}(\mathbb{R}^{n+1})$, $|\beta| = 2$, имеет место оценка

$$\gamma \| (|\xi|^2 + a) (|\eta| + \gamma + |\xi|) \hat{u}_\gamma(\eta, \xi), L_2(\mathbb{R}^{n+1}) \| \leq c \| \mathcal{L}_a^1(D_t, D_x) u(t, x), L_{2,\gamma}(\mathbb{R}^{n+1}) \|$$

с константой $c > 0$, не зависящей от $u(t, x)$.

3. ЭНЕРГЕТИЧЕСКИЕ ОЦЕНКИ ДЛЯ ОПЕРАТОРОВ С ПЕРЕМЕННЫМИ КОЭФФИЦИЕНТАМИ

Докажем энергетические оценки для строго гиперболических операторов $\mathcal{L}(x; D_t, D_x)$ вида (1.3) с переменными коэффициентами.

Будем предполагать, что переменные коэффициенты $\alpha_{\beta}^k(x)$ оператора (1.6) вместе со своими производными до третьего порядка включительно достаточно малы. Рассмотрим аналог формы (2.17)

$$\mathcal{M}u = -\operatorname{Im} \int_{\mathbb{R}^{n+1}} e^{-\gamma t} \mathcal{L}(x; D_t, D_x) u(t, x) \overline{\left(e^{-\gamma t} \mathcal{L}_{a,1}^1(D_t, D_x) u(t, x) \right)} dz,$$

где $u(t, x) \in C_0^\infty(\mathbb{R}^{n+1})$, оператор $\mathcal{L}_{a,1}^1(D_t, D_x)$ имеет вид (2.16). Ее можно записать в следующем виде:

$$\mathcal{M}u = -\operatorname{Im} \int_{\mathbb{R}^{n+1}} e^{-\gamma t} \left(\mathcal{L}_a^1(D_t, D_x) u(t, x) + \mathcal{L}_0^2(x; D_t, D_x) u(t, x) \right) \overline{\left(e^{-\gamma t} \mathcal{L}_{a,1}^1(D_t, D_x) u(t, x) \right)} dz = \mathcal{M}_a^1 u + \mathcal{M}_a^2 u, \quad (3.1)$$

где $\mathcal{L}_0^2(x; D_t, D_x)$ — оператор из (1.6) при $a = 0$, $\mathcal{M}_a^1 u$ — форма (2.17),

$$\begin{aligned} \mathcal{M}_a^2 u &= -\operatorname{Im} \int_{\mathbb{R}^{n+1}} e^{-\gamma t} \mathcal{L}_0^2(x; D_t, D_x) u(t, x) \overline{\left(e^{-\gamma t} \mathcal{L}_{a,1}^1(D_t, D_x) u(t, x) \right)} dz = \\ &= -\operatorname{Im} \int_{\mathbb{R}^{n+1}} \mathcal{L}_0^2(x; D_t + \gamma I, D_x) u_\gamma(t, x) \overline{\left(\mathcal{L}_{a,1}^1(D_t + \gamma I, D_x) u_\gamma(t, x) \right)} dz. \end{aligned}$$

В дальнейшем нам удобно будет использовать следующее обозначение для скалярного произведения в $L_2(\mathbb{R}^{n+1})$:

$$\langle v, w \rangle = \int_{\mathbb{R}^{n+1}} v(z) \overline{w(z)} dz$$

и для сокращения записи будем писать

$$\begin{aligned} \mathcal{L}_0^2(x) &= \mathcal{L}_0^2(x; D_t + \gamma I, D_x) = L_{0,0}^2(x, D_x)(D_t + \gamma I)^2 + L_1^2(x, D_x)(D_t + \gamma I) + L_2^2(x, D_x), \\ \mathcal{L}_{a,1}^1 &= \mathcal{L}_{a,1}^1(D_t + \gamma I, D_x) = 2i(L_0^1(D_x) + aI)(D_t + \gamma I) + iL_1^1(D_x). \end{aligned}$$

Тогда форму $\mathcal{M}_a^2 u$ можно представить в виде

$$\begin{aligned} \mathcal{M}_a^2 u &= -\operatorname{Im} \langle \mathcal{L}_0^2(x) u_\gamma, \mathcal{L}_{a,1}^1 u_\gamma \rangle = -\frac{1}{2i} \left(\langle \mathcal{L}_0^2(x) u_\gamma, \mathcal{L}_{a,1}^1 u_\gamma \rangle - \langle \mathcal{L}_{a,1}^1 u_\gamma, \mathcal{L}_0^2(x) u_\gamma \rangle \right) = \\ &= -\frac{1}{2i} \langle \left((\mathcal{L}_{a,1}^1)^* \mathcal{L}_0^2(x) - (\mathcal{L}_0^2(x))^* \mathcal{L}_{a,1}^1 \right) u_\gamma, u_\gamma \rangle. \end{aligned} \quad (3.2)$$

В силу определения операторов $\mathcal{L}_{a,1}^1(D_t, D_x)$, $\mathcal{L}_0^2(x, D_t, D_x)$, учитывая, что все коэффициенты вещественные, для сопряженных операторов имеем

$$\begin{aligned} (\mathcal{L}_{a,1}^1)^* &= 2i(D_t - \gamma I)(L_0^1(D_x) + aI) + iL_1^1(D_x), \\ (\mathcal{L}_0^2(x))^* v(z) &= (D_t - \gamma I)^2 \left(\sum_{|\beta|=2} D_x^\beta (\alpha_\beta^0(x) v(z)) \right) + \\ &+ (D_t - \gamma I) \sum_{|\beta|=3} D_x^\beta (\alpha_\beta^1(x) v(z)) + \sum_{|\beta|=4} D_x^\beta (\alpha_\beta^2(x) v(z)), \quad v(z) \in C_0^\infty(\mathbb{R}^{n+1}). \end{aligned}$$

Введем дифференциальный оператор

$$\mathcal{P}(x; D_t, D_x, \gamma) = -\frac{1}{2i} \left((\mathcal{L}_{a,1}^1)^* \mathcal{L}_0^2(x) - (\mathcal{L}_0^2(x))^* \mathcal{L}_{a,1}^1 \right).$$

Тогда выражение (3.2) будет иметь вид

$$\mathcal{M}_a^2 u = \langle \mathcal{P}(x; D_t, D_x, \gamma) u_\gamma, u_\gamma \rangle. \quad (3.3)$$

Учитывая гладкость коэффициентов операторов $L_{0,0}^2(x; D_x)$, $L_1^2(x; D_x)$, $L_2^2(x; D_x)$, дифференциальный оператор $\mathcal{P}(x; D_t, D_x, \gamma)$ можно представить в виде суммы старшей его части $P(x; D_t, D_x, \gamma)$ и операторов меньшего порядка $p(x; D_t, D_x, \gamma)$, при этом будем считать умножение на параметр γ^k оператором k -го порядка, т. е.

$$\mathcal{P}(x; D_t, D_x, \gamma) = P(x; D_t, D_x, \gamma) + p(x; D_t, D_x, \gamma), \quad (3.4)$$

где оператор $P(x; D_t, D_x, \gamma)$ имеет седьмой порядок, а оператор $p(x; D_t, D_x, \gamma)$ — шестой порядок. Тогда (3.3) можно записать в виде

$$\mathcal{M}_a^2 u = \langle P(x; D_t, D_x, \gamma) u_\gamma, u_\gamma \rangle + \langle p(x; D_t, D_x, \gamma) u_\gamma, u_\gamma \rangle. \quad (3.5)$$

В силу гладкости коэффициентов оператора (1.6), нетрудно получить, что

$$\begin{aligned} P(x; D_t, D_x, \gamma) = & -(D_t - \gamma I)(D_t + \gamma I)L_{0,0}^2(x; D_x)L_0^1(D_x) - (D_t^2 - \gamma^2 I)L_1^2(x; D_x)L_0^1(D_x) - \\ & - (D_t - \gamma I)L_2^2(x; D_x)L_0^1(D_x) - \frac{1}{2}(D_t + \gamma I)^2 L_{0,0}^2(x; D_x)L_1^1(D_x) - \\ & - \frac{1}{2}(D_t + \gamma I)L_1^2(x; D_x)L_1^1(D_x) - \frac{1}{2}L_2^2(x; D_x)L_1^1(D_x) + \\ & + (D_t - \gamma I)^2(D_t + \gamma I)L_{0,0}^2(x; D_x)L_0^1(D_x) + \frac{1}{2}(D_t - \gamma I)^2 L_{0,0}^2(x; D_x)L_1^1(D_x) + \\ & + (D_t - \gamma I)(D_t + \gamma I)L_1^2(x; D_x)L_0^1(D_x) + \frac{1}{2}(D_t - \gamma I)L_1^2(x; D_x)L_1^1(D_x) + \\ & + (D_t + \gamma I)L_2^2(x; D_x)L_0^1(D_x) + \frac{1}{2}L_2^2(x; D_x)L_1^1(D_x). \end{aligned}$$

А этот оператор, очевидно, имеет вид

$$\begin{aligned} P(x; D_t, D_x, \gamma) = & -2\gamma \left((D_t^2 - \gamma^2 I)L_{0,0}^2(x; D_x)L_0^1(D_x) + \right. \\ & \left. + D_t L_{0,0}^2(x; D_x)L_1^1(D_x) + \frac{1}{2}L_1^2(x; D_x)L_1^1(D_x) - L_2^2(x; D_x)L_0^1(D_x) \right). \end{aligned} \quad (3.6)$$

Поэтому, подставляя оператор (3.6) в первое слагаемое в правой части (3.5) и учитывая, что $u(t, x) \in C_0^\infty(\mathbb{R}^{n+1})$, то после интегрирования по частям его можно записать следующим образом

$$\begin{aligned} \langle P(x; D_t, D_x, \gamma) u_\gamma, u_\gamma \rangle = & 2\gamma \left(\langle (D_t - \gamma I)L_0^1(D_x) u_\gamma, (D_t - \gamma I)(L_{0,0}^2(x; D_x))^* u_\gamma \rangle + \right. \\ & \left. + \langle L_1^1(D_x) u_\gamma, D_t(L_{0,0}^2(x; D_x))^* u_\gamma \rangle - \frac{1}{2} \langle L_1^1(D_x) u_\gamma, (L_1^2(x; D_x))^* u_\gamma \rangle + \sum_{j=1}^n \sum_{|\beta| \leq 3} \langle D_{x_j} L_0^1(D_x) u_\gamma, b_{j,\beta}(x) D_x^\beta u_\gamma \rangle \right). \end{aligned}$$

Тогда, очевидно, имеем оценку

$$\begin{aligned} |\langle P(x; D_t, D_x, \gamma) u_\gamma, u_\gamma \rangle| \leq & d\gamma \| |\xi|^2 (|\eta| + \gamma + |\xi|) \widehat{u}_\gamma(\eta, \xi), L_2(\mathbb{R}^{n+1}) \| \times \left(\| (D_t - \gamma I)(L_{0,0}^2(x; D_x))^* u_\gamma, L_2(\mathbb{R}^{n+1}) \| + \right. \\ & \left. + \| D_t(L_{0,0}^2(x; D_x))^* u_\gamma, L_2(\mathbb{R}^{n+1}) \| + \| (L_1^2(x; D_x))^* u_\gamma, L_2(\mathbb{R}^{n+1}) \| + \sum_{j=1}^n \sum_{|\beta| \leq 3} \| b_{j,\beta}(x) D_x^\beta u_\gamma, L_2(\mathbb{R}^{n+1}) \| \right). \end{aligned} \quad (3.7)$$

Обозначим выражение в скобках через J .

Напомним, что коэффициенты операторов $L_{0,0}^2(x; D_x)$, $L_1^2(x; D_x)$, $L_2^2(x; D_x)$ являются бесконечно дифференцируемыми и финитными, при этом сейчас мы пока предполагаем, что они вместе со своими производными до третьего порядка включительно достаточно малы. Тогда для J можно выписать следующую оценку:

$$J \leq \varepsilon \left(\| (|\xi|^2 + 1)(|\eta| + \gamma) \widehat{u}_\gamma(\eta, \xi), L_2(\mathbb{R}^{n+1}) \| + \| (|\xi|^3 + 1) \widehat{u}_\gamma(\eta, \xi), L_2(\mathbb{R}^{n+1}) \| \right),$$

где величина малой константы $\varepsilon > 0$ полностью определяется малостью коэффициентов $\alpha_\beta^k(x)$ и их производных до третьего порядка.

Из этого неравенства, очевидно, следует, что при $a > 0$ существует $\gamma_2 > 0$ такое, что при $\gamma \geq \gamma_2$ выполняется следующая оценка:

$$J \leq c_1(a, \gamma_2) \varepsilon \|(|\xi|^2 + a)(|\eta| + \gamma + |\xi|) \widehat{u}_\gamma(\eta, \xi), L_2(\mathbb{R}^{n+1})\|,$$

где константа $c_1(a, \gamma_2) > 0$ зависит от a и γ_2 . Отсюда в силу (3.7) получаем неравенство

$$|\langle P(x; D_t, D_x, \gamma) u_\gamma, u_\gamma \rangle| \leq dc_1(a, \gamma_2) \varepsilon \gamma \|(|\xi|^2 + a)(|\eta| + \gamma + |\xi|) \widehat{u}_\gamma(\eta, \xi), L_2(\mathbb{R}^{n+1})\|^2.$$

Учитывая еще, что порядок оператора $p(x; D_t, D_x, \gamma)$ меньше, чем порядок оператора $P(x; D_t, D_x, \gamma)$, получим также

$$|\langle p(x; D_t, D_x, \gamma) u_\gamma, u_\gamma \rangle| \leq c_2(a, \gamma_2) \|(|\xi|^2 + a)(|\eta| + \gamma + |\xi|) \widehat{u}_\gamma(\eta, \xi), L_2(\mathbb{R}^{n+1})\|^2,$$

где константа $c_2(a, \gamma_2) > 0$ зависит от a и γ_2 . Из этих двух оценок и (3.5) при $\gamma \geq \gamma_2$ вытекает неравенство

$$|\mathcal{M}_a^2 u| \leq (dc_1(a, \gamma_2) \varepsilon \gamma + c_2(a, \gamma_2)) \|(|\xi|^2 + a)(|\eta| + \gamma + |\xi|) \widehat{u}_\gamma(\eta, \xi), L_2(\mathbb{R}^{n+1})\|^2. \quad (3.8)$$

Используя теперь оценки (2.18) и (3.8), форму $\mathcal{M}u$ из (3.1) можно оценить снизу. Вначале отметим, что из оценки (2.18) вытекает неравенство

$$\mathcal{M}_a^1 u \geq \rho_1 \gamma \|(|\xi|^2 + a)(|\eta| + \gamma + |\xi|) \widehat{u}_\gamma(\eta, \xi), L_2(\mathbb{R}^{n+1})\|^2, \quad \gamma \geq \gamma_1,$$

где константа $\rho_1 > 0$ зависит от r_1, q_1, a . Тогда, учитывая (3.8), при $\gamma \geq \max\{\gamma_1, \gamma_2\}$ получим

$$\mathcal{M}u = \mathcal{M}_a^1 u + \mathcal{M}_a^2 u \geq \mathcal{M}_a^1 u - |\mathcal{M}_a^2 u| \geq \gamma \left(\rho_1 - dc_1(a, \gamma_2) \varepsilon - \frac{c_2(a, \gamma_2)}{\gamma} \right) \|(|\xi|^2 + a)(|\eta| + \gamma + |\xi|) \widehat{u}_\gamma(\eta, \xi), L_2(\mathbb{R}^{n+1})\|^2.$$

Пусть $\gamma_3 = 4c_2(a, \gamma_2)/\rho_1$ и предположим, что коэффициенты оператора (1.6) вместе с производными до третьего порядка включительно настолько малы, что

$$dc_1(a, \gamma_2) \varepsilon \leq \rho_1/4.$$

Тогда из последнего неравенства при $\gamma \geq \max\{\gamma_1, \gamma_2, \gamma_3\}$ получаем оценку

$$\mathcal{M}u \geq \gamma \frac{\rho_1}{2} \|(|\xi|^2 + a)(|\eta| + \gamma + |\xi|) \widehat{u}_\gamma(\eta, \xi), L_2(\mathbb{R}^{n+1})\|^2, \quad u(t, x) \in C_0^\infty(\mathbb{R}^{n+1}). \quad (3.9)$$

Оценим теперь форму $\mathcal{M}u$ сверху. Из определения (3.1), очевидно, имеем

$$\begin{aligned} \mathcal{M}u &\leq \|e^{-\gamma t} \mathcal{L}(x; D_t, D_x) u(t, x), L_2(\mathbb{R}^{n+1})\| \|e^{-\gamma t} \mathcal{L}_{a,1}^1(D_t, D_x) u(t, x), L_2(\mathbb{R}^{n+1})\| = \\ &= \|\mathcal{L}(x; D_t, D_x) u(t, x), L_{2,\gamma}(\mathbb{R}^{n+1})\| \|\mathcal{L}_{a,1}^1(i\eta + \gamma, i\xi) \widehat{u}_\gamma(\eta, \xi), L_2(\mathbb{R}^{n+1})\|. \end{aligned}$$

А поскольку из определения (2.16) вытекает

$$|\mathcal{L}_{a,1}^1(i\eta + \gamma, i\xi)| \leq c(|\xi|^2 + a)(|\eta| + \gamma + |\xi|), \quad \gamma > 0,$$

то

$$\mathcal{M}u \leq c \|\mathcal{L}(x; D_t, D_x) u(t, x), L_{2,\gamma}(\mathbb{R}^{n+1})\| \|(|\xi|^2 + a)(|\eta| + \gamma + |\xi|) \widehat{u}_\gamma(\eta, \xi), L_2(\mathbb{R}^{n+1})\|.$$

Учитывая это неравенство и (3.9), а также теорему о всюду плотности $C_0^\infty(\mathbb{R}^{n+1})$ в соболевском пространстве $W_{2,\gamma}^{2,4}(\mathbb{R}^{n+1})$, при $\gamma \geq \max\{\gamma_1, \gamma_2, \gamma_3\}$ получаем энергетическую оценку (1.7) для псевдогиперболического оператора (1.3) при достаточно малых возмущениях коэффициентов.

Теорема доказана.

Замечание. Учитывая доказанную теорему и используя теорему о разбиении единицы, нетрудно установить энергетическую оценку (1.7) для операторов вида (1.3), являющихся строго псевдогиперболическими при любом $x_0 \in \mathbb{R}^n$ и имеющих достаточно гладкие коэффициенты, постоянные вне компакта.

СПИСОК ЛИТЕРАТУРЫ

1. Демиденко Г.В., Успенский С.В. Уравнения и системы, не разрешенные относительно старшей производной. Новосибирск: Научная книга, 1998.
2. Соболев С.Л. Избранные труды. Новосибирск: Изд-во Ин-та математики, Филиал “Гео” Изд-ва СО РАН. Т. I. 2003. Т. II. 2006.
3. Свеишников А.Г., Альшин А.Б., Корпусов М.О., Плетнер Ю.Д. Линейные и нелинейные уравнения соболевского типа. М.: Физматлит, 2007.
4. Demidenko G. The Cauchy problem for pseudohyperbolic equations // Selcuk J. Appl. Math. 2001. V. 1. № 1. P. 47–62.
5. Демиденко Г.В. Условия разрешимости задачи Коши для псевдогиперболических уравнений // Сиб. матем. ж. 2015. Т. 56. № 6. С. 1289–1303.
6. Fedotov I., Volevich L.R. The Cauchy problem for hyperbolic equations not resolved with respect to the highest time derivative // Russian J. Math. Physics. 2006. V. 13. № 3. P. 278–292.
7. Bondar L.N., Demidenko G.V. On the Cauchy problem for pseudohyperbolic equations with lower order terms // Mathematics. 2023. V. 11. 3943.
8. Бондарь Л.Н., Демиденко Г.В. О корректности задачи Коши для псевдогиперболических уравнений в весовых соболевских пространствах // Сиб. матем. ж. 2023. Т. 64. № 5. С. 895–911.
9. Власов В.З. Тонкостенные упругие стержни. 2-е изд., перераб. и доп. М.: Физматгиз, 1959.
10. Герасимов С.И., Ерофеев В.И. Задачи волновой динамики элементов конструкций. Саров: ФГУП РФЯЦ-ВНИИЭФ, 2014.
11. Bishop R.E.D. Longitudinal waves in beams // Aeronautical Quarterly. 1952. V. 3. № 4. P. 280–293.
12. Rao J.S. Advanced Theory of Vibration. New York: Wiley, 1992.
13. Fedotov I., Shatalov M., Marais J. Hyperbolic and pseudo-hyperbolic equations in the theory of vibration // Acta Mechanica. 2016. V. 227. № 11. P. 3315–3324.
14. Петровский И.Г. Избранные труды. Системы уравнений с частными производными. Алгебраическая геометрия. М.: Наука, 1986.
15. Лере Ж. Гиперболические дифференциальные уравнения. М.: Наука, 1984.

ENERGY ESTIMATES FOR A CLASS OF PSEUDO-HYPERBOLIC OPERATORS WITH VARIABLE COEFFICIENTS

G. V. Demidenko*

Novosibirsk State University, Pirogov St., 2, Novosibirsk, 630090 Russia

*e-mail: demidenk@math.nsc.ru

Received: 10 March, 2024

Revised: 10 March, 2024

Accepted: 05 May, 2024

Abstract. A class of strictly pseudo-hyperbolic fourth-order operators with variable coefficients is considered. Energy estimates are established under certain conditions on the coefficients. These estimates imply the uniqueness of the solution to the Cauchy problem, as well as a priori estimates.

Keywords: equations not resolved with respect to the highest derivative, pseudo-hyperbolic operators, weighted Sobolev spaces, energy estimates.

СТРУКТУРИРОВАННЫЕ ПСЕВДОСПЕКТРЫ В ЗАДАЧАХ ПРОСТРАНСТВЕННОЙ УСТОЙЧИВОСТИ ПОГРАНИЧНЫХ СЛОЕВ¹⁾

© 2024 г. К. В. Демьянко^{1,2}, Г. В. Засько^{1,2}, Ю. М. Нечепуренко^{1,2,*},

¹119333 Москва, ул. Губкина, 8, Институт вычислительной математики им. Г.И. Марчука РАН

²125047 Москва, Миусская пл., 4, Институт прикладной математики им. М.В. Келдыша РАН, Россия

*e-mail: yumnec@yandex.ru

Поступила в редакцию 23.02.2024 г.

Переработанный вариант 23.02.2024 г.

Принята к публикации 02.05.2024 г.

Работа посвящена численному анализу чувствительности характеристик пространственной устойчивости пограничных слоев к погрешностям, с которыми задано основное течение. Предлагается использовать для этого структурированные псевдоспектры. Показано, что полученные оценки значительно точнее оценок на основе неструктурированного псевдоспектра. Изложение ведется на примере течения вязкой несжимаемой жидкости над вогнутой поверхностью малой кривизны при параметрах течения благоприятных для развития вихрей Гертлера и волн Толлмина–Шлихтинга. Библ. 28. Фиг. 2.

Ключевые слова: структурированные псевдоспектры, резольвента, пространственная устойчивость, пограничный слой, вихри Гертлера, волны Толлмина–Шлихтинга.

DOI: 10.31857/S0044466924080125, EDN: YAGHYQ

1. ВВЕДЕНИЕ

Теория гидродинамической устойчивости рассматривает внешние возмущения основного течения, устойчивость которого представляет интерес. Она изучает их распространение в пространстве или эволюцию во времени. (см. [1]–[3]). Характеристики устойчивости, такие как инкременты нарастания, фазовые и групповые скорости, описывают развитие наиболее неустойчивых возмущений. Вычисление характеристик устойчивости сводят к вычислению собственных значений матриц, зависящих от параметров, отбору ведущих собственных значений и вычислению отвечающих им собственных векторов. Наряду с внешними возмущениями, развивающимися на фоне основного течения, возможны внутренние возмущения основного течения, которые следует рассматривать как возмущения параметров задачи устойчивости, а значит элементов матриц, с помощью которых вычисляют характеристики устойчивости. Настоящая работа посвящена именно таким возмущениям.

Первостепенной задачей, возникающей при постановке проблем собственных значений и выборе алгоритмов их решения в случае возможного возмущения матриц, является оценка того, на сколько могут отличаться собственные значения возмущенной матрицы от собственных значений исходной. В настоящее время используют три типа оценок: мажорантные оценки, асимптотические оценки и псевдоспектры. Каждый из них имеет свою область применимости.

Примером мажорантной оценки является неравенство Бауэра–Файка [4], в котором возмущение собственных значений матрицы простой структуры оценивается через число обусловленности матрицы собственных векторов. Эта и другие подобные оценки (см. [5]) полезны в теоретических исследованиях, однако на практике, за исключением специального случая, когда матрица является нормальной или близкой к таковой, они оказываются слишком завышенными.

При получении асимптотических оценок (см. [6]) предполагается, что норма возмущений мала и для возмущения собственного значения выводится достижимая оценка члена первого порядка малости. Такие оценки хорошо вписываются в вычислительные схемы известных алгоритмов (например, QR-алгоритма) и полезны при численном анализе возможного влияния погрешностей округления. Однако, если требуется оценить возмущение спектра при возмущениях матрицы, норма которых значительно больше погрешностей округления,

¹⁾Работа выполнена при финансовой поддержке РНФ (грант № 22-11-00025).

эти оценки не применимы. Такая проблема возникает, например, когда все, либо некоторые элементы матрицы получены в результате физического эксперимента или численного моделирования и имеют достаточно большую погрешность.

Для произвольной матрицы $A \in \mathbb{C}^{n \times n}$ через $\Lambda_\varepsilon(A)$ будем обозначать объединение спектров всех возмущенных матриц вида $A + \Delta$ при норме возмущения не превосходящей ε :

$$\Lambda_\varepsilon(A) = \bigcup_{\|\Delta\|_2 \leq \varepsilon} \lambda(A + \Delta), \quad \Delta \in \mathbb{C}^{n \times n}.$$

Иначе говоря, $z \in \Lambda_\varepsilon(A)$ в том, и только том случае, когда существует матрица $\Delta \in \mathbb{C}^{n \times n}$ такая, что $\|\Delta\|_2 \leq \varepsilon$ и $z \in \lambda(A + \Delta)$. Множество $\Lambda_\varepsilon(A)$ называют *псевдоспектром* (или ε -*спектром*) матрицы A . Псевдоспектр обладает следующим замечательным свойством:

$$\Lambda_\varepsilon(A) = \{z : \|(zI - A)^{-1}\|_2 \geq 1/\varepsilon\}, \quad (1.1)$$

где I означает единичную матрицу порядка n .

По своему определению, псевдоспектр является достижимой оценкой возможного перемещения собственных значений при норме возмущений не превосходящих заданного ε . Нарисовав в заданной области комплексной плоскости линии уровня нормы резольвенты мы получим так называемый *спектральный портрет* матрицы A , дающий в силу (1.1) псевдоспектр для любого заданного ε . Если задано множество допустимых возмущений $\mathcal{M} \subset \mathbb{C}^{n \times n}$, то можно оценить сверху величину $\|\Delta\|_2$, где $\Delta \in \mathcal{M}$, и выбрать эту верхнюю оценку в качестве ε . Тогда мы получим верхнюю оценку возможного смещения собственных значений при допустимых возмущениях матрицы:

$$\bigcup_{\Delta \in \mathcal{M}} \lambda(A + \Delta) \subset \Lambda_\varepsilon(A).$$

Понятие псевдоспектра было независимо введено в работах С.К. Годунова (см. [7], [8]) и Л. Трефетена (см. [9], [10]). С этим понятием тесно связан ряд математических результатов, в которых фигурирует резольвента и различные ее нормы, например, оценка чувствительности спектрального проектора на инвариантное подпространство, отвечающее заданному подмножеству спектра [11]–[14], оценка нормы матрицы Грина [15], оценка нормы матричной экспоненты [10]. Следует отметить, что оценки нормы матричной экспоненты, полученные на основе псевдоспектров, завышены. Более точные оценки предложены, например, в работах [16], [17] и основаны на решениях уравнений Ляпунова.

Псевдоспектры используют в задачах гидродинамической устойчивости (см., например, [3], [18], [19]). Однако полученные результаты скорее позволяют судить о кластеризации спектра, чем о возможном смещении собственных значений при возмущении параметров течений. Связано это с тем, что использование псевдоспектра для оценки чувствительности спектра к возмущению элементов матрицы имеет один большой недостаток: псевдоспектр не позволяет учесть специфику возмущения. Из-за этого, полученные оценки на практике могут оказаться сильно завышенными. Более того, если A — это конечномерный аналог неограниченного оператора, то при параметре дискретизации стремящемся к бесконечности максимальная норма допустимого возмущения может стремиться к бесконечности, и, как следствие, псевдоспектр будет стремиться ко всей комплексной плоскости. При этом минимальные по абсолютной величине собственные значения матрицы (а именно они как правило представляют наибольший интерес) могут возмущаться не слишком сильно.

Эту проблему в значительной мере позволяет решить обобщение псевдоспектра, называемое структурированным псевдоспектром [20], [21], ориентированное на возмущения, представимые в виде $\Delta = BSC$, где B и C фиксированные квадратные (структурирующие) матрицы порядка n , $S \in \mathbb{C}^{n \times n}$ и справедливо вложение

$$\mathcal{M} \subset \bigcup_{\|S\|_2 \leq \varepsilon} BSC.$$

Выбирая подходящим образом структурирующие матрицы, можно более точно локализовать возмущения и соответственно множество спектров возмущенных матриц, чем это позволяет сделать неструктурированный псевдоспектр [22].

В настоящей работе мы покажем как работают структурированные псевдоспектры в задачах гидродинамической устойчивости на примере проблем собственных значений, возникающих при исследовании пространственной устойчивости течения вязкой несжимаемой жидкости над вогнутой поверхностью малой кривизны (основное течение) при параметрах течения благоприятных для развития вихрей Гертлера и волн Толлмина–Шлихтинга. Во втором разделе мы дадим определение структурированного псевдоспектра и приведем его основные свойства. В третьем разделе опишем рассматриваемую задачу гидродинамической устойчивости и вы-

бор структурирующих матриц. В четвертом разделе приведем и обсудим результаты исследования чувствительности собственных значений к возмущению поля скорости основного течения. Итог работы подведем в пятом разделе.

Всюду далее мы будем использовать следующие обозначения: $\|\cdot\|_2$ и $\|\cdot\|_\infty$ — соответственно вторая и бесконечная нормы матриц и векторов, $\text{diag}(d)$ — диагональная матрица с диагональными элементами равными компонентам вектора d .

2. СТРУКТУРИРОВАННЫЕ ПСЕВДОСПЕКТРЫ

Структурированным псевдоспектром матрицы $A \in \mathbb{C}^{n \times n}$ называют множество вида

$$\Lambda_\varepsilon(A; B, C) = \bigcup_{\|S\|_2 \leq \varepsilon} \lambda(A + BSC), \quad S \in \mathbb{C}^{n \times n},$$

где B и $C \in \mathbb{C}^{n \times n}$ — фиксированные матрицы. В частности, $\Lambda_\varepsilon(A; I, I) = \Lambda_\varepsilon(A)$.

Конструктивное определение структурированного псевдоспектра дает следующая формула:

$$\Lambda_\varepsilon(A; B, C) = \lambda(A) \bigcup \{z : \|C(zI - A)^{-1}B\|_2 \geq 1/\varepsilon\}.$$

При ее обосновании используют одно утверждение, представляющее самостоятельный интерес. А именно: если комплексное число z не принадлежит спектру матрицы A и норма структурированной резольвенты в точке z равна $s > 0$, то z является собственным значением возмущенной матрицы $A + BSC$, где $S = vu^*/s$, а v и u соответственно нормированные правый и левый сингулярные векторы структурированной резольвенты, отвечающие ее максимальному сингулярному числу s , то есть $C(zI - A)^{-1}Bv = su$. Отметим, что указанная матрица S — это минимальная по норме матрица, делающая z собственным значением возмущенной матрицы $A + BSC$.

Построение структурированного спектрального портрета в заданной области комплексной плоскости сводится к вычислению линий уровня максимального сингулярного числа структурированной резольвенты $R(z) = C(zI - A)^{-1}B$ в узлах заданной прямоугольной сетки по z . Если для этого использовать метод Ланцоша, то вычислительные затраты при каждом значении z составят несколько десятков умножений матриц $R(z)$ и $R(z)^*$ на вектор. В случае плотных матриц небольшого порядка ($n \leq 10^4$ для вычислений на современном персональном компьютере) целесообразно выполнить разложение Шура $A = QTQ^*$, где Q — унитарная матрица, а T — верхняя треугольная (форма Шура), вычислить матрицы $\tilde{C} = CQ$ и $\tilde{B} = Q^*B$ и использовать тождество $R(z) = \tilde{C}(zI - T)^{-1}\tilde{B}$. Это позволит умножать на вектор матрицы $R(z)$ и $R(z)^*$ за $O(n^2)$ арифметических операций с небольшой мультипликативной константой. Суммарные вычислительные затраты на построение спектрального портрета составят $O(n^3) + O(Nn^2)$ арифметических операций, где N — число узлов сетки по z .

3. РАССМАТРИВАЕМЫЕ ЗАДАЧИ

Следуя работам [23], [24], рассмотрим типичную задачу пространственной устойчивости пограничного слоя, а именно течение вязкой несжимаемой жидкости над вогнутой пластиной бесконечного размаха и постоянного радиуса кривизны, помещенной под нулевым углом атаки и с нулевым углом скольжения в однородный набегающий поток. Будем использовать ортогональные криволинейные координаты (x, y, z) , где y — расстояние до пластины по нормали к поверхности, x — расстояние от передней кромки пластины до основания этой нормали вдоль поверхности пластины, z — поперечная координата. Будем предполагать, что требуется исследовать устойчивость основного течения к внешним возмущениям вида

$$(u', v', w', p') = \text{Real}\{(u, v, w, p)e^{i(\beta z - \omega t)}\}, \quad (3.1)$$

которые сгенерированы гармоническим источником [23], расположенным на некотором расстоянии от передней кромки пластины. Здесь u, v, w и p — зависящие от x и y комплекснозначные амплитуды продольной, нормальной и поперечной компонент скорости и давления соответственно, t — время, β и ω — поперечное волновое число и угловая частота соответственно, а i — мнимая единица.

Предполагая толщину пограничного слоя малой по сравнению с радиусом кривизны пластины, обезразмеренные линеаризованные уравнения распространения амплитуд возмущений можно записать в следующем виде:

$$\begin{aligned}
U \frac{\partial u}{\partial x} - \frac{\partial V}{\partial y} u + V \frac{\partial u}{\partial y} + \frac{\partial U}{\partial y} v + \frac{\partial p}{\partial x} &= \frac{1}{Re} \frac{\partial^2 u}{\partial x^2} + L_{\beta\omega} u, \\
U \frac{\partial v}{\partial x} + \frac{\partial V}{\partial x} u + \frac{\partial V}{\partial y} v + V \frac{\partial v}{\partial y} + 2G\ddot{\sigma}^2 U u + Re \frac{\partial p}{\partial y} &= \frac{1}{Re} \frac{\partial^2 v}{\partial x^2} + L_{\beta\omega} v, \\
U \frac{\partial w}{\partial x} + V \frac{\partial w}{\partial y} + i\beta Re p &= \frac{1}{Re} \frac{\partial^2 w}{\partial x^2} + L_{\beta\omega} w, \\
\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + i\beta w &= 0.
\end{aligned} \tag{3.2}$$

Здесь

$$L_{\beta\omega} = \frac{\partial^2}{\partial y^2} - \beta^2 + i\omega,$$

$U(x, y)$ и $V(x, y)$ — заданные продольная и нормальная компоненты скорости основного течения соответственно, Re — число Рейнольдса, $G\ddot{\sigma}$ — число Гертлера (случай $G\ddot{\sigma} = 0$ соответствует плоской пластине). Для возмущений предполагается прилипание на поверхности пластины и затухание вдали от пограничного слоя:

$$u = v = w = 0, \quad \text{при } y = 0, y = +\infty. \tag{3.3}$$

Следуя работе [24], аппроксимируем уравнения (3.2) с краевыми условиями (3.3) по y методом Галеркина—коллокаций [25] с базисными и пробными функциями, представляющими собой модификацию функций предложенных в работах [26], [27], а именно,

$$\psi_i(y) = \ell_i(y)e^{-(y-y_i)/2}, \quad \phi_i(y) = \tilde{\ell}_i(y)e^{-(y-y_i)/2},$$

где $\ell_i(y)$ — это элементарные интерполяционные многочлены Лагранжа для сетки Лагерра—Гаусса—Радона $0 = y_0 < y_1 < \dots < y_{n_y}$, ненулевые узлы которой являются корнями производной многочленов Лагерра степени $n_y + 1$, а $\tilde{\ell}_i(y)$ — это элементарные интерполяционные многочлены Лагранжа для сетки $y_1 < \dots < y_{n_y}$. Функции $\psi_i(y)$ будем использовать в качестве базисных функций для компонент скорости и пробных функций для уравнений переноса импульса. Функции $\phi_i(y)$ будем использовать в качестве базисных функций для давления и пробных функций для уравнения неразрывности. Через D_{y0} и D_y будем обозначать прямоугольную матрицу размера $(n_y + 1) \times n_y$ и квадратную порядка n_y , которые при умножении на вектор значений в узлах $y_1 < \dots < y_{n_y}$ функции, принимающей нулевое значение в узле y_0 , дают вектор приближенных значений первой производной этой функции в узлах $y_0 < \dots < y_{n_y}$ и $y_1 < \dots < y_{n_y}$, соответственно. Через K_0 обозначим квадратную диагональную матрицу порядка $n_y + 1$ значений квадратных корней весов квадратурной формулы Лагерра—Гаусса—Радона, а через K — ее подматрицу отвечающую внутренним узлам сетки. Через Π обозначим прямоугольную матрицу размера $(n_y + 1) \times n_y$ проектирования с сетки $y_1 < \dots < y_{n_y}$ на сетку $y_0 < \dots < y_{n_y}$, через u, v, w, p — вектора значений соответствующих функций во внутренних узлах расчетной сетки.

Вводя дополнительные переменные

$$\tilde{v} = \frac{1}{Re} \frac{dv}{dx}, \quad \tilde{w} = \frac{1}{Re} \frac{dw}{dx},$$

полученную после аппроксимации системы (3.2) систему обыкновенных дифференциальных уравнений второго порядка по x сведем к следующей системе дифференциальных уравнений первого порядка:

$$\frac{dq}{dx} = A(x)q \tag{3.4}$$

относительно неизвестной вектор-функции $q = (u^T, v^T, w^T, \tilde{v}^T, \tilde{w}^T, p^T)^T$, где

$$A = \begin{bmatrix}
0 & -F_y & -i\beta I & 0 & 0 & 0 \\
0 & 0 & 0 & ReI & 0 & 0 \\
0 & 0 & 0 & 0 & ReI & 0 \\
V_x + 2G\ddot{\sigma}^2 U & V_y + V_D - L_{\beta\omega} & 0 & ReU & 0 & ReG_y \\
0 & 0 & V_D - L_{\beta\omega} & 0 & ReU & i\beta ReI \\
V_y - V_D + L_{\beta\omega} & UF_y - U_y & i\beta U & -F_y & -i\beta I & 0
\end{bmatrix}$$

является квадратной блочной матрицей блочного порядка 6, состоящей из квадратных блоков порядка n_y , некоторые из которых зависят от x . Здесь U, V, U_y, V_x, V_y — диагональные матрицы значений соответствующих

компонент скорости основного течения и их производных во внутренних узлах сетки по y , $V_D = KVD_y K^{-1}$ — дискретный аналог оператора $V\partial/\partial y$, $L_{\beta\omega} = L_\beta + i\omega I$ — дискретный аналог оператора $L_{\beta\omega}$, где $L_\beta = -K^{-1}D_{y0}^T K_0^2 D_{y0} K^{-1} - \beta^2 I$ — дискретный аналог оператора $\partial^2/\partial y^2 - \beta^2$, матрицы $F_y = K^{-1}\Pi^T K_0^2 D_{y0} K^{-1}$ и $G_y = -F_y^T$ — дискретные аналоги оператора $\partial/\partial y$ соответственно в уравнении неразрывности (действующего на компоненту скорости) и в уравнениях движения (действующего на давление).

Значения компонент скорости основного течения будем считать изначально заданными в узлах некоторой сетки по x и y , отличной от сетки используемой для аппроксимации уравнений распространения амплитуд возмущений. В численных экспериментах, описанных в данной работе, эти значения рассчитывались на основе уравнения Блазиуса [1]. Значения производных компонент скорости основного течения будем считать вычисленными с помощью некоторых сеточных операторов дифференцирования. В качестве элементов диагональных матриц U , V , U_y , V_x и V_y будем брать результаты линейной интерполяции соответствующих сеточных функций в точки $(x, y_1), \dots, (x, y_{n_y})$ при рассматриваемом значении x .

Перейдем к описанию внутренних возмущений основного течения, как параметра задачи устойчивости. Будем предполагать, что компоненты скорости основного течения задаются с некоторой погрешностью, такой что их относительные возмущения — это независимые случайные величины, принимающие значения в интервале $[-\delta, \delta]$, где $0 < \delta < 1$. Векторы погрешностей основного течения и их сеточных производных в точках $(x, y_1), \dots, (x, y_{n_y})$, полученные так же, как для самого основного течения, будем обозначать через δ_U , δ_V , δ_{U_y} , δ_{V_x} , δ_{V_y} и предполагать, что справедливы следующие неравенства:

$$\|\delta_*\|_\infty \leq c_*\delta, \|\delta_{*x}\|_\infty \leq c_{*x}\delta, \|\delta_{*y}\|_\infty \leq c_{*y}\delta,$$

где c_* , c_{*x} , c_{*y} — некоторые неотрицательные константы. Эти константы не зависят от сетки $y_1 < \dots < y_{n_y}$, выбираемой для амплитуд возмущений, и могут быть легко вычислены, если известны матрицы операторов сеточного дифференцирования для компонент основного течения.

При сделанных предположениях матрица возмущения матрицы A будет иметь следующий вид:

$$\Delta = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \Delta_{41} & \Delta_{42} & 0 & \Delta_{44} & 0 & 0 \\ 0 & 0 & \Delta_{53} & 0 & \Delta_{55} & 0 \\ \Delta_{61} & \Delta_{62} & \Delta_{63} & 0 & 0 & 0 \end{bmatrix}, \quad (3.5)$$

где

$$\begin{aligned} \Delta_{41} &= \text{diag}(\delta_{V_x}) + 2G\ddot{o}^2 \text{diag}(\delta_U), \quad \Delta_{42} = \text{diag}(\delta_{V_y}) + K \text{diag}(\delta_V) D_y K^{-1}, \\ \Delta_{44} &= \text{Re} \text{diag}(\delta_U), \quad \Delta_{53} = K \text{diag}(\delta_V) D_y K^{-1}, \quad \Delta_{55} = \Delta_{44}, \quad \Delta_{63} = i\beta \text{diag}(\delta_U), \\ \Delta_{61} &= \text{diag}(\delta_{V_y}) - K \text{diag}(\delta_V) D_y K^{-1}, \quad \Delta_{62} = \text{diag}(\delta_U) F_y - \text{diag}(\delta_{U_y}), \end{aligned}$$

а для норм блоков матрицы Δ будут справедливы следующие верхние оценки:

$$\begin{aligned} \|\Delta_{41}\|_2/\delta &\leq h_{41} = c_{V_x} + 2c_U G\ddot{o}^2, \quad \|\Delta_{42}\|_2/\delta \leq h_{42} = c_{V_y} + c_V \|KD_y K^{-1}\|_2, \\ \|\Delta_{44}\|_2/\delta &\leq h_{44} = c_U \text{Re}, \quad \|\Delta_{53}\|_2/\delta \leq h_{53} = c_V \|KD_y K^{-1}\|_2, \quad \|\Delta_{55}\|_2/\delta \leq h_{55} = h_{44}, \\ \|\Delta_{61}\|_2/\delta &\leq h_{61} = h_{42}, \quad \|\Delta_{62}\|_2/\delta \leq h_{62} = c_{U_y} + c_U \|F_y\|_2, \quad \|\Delta_{63}\|_2/\delta \leq h_{63} = c_U |\beta|. \end{aligned}$$

Следовательно, $\|\Delta\|_2 \leq \delta \|H\|_2$, где H — квадратная матрица порядка 6 с указанными выше ненулевыми элементами h_{kl} .

Для структурирования возмущений введем в рассмотрение две блочно-диагональные матрицы вида

$$B = \text{diag}(0, 0, 0, b_4 I, b_5 I, b_6 I), \quad C = \text{diag}(c_1 M, c_2 M, c_3 M, c_4 I, c_5 I, 0)$$

и псевдообратные к ним матрицы

$$B^+ = \text{diag}(0, 0, 0, I/b_4, I/b_5, I/b_6), \quad C^+ = \text{diag}(M^{-1}/c_1, M^{-1}/c_2, M^{-1}/c_3, I/c_4, I/c_5, 0).$$

Здесь M — невырожденная матрица, обратная к которой имеет единичную вторую норму: $\|M^{-1}\|_2 = 1$, а b_* и c_* — положительные числа. Тогда матрица $S = B^+ \Delta C^+$ будет иметь ту же блочную структуру, что и матрица Δ , причем $\Delta = BSC$, а нормы ненулевых блоков матрицы S будут оцениваться следующим образом:

$$\begin{aligned} \|S_{41}\|_2/\delta &\leq \tilde{h}_{41} = h_{41}/(b_4 c_1), \quad \|S_{42}\|_2/\delta \leq \tilde{h}_{42} = (c_{V_y} + c_V \|KD_y K^{-1} M^{-1}\|_2)/(b_4 c_2), \\ \|S_{44}\|_2/\delta &\leq \tilde{h}_{44} = h_{44}/(b_4 c_4), \quad \|S_{53}\|_2/\delta \leq \tilde{h}_{53} = c_V \|KD_y K^{-1} M^{-1}\|_2/(b_5 c_3), \\ \|S_{55}\|_2/\delta &\leq \tilde{h}_{55} = h_{55}/(b_5 c_5), \quad \|S_{61}\|_2/\delta \leq \tilde{h}_{61} = (c_{V_y} + c_V \|KD_y K^{-1} M^{-1}\|_2)/(b_6 c_1), \\ \|S_{62}\|_2/\delta &\leq \tilde{h}_{62} = (c_{U_y} + c_U \|F_y M^{-1}\|_2)/(b_6 c_2), \quad \|S_{63}\|_2/\delta \leq \tilde{h}_{63} = h_{63}/(b_6 c_3). \end{aligned}$$

Следовательно, $\|S\|_2 \leq \delta \|\tilde{H}\|_2$, где $\tilde{H}$ — квадратная матрица порядка b с указанными выше ненулевыми элементами $\tilde{h}_{kl}$.

В качестве матрицы M будем выбирать эрмитову положительно определенную матрицу $(-L_\beta)^{-1/2}$, нормированную на ее минимальное собственное значение. При таком выборе матрицы M нормы матриц $F_y M^{-1}$ и $K D_y K^{-1} M^{-1}$ будут оставаться ограниченными с ростом числа узлов сетки Лагерра—Гаусса—Радо, что обеспечит ограниченность нормы матрицы S . Параметры структурирования b_* и c_* будем выбирать решая задачу минимизации:

$$\|\tilde{H}\|_2 \|C(z_0 I - A)^{-1} B\|_2 \rightarrow \min, \quad (3.6)$$

по этим параметрам, где z_0 — некоторая контрольная точка на комплексной плоскости, выбираемая в области представляющих интерес собственных значений матрицы A .

4. РЕЗУЛЬТАТЫ

В этом разделе мы обсудим результаты расчета спектра возмущенной матрицы A системы (3.4) и его локализации с помощью псевдоспектров при $x = 2$ и $n_y = 32$, обеспечивающем сходимость результатов по шагу сетки. Будем использовать следующие два набора значений параметров задачи (3.4): $Re = 2.003 \times 10^5$, $G\ddot{o} = 3.66$, $\beta = 0.2106$, $\omega = 0.1885$ и $Re = 4.674 \times 10^5$, $G\ddot{o} = 6.92$, $\beta = 6.433 \times 10^{-3}$, $\omega = 21.99$, взятых из работ [23], [24] и благоприятных для развития вихрей Гертлера и волн Толлмина—Шлихтинга соответственно. Конфигурации, соответствующие этим наборам, будем обозначать через G и TS .

Будем возмущать поле скоростей основного течения как описано в предыдущем разделе, полагая $\delta = 10^{-3}$. Собственные значения возмущенных матриц $A + \Delta$ будем обозначать на рисунках белыми точками, собственные значения исходной матрицы A — черными. Примерно половина собственных значений матрицы A отвечает собственным модам, амплитуда которых сильно нарастает вниз по потоку, а сами эти моды распространяются вверх по потоку [28]. Такие моды удовлетворяют уравнениям вязкой несжимаемой жидкости, но их вклад в решения наблюдаемые на практике, пренебрежимо мал. Поэтому, эти моды исключают из рассмотрения (см. [23], [24], [28]). Оставшиеся, физически значимые собственные значения при рассмотренном значении x включают одно ведущее (с наибольшей вещественной частью) собственное значение, представляющее наибольший интерес (на рисунках оно отмечено косым черным крестом). Мода, отвечающая этому собственному значению, для набора G является вихрем Гертлера, а для набора TS — волной Толлмина—Шлихтинга. Обе эти моды неустойчивые, т.е. вещественные части соответствующих им собственных значений положительны. Естественный ламинарно—турбулентный переход происходит вследствие развития таких мод вниз по потоку.

Мы локализуем спектры возмущенных матриц с помощью неструктурированного псевдоспектра и с помощью структурированного псевдоспектра. Для этого мы строим в областях собственных значений, представляющих наибольший интерес, неструктурированный и структурированный спектральные портреты, то есть линии уровня функций

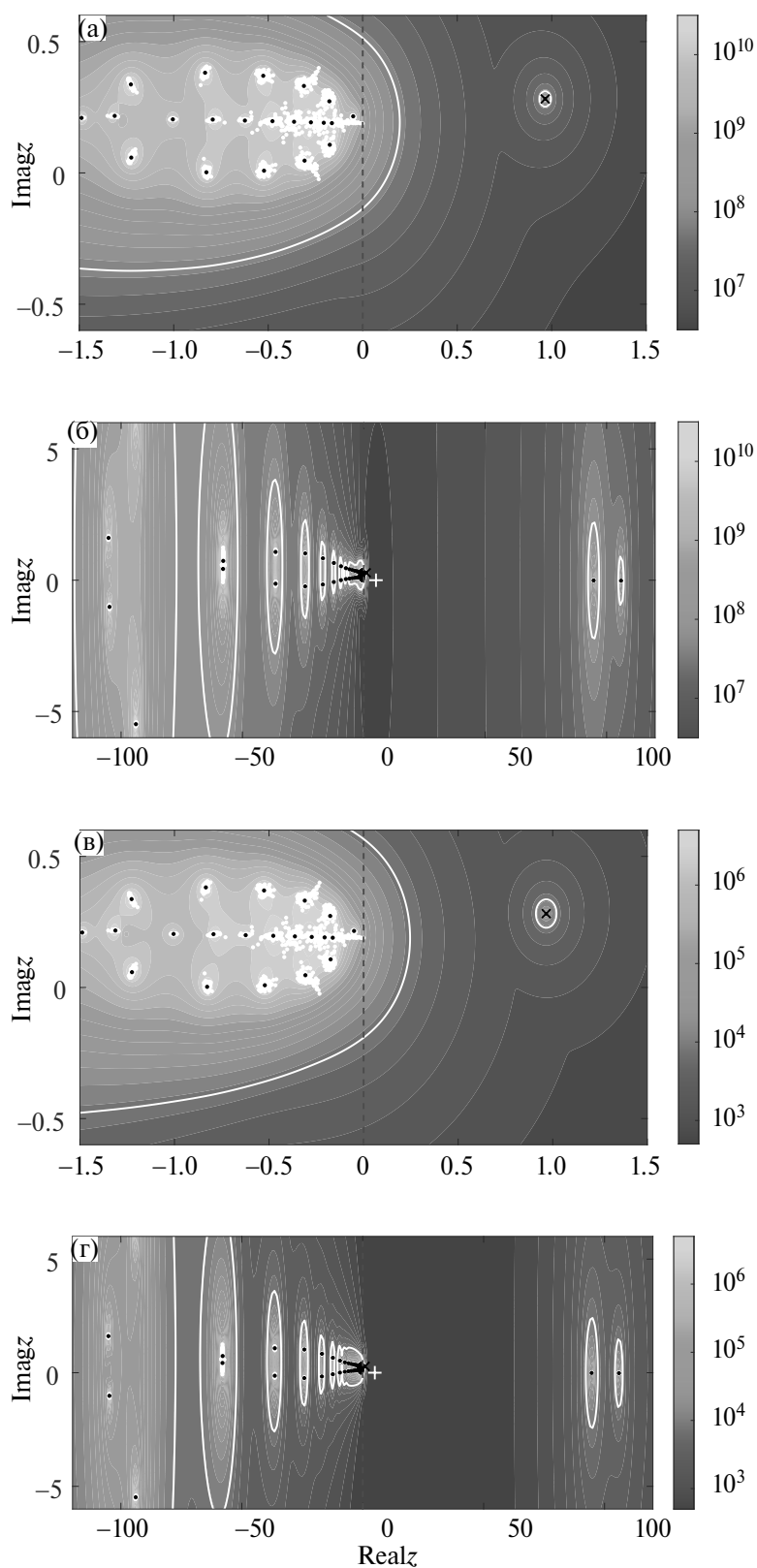
$$f(z) = \|H\|_2 \| (zI - A)^{-1} \|_2$$

и

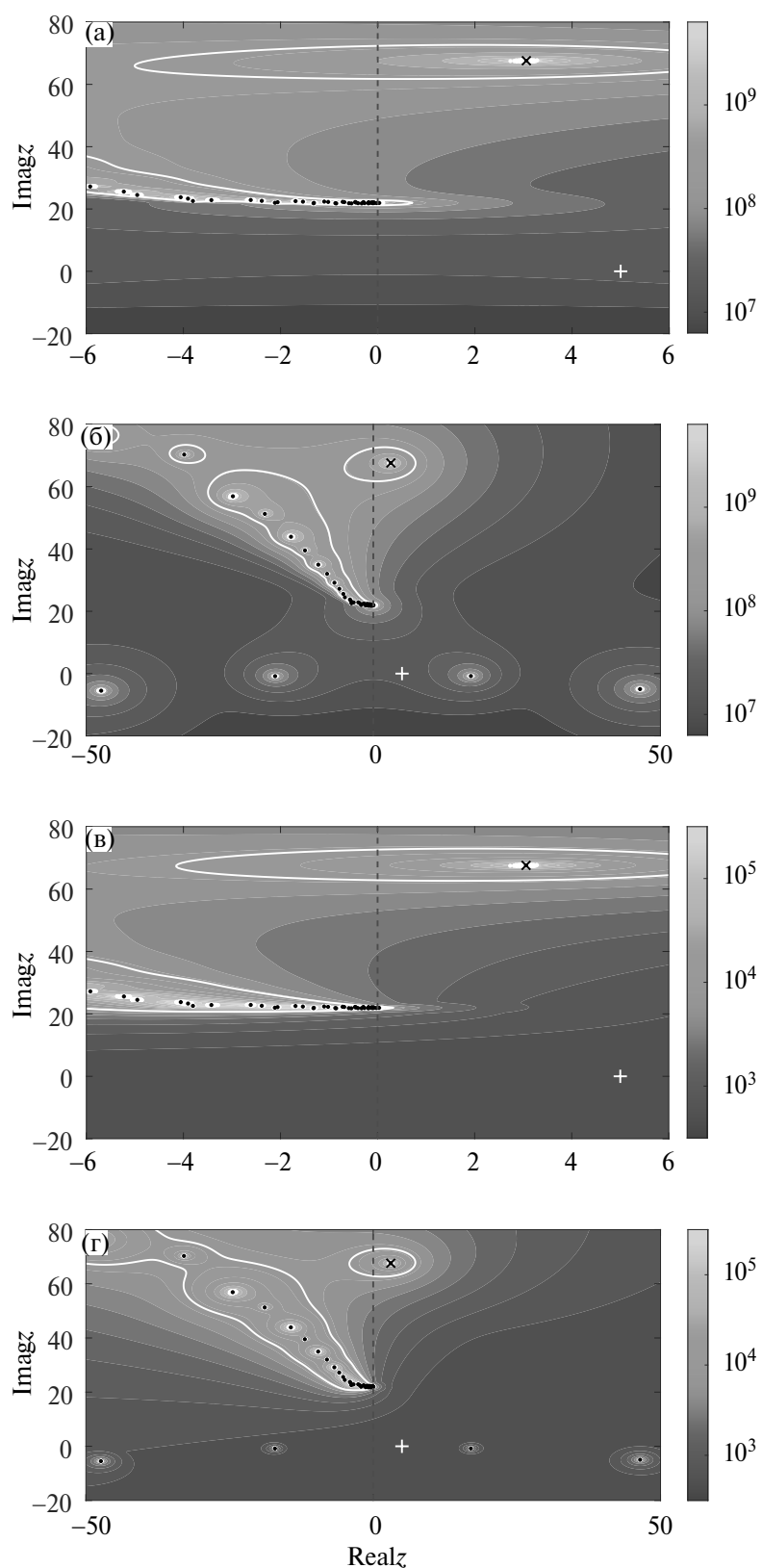
$$\tilde{f}(z) = \|\tilde{H}\|_2 \|C(zI - A)^{-1} B\|_2$$

соответственно. Если для некоторого δ справедливо неравенство $1/f(z) > \delta$, либо $1/\tilde{f}(z) > \delta$, то точка z заведомо не является собственным значением возмущенной матрицы при этом значении δ . На каждый спектральный портрет мы будем наносить (белым) линию уровня, отвечающую некоторому значению нормы резольвенты. При δ , меньших обратной величине значения на этой линии, возмущения основного течения не оказывают существенного влияния на ведущую моду и не приводят к перемешиванию физически значимых и незначимых собственных значений, то есть не нарушают работу алгоритмов, обычно используемых для исследования гидродинамической устойчивости.

На фиг. 1 для конфигурации G неструктурированный спектральный портрет матрицы A (подрисунки (а) и (б)) сравнивается со структурированным (подрисунки (в), (г)). Кроме того, собственные значения матрицы A сравниваются с собственными значениями возмущенных матриц $A + \Delta$ при ста различных возмущениях основного течения. При этом подрисунки (а) и (в) демонстрируют спектральные портреты вблизи физический значимого ведущего собственного значения, а подрисунки (б) и (г) — в более широкой области комплексной плоскости, где видны два собственных значения с действительной частью, примерно равной 100, которые хорошо отделены от остальной части спектра и отвечают физически незначимым модам. Белая линия для неструктурированного спектрального портрета отвечает значению $f(z) = 10^8$, а для структурированного — $\tilde{f}(z) = 10^4$, что есть $\delta = 10^{-8}$ и $\delta = 10^{-4}$ соответственно. При этом эти линии геометрически близки. Следовательно, учет специфики возмущений основного течения в рамках структурированного псевдоспектра позволил получить значительно (на четыре порядка) более точную оценку чувствительности спектра к возмущениям основного



Фиг. 1. Неструктурированные (а), (б) и структурированные (в), (г) спектральные портреты для конфигурации G. Черная пунктирная линия — мнимая ось, черные точки — спектр матрицы A, черный косой крест — ее ведущее собственное значение, отвечающее вихрю Гертлера, белые точки — спектры ста возмущенных матриц при $\delta = 10^{-3}$, белый прямой крест — точка z_0 , используемая при оптимизации (3.6), белая линия на (а) и (б) — линия уровня $f(z) = 10^8$, на (в) и (г) — линия уровня $\hat{f}(z) = 10^4$.



Фиг. 2. Неструктурированные (а), (б) и структурированные (в), (г) спектральные портреты для конфигурации TS. Черная пунктирная линия — мнимая ось, черные точки — спектр матрицы A , черный косой крест — ее ведущее собственное значение, отвечающее волне Толлмина–Шлихтинга, белые точки — спектры ста возмущенных матриц при $\delta = 10^{-3}$, белый прямой крест — точка z_0 , используемая при оптимизации (3.6), белая линия на (а) и (б) — линия уровня $f(z) = 10^{8.5}$, на (в) и (г) — линия уровня $f(z) = 10^4$.

течения. Вместе с тем, собственные значения ста возмущенных матриц, не достигают белой линии, нанесенной на структурированный псевдоспектр, несмотря на то, что они вычислены при $\delta = 10^{-3}$, то есть для возмущений на порядок больших. Это позволяет предполагать, что оценка на основе структурированного псевдоспектра все еще остается завышенной и может быть улучшена.

На фиг. 2 с использованием тех же обозначений, что и для фиг. 1, представлены результаты расчетов для конфигурации TS. В данном случае белая линия на неструктурированном спектральном портрете отвечает значению $f(z) = 10^{8.5}$, а на структурированном — $\tilde{f}(z) = 10^4$. То есть и для этой конфигурации структурированный псевдоспектр дает значительно (более, чем на четыре порядка) более точную оценку чувствительности спектра возмущенной матрицы, чем обычный псевдоспектр. Следует отметить, что в отличие от конфигурации G, для конфигурации TS и неструктурированный, и структурированный псевдоспектры допускают смещение ведущего физически значимого собственного значения в левую полуплоскость. Кроме того, ведущие физически значимые собственные значения ста возмущенных матриц заметно сильнее отличаются от ведущего физически значимого собственного значения исходной матрицы. Тем не менее, эти собственные значения возмущенных матриц лежат в правой полуплоскости, имея достаточно большие положительные действительные части. Это также позволяет предполагать, что оценка на основе структурированного псевдоспектра все еще завышена и для конфигурации TS.

5. ЗАКЛЮЧЕНИЕ

Данная работа посвящена оценке чувствительности собственных значений к внутренним возмущениям поля основного течения в задачах гидродинамической устойчивости. Предложено использовать структурированные псевдоспектры вместо широко используемых обычных псевдоспектров. На примере течения вязкой несжимаемой жидкости над вогнутой поверхностью малой кривизны при параметрах течения благоприятных для развития вихрей Гёртлера и волн Толлмина—Шлихтинга, было показано, что структурирование можно выбрать так, что оценки на основе структурированного псевдоспектра будут значительно более точными, чем оценки, полученные на основе неструктурированного псевдоспектра.

СПИСОК ЛИТЕРАТУРЫ

1. *Schlichting H., Gersten K.* Boundary-Layer Theory (9th ed.). Berlin/Heidelberg: Springer-Verlag, 2016.
2. *Boiko A.V., Dovgal A.V., Grek G.R., Kozlov V.V.* Physics of Transitional Shear Flows: instability and laminar-turbulent transition in incompressible near-wall shear layers. Berlin: Springer-Verlag, 2011. 272 p.
3. *Schmid P.J., Henningson D.S.* Stability and Transition in Shear Flows. New York: Springer New York, 2000. 558 p.
4. *Bauer F.L., Fike C.T.* Norms and exclusion theorems // Numer. Math. 1960. V. 2. N. 3. P. 137–141.
5. *Уилкинсон Дж.Х.* Алгебраическая проблема собственных значений. М.: Наука, 1970.
6. *Голуб Дж., Ван Лоун Ч.* Матричные вычисления. М.: Мир, 1999.
7. *Годунов С.К., Кирилук О.П., Костин В.И.* Спектральные портреты матриц // Препринт № 3, Институт Математики СО АН СССР, 1990.
8. *Годунов С.К.* Современные аспекты линейной алгебры. Новосибирск: Научная Книга, 1997. 390 с.
9. *Trefethen L.N.* Pseudospectra of matrices // Numerical Analysis (ed. by D.F. Griffiths and G.A. Watson), 1991. P. 234–266.
10. *Trefethen L.N., Embree M.* Spectra and Pseudospectra: The Behavior of Non-normal Matrices and Operators. Princeton: Princeton Univer. Press, 2005. 606 p.
11. *Годунов С.К.* Задача о дихотомии спектра матрицы // Сиб. матем. журн. 1986. Т. 27. № 5. С. 24–37.
12. *Булгаков А.Я., Годунов С.К.* Круговая дихотомия матричного спектра // Сиб. матем. журн. 1988. Т. 29. № 5. С. 59–70.
13. *Godunov S.K., Sadkane M.* Computation of pseudospectra via spectral projectors // Lin. Alg. Appl., 1998. V. 279. P. 163–175.
14. *Нечепуренко Ю.М.* Интегральные критерии качества дихотомии замкнутым контуром // Матем. заметки. 2005. Т. 78. № 5. С. 718–726.

15. Нечепуренко Ю.М. Оценка нормы матрицы Грина через интегральный критерий качества дихотомии и границы хаусдорфова множества // Матем. заметки. 2002. Т. 71. № 2. С.232–238.
16. Нечепуренко Ю.М. Оценка нормы матричной экспоненты через норму решения уравнения Ляпунова и границы хаусдорфова множества // Ж. вычисл. матем. и матем. физ. 2002. Т. 42. № 2, С. 131–141.
17. *Nechepurenko Yu.M., Zasko G.V.* Constant upper bounds on the matrix exponential norm. // *Russ. J. Num. Anal. Math. Model.* 2022. V. 37. № 1. P. 15–23.
18. *Trefethen L.N., Trefethen A.E., Reddy S.C., Driscoll T.A.* Hydrodynamic stability without eigenvalues // *Science*. 1993. V. 261. P. 578–584.
19. *Foster R.C.* Structure and energetics of optimal Ekman layer perturbations // *J. Fluid Mech.* 1997. V. 333. P. 97–123.
20. *Hinrichsen D., Kelb B.* Spectral value sets: a graphical tool for robustness analysis // *Systems Control Lett.* 1993. V. 21. P. 127–136.
21. *Gallestay E., Hinrichsen D., Pritchard A.J.* Spectral value sets of closed linear operators // *Proc. Roy. Soc. London Ser. A*. 2000. V. 456. P. 1397–1418.
22. *Nechepurenko Yu.M.* The regularly structured pseudospectrum // *Russ. J. Numer. Anal. Math. Model.* 2004. V. 19. № 3. P. 265–288.
23. Бойко А.В., Демьянко К.В., Засько Г.В., Нечепуренко Ю.М. О параболизации уравнений распространения малых возмущений в двумерных пограничных слоях // Теплофизика и Аэромеханика. 2024. Т. 31. № 3. С. 423–440. sibran.ru/journals/issue.php?ID=189093&ARTICLE_ID=189094
24. *Zasko G.V., Boiko A.V., Demyanko K.V., Nechepurenko Yu.M.* Simulating the propagation of boundary-layer disturbances by solving boundary-value and initial-value problems // *Russ. J. Num. Anal. Math. Model.* 2024. V. 39. № 1. P. 47–59.
25. *Canuto C., Hussaini M.Y., Quarteroni A., Zang T.A.* Spectral Methods: Fundamentals in Single Domain. Berlin/Heidelberg: Springer-Verlag. 2006. 563 p.
26. *Shen J.* Stable and efficient spectral methods in unbounded domains using Laguerre functions // *SIAM J. Numer. Anal.* 2000. V. 28. P. 1113–1133.
27. *Shen J., Wang L.-L.* Some recent advances on spectral methods for unbounded domains // *Comm. Comp. Phys.* 2009. V. 5. P. 195–241.
28. *Tumin A., Reshotko E.* Spatial theory of optimal disturbances in boundary layers // *Phys. Fluids*, 2001. V. 13. № 7. P. 2097–2104.

STRUCTURED PSEUDOSPECTRA IN PROBLEMS OF SPATIAL STABILITY OF BOUNDARY LAYERS

K. V. Demyanko^{a,b}, Yu. M. Nechepurenko^{a,b}, G. V. Zasko^{a,b,*}

^a *Marchuk Institute of Numerical Mathematics, Russian Academy of Sciences, Gubkina St., 8, Moscow, 119333 Russia*

^b *Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, Miusskaya sq., 4, Moscow, 125047 Russia*

**e-mail: yumnech@yandex.ru*

Received 23 February, 2024

Revised 23 February, 2024

Accepted 02 May, 2024

Abstract. The paper is devoted to the numerical analysis of the sensitivity of the characteristics of spatial stability of boundary layers to the errors with which the main flow is specified. It is proposed to use structured pseudospectra for this purpose. It is shown that the obtained estimates are significantly more accurate than the estimates based on the unstructured pseudospectrum. The presentation is carried out using the example of a viscous incompressible fluid flow over a concave surface of small curvature with flow parameters favorable for the development of Goertler vortices and Tollmien–Schlichting waves.

Keywords: structured pseudospectra, resolvent, spatial stability, boundary layer, Goertler vortices, Tollmien–Schlichting waves.

ОЦЕНКИ РЕШЕНИЙ КЛАССА НЕАВТОНОМНЫХ СИСТЕМ НЕЙТРАЛЬНОГО ТИПА С СОСРЕДОТОЧЕННЫМ И РАСПРЕДЕЛЕННЫМ ЗАПАЗДЫВАНИЯМИ¹⁾

© 2024 г. И. И. Матвеева^{1,*}¹ 630090 Новосибирск, ул. Пирогова, 1, Новосибирский государственный университет, Россия

*e-mail: i.matveeva@g.nsu.ru

Поступила в редакцию 02.04.2024 г.

Переработанный вариант 02.04.2024 г.

Принята к публикации 02.05.2024 г.

Рассмотрен класс систем неавтономных дифференциальных уравнений нейтрального типа с сосредоточенным и распределенным запаздываниями. Используя функционал Ляпунова–Красовского, установлены оценки, которые позволяют сделать вывод об устойчивости решений. В случае экспоненциальной устойчивости указаны оценки на скорость стабилизации решений на бесконечности. Библ. 31.

Ключевые слова: системы нейтрального типа, переменные коэффициенты, оценки решений, устойчивость, функционал Ляпунова–Красовского.

DOI: 10.31857/S0044466924080139, EDN: YAEUDG

1. ВВЕДЕНИЕ

Мы рассматриваем системы дифференциальных уравнений с запаздыванием следующего вида:

$$\begin{aligned} \frac{d}{dt}y(t) = A(t)y(t) + B(t)y(t - \tau) + C(t)\frac{d}{dt}y(t - \tau) + \int_{t-\tau}^t D(t, t-s)y(s) ds, \\ + F\left(t, y(t), y(t - \tau), \frac{d}{dt}y(t - \tau)\right), \quad t \geq 0, \end{aligned} \quad (1.1)$$

где $A(t)$, $B(t)$, $C(t)$, $D(t, s)$ — матрицы размера $n \times n$ с непрерывными вещественнозначными элементами; т.е.

$$a_{ij}(t), b_{ij}(t), c_{ij}(t) \in C(\overline{\mathbb{R}}_+), \quad d_{ij}(t, s) \in C(\overline{\mathbb{R}}_+ \times [0, \tau]), \quad i, j = 1, \dots, n,$$

$\tau > 0$ — параметр запаздывания, $F(t, u_1, u_2, u_3)$ — непрерывная вещественнозначная вектор-функция. Мы предполагаем, что $F(t, u_1, u_2, u_3)$ липшицева по u_1 на любом компакте $G \subset [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ и удовлетворяет неравенству

$$\|F(t, u_1, u_2, u_3)\| \leq q_1\|u_1\| + q_2\|u_2\|, \quad t \geq 0, \quad u_j \in \mathbb{R}^n, \quad q_j \geq 0. \quad (1.2)$$

Наша цель — получить оценки для решений системы (1.1) на всей полуоси $\{t \geq 0\}$, на основе которых можно сделать вывод об устойчивости решений и указать скорость стабилизации.

Существует большое количество работ, посвященных изучению устойчивости решений дифференциальных уравнений с запаздыванием (например, см. [1]–[12] и библиографию в них). Исследователи часто используют функционалы Ляпунова–Красовского для получения условий устойчивости. Однако не каждый функционал Ляпунова–Красовского позволяет получить оценки, характеризующие скорость убывания на бесконечности. В последние годы исследования в этом направлении активно развиваются. Много работ посвящено дифференциальным уравнениям с запаздыванием и постоянными коэффициентами. В неавтономном случае количество соответствующих статей значительно меньше.

¹⁾ Работа выполнена при финансовой поддержке РНФ (код проекта № 24-21-00367), <https://rscf.ru/project/24-21-00367/>.

Эта статья продолжает наши исследования устойчивости решений неавтономных дифференциальных уравнений с запаздыванием (например, см. [13]–[19]). Мы исследовали системы с запаздыванием и периодически-ми коэффициентами в линейных членах. Были установлены условия экспоненциальной устойчивости нулевого решения и получены оценки экспоненциального убывания решений на бесконечности с использованием подходящих функционалов Ляпунова–Красовского. В [20] были изучены некоторые нелинейные системы с переменными коэффициентами и переменным сосредоточенным запаздыванием. В случае $C(t) \equiv 0$ системы вида (1.1) с переменными сосредоточенным и распределенным запаздываниями исследованы в [21].

В этой статье мы рассмотрим неавтономные системы вида (1.1) с сосредоточенным и распределенным запаздываниями в случае $C(t) \not\equiv 0$; иными словами, мы имеем дело с системами нейтрального типа. Мы устанавливаем оценки для решений, которые позволят нам сделать вывод о том, являются ли решения устойчивыми. В случае экспоненциальной устойчивости мы указываем оценки на скорость стабилизации решений на бесконечности. Во втором параграфе мы устанавливаем оценки для решений систем в линейном случае ($F(t, u_1, u_2, u_3) \equiv 0$), в третьем параграфе — для решений нелинейных систем вида (1.1). Некоторые примеры даны в четвертом параграфе.

2. ОЦЕНКИ РЕШЕНИЙ ЛИНЕЙНЫХ СИСТЕМ

В этом параграфе мы рассмотрим линейные системы вида (1.1) ($F(t, u_1, u_2, u_3) \equiv 0$). Вначале введем некоторые обозначения. Определим $(2n \times 2n)$ -матрицу

$$\mathcal{H}(t) = \begin{pmatrix} H_1(t) & H_2(t) \\ H_2^*(t) & H_3(t) \end{pmatrix}$$

такую, что

$$\mathcal{H}(t) \in C^1(\overline{\mathbb{R}}_+), \quad \mathcal{H}(t) = \mathcal{H}^*(t), \quad t \geq 0, \quad (2.1)$$

$$\left\langle \mathcal{H}(t) \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \right\rangle \geq h(t) \|u_1 - G(t)u_2\|^2, \quad u_1, u_2 \in \mathbb{R}^n, \quad t \geq 0, \quad (2.2)$$

где $h(t) \in C([0, \infty))$, $h(t) \geq h_0 > 0$, $G(t)$ — матрица размера $(n \times n)$ с непрерывными элементами. Определим матрицу $K(t, s)$ размера $(n \times n)$ такую, что

$$K(t, s) \in C^1(\overline{\mathbb{R}}_+ \times [0, \tau]), \quad K(t, s) = K^*(t, s), \quad K(t, s) > 0, \quad t \geq 0, \quad s \in [0, \tau]. \quad (2.3)$$

Здесь и далее матричное неравенство $S > 0$ (или $S < 0$) означает, что S — положительно (отрицательно) определенная эрмитова матрица. Для матриц мы используем спектральную норму.

Определим матрицу

$$Q(t, s) = \begin{pmatrix} Q_{11}(t, s) & Q_{12}(t, s) & Q_{13}(t, s) & Q_{14}(t, s) \\ Q_{12}^*(t, s) & Q_{22}(t, s) & Q_{23}(t, s) & Q_{24}(t, s) \\ Q_{13}^*(t, s) & Q_{23}^*(t, s) & Q_{33}(t, s) & Q_{34}(t, s) \\ Q_{14}^*(t, s) & Q_{24}^*(t, s) & Q_{34}^*(t, s) & Q_{44}(t, s) \end{pmatrix} \quad (2.4)$$

с элементами

$$\begin{aligned} Q_{11}(t, s) &= -\frac{d}{dt}H_1(t) - H_1(t)A(t) - A^*(t)H_1(t) - K(t, 0), \\ Q_{12}(t, s) &= -\frac{d}{dt}H_2(t) - A^*(t)H_2(t) - H_1(t)B(t), \\ Q_{13}(t, s) &= -H_1(t)C(t) - H_2(t), \\ Q_{14}(t, s) &= -\tau H_1(t)D(t, s), \\ Q_{22}(t, s) &= -\frac{d}{dt}H_3(t) - H_2^*(t)B(t) - B^*(t)H_2(t) + K(t, \tau), \\ Q_{23}(t, s) &= -H_2^*(t)C(t) - H_3(t), \\ Q_{24}(t, s) &= -\tau H_2^*(t)D(t, s), \\ Q_{33}(t, s) &= 0, \quad Q_{34}(t, s) = 0, \\ Q_{44}(t, s) &= -\tau \left(\frac{\partial}{\partial t}K(t, s) + \frac{\partial}{\partial s}K(t, s) \right). \end{aligned} \quad (2.5)$$

Рассмотрим начальную задачу для системы (1.1)

$$\begin{aligned} \frac{d}{dt}y(t) &= A(t)y(t) + B(t)y(t-\tau) + C(t)\frac{d}{dt}y(t-\tau) + \int_{t-\tau}^t D(t, t-s)y(s)ds, \quad t \geq 0, \\ y(t) &= \varphi(t), \quad t \in [-\tau, 0], \\ y(+0) &= \varphi(0), \end{aligned} \quad (2.6)$$

где $\varphi(t) \in C^1([-\tau, 0])$ — заданная вещественнозначная вектор-функция. Ниже мы устанавливаем оценки для решений задачи (2.6).

Теорема 1. *Предположим, что существуют матрицы $\mathcal{H}(t)$, $K(t, s)$, удовлетворяющие условиям (2.1)–(2.3) и такие, что*

$$\left\langle Q(t, s) \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} \right\rangle \geq p(t) \left\langle \mathcal{H}(t) \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \right\rangle + \tau k(t) \langle K(t, s)u_4, u_4 \rangle, \quad (2.7)$$

$$u_j \in \mathbb{R}^n, \quad t \geq 0, \quad s \in [0, \tau],$$

где $p(t), k(t) \in C(\overline{\mathbb{R}}_+)$. Тогда для решения $y(t)$ задачи (2.6) имеет место оценка

$$\|y(t) - G(t)y(t-\tau)\| \leq \sqrt{\frac{V(0, \varphi)}{h(t)}} \exp\left(-\frac{1}{2} \int_0^t \gamma(\xi) d\xi\right), \quad t > 0, \quad (2.8)$$

где

$$V(0, \varphi) = \left\langle \mathcal{H}(0) \begin{pmatrix} \varphi(0) \\ \varphi(-\tau) \end{pmatrix}, \begin{pmatrix} \varphi(0) \\ \varphi(-\tau) \end{pmatrix} \right\rangle + \int_{-\tau}^0 \langle K(0, -s)\varphi(s), \varphi(s) \rangle ds, \quad (2.9)$$

$$\gamma(t) = \min\{p(t), k(t)\}. \quad (2.10)$$

Доказательство. Пусть $y(t)$ — решение задачи (2.6). В [18] был введен достаточно широкий класс функционалов Ляпунова–Красовского. Используя матрицы $\mathcal{H}(t)$, $K(t, s)$, удовлетворяющие условиям теоремы 1, рассмотрим на $y(t)$ следующий функционал Ляпунова–Красовского из этого класса

$$V(t, y) = \left\langle \mathcal{H}(t) \begin{pmatrix} y(t) \\ y(t-\tau) \end{pmatrix}, \begin{pmatrix} y(t) \\ y(t-\tau) \end{pmatrix} \right\rangle + \int_{t-\tau}^t \langle K(t, t-s)y(s), y(s) \rangle ds. \quad (2.11)$$

Дифференцируя этот функционал, имеем

$$\begin{aligned} \frac{d}{dt}V(t, y) &= \left\langle \frac{d}{dt}\mathcal{H}(t) \begin{pmatrix} y(t) \\ y(t-\tau) \end{pmatrix}, \begin{pmatrix} y(t) \\ y(t-\tau) \end{pmatrix} \right\rangle + \left\langle \mathcal{H}(t) \begin{pmatrix} \frac{d}{dt}y(t) \\ \frac{d}{dt}y(t-\tau) \end{pmatrix}, \begin{pmatrix} y(t) \\ y(t-\tau) \end{pmatrix} \right\rangle + \\ &\quad + \left\langle \mathcal{H}(t) \begin{pmatrix} y(t) \\ y(t-\tau) \end{pmatrix}, \begin{pmatrix} \frac{d}{dt}y(t) \\ \frac{d}{dt}y(t-\tau) \end{pmatrix} \right\rangle + \\ &\quad + \langle K(t, 0)y(t), y(t) \rangle - \langle K(t, \tau)y(t-\tau), y(t-\tau) \rangle + \int_{t-\tau}^t \left\langle \frac{d}{dt}K(t, t-s)y(s), y(s) \right\rangle ds. \end{aligned}$$

Обозначим

$$z(t) = A(t)y(t) + B(t)y(t-\tau) + C(t)\frac{d}{dt}y(t-\tau) + \int_{t-\tau}^t D(t, t-s)y(s)ds.$$

Поскольку $y(t)$ удовлетворяет (2.6), получаем

$$\begin{aligned} \frac{d}{dt}V(t, y) = & \left\langle \frac{d}{dt}\mathcal{H}(t) \begin{pmatrix} y(t) \\ y(t-\tau) \end{pmatrix}, \begin{pmatrix} y(t) \\ y(t-\tau) \end{pmatrix} \right\rangle + \left\langle \mathcal{H}(t) \begin{pmatrix} z(t) \\ \frac{d}{dt}y(t-\tau) \end{pmatrix}, \begin{pmatrix} y(t) \\ y(t-\tau) \end{pmatrix} \right\rangle + \\ & + \left\langle \mathcal{H}(t) \begin{pmatrix} y(t) \\ y(t-\tau) \end{pmatrix}, \begin{pmatrix} z(t) \\ \frac{d}{dt}y(t-\tau) \end{pmatrix} \right\rangle + \\ & + \langle K(t, 0)y(t), y(t) \rangle - \langle K(t, \tau)y(t-\tau), y(t-\tau) \rangle + \int_{t-\tau}^t \left\langle \frac{d}{dt}K(t, t-s)y(s), y(s) \right\rangle ds. \end{aligned}$$

Используя матрицу $Q(t, s)$, определенную в (2.4) и (2.5), имеем

$$\frac{d}{dt}V(t, y) = -\frac{1}{\tau} \int_{t-\tau}^t \left\langle Q(t, t-s) \begin{pmatrix} y(t) \\ y(t-\tau) \\ \frac{d}{dt}y(t-\tau) \end{pmatrix}, \begin{pmatrix} y(t) \\ y(t-\tau) \\ \frac{d}{dt}y(t-\tau) \end{pmatrix} \right\rangle ds.$$

В силу (2.7) мы приходим к неравенству

$$\frac{d}{dt}V(t, y) \leq -p(t) \left\langle \mathcal{H}(t) \begin{pmatrix} y(t) \\ y(t-\tau) \end{pmatrix}, \begin{pmatrix} y(t) \\ y(t-\tau) \end{pmatrix} \right\rangle - k(t) \int_{t-\tau}^t \langle K(t, t-s)y(s), y(s) \rangle ds.$$

Согласно определению функционала Ляпунова–Красовского (2.11), мы имеем

$$\frac{d}{dt}V(t, y) \leq -\gamma(t)V(t, y),$$

где функция $\gamma(t)$ определена в (2.10). Это дифференциальное неравенство дает оценку

$$V(t, y) \leq V(0, \varphi) \exp \left(- \int_0^t \gamma(\xi) d\xi \right),$$

где $V(0, \varphi)$ определено в (2.9). Используя (2.2), имеем

$$\|y(t) - G(t)y(t-\tau)\|^2 \leq \frac{1}{h(t)} \left\langle \mathcal{H}(t) \begin{pmatrix} y(t) \\ y(t-\tau) \end{pmatrix}, \begin{pmatrix} y(t) \\ y(t-\tau) \end{pmatrix} \right\rangle \leq \frac{V(t, y)}{h(t)} \leq \frac{V(0, \varphi)}{h(t)} \exp \left(- \int_0^t \gamma(\xi) d\xi \right),$$

откуда следует (2.8).

Теорема 1 доказана.

Следствие 1. Пусть $G(t) \equiv 0$ и выполнены условия теоремы 1.

А. Если $\tilde{\gamma}(t) = \int_0^t \gamma(s) ds \geq 0$, тогда нулевое решение системы (1.1) устойчиво, при этом решение $y(t)$ задачи (2.6) удовлетворяет оценке

$$\|y(t)\| \leq \sqrt{\frac{V(0, \varphi)}{h_0}}, \quad t > 0.$$

В. Если $\tilde{\gamma}(t) \rightarrow \infty$ при $t \rightarrow \infty$, тогда нулевое решение системы (1.1) асимптотически устойчиво, при этом скорость стабилизации определяется функцией

$$\exp \left(-\frac{1}{2} \int_0^t \gamma(\xi) d\xi \right).$$

С. Если $\tilde{\gamma}(t) \geq \gamma_1 t + \gamma_2$ с $\gamma_1 > 0$, тогда нулевое решение системы (1.1) экспоненциально устойчиво, при этом решение $y(t)$ задачи (2.6) удовлетворяет оценке

$$\|y(t)\| \leq \sqrt{\frac{V(0, \varphi)}{h_0}} \exp\left(-\frac{\gamma_1 t}{2} - \frac{\gamma_2}{2}\right), \quad t > 0.$$

Рассмотрим следующее функционально-разностное уравнение

$$x(t) = G(t)x(t - \tau), \quad t \geq 0, \quad (2.12)$$

где $\tau > 0$ — параметр запаздывания, $G(t)$ — матрица размера $(n \times n)$ с непрерывными элементами. Системы вида (2.12) в литературе часто называют непрерывными по времени разностными системами с запаздыванием (см., например, [22]–[24]). Предположим, что нулевое решение системы (2.12) экспоненциально устойчиво; т.е. справедлива следующая оценка

$$\|x(t)\| \leq a_1 e^{-a_2 t} \max_{s \in [-\tau, 0]} \|x_0(s)\|, \quad t \geq 0, \quad (2.13)$$

где $a_1, a_2 > 0$, $x_0(t) \in C([-\tau, 0])$ — заданная начальная вектор-функция, причем $x(t) = x_0(t)$ при $t \in [-\tau, 0]$. Тогда имеет место следующий результат.

Теорема 2. Пусть $G(t) \neq 0$, выполнены условия теоремы 1 и нулевое решение системы (2.12) экспоненциально устойчиво. Если

$$\int_0^t \gamma(s) ds \geq \gamma_1 t + \gamma_2, \quad \gamma_1 > 0, \quad (2.14)$$

тогда нулевое решение системы (1.1) экспоненциально устойчиво.

Доказательство. В силу (2.8) для решения $y(t)$ задачи (2.6) справедливо неравенство

$$\|y(t) - G(t)y(t - \tau)\| \leq \sqrt{\frac{V(0, \varphi)}{h_0}} \exp\left(-\frac{\gamma_1 t}{2} - \frac{\gamma_2}{2}\right), \quad t > 0.$$

Согласно определению

$$\sqrt{V(0, \varphi)} \leq \sqrt{2\|\mathcal{H}(0)\| + \tau \max_{s \in [0, \tau]} \|K(0, s)\|} \max_{s \in [-\tau, 0]} \|\varphi(s)\|.$$

Следовательно, чтобы получить оценку для $\|y(t)\|$, достаточно получить соответствующую оценку для решения следующей начальной задачи

$$\begin{aligned} y(t) - G(t)y(t - \tau) &= F(t), \quad t \geq 0, \\ y(t) &= \varphi(t), \quad t \in [-\tau, 0], \end{aligned} \quad (2.15)$$

где вектор-функция $F(t) \in C(\overline{\mathbb{R}}_+)$ удовлетворяет неравенству

$$\|F(t)\| \leq \frac{\sqrt{2\|\mathcal{H}(0)\| + \tau \max_{s \in [0, \tau]} \|K(0, s)\|}}{\sqrt{h_0}} \exp\left(-\frac{\gamma_1 t}{2} - \frac{\gamma_2}{2}\right) \max_{s \in [-\tau, 0]} \|\varphi(s)\|.$$

Рассмотрим начальную задачу для неоднородного функционально-разностного уравнения

$$\begin{aligned} z(t) &= G(t)z(t - \tau) + f(t), \quad t \geq 0, \\ z(t) &= \psi(t), \quad t \in [-\tau, 0], \end{aligned} \quad (2.16)$$

где $\psi(t) \in C([-\tau, 0])$ — заданная вектор-функция, $f(t) \in C(\overline{\mathbb{R}}_+)$ удовлетворяет оценке

$$\|f(t)\| \leq b_1 e^{-b_2 t}, \quad t \geq 0, \quad (2.17)$$

$b_1, b_2 > 0$. Пусть $t \in [(k-1)\tau, k\tau]$, $k = 1, 2, \dots$. Очевидно, решение задачи (2.16) имеет вид

$$z(t) = G(t) \cdots G(t - (k-1)\tau) \psi(t - k\tau) + \sum_{j=0}^{k-1} G(t) \cdots G(t - (j-1)\tau) f(t - j\tau). \quad (2.18)$$

Нетрудно получить следующую оценку:

$$\|G(t) \cdots G(t - (j-1)\tau)\| \leq a_1 e^{-a_2 \tau(j-1)}, \quad j = 1, 2, \dots \quad (2.19)$$

Действительно, решение уравнения (2.12) может быть записано в виде

$$x(t) = G(t) \cdots G(t - (j-1)\tau) x_0(t - j\tau)$$

при $t \in [(j-1)\tau, j\tau]$, $j = 1, 2, \dots$. В силу (2.13) мы имеем неравенство

$$\|x(t)\| = \|G(t) \cdots G(t - (j-1)\tau) x_0(t - j\tau)\| \leq a_1 e^{-a_2 t} \max_{s \in [-\tau, 0]} \|x_0(s)\| \leq a_1 e^{-a_2 \tau(j-1)} \max_{s \in [-\tau, 0]} \|x_0(s)\|,$$

которое дает нам (2.19). Из (2.18) мы имеем

$$\begin{aligned} \|z(t)\| &\leq \|G(t) \cdots G(t - (k-1)\tau) \psi(t - k\tau)\| + \sum_{j=0}^{k-1} \|G(t) \cdots G(t - (j-1)\tau) f(t - j\tau)\| \leq \\ &\leq a_1 e^{-a_2 t} \max_{s \in [-\tau, 0]} \|\psi(s)\| + a_1 \sum_{j=0}^{k-1} e^{-a_2 \tau(j-1)} \|f(t - j\tau)\| \end{aligned}$$

при $t \in [(k-1)\tau, k\tau]$, $k = 1, 2, \dots$. Используя (2.17), получаем

$$\|z(t)\| \leq a_1 e^{-a_2 t} \max_{s \in [-\tau, 0]} \|\psi(s)\| + a_1 b_1 e^{a_2 \tau} \sum_{j=0}^{k-1} e^{-a_2 \tau j} e^{-b_2(t-j\tau)} \quad (2.20)$$

при $t \in [(k-1)\tau, k\tau]$, $k = 1, 2, \dots$.

Для дальнейших рассуждений нам понадобится вспомогательная лемма.

Лемма. Пусть

$$S = \sum_{j=0}^{k-1} \beta^j e^{-b(t-j\tau)}, \quad 0 < \beta < 1, \quad b > 0.$$

Тогда при $t \in [(k-1)\tau, k\tau]$, $k = 1, 2, \dots$, имеет место оценка

$$S \leq \begin{cases} (1 - \beta e^{b\tau})^{-1} e^{-bt} & \text{при } \beta e^{b\tau} < 1; \\ (\frac{t}{\tau} + 1) e^{-bt} & \text{при } \beta e^{b\tau} = 1; \\ (\beta e^{b\tau} - 1)^{-1} e^{b(1-r)\tau} e^{-r bt} & \text{при } \beta e^{b\tau} > 1, \quad 0 < r < -\frac{1}{b\tau} \ln \beta < 1. \end{cases}$$

Доказательство. Рассмотрим первый случай $\beta e^{b\tau} < 1$. Тогда

$$S = e^{-bt} \sum_{j=0}^{k-1} (\beta e^{b\tau})^j \leq e^{-bt} \sum_{j=0}^{\infty} (\beta e^{b\tau})^j = e^{-bt} (1 - \beta e^{b\tau})^{-1}.$$

Рассмотрим второй случай $\beta e^{b\tau} = 1$. Очевидно,

$$S = e^{-bt} \sum_{j=0}^{k-1} (1)^j = k e^{-bt} \leq \left(\frac{t}{\tau} + 1\right) e^{-bt}.$$

Рассмотрим третий случай $\beta e^{b\tau} > 1$. Очевидно,

$$S = e^{-bt} \sum_{j=0}^{k-1} (\beta e^{b\tau})^j = (\beta e^{b\tau} - 1)^{-1} e^{-bt} (\beta^k e^{kb\tau} - 1) \leq (\beta e^{b\tau} - 1)^{-1} e^{-bt} (\beta e^{b\tau})^k.$$

Поскольку $\beta < 1$, то существует $r > 0$ такое, что $\beta e^{rb\tau} < 1$. Следовательно,

$$S \leq (\beta e^{b\tau} - 1)^{-1} e^{-bt} e^{b(1-r)k\tau}.$$

Тогда при $t \in [(k-1)\tau, k\tau]$, $k = 1, 2, \dots$, имеем

$$S \leq (\beta e^{b\tau} - 1)^{-1} e^{-bt} e^{b(1-r)(k-1)\tau} e^{b(1-r)\tau} \leq (\beta e^{b\tau} - 1)^{-1} e^{b(1-r)\tau} e^{-rbt}.$$

Отметим, что r может быть выбрано из неравенства

$$0 < r < -\frac{1}{b\tau} \ln \beta.$$

Лемма доказана.

Пусть $\beta = e^{-a_2\tau}$, $b = b_2$. Используя эту лемму, в силу (2.20) решение $z(t)$ задачи (2.16) экспоненциально стремится к нулю. Следовательно, решение $y(t)$ задачи (2.15) также экспоненциально стремится к нулю. Тогда нулевое решение системы (1.1) экспоненциально устойчиво, причем мы имеем оценки, характеризующие скорость убывания на бесконечности. Чтобы выписать эти оценки, достаточно воспользоваться неравенством (2.20) и леммой, доказанной выше, при

$$\beta = e^{-a_2\tau}, \quad b = b_2 = \frac{\gamma_1}{2}, \quad b_1 = \frac{\sqrt{2\|\mathcal{H}(0)\| + \tau \max_{s \in [0, \tau]} \|K(0, s)\|}}{\sqrt{h_0}} e^{-\gamma_2/2} \max_{s \in [-\tau, 0]} \|\varphi(s)\|.$$

Теорема доказана.

Следствие 2. Пусть выполнены условия теоремы 2. Если $\gamma_1 - 2a_2 < 0$, тогда для решения задачи (2.6) имеет место оценка

$$\|y(t)\| \leq \left(a_1 e^{-a_2 t} + d \left(1 - e^{(\gamma_1/2 - a_2)\tau} \right)^{-1} e^{-\gamma_1 t/2} \right) \max_{s \in [-\tau, 0]} \|\varphi(s)\|,$$

где

$$d = a_1 e^{a_2\tau - \gamma_2/2} \frac{\sqrt{2\|\mathcal{H}(0)\| + \tau \max_{s \in [0, \tau]} \|K(0, s)\|}}{\sqrt{h_0}}.$$

Следствие 3. Пусть выполнены условия теоремы 2. Если $\gamma_1 - 2a_2 = 0$, тогда для решения задачи (2.6) имеет место оценка

$$\|y(t)\| \leq \left(a_1 e^{-a_2 t} + d \left(\frac{t}{\tau} + 1 \right) e^{-\gamma_1 t/2} \right) \max_{s \in [-\tau, 0]} \|\varphi(s)\|.$$

Следствие 4. Пусть выполнены условия теоремы 2. Если $\gamma_1 - 2a_2 > 0$, тогда для решения задачи (2.6) имеет место оценка

$$\|y(t)\| \leq \left(a_1 e^{-a_2 t} + d \left(e^{(\gamma_1/2 - a_2)\tau} - 1 \right)^{-1} e^{\gamma_1(1-r)\tau/2} e^{-r\gamma_1 t/2} \right) \max_{s \in [-\tau, 0]} \|\varphi(s)\|,$$

где $0 < r < \frac{2a_2}{\gamma_1}$.

3. ОЦЕНКИ РЕШЕНИЙ НЕЛИНЕЙНЫХ СИСТЕМ

В этом разделе мы исследуем нелинейные системы вида (1.1) при условии, что вектор-функция $F(t, u_1, u_2, u_3)$, определяющая нелинейные слагаемые, удовлетворяет условию (1.2). Рассмотрим начальную задачу

$$\begin{aligned} \frac{d}{dt}y(t) &= A(t)y(t) + B(t)y(t-\tau) + C(t)\frac{d}{dt}y(t-\tau) + \\ &+ \int_{t-\tau}^t D(t, t-s)y(s)ds + F\left(t, y(t), y(t-\tau), \frac{d}{dt}y(t-\tau)\right), \quad t \geq 0, \\ y(t) &= \varphi(t), \quad t \in [-\tau, 0], \\ y(+0) &= \varphi(0), \end{aligned} \quad (3.1)$$

где $\varphi(t) \in C^1([-\tau, 0])$ — заданная вещественнозначная вектор-функция. Ниже мы устанавливаем оценки для решений задачи (3.1).

Пусть $\alpha(t) > 0$ — произвольная непрерывная функция. Введем функции

$$\alpha_1(t) = 2q_1\|H_1(t)\| + \frac{(q_1\|H_2(t)\| + q_2\|H_1(t)\|)^2}{\alpha(t)}, \quad \alpha_2(t) = \alpha(t) + 2q_2\|H_2(t)\| \quad (3.2)$$

и матрицу

$$Q_\alpha(t, s) = Q(t, s) - \begin{pmatrix} \alpha_1(t)I & 0 & 0 & 0 \\ 0 & \alpha_2(t)I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (3.3)$$

где матрица $Q(t, s)$ определена в (2.4) и (2.5), I — единичная матрица.

Теорема 3. *Предположим, что существуют матрицы $\mathcal{H}(t)$, $K(t, s)$, удовлетворяющие условиям (2.1)–(2.3) и такие, что*

$$\left\langle Q_\alpha(t, s) \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} \right\rangle \geq p_\alpha(t) \left\langle \mathcal{H}(t) \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \right\rangle + \tau k_\alpha(t) \langle K(t, s) u_4, u_4 \rangle, \quad (3.4)$$

$$u_j \in \mathbb{R}^n, \quad t \geq 0, \quad s \in [0, \tau],$$

где $p_\alpha(t)$, $k_\alpha(t) \in C(\overline{\mathbb{R}}_+)$. Тогда для решения $y(t)$ задачи (3.1) имеет место оценка

$$\|y(t) - G(t)y(t-\tau)\| \leq \sqrt{\frac{V(0, \varphi)}{h(t)}} \exp\left(-\frac{1}{2} \int_0^t \gamma_\alpha(\xi) d\xi\right), \quad t > 0, \quad (3.5)$$

где $V(0, \varphi)$ определено в (2.9), $\gamma_\alpha(t) = \min\{p_\alpha(t), k_\alpha(t)\}$.

Доказательство. Пусть $y(t)$ — решение задачи (3.1). Рассмотрим на $y(t)$ функционал Ляпунова–Красовского (2.11). Дифференцируя этот функционал и повторяя рассуждения как при доказательстве теоремы 1, имеем

$$\frac{d}{dt}V(t, y) = -\frac{1}{\tau} \int_{t-\tau}^t \left\langle Q(t, t-s) \begin{pmatrix} y(t) \\ y(t-\tau) \\ \frac{d}{dt}y(t-\tau) \\ y(s) \end{pmatrix}, \begin{pmatrix} y(t) \\ y(t-\tau) \\ \frac{d}{dt}y(t-\tau) \\ y(s) \end{pmatrix} \right\rangle ds + W(t), \quad (3.6)$$

где матрица $Q(t, s)$ определена в (2.4) и (2.5),

$$\begin{aligned} W(t) &= \left\langle \mathcal{H}(t) \begin{pmatrix} F\left(t, y(t), y(t-\tau), \frac{d}{dt}y(t-\tau)\right) \\ 0 \end{pmatrix}, \begin{pmatrix} y(t) \\ y(t-\tau) \end{pmatrix} \right\rangle + \\ &+ \left\langle \mathcal{H}(t) \begin{pmatrix} y(t) \\ y(t-\tau) \end{pmatrix}, \begin{pmatrix} F\left(t, y(t), y(t-\tau), \frac{d}{dt}y(t-\tau)\right) \\ 0 \end{pmatrix} \right\rangle. \end{aligned}$$

В силу определения матрицы $\mathcal{H}(t)$ имеем

$$W(t) \leq 2(\|H_1(t)\| \|y(t)\| + \|H_2(t)\| \|y(t-\tau)\|) \left\| F \left(t, y(t), y(t-\tau), \frac{d}{dt} y(t-\tau) \right) \right\|.$$

Учитывая условие (1.2) получаем

$$W(t) \leq \beta_1(t) \|y(t)\|^2 + \beta_2(t) \|y(t)\| \|y(t-\tau)\| + \beta_3(t) \|y(t-\tau)\|^2, \quad (3.7)$$

где

$$\beta_1(t) = 2q_1 \|H_1(t)\|, \quad \beta_2(t) = 2(q_1 \|H_2(t)\| + q_2 \|H_1(t)\|), \quad \beta_3(t) = 2q_2 \|H_2(t)\|.$$

Очевидно, что для любого $\alpha > 0$ справедливо неравенство

$$b_1 u_1^2 + b_2 u_1 u_2 \leq \left(b_1 + \frac{b_2^2}{4\alpha} \right) u_1^2 + \alpha u_2^2.$$

Тогда правую часть неравенства (3.7) можно оценить следующим образом:

$$W(t) \leq \alpha_1(t) \|y(t)\|^2 + \alpha_2(t) \|y(t-\tau)\|^2,$$

где функции $\alpha_j(t)$, $j = 1, 2$, определены в (3.2), при этом $\alpha(t) > 0$ — произвольная непрерывная функция, которой можно управлять. С учетом этой оценки из (3.6) получаем

$$\frac{d}{dt} V(t, y) \leq -\frac{1}{\tau} \int_{t-\tau}^t \left\langle Q_\alpha(t, t-s) \begin{pmatrix} y(t) \\ y(t-\tau) \\ \frac{d}{dt} y(t-\tau) \\ y(s) \end{pmatrix}, \begin{pmatrix} y(t) \\ y(t-\tau) \\ \frac{d}{dt} y(t-\tau) \\ y(s) \end{pmatrix} \right\rangle ds,$$

где матрица $Q_\alpha(t, s)$ определена в (3.3). Тогда, используя условие (3.4) и повторяя рассуждения из доказательства теоремы 1, мы получаем оценку (3.5).

Теорема доказана.

Справедлив аналог теоремы 2.

Теорема 4. Пусть $G(t) \neq 0$, выполнены условия теоремы 3 и нулевое решение системы (2.12) экспоненциально устойчиво. Если выполнено условие (2.14), тогда нулевое решение нелинейной системы (1.1) экспоненциально устойчиво.

По аналогии с предыдущим параграфом можно сформулировать следствия и указать оценки экспоненциального убывания решений нелинейной системы (1.1).

4. ПРИМЕРЫ

Рассмотрим несколько примеров систем с запаздыванием вида (1.1).

Пример 1. Пусть $C(t) \equiv 0$. Если $A(t), B(t)$ есть T -периодические матрицы, $D(t, s) \equiv 0$, линейные и нелинейные системы с запаздыванием и периодическими линейными членами исследовались в [13]. Мы использовали функционал Ляпунова–Красовского вида (2.11), где

$$\mathcal{H}(t) = \begin{pmatrix} H(t) & 0 \\ 0 & 0 \end{pmatrix}, \quad K(t, s) = K(s),$$

причем $H(t) = H^*(t) > 0$ есть T -периодическая матрица, $K(s) = K^*(s) > 0$, $s \in [0, \tau]$. Оценки экспоненциального убывания решений на бесконечности и оценки для областей притяжения были получены в явном виде.

Системы вида (1.1) с переменными коэффициентами и двумя видами запаздываний (сосредоточенное и распределенное) при $C(t) \equiv 0$ изучались в [21]. В этом случае достаточно взять

$$\mathcal{H}(t) = \begin{pmatrix} H(t) & 0 \\ 0 & 0 \end{pmatrix}, \quad H(t) \in C^1(\overline{\mathbb{R}}_+), \quad H(t) = H^*(t) > 0, \quad t \geq 0.$$

Очевидно, неравенство (2.2) выполнено при $G(t) \equiv 0$, где $h(t)$ — минимальное собственное значение матрицы $H(t)$. Следовательно, используя теорему 1 (или теорему 3) и следствие 1, мы имеем оценки для решений задачи (2.6) (или задачи (3.1)) на всей полуоси $\{t \geq 0\}$, на основании которых можно сделать выводы об устойчивости решений (в частности, экспоненциальной или асимптотической устойчивости) и указать оценки на скорость стабилизации. Следует отметить, что в [21] мы также исследовали нелинейные системы вида

$$\frac{d}{dt}y(t) = A(t)y(t) + B(t)y(t - \tau) + \int_{t-\tau}^t D(t, t-s)y(s) ds + F\left(t, y(t), y(t - \tau), \int_{t-\tau}^t D(t, t-s)y(s) ds\right), \quad t \geq 0,$$

где $F(t, u_1, u_2, u_3)$ — непрерывная вещественнозначная вектор-функция, удовлетворяющая условию Липшица по u_1, u_3 и неравенству

$$\|F(t, u_1, u_2, u_3)\| \leq q\|u_1\|^{1+\omega}, \quad t \geq 0, \quad u_j \in \mathbb{R}^n, \quad q, \omega \geq 0.$$

Отметим, что функционалы вида (2.11) могут быть использованы для исследования моделей, описываемых системами дифференциальных уравнений с запаздыванием (см., например, [25], [26]).

Пример 2. Пусть $C(t)$ — ненулевая постоянная матрица; т. е. $C(t) \equiv C$. Если $A(t), B(t)$ суть T -периодические матрицы, $D(t, s) \equiv 0$, то линейные и нелинейные системы нейтрального типа с периодическими линейными членами были изучены в [14, 17]. Мы использовали функционал Ляпунова — Красовского вида (2.11) с

$$\mathcal{H}(t) = \begin{pmatrix} H(t) & -H(t)C \\ -C^*H(t) & C^*H(t)C \end{pmatrix}, \quad K(t, s) = K(s),$$

где $H(t) = H^*(t) > 0$ есть T -периодическая матрица, $K(s) = K^*(s) > 0, s \in [0, \tau]$. Оценки экспоненциального убывания решений на бесконечности и оценки для областей притяжения были получены в явном виде. Очевидно,

$$\left\langle \mathcal{H}(t) \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \right\rangle = \langle H(t)(u_1 - Cu_2), (u_1 - Cu_2) \rangle$$

и (2.2) выполнено при $G(t) \equiv C$, где $h(t) > 0$ — минимальное собственное значение матрицы $H(t)$. Чтобы использовать теоремы 2, 4 и соответствующие следствия, нам нужно определить a_1, a_2 в (2.13). Из результатов в [27], [28] вытекают следующие теоремы.

Теорема 5. Нулевое решение системы (2.12) при $G(t) \equiv C$ экспоненциально устойчиво тогда и только тогда, когда существует эрмитово решение $L > 0$ дискретного уравнения Ляпунова

$$L - C^*LC = I. \quad (4.1)$$

Теорема 6. Пусть $L = L^* > 0$ — решение уравнения (4.1). Для решений системы (2.12) при $G(t) \equiv C$ имеет место следующая оценка:

$$\|x(t)\| \leq \sqrt{\mu(L)} \exp\left(\frac{\ln\left(1 - \frac{1}{\|L\|}\right)}{2\tau}t\right) \max_{s \in [-\tau, 0]} \|x_0(s)\|,$$

где $\mu(L) = l_{\min}^{-1}\|L\|$, $l_{\min} \geq 1$ — минимальное собственное число матрицы L .

Следовательно, мы можем взять $a_1 = \sqrt{\mu(L)}$, $a_2 = -\frac{\ln(1 - \frac{1}{\|L\|})}{2\tau}$. В [14, 17] мы использовали значения, зависящие от $\|C^m\|$, где $m > 0$ — минимальное целое число, при котором $\|C^m\| < 1$. Теорема 5 гарантирует, что спектр матрицы C принадлежит единичному кругу $\{\lambda \in \mathbb{C} : |\lambda| < 1\}$. Поэтому такое m существует. Отметим, что (2.14) выполнено, если $\gamma_1 = \int_0^T \gamma(s)ds > 0$. Действительно, в этом случае

$$\int_0^t \gamma(s)ds \geq \gamma_1 t + \left(\min_{s \in [0, T]} \gamma(s) - \gamma_1\right)T.$$

Поэтому мы можем взять $\gamma_2 = \left(\min_{s \in [0, T]} \gamma(s) - \gamma_1\right)T$.

Пример 3. Пусть $A(t)$, $B(t)$, $C(t)$ суть T -периодические матрицы и $D(t, s)$ есть T -периодическая матрица по t . Рассмотрим функционал Ляпунова–Красовского вида (2.11) с

$$\mathcal{H}(t) = \begin{pmatrix} H(t) & -H(t)C(t) \\ -C^*(t)H(t) & C^*(t)H(t)C(t) \end{pmatrix},$$

где $H(t) = H^*(t) > 0$. Очевидно,

$$\left\langle \mathcal{H}(t) \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \right\rangle = \langle H(t)(u_1 - C(t)u_2), (u_1 - C(t)u_2) \rangle$$

и (2.2) выполнено при $G(t) = C(t)$, где $h(t) > 0$ — минимальное собственное значение матрицы $H(t)$. Чтобы использовать теоремы 2, 4 и соответствующие следствия, нужно определить a_1 , a_2 в (2.13). Из результатов в [29]–[31] вытекают следующие теоремы.

Теорема 7. Нулевое решение системы (2.12) при $G(t) \equiv C(t)$ экспоненциально устойчиво тогда и только тогда, когда существует эрмитово T -периодическое непрерывное решение $L(t) > 0$ функционально-разностного уравнения Ляпунова

$$L(t - \tau) - C^*(t)L(t)C(t) = I, \quad t \geq 0. \quad (4.2)$$

Теорема 8. Пусть $L(t) = L^*(t) > 0$ есть T -периодическое непрерывное решение уравнения (4.2). Для решений системы (2.12) с $G(t) \equiv C(t)$ справедлива следующая оценка:

$$\|x(t)\| \leq \sqrt{\mu(L)} \exp \left(\frac{\ln \left(1 - \frac{1}{L_{\max}} \right)}{2\tau} t \right) \max_{s \in [-\tau, 0]} \|x_0(s)\|,$$

где $\mu(L) = l_{\min} L_{\max}$, $l_{\min} = \left(\min_{s \in [0, T]} l_{\min}(s) \right)^{-1}$, $L_{\max} = \max_{s \in [0, T]} \|L(s)\|$, $l_{\min}(t) \geq 1$ — минимальное собственное значение матрицы $L(t)$.

Следовательно, мы можем взять $a_1 = \sqrt{\mu(L)}$, $a_2 = -\frac{\ln(1 - \frac{1}{L_{\max}})}{2\tau}$.

Замечание 1. Если $A(t)$, $C(t)$ суть T -периодические матрицы, $B(t) = \frac{d}{dt}C(t)$, и $D(t, s)$ есть T -периодическая матрица по t , то системы вида (1.1) изучались в [31], при этом использовался функционал Ляпунова–Красовского вида (2.11) при

$$\mathcal{H}(t) = \begin{pmatrix} H(t) & -H(t)C(t) \\ -C^*(t)H(t) & C^*(t)H(t)C(t) \end{pmatrix}, \quad K(t, s) = (\tau - s)P(s) + M(s, t - s).$$

Замечание 2. Если $A(t)$, $B(t)$, $C(t)$ суть T -периодические матрицы и $D(t, s) \equiv 0$, то линейные и нелинейные системы нейтрального типа с периодическими линейными членами изучались в [15, 16, 18, 19]. Используя специальные функционалы Ляпунова–Красовского, в явном виде были получены оценки экспоненциального убывания решений на бесконечности и оценки для областей притяжения.

Пример 4. Рассмотрим системы вида (1.1), где $A(t)$, $B(t)$, $C(t)$, $D(t, s)$ — матрицы с непрерывными вещественнозначными элементами. Пусть выполнены условия теоремы 1 (или теоремы 3). Чтобы использовать теорему 2 (или теорему 4) и соответствующие следствия, нужно определить a_1 , a_2 в (2.13). Имеют место следующие утверждения, доказательство которых проводится по классической схеме.

Теорема 9. Нулевое решение системы (2.12) экспоненциально устойчиво тогда и только тогда, когда существует эрмитово непрерывное ограниченное решение $L(t) > 0$ функционально-разностного уравнения Ляпунова

$$L(t - \tau) - G^*(t)L(t)G(t) = I, \quad t \geq 0. \quad (4.3)$$

Теорема 10. Пусть $L(t) = L^*(t) > 0$ — непрерывное решение уравнения (4.3), причем $\sup_{t \geq 0} \|L(t)\| < \infty$. Для решений системы (2.12) справедлива следующая оценка:

$$\|x(t)\| \leq \sqrt{\mu(L)} \exp \left(\frac{\ln \left(1 - \frac{1}{L_{\sup}} \right)}{2\tau} t \right) \max_{s \in [-\tau, 0]} \|x_0(s)\|,$$

где $\mu(L) = l_{\min} L_{\sup}$, $l_{\min} = \left(\min_{s \geq 0} l_{\min}(s) \right)^{-1}$, $L_{\sup} = \sup_{s \geq 0} \|L(s)\|$, $l_{\min}(t) \geq 1$ — минимальное собственное значение матрицы $L(t)$.

Следовательно, мы можем взять $a_1 = \sqrt{\mu(L)}$, $a_2 = -\frac{\ln \left(1 - \frac{1}{L_{\sup}} \right)}{2\tau}$.

Автор выражает благодарность Г.В. Демиденко за полезные дискуссии.

СПИСОК ЛИТЕРАТУРЫ

1. Эльсгольц Л.Э., Норкин С.Б. Введение в теорию дифференциальных уравнений с отклоняющимся аргументом. М.: Наука, 1971.
2. Хейл Дж. Теория функционально-дифференциальных уравнений. М.: Мир, 1984.
3. Корневский Д.Г. Устойчивость динамических систем при случайных возмущениях параметров. Алгебраические критерии. Киев: Наукова думка, 1989.
4. Избранные труды Н.В. Азбелева. М.-Ижевск: Институт компьютерных исследований, 2012.
5. Долгий Ю.Ф. Устойчивость периодических дифференциально-разностных уравнений. Екатеринбург: Изд-во Уральского ун-та, 1996.
6. Хусаинов Д.Я., Шатырко А.В. Метод функций Ляпунова в исследовании устойчивости дифференциально-функциональных систем. Киев: Изд-во Киев. ун-та, 1997.
7. Kolmanovskii V.B., Myshkis A.D. Introduction to the theory and applications of functional differential equations // Math. Appl. V. 463. Dordrecht: Kluwer Acad. Publ., 1999.
8. Michiels W., Niculescu S.I. Stability, control, and computation for time-delay systems. An eigenvalue-based approach, Advances in Design and Control, vol. 27. Philadelphia, Society for Industrial and Applied Mathematics, 2014.
9. Agarwal R.P., Berezansky L., Braverman E., Domoshnitsky A. Nonoscillation theory of functional differential equations with applications. New York: Springer, 2012.
10. Kharitonov V.L. Time-delay systems. Lyapunov functionals and matrices, Control Engineering. New York: Birkhauser, Springer, 2013.
11. Gil' M.I. Stability of neutral functional differential equations, Atlantis Studies in Differential Equations, vol. 3. Paris: Atlantis Press, 2014.
12. Park J.H., Lee T.H., Liu Y., Chen J. Dynamic systems with time delays: stability and control, Springer, Singapore, 2019.
13. Демиденко Г.В., Матвеева И.И. Устойчивость решений дифференциальных уравнений с запаздывающим аргументом и периодическими коэффициентами в линейных членах // Сиб. матем. ж. 2007. Т. 48. № 5. С. 1025–1040.
14. Демиденко Г.В., Матвеева И.И. Об оценках решений систем дифференциальных уравнений нейтрального типа с периодическими коэффициентами // Сиб. матем. ж. 2014. Т. 55. № 5. С. 1059–1077.
15. Матвеева И.И. Об экспоненциальной устойчивости решений периодических систем нейтрального типа // Сиб. матем. ж. 2017. Т. 58. № 2. С. 344–352.

16. *Матвеева И.И.* Об экспоненциальной устойчивости решений периодических систем нейтрального типа с несколькими запаздываниями // Дифференц. ур-ния. 2017. Т. 53. № 6. С. 730–740.
17. *Демиденко Г.В., Матвеева И.И., Скворцова М.А.* Оценки решений дифференциальных уравнений нейтрального типа с периодическими коэффициентами в линейных членах // Сиб. матем. ж. 2019. Т. 60. № 5. С. 1063–1079.
18. *Матвеева И.И.* Оценки экспоненциального убывания решений линейных систем нейтрального типа с периодическими коэффициентами // Сиб. журн. индустр. матем. 2019. Т. 22. № 3. С. 96–103.
19. *Матвеева И.И.* Оценки экспоненциального убывания решений одного класса нелинейных систем нейтрального типа с периодическими коэффициентами // Ж. вычисл. матем. и матем. физ. 2020. Т. 60. № 4. С. 612–620.
20. *Матвеева И.И.* Оценки решений класса неавтономных систем нейтрального типа с неограниченным запаздыванием // Сиб. матем. журн. 2021. Т. 62. № 3. С. 579–594.
21. *Matveeva I.I.* Estimates for solutions to one class of nonlinear nonautonomous systems with time-varying concentrated and distributed delays // Сиб. электрон. матем. изв. 2021. Т. 18. № 2. С. 1689–1697.
22. *Gu K.* Stability problem of systems with multiple delay channels // Automatica. 2010. V. 46. P. 743–751.
23. *Damak S., Di Loreto M., Mondie S., Brun X.* Exponential stability with decay rate estimation for linear difference equations // IEEE Trans. Automat. Control. 2016. V. 61. № 1. P. 252–257.
24. *Melchor-Aguilar D.* On Lyapunov functionals for linear functional difference equations // Systems & Control Letters. 2019. V. 127. P. 1–5.
25. *Скворцова М.А., Ыскак Т.* Асимптотическое поведение решений в одной модели “хищник-жертва” с запаздыванием // Сиб. матем. ж. 2021. Т. 62. № 2. С. 402–416.
26. *Скворцова М.А., Ыскак Т.* Оценки решений дифференциальных уравнений с распределенным запаздыванием, описывающих конкуренцию нескольких видов микроорганизмов // Сиб. журн. индустр. матем. 2022. Т. 25. № 4. С. 193–205.
27. *Далецкий Ю.Л., Крейн М.Г.* Устойчивость решений дифференциальных уравнений в банаховом пространстве. М.: Наука, 1970.
28. *Годунов С.К.* Современные аспекты линейной алгебры. Новосибирск: Научная книга, 1997.
29. *Demidenko G.V.* Stability of solutions to difference equations with periodic coefficients in linear terms // J. Comput. Math. Optim. 2010. V. 6. № 1. P. 1–12.
30. *Demidenko G.V., Matveeva I.I.* On estimates of solutions to one class of functional difference equations with periodic coefficients // In: Continuum Mechanics, Applied Mathematics and Scientific Computing: Godunov’s Legacy – A Liber Amicorum to Professor Godunov (Editors: Demidenko G.V., Romenski E., Toro E., Dumbser M.). Cham, Switzerland: Springer Nature, 2020. P. 101–109.
31. *Ыскак Т.* Оценки решений одного класса систем уравнений нейтрального типа с распределенным запаздыванием // Сиб. электрон. матем. изв. 2020. Т. 17. С. 416–427.

ESTIMATES OF SOLUTIONS FOR A CLASS OF NONAUTONOMOUS SYSTEMS OF NEUTRAL TYPE WITH CONCENTRATED AND DISTRIBUTED DELAYS

I. I. Matveeva*

Novosibirsk State University, Pirogov St., 1, Novosibirsk, 630090 Russia

*e-mail: i.matveeva@g.nsu.ru

Received: 02 April, 2024

Revised: 02 April, 2024

Accepted: 02 May, 2024

Abstract. A class of systems of nonautonomous differential equations of neutral type with concentrated and distributed delays is considered. By using a Lyapunov–Krasovskii functional, estimates are established imply whether the solutions are stable. In the case of exponential stability, estimates for the stabilization rate of the solutions at infinity are given.

Keywords: systems of neutral type, variable coefficients, estimates of solutions, stability, Lyapunov–Krasovskii functional.

МОДЕЛЬ ГРАДИЕНТА ПЛОТНОСТИ В СФЕРИЧЕСКИ-СИММЕТРИЧНОЙ ПОСТАНОВКЕ И ЕЕ ЯВНО-НЕЯВНАЯ ДИССИПАТИВНАЯ ДИСКРЕТИЗАЦИЯ ДЛЯ ИССЛЕДОВАНИЯ ДИНАМИКИ МЕЖФАЗНОЙ ГРАНИЦЫ

© 2024 г. В. А. Балашов^{1,*}, Е. А. Павлишина^{2,**}, Е. Б. Савенков^{1,***}

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

² 141701 Долгопрудный, М.о., Институтский пер., 9, НИУ МФТИ, Россия

*e-mail: vladislav.balashov@gmail.com

**e-mail: pavlishina.ea@phystech.edu

***e-mail: savenkov@keldysh.ru

Поступила в редакцию 02.04.2024 г.

Переработанный вариант 02.04.2024 г.

Принята к публикации 02.05.2024 г.

Работа посвящена разработке безусловно градиентно-устойчивого (диссипативного) численного метода для решения консервативной модели градиента плотности в сферически-симметричной постановке. Для построения алгоритма использован метод Эйра на основе выпуклого расщепления свободной энергии системы. Доказывается градиентная устойчивость построенного алгоритма в полудискретном и полностью дискретном случаях. Теоретические результаты подтверждены рядом тестовых расчетов. Предложенный численный метод применен для анализа влияния способа задания диффузионной подвижности на характер эволюции межфазной границы. Библ. 23. Фиг. 12. Табл. 2.

Ключевые слова: теория градиента плотности, диссипативный метод, явно-неявная аппроксимация, выпуклое расщепление, сферически-симметричная постановка.

DOI: 10.31857/S0044466924080148, EDN: XZXYEB

1. ВВЕДЕНИЕ

Задача математического моделирования двухфазных микротечений часто возникает при анализе целого ряда естественных и техногенных систем, например, в рамках разработки технологий «lab-on-a-chip» («лаборатория на чипе») или описании микротечений в поровых пространствах образцов горных пород. Поэтому создание новых и совершенствование существующих математических моделей, равно как и разработка вычислительных методов для их анализа, являются важными задачами, решение которых дает возможность исследования и оптимизации соответствующих гидродинамических процессов.

Содержательные постановки целого ряда задач из указанного выше класса предполагают прямое и явное разрешение границ раздела фаз и их эволюции, а также учета капиллярных эффектов. В настоящее время известно несколько классов математических моделей, допускающих такое описание.

Первая широко распространенная группа моделей — так называемые модели с четкой границей — предполагает, что межфазная граница является математической поверхностью «нулевой толщины». В рамках этого допущения известны такие методы, как метод «жидкого объема» («volume-of-fluid») (см. [1]) (в отечественной литературе более известный как метод концентраций, см. [2]), метод множеств уровня («level set») (см. [3]), метод решеточных уравнений Больцмана (впрочем в работе [4] показано, что данный подход эквивалентен некоторому многошаговому конечно-разностному методу).

Ко второму часто применяемому классу моделей относятся модели типа фазового поля (их также называют моделями с «диффузной» границей) (см. [5]). В отличие от моделей с четкой границей в них предполагается, что межфазная граница представляет собой тонкий слой конечной толщины, в пределах которого действуют

¹⁾ Работа выполнена при финансовой поддержке РНФ (проект № 22-11-00203).

силы межфазного натяжения. Примерами таких моделей являются модель Кана—Хилларда (см. [6]), Кортвега (см. [7]), Алена—Кана (см. [8]) и модели, основанные на теории градиента плотности (Density Gradient Theory, DGT) (см. [9]–[12]).

Основным элементом моделей типа фазового поля является предположение о том, что (а) свободная энергия Гельмгольца системы (или другой термодинамический потенциал) зависит не только от фазового поля (в качестве которого может выступать какая-либо искусственная величина, концентрация или, в случае DGT, плотности компонентов смеси), но также от их пространственных производных, и (б) зависимость описывающего состояние системы термодинамического потенциала от параметров порядка является *невыпуклой* функцией. В комбинации эти два фактора позволяют учесть как поверхностное натяжение (энергию межфазной границы), так и обеспечить конечную толщину межфазного слоя, в том числе, в случае сложной геометрической и топологической эволюции границы раздела фаз. Модели типа фазового поля являются наиболее физически обоснованными, но в то же время и наиболее «тяжелыми» с вычислительной точки зрения: они включают в себя пространственные производные высоких порядков (например, баланс массы в модели Кана—Хилларда является уравнением 4-го порядка) и являются сильно нелинейными.

В настоящей работе рассматривается консервативная эволюционная модель, основанная на теории градиента плотности (см. [9]–[11]), в пространственно-одномерной сферически-симметричной постановке. Для модели разработан полудискретный (непрерывный по времени и дискретный по пространству) диссипативный («энергетически-устойчивый») метод. С использованием метода выпуклого расщепления (convex-splitting method) полной энергии Гельмгольца, предложенного в работе [13], построен *полностью* дискретный *безусловно* диссипативный метод.

Суть метода выпуклого расщепления заключается в представлении исходного функционала свободной энергии Ψ в виде суммы выпуклого и вогнутого:

$$\Psi = \Psi^c + (-\Psi^e) \equiv \Psi^c - \Psi^e.$$

При построении дискретизации по времени выпуклая часть Ψ^c аппроксимируется неявным образом, а вогнутая $-\Psi^e$ — явным. Верхние индексы «с» и «е» являются сокращениями от «contraction» (сжатие) и «expansion» (расширение) соответственно. Такая терминология связана с тем, что модели типа фазового поля тесно связаны с моделями «градиентных течений» (gradient flow models, см., например, [14]) вида

$$\frac{du}{dt} = -\nabla \mathcal{F}(u(t)),$$

и если $\mathcal{F}$ является выпуклым, то пучок кривых $u(t)$ сходится, а если вогнутым, то расходится.

По всей видимости, метод выпуклого расщепления является единственным из известных в настоящее время, который позволяет добиться безусловной градиентной устойчивости (диссипативности) при дискретизации по времени в случае, если пространственная дискретизация, т.е. соответствующий полудискретный метод, обладает этим свойством.

Здесь под диссипативностью численного метода подразумевается наследование разностной схемой свойства невозрастания полной энергии Гельмгольца, выполненного для исходной («непрерывной») модели. Часто это свойство также называют градиентной (или энергетической) устойчивостью численного метода. Градиентная устойчивость метода гарантирует термодинамическую корректность результатов моделирования, отсутствие таких численных артефактов, как «паразитные токи» (см., например, [15]), и обеспечивает численную устойчивость метода в классическом понимании этого термина.

В настоящей работе дискретизация по пространству выполнена аналогично более ранним работам авторов [16]–[18], где рассматривались многофазные гидродинамические регуляризованные (согласно квазигидродинамической методике) модели типа фазового поля. Рассмотренные в этих работах вычислительные алгоритмы используют чисто явные эйлеровы аппроксимации по времени. Это накладывает серьезные ограничения на величину временного шага. Поэтому в рамках настоящей работы разработан *полностью* дискретный *безусловно* диссипативный метод (т.е. метод, являющийся диссипативным при любой величине шага по времени).

Рассматриваемая в настоящей работе математическая модель является упрощенным одномерным сферически-симметричным вариантом DGT-моделей из [16]–[18], однако отражает многие их характерные особенности, основные из которых — наличие высоких производных по пространству и невыпуклость энергии Гельмгольца системы. Отметим, что обычно в литературе рассматриваются одномерные DGT-модели, соответствующие плоско-параллельному случаю. При построении аппроксимаций чаще всего рассматривается их запись в неконсервативной форме. Сферически-симметричная постановка рассмотрена в [10], но и в этой работе используется неконсервативная форма записи уравнений. Новизна настоящей работы, с одной стороны, состоит в выбранной постановке: рассматривается сферически-симметричная *консервативная* модель DGT. С другой — в использовании невыпуклой части свободной энергии полиномиального

вида. Обычно в рамках DGT-моделей рассматриваются более сложные реалистичные уравнения состояния, которые часто оказываются избыточно сложными в рамках постановок, рассматриваемых при моделировании нестационарных гидродинамических течений с капиллярными эффектами.

Модели типа диффузной границы в целом и DGT-модели в частности являются феноменологическими и требуют для своего замыкания ряд зависимостей, вид которых, с одной стороны, не может быть получен априори, а с другой — существенно влияет на характер эволюции системы. Типичным примером является влияние способа задания диффузионной подвижности на так называемый эффект сжатия капли («droplet shrinkage effect») (см. [19]). Это явление характерно для всех моделей типа фазового поля и состоит в следующем: в зависимости от радиуса капли, толщины межфазной границы и объема окружающей (объемлющей) фазы при стремлении к равновесию одиночная капля может уменьшаться в размерах вплоть до своего полного растворения во вмещающей фазе или сохранять свой объем на больших временах. В настоящей работе этот эффект исследуется численно с применением разработанных вычислительных алгоритмов. При этом сферическая симметрия постановки задачи необходима для аккуратного представления формы капли, а градиентная устойчивость схемы, помимо увеличения шага по времени, позволяет считать, что наблюдаемые в расчетах эффекты связаны с термодинамикой двухфазного равновесия, а не с численными артефактами.

2. ОБОЗНАЧЕНИЯ

В настоящей работе использованы следующие обозначения: Ω — ограниченная область с кусочно-гладкой границей $\partial\Omega$; t и r — временная и радиальная координаты; $\alpha, \beta = 1, 2$ — индексы, используемые для обозначения номера компонента смеси; $\partial_r := \partial/\partial r$ и $\partial_t := \partial/\partial t$ — частные производные по радиальной координате и времени соответственно; δ_r, δ_r^* — разностные отношения (аппроксимации производной ∂_r). Жирный шрифт используется для обозначения векторов или наборов переменных (например, $\mathbf{p} := (\rho_1, \rho_2)$).

Для произвольных гладких скалярных функций f и g , произвольных гладких векторных полей $\mathbf{u}$ и $\mathbf{v}$: ∇f — градиент f , $\partial_g f := \partial f / \partial g$ — частная производная по g (если $f = f(g, \cdot)$), $\mathbf{u} \cdot \mathbf{v}$ — скалярное произведение. Для некоторого функционала $\mathcal{F} = \mathcal{F}[f, \cdot]$: $\delta \mathcal{F}$ — его первая вариация, $\delta \mathcal{F} / \delta f$ — вариационная производная функционала.

По повторяющимся индексам α и β предполагается суммирование от 1 до 2.

3. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ДВУХФАЗНОЙ ДВУХКОМПОНЕНТНОЙ ЖИДКОСТИ

3.1. Общее описание модели

Для описания двухфазной двухкомпонентной смеси будем использовать модель, основанную на теории градиента плотности (см. [9]–[12]). В рамках нее предполагается, что полная свободная энергия Гельмгольца смеси Ψ зависит не только от пространственных распределений плотностей компонентов $\mathbf{p}(\mathbf{x}) := (\rho_1(\mathbf{x}), \rho_2(\mathbf{x}))$, но также от распределений их градиентов $\nabla \rho_\alpha(\mathbf{x})$, $\alpha = 1, 2$. Функционал Ψ имеет следующий вид:

$$\Psi[\mathbf{p}] := \int_{\Omega} \left\{ \psi_0(\mathbf{p}) + \frac{1}{2} \lambda_{\alpha\beta} \nabla \rho_\alpha \cdot \nabla \rho_\beta \right\} d\Omega. \quad (3.1)$$

Здесь ψ_0 — объемная плотность локальной (так как зависит только от ρ_α) части свободной энергии Гельмгольца. Поскольку смесь предполагается двухфазной (далее фазы будем обозначать буквами A и B), функцию $\psi_0(\mathbf{p})$ полагают невыпуклой и обладающей двумя (по числу фаз) минимумами. В настоящей работе зададим ее в виде

$$\psi_0(\rho_1, \rho_2) = A_\psi \psi^A \psi^B, \quad (3.2)$$

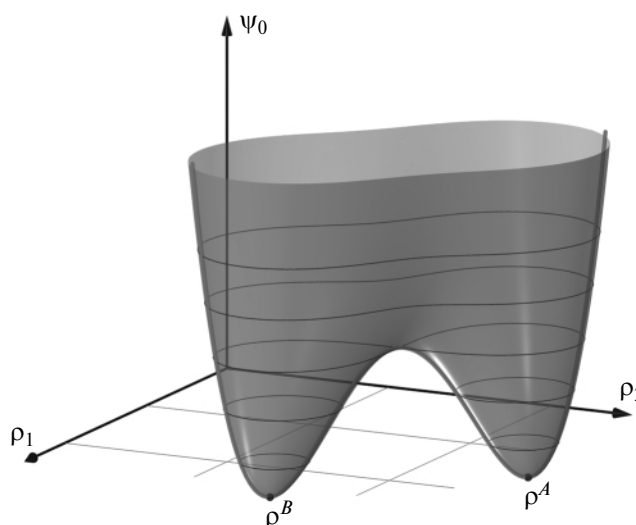
$$\psi^A(\rho_1, \rho_2) = (\rho_1 - \rho_1^A)^2 + (\rho_2 - \rho_2^A)^2, \quad (3.3)$$

$$\psi^B(\rho_1, \rho_2) = (\rho_1 - \rho_1^B)^2 + (\rho_2 - \rho_2^B)^2, \quad (3.4)$$

где $A_\psi > 0$ — некоторая константа, $\rho_\alpha^A, \rho_\alpha^B$ — значения плотностей компонентов в равновесном состоянии. На фиг. 1 приведен характерный вид соответствующей поверхности.

В соотношении (3.1) слагаемое, зависящее от $\nabla \rho_\alpha$, представляет собой градиентную энергию. Она существенно отличается от нуля только в межфазном слое. Вместе с невыпуклостью ψ_0 оно позволяет учесть поверхностное натяжение и обеспечить конечную толщину межфазного слоя. Параметры влияния (градиентные параметры) $\lambda_{\alpha\beta}$ образуют положительно-определенную симметричную матрицу 2×2 (по количеству компонентов). В настоящей работе будем считать, что $\lambda_{\alpha\beta}$ являются постоянными.

Параметры $A_\psi, \rho_\alpha^A, \rho_\alpha^B$ и $\lambda_{\alpha\beta}$ полностью определяют структуру межфазного слоя и, в частности, поверхностное (межфазное) натяжение и толщину межфазной границы (слоя).



Фиг. 1. График поверхности свободной энергии, заданной соотношениями (3.2) – (3.4). Красной линией отмечено сечение поверхности плоскостью, проходящей через точки ρ^A и ρ^B и перпендикулярной плоскости $\rho_1\rho_2$.

Равновесные распределения ρ_α таковы, что на них функционал (3.1) достигает минимума. Необходимым условием для этого, как известно, является выполнение уравнений Эйлера–Лагранжа, которые следуют из равенства нулю первой вариации $\delta\Psi$ и имеют вид

$$\frac{\delta\Psi}{\delta\rho_\alpha} = 0, \quad \alpha = 1, 2. \quad (3.5)$$

Здесь в левой части стоит вариационная производная функционала Ψ по функции ρ_α . Напомним, что вариационная производная определяется из равенства (см., например, [20, гл. 2, разд. 35], [21, разд. II, гл. 3, § 3], [22, гл. 1, § 6])

$$\delta\Psi[\rho, u] := \int_{\Omega} \frac{\delta\Psi}{\delta\rho_\alpha} u_\alpha d\Omega, \quad (3.6)$$

если оно выполнено для произвольных функций приращений $u := (u_1, u_2)$, таких что $u_\alpha = 0$ на $\partial\Omega$. В свою очередь первая вариация определяется как (см. [20]–[22])

$$\delta\Psi[\rho, u] := \left. \frac{d}{d\varepsilon} \Psi[\rho + \varepsilon u] \right|_{\varepsilon=0}.$$

Поскольку на границе области $u_\alpha = 0$, имеем

$$\frac{\delta\Psi}{\delta\rho_\alpha} = \mu_\alpha - \lambda_{\alpha\beta} \Delta\rho_\beta, \quad \mu_\alpha = \partial_{\rho_\alpha} \Psi_0, \quad (3.7)$$

где μ_α — (классический) химический потенциал компонента с номером α .

Далее будем предполагать, что эволюция состояния смеси описывается уравнениями баланса массы компонентов

$$\partial_t \rho_\alpha = -\operatorname{div} b_\alpha, \quad \alpha = 1, 2, \quad (3.8)$$

где векторы b_α — диффузионные потоки компонентов:

$$b_\alpha = -M_{\alpha\beta} \nabla \hat{\mu}_\beta, \quad \alpha = 1, 2. \quad (3.9)$$

Здесь $M_{\alpha\beta}(\rho_1, \rho_2)$ — коэффициенты диффузионной подвижности, которые образуют положительно-полуопределенную симметричную матрицу и, вообще говоря, зависят от плотностей компонентов ρ_α . Символом $\hat{\mu}_\alpha$ обозначен обобщенный химический потенциал, по определению, равный вариационной производной Ψ :

$$\hat{\mu}_\alpha := \frac{\delta\Psi}{\delta\rho_\alpha} = \mu_\alpha - \lambda_{\alpha\beta} \Delta\rho_\beta, \quad \mu_\alpha = \partial_{\rho_\alpha} \Psi_0. \quad (3.10)$$

Из (3.5) и (3.10) следует, что в равновесии $\hat{\mu}_1 = \hat{\mu}_2 = 0$. Обратим внимание, что из (3.5) и (3.7) видно, что параметры диффузионной подвижности $M_{\alpha\beta}$ не влияют на равновесное состояние. Однако, поскольку $M_{\alpha\beta}$ входят в эволюционное уравнение (3.8), они определяют *как* и с какой скоростью смесь стремится к равновесному состоянию. Это хорошо видно из следующего соотношения, справедливого для рассматриваемой модели:

$$\frac{d\Psi}{dt} = - \int_{\Omega} M_{\alpha\beta} \nabla \hat{\mu}_{\alpha} \cdot \nabla \hat{\mu}_{\beta} d\Omega \leq 0. \quad (3.11)$$

Неравенство (3.11) выражает свойство диссипативности рассматриваемой модели: при стремлении к равновесию полная энергия Гельмгольца является невозрастающей функцией времени. Применительно к решениям уравнений модели, (3.11) выражает свойство их градиентной (или, что то же, энергетической) устойчивости.

Важной характеристикой межфазной границы является поверхностное (межфазное) натяжение, величина которого описывается коэффициентом σ . Если межфазная граница является плоской и перпендикулярна оси Ox , то σ определяется следующим образом:

$$\sigma = \int_{-\infty}^{\infty} \lambda_{\alpha\beta} (\partial_x \rho_{\alpha}) (\partial_x \rho_{\beta}) dx. \quad (3.12)$$

3.2. Сферически-симметричная одномерная постановка

В настоящей работе рассмотрен случай одномерной сферически-симметричной постановки в области $\Omega = (R_0, R)$, где $R_0 \geq 0$, $R > R_0$. Радиальная координата обозначается буквой r . В данном случае полная энергия Гельмгольца (3.1) и уравнения модели (3.8)–(3.10) примут соответственно вид

$$\Psi[\rho] = \int_{R_0}^R \left\{ \Psi_0(\rho) + \frac{1}{2} \lambda_{\alpha\beta} \nabla \rho_{\alpha} \cdot \nabla \rho_{\beta} \right\} r^2 dr \quad (3.13)$$

и

$$\partial_t \rho_{\alpha} = - \frac{1}{r^2} \partial_r (r^2 b_{\alpha}), \quad (3.14)$$

$$b_{\alpha} = -M_{\alpha\beta} \partial_r \hat{\mu}_{\beta}, \quad (3.15)$$

$$\hat{\mu}_{\alpha} = \mu_{\alpha} - \lambda_{\alpha\beta} \frac{1}{r^2} \partial_r (r^2 \partial_r \rho_{\beta}), \quad \mu_{\alpha} = \partial_{\rho_{\alpha}} \Psi_0. \quad (3.16)$$

На границах области зададим следующие условия:

$$\partial_r \rho_{\alpha}|_{r=R_0} = \partial_r \rho_{\alpha}|_{r=R} = 0, \quad (3.17)$$

$$b_{\alpha}|_{r=R_0} = b_{\alpha}|_{r=R} = 0. \quad (3.18)$$

Если же $R_0 = 0$, то при $r = 0$ дополнительно наложим условия симметрии

$$\partial_r b_{\alpha} = 0, \quad \partial_{rr}^2 \rho_{\alpha} = 0.$$

Отметим, что в этом случае при $r \rightarrow 0$ в уравнении (3.14) не возникает неустраняемой особенности. Действительно,

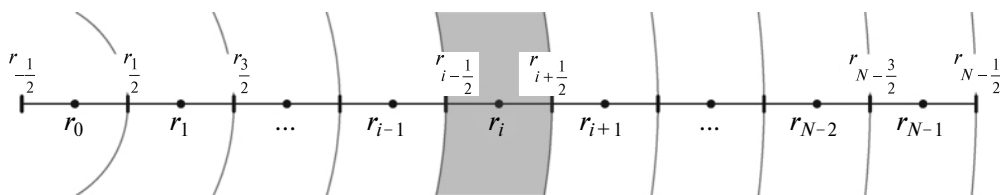
$$\lim_{r \rightarrow 0} \frac{1}{r^2} \partial_r (r^2 b_{\alpha}) = \lim_{r \rightarrow 0} \left(\frac{2b_{\alpha}}{r} + \partial_r b_{\alpha} \right) = 2 \lim_{r \rightarrow 0} \frac{b_{\alpha}}{r} = 2 \lim_{r \rightarrow 0} \partial_r b_{\alpha} = 0,$$

где в предпоследнем равенстве применено правило Лопиталя. Аналогично показывается корректность для соотношений (3.16).

В рассматриваемом случае сферической симметрии коэффициент поверхностного натяжения σ будем вычислять по формуле (вместо (3.12))

$$\sigma = \int_{R_0}^R \lambda_{\alpha\beta} (\partial_r \rho_{\alpha}) (\partial_r \rho_{\beta}) dr. \quad (3.19)$$

Напомним, по повторяющимся индексам α и β подразумевается суммирование от 1 до 2.



Фиг. 2. Иллюстрация используемых пространственных сеток вдоль радиального направления r . Каждая ячейка представляет собой сферический слой.

4. ДИСКРЕТИЗАЦИЯ

В настоящем разделе приведен полудискретный (непрерывный по времени и дискретный по пространству) конечно-разностный метод для аппроксимации системы (3.14) – (3.16) в области $\bar{\Omega} = [R_0, R]$. Показано, что данная дискретизация наследует свойство диссипативности исходной («непрерывной») математической модели (см. (3.11)).

Разобьем отрезок $[R_0, R]$ на N ячеек (см. фиг. 2). Введем одномерные пространственные внутренние сетки ω_h и ω_h^* , а также соответствующие расширенные сетки $\bar{\omega}_h$ и $\bar{\omega}_h^*$:

$$\begin{aligned}\omega_h^* &:= \{r_i\}_{i=0}^{N-1}, & \omega_h &:= \{r_{i-\frac{1}{2}}\}_{i=1}^{N-1}, \\ \bar{\omega}_h^* &:= \{r_{-1}\} \cup \omega_h \cup \{r_N\}, & \bar{\omega}_h &:= \{r_{-\frac{1}{2}}\} \cup \omega_h \cup \{r_{N-\frac{1}{2}}\},\end{aligned}$$

где $r_i := h(i + 1/2)$, $r_{i-1/2} := hi$, $h = (R - R_0)/N$ — шаг по пространству. Таким образом, элементами сетки ω_h^* являются центры ячеек, а ω_h — их грани (для расширенных сеток аналогично).

Пусть $H(\omega)$ — пространство функций, определенных на сетке ω . Для функций $v \in H(\bar{\omega}_h)$ и $f \in H(\bar{\omega}_h^*)$ определим операторы усреднений и разностных отношений:

$$\begin{aligned}(s_r^* f)_{i-\frac{1}{2}} &= \frac{1}{2}(f_i + f_{i-1}), & (s_r v)_i &= \frac{1}{2}(v_{i+\frac{1}{2}} + v_{i-\frac{1}{2}}), \\ (\delta_r^* f)_{i-\frac{1}{2}} &= \frac{1}{h}(f_i - f_{i-1}), & (\delta_r v)_i &= \frac{1}{h}(v_{i+\frac{1}{2}} - v_{i-\frac{1}{2}}).\end{aligned}$$

Ясно, что $\delta_r^*, s_r^* : H(\bar{\omega}_h^*) \rightarrow H(\bar{\omega}_h)$ и $\delta_r, s : H(\bar{\omega}_h) \rightarrow H(\bar{\omega}_h^*)$.

Пусть $u \in H(\bar{\omega}_h)$, $g \in H(\bar{\omega}_h^*)$. На сетках ω_h и $\bar{\omega}_h^*$ введем скалярные произведения сеточных функций:

$$(u, v) := h \sum_{i=0}^N u_{i-\frac{1}{2}} v_{i-\frac{1}{2}} r_{i-\frac{1}{2}}^2, \quad (f, g)_* := h \sum_{i=0}^{N-1} f_i g_i \tilde{r}_i^2, \quad (4.1)$$

где используется обозначение

$$\tilde{r}_i^2 := \frac{1}{3} \left(r_{i+\frac{1}{2}}^2 + r_{i+\frac{1}{2}} r_{i-\frac{1}{2}} + r_{i-\frac{1}{2}}^2 \right) = r_i^2 + \frac{1}{12} h^2. \quad (4.2)$$

Таким образом, $\tilde{r}_i^2$ можно рассматривать как аппроксимацию величины r_i^2 .

Для введенных скалярных произведений несложно доказать тождество, являющееся дискретным аналогом формулы интегрирования по частям:

$$\left(\frac{1}{\tilde{r}^2} \delta_r(r^2 u), f \right)_* = f_{N+1} r_{N-\frac{1}{2}}^2 u_{N-\frac{1}{2}} - f_{-1} r_{-\frac{1}{2}}^2 u_{-\frac{1}{2}} - (u, \delta_r^* f),$$

которое в частном случае $u_{N-1/2} = u_{-1/2} = 0$ принимает вид

$$\left(\frac{1}{\tilde{r}^2} \delta_r(r^2 u), f \right)_* = -(u, \delta_r^* f). \quad (4.3)$$

Используя введенные выше обозначения, определим дискретный аналог полной энергии Гельмгольца (3.13) следующим образом:

$$\Psi_h[\mathbf{p}(t)] := (\psi_0, 1)_* + \frac{1}{2} \lambda_{\alpha\beta} (\delta_r^* \rho_\alpha, \delta_r^* \rho_\beta). \quad (4.4)$$

Вычислим первую вариацию функционала $\Psi_h[\rho]$:

$$\begin{aligned}\delta\Psi_h[\rho, u] &:= \lim_{\varepsilon \rightarrow 0} \frac{\Psi_h[\rho + \varepsilon u] - \Psi_h[\rho]}{\varepsilon} = \\ &= (\partial_{\rho_\alpha} \Psi_0, u_\alpha)_* + \frac{1}{2} \lambda_{\alpha\beta} (\delta_r^* \rho_\alpha, \delta_r^* u_\beta) + \frac{1}{2} \lambda_{\alpha\beta} (\delta_r^* u_\alpha, \delta_r^* \rho_\beta) = \\ &= (\partial_{\rho_\alpha} \Psi_0, u_\alpha)_* + \lambda_{\alpha\beta} (\delta_r^* \rho_\beta, \delta_r^* u_\alpha) =\end{aligned}\quad (4.5)$$

$$\begin{aligned}&= (\partial_{\rho_\alpha} \Psi_0, u_\alpha)_* - \lambda_{\alpha\beta} \left(\frac{1}{\tilde{r}^2} \delta_r (r^2 \delta_r^* \rho_\beta), u_\alpha \right)_* = \\ &= \left(\partial_{\rho_\alpha} \Psi_0 - \lambda_{\alpha\beta} \frac{1}{\tilde{r}^2} \delta_r (r^2 \delta_r^* \rho_\beta), u_\alpha \right)_* \equiv \left(\frac{\delta\Psi_h}{\delta\rho_\alpha}, u_\alpha \right)_*,\end{aligned}\quad (4.6)$$

где при переходе к равенству (4.5) использована симметричность $\lambda_{\alpha\beta}$, в равенстве (4.6) — тождество (4.3). Отсюда видно, что вариационная производная, а значит, и дискретизация обобщенного химического потенциала $\hat{\mu}_\alpha$ имеют вид

$$\hat{\mu}_\alpha := \frac{\delta\Psi_h}{\delta\rho_\alpha} = \partial_{\rho_\alpha} \Psi_0 - \lambda_{\alpha\beta} \frac{1}{\tilde{r}^2} \delta_r (r^2 \delta_r^* \rho_\beta). \quad (4.7)$$

4.1. Полудискретный метод

В данном пункте рассмотрим полудискретный (непрерывный по времени и дискретный по пространству) метод для решения уравнений модели (3.14) — (3.16). Воспользуемся методом конечных объемов (следуя, например, [23, гл. XI, § 1, разд. 7]). При этом конечный объем в силу специфики рассматриваемой постановки представляет собой сферический слой. Предполагая, что плотность компонентов в пределах сферического слоя не зависит от r , запишем балансовые соотношения для i -го слоя (см. фиг. 2, где этот слой выделен цветом):

$$\frac{4\pi}{3} \left(r_{i+\frac{1}{2}}^3 - r_{i-\frac{1}{2}}^3 \right) \partial_t \rho_\alpha = -4\pi \left(r_{i+\frac{1}{2}}^2 b_{\alpha, i+\frac{1}{2}} - r_{i-\frac{1}{2}}^2 b_{\alpha, i-\frac{1}{2}} \right). \quad (4.8)$$

Деля обе части равенства на 4π и на $r_{i+\frac{1}{2}} - r_{i-\frac{1}{2}} \equiv h$, а также учитывая равенство (4.2), окончательно получаем

$$\partial_t \rho_\alpha = -\frac{1}{\tilde{r}^2} \delta_r (r^2 b_\alpha), \quad (4.9)$$

$$b_\alpha = -M_{\alpha\beta} \delta_r^* \hat{\mu}_\beta. \quad (4.10)$$

Отсюда, в частности, видно, что аппроксимация (4.2) для r_i^2 возникает естественным образом.

Дискретизация для обобщенного химического потенциала имеет вид (4.7).

Дополним (4.7), (4.9) и (4.10) следующими граничными условиями:

$$b_{\alpha, -\frac{1}{2}} = 0, \quad b_{\alpha, N-\frac{1}{2}} = 0, \quad (4.11)$$

$$(\delta_r \rho_\alpha)_{-\frac{1}{2}} = (\delta_r \rho_\alpha)_{N-\frac{1}{2}} = 0. \quad (4.12)$$

Заметим, что в данной дискретной постановке (поскольку в ее основе лежит метод конечных объемов) не возникает необходимости дополнительно вводить условия, обеспечивающие регулярность задачи при $R_0 = 0$: в этом случае $r_{-1/2} = 0$ и $\tilde{r}_0^2 = \frac{1}{3}h^2 > 0$.

Также система дополняется начальными условиями

$$\rho_{\alpha, i} = \rho_{\alpha, i}^0 := \rho_\alpha^0(r_i), \quad t = 0.$$

Для корректного использования введенных конечно-разностных операторов продолжим функции ρ_α и $\hat{\mu}_\alpha$ на расширенную сетку $\bar{\omega}_h^*$ симметрично:

$$\rho_{\alpha, -1} = \rho_{\alpha, 0}, \quad \rho_{\alpha, N} = \rho_{\alpha, N-1}, \quad (4.13)$$

$$\hat{\mu}_{\alpha, -1} = \hat{\mu}_{\alpha, 0}, \quad \hat{\mu}_{\alpha, N} = \hat{\mu}_{\alpha, N-1}. \quad (4.14)$$

При таком способе продолжения выполнены условия (4.11) и (4.12).

Дискретный аналог интегрального баланса массы компонента имеет вид

$$(\rho_\alpha, 1)_* = (\rho_\alpha^0, 1)_*, \quad t \geq 0.$$

Для рассматриваемой полудискретной постановки выполнено свойство диссипативности по полной энергии (полудискретный аналог неравенства (3.11)):

Теорема 4.1. Полудискретный метод (4.7), (4.9), (4.10), дополненный условиями (4.11) и (4.12), наследует свойство диссипативности исходной (непрерывной) модели. Именно,

$$\frac{d\Psi_h[\mathbf{p}(t)]}{dt} = -(M_{\alpha\beta}\delta_r^*\hat{\mu}_\beta, \delta_r^*\hat{\mu}_\alpha) \leq 0. \quad (4.15)$$

Доказательство. Скалярно умножим (4.9) на (4.7) и просуммируем по α :

$$(\hat{\mu}_\alpha, \partial_t \rho_\alpha)_* = - \left(\hat{\mu}_\alpha, \frac{1}{\tilde{r}} \delta_r^*(r^2 b_\alpha) \right)_*. \quad (4.16)$$

Рассмотрим левую часть равенства (4.16):

$$\begin{aligned} (\hat{\mu}_\alpha, \partial_t \rho_\alpha)_* &= (\partial_{\rho_\alpha} \Psi_0, \partial_t \rho_\alpha)_* - \lambda_{\alpha\beta} \left(\frac{1}{\tilde{r}^2} \delta_r(r^2 \delta_r^* \rho_\beta), \partial_t \rho_\alpha \right)_* = \\ &= (\partial_t \Psi_0, 1)_* + \lambda_{\alpha\beta} (\delta_r^* \rho_\beta, \delta_r^* \partial_t \rho_\alpha) = \\ &= (\partial_t \Psi_0, 1)_* + \frac{1}{2} \lambda_{\alpha\beta} \partial_t (\delta_r^* \rho_\alpha, \delta_r^* \rho_\beta) = \frac{d\Psi_h}{dt}. \end{aligned} \quad (4.17)$$

Здесь использованы формула (4.3), равенства $\delta_r^* \partial_t = \partial_t \delta_r^*$, а также симметричность постоянных коэффициентов $\lambda_{\alpha\beta}$ по индексам.

Преобразуем правую часть равенства (4.16) с помощью тождества (4.3):

$$- \left(\hat{\mu}_\alpha, \frac{1}{\tilde{r}} \delta_r^*(r^2 b_\alpha) \right)_* = (\delta_r^* \hat{\mu}_\alpha, b_\alpha) = -(M_{\alpha\beta} \delta_r^* \hat{\mu}_\beta, \delta_r^* \hat{\mu}_\alpha). \quad (4.18)$$

Подставим (4.17) и (4.18) в (4.16), учтем, что матрица $M_{\alpha\beta}$ является положительно-полуопределенной, и окончательно получим (4.15).

4.2. Полностью дискретный метод

Для построения полностью дискретного метода, обладающего свойством диссипативности, воспользуемся методом выпуклого расщепления (convex-splitting method) свободной энергии (расщепление Эйра), предложенным в работе [13].

Его суть заключается в представлении исходного функционала свободной энергии Ψ в виде суммы выпуклого и вогнутого функционалов:

$$\Psi = \Psi^c + (-\Psi^e) \equiv \Psi^c - \Psi^e.$$

При построении дискретизации по времени выпуклая часть Ψ^c аппроксимируется неявным образом, а вогнутая $-\Psi^e$ — явным.

Представим объемную плотность однородной части свободной энергии в виде

$$\Psi_0(\rho_1, \rho_2) = \Psi^c(\rho_1, \rho_2) - \Psi^e(\rho_1, \rho_2), \quad (4.19)$$

$$\Psi^c(\rho_1, \rho_2) := A_\Psi [(\rho_1 - \rho_1^S)^2 + (\rho_2 - \rho_2^S)^2 + \frac{1}{4} l_{AB}^2]^2, \quad (4.20)$$

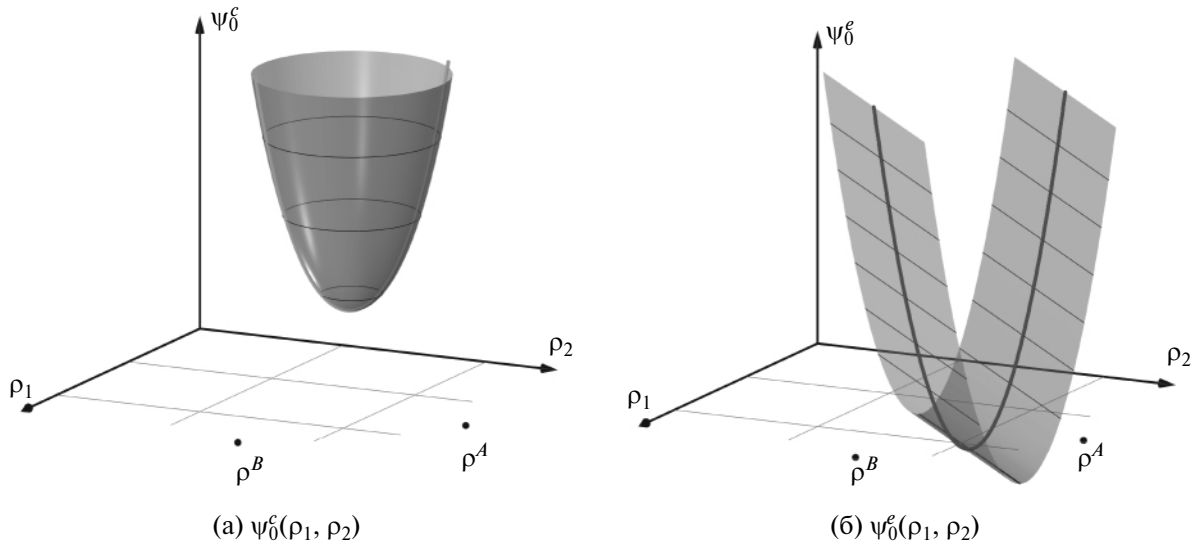
$$\Psi^e(\rho_1, \rho_2) := 4A_\Psi [(\rho_1 - \rho_1^S)k_1 + (\rho_2 - \rho_2^S)k_2]^2, \quad (4.21)$$

где введены обозначения

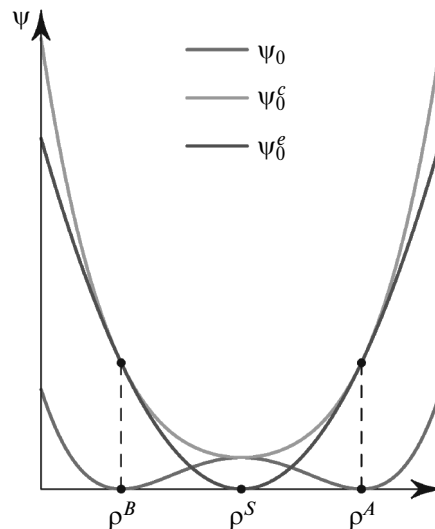
$$\rho_\alpha^S := \frac{1}{2}(\rho_\alpha^A + \rho_\alpha^B), \quad k_\alpha := \frac{1}{2}(\rho_\alpha^B - \rho_\alpha^A), \quad l_{AB}^2 := (\rho_1^B - \rho_1^A)^2 + (\rho_2^B - \rho_2^A)^2. \quad (4.22)$$

На фиг. 3а и фиг. 3б приведены поверхности (4.20) и (4.21), соответствующие Ψ_0 , представленной на фиг. 1. Для наглядности на фиг. 4 приведены графики сечений поверхностей в плоскости, проходящей через точки $\mathbf{p}^A$ и $\mathbf{p}^B$ и перпендикулярной плоскости $\rho_1 \rho_2$. Видно, что поверхности Ψ_0^c и Ψ_0^e касаются в точках $(\rho_1^A, \rho_2^A, \Psi_0^{ce})$ и $(\rho_1^B, \rho_2^B, \Psi_0^{ce})$, где

$$\begin{aligned} \Psi_0^{ce} &= \Psi_0^c(\mathbf{p}^A) = \Psi_0^c(\mathbf{p}^B) = \\ &= \Psi_0^e(\mathbf{p}^A) = \Psi_0^e(\mathbf{p}^B). \end{aligned}$$



Фиг. 3. Выпуклое расщепление функции ψ_0 , график которой представлен на фиг. 1. Зеленой и синей линиями отмечено сечение поверхности плоскостью, проходящей через точки ρ^A и ρ^B и перпендикулярной плоскости $\rho_1\rho_2$.



Фиг. 4. Графики сечений поверхностей плоскостью, проходящей через точки ρ^A и ρ^B и перпендикулярной плоскости $\rho_1\rho_2$. Цвета совпадают с цветами сечений на фиг. 1, 3а и 3б.

Таким образом, с учетом формулы (4.4) имеем

$$\Psi_h^c[\rho] = (\psi_0^c, 1)_* + \frac{1}{2}\lambda_{\alpha\beta} (\delta_r^* \rho_\alpha, \delta_r^* \rho_\beta), \quad (4.23)$$

$$\Psi_h^e[\rho] = (\psi_0^e, 1)_*. \quad (4.24)$$

Функционал Ψ_h^c является выпуклым (строго): первое слагаемое в (4.23) представляет собой выпуклый функционал, поскольку ψ_0^c — строго выпуклая функция, а второе слагаемое является выпуклым функционалом, так как коэффициенты влияния $\lambda_{\alpha\beta}$ образуют симметричную положительно-определенную матрицу. Функционал Ψ_h^e является выпуклым (в нестрогом смысле), так как ψ_0^e выпукла (нестрого).

Аппроксимация по времени обобщенного химического потенциала имеет вид

$$\hat{\mu}_\alpha := \hat{\mu}_\alpha^{c,n+1} - \hat{\mu}_\alpha^{e,n}, \quad (4.25)$$

$$\hat{\mu}_\alpha^{c,n+1} := \frac{\delta \Psi^c}{\delta \rho_\alpha}[\mathbf{p}^{n+1}] = \partial_{\rho_\alpha} \Psi_0^c(\mathbf{p}^{n+1}) - \lambda_{\alpha\beta} \frac{1}{\tilde{r}^2} \delta_r \left(r^2 \delta_r^* \rho_\beta^{n+1} \right), \quad (4.26)$$

$$\hat{\mu}_\alpha^{e,n} := \frac{\delta \Psi^e}{\delta \rho_\alpha}[\mathbf{p}^n] = \partial_{\rho_\alpha} \Psi_0^e(\mathbf{p}^n). \quad (4.27)$$

Обратим внимание, что градиентные слагаемые учитываются на верхнем временном слое (неявным образом). Это гарантирует, что функционал Ψ_h^c является выпуклым.

Аппроксимируя в (4.9) производную по времени разностью вперед (с шагом Δt), с учетом соотношений (4.25) – (4.27), приходим к следующей конечномерной задаче:

$$\frac{\rho_\alpha^{n+1} - \rho_\alpha^n}{\Delta t} = -\frac{1}{\tilde{r}^2} \delta_r(r^2 b_\alpha), \quad (4.28)$$

$$b_\alpha = -M_{\alpha\beta}(s_r^* \mathbf{p}^n) \delta_r^* \hat{\mu}_\beta \equiv -M_{\alpha\beta}(s_r^* \mathbf{p}^n) \delta_r^* \left(\hat{\mu}_\beta^{c,n+1} - \hat{\mu}_\beta^{e,n} \right). \quad (4.29)$$

Опираясь на работу [14] и теорему 4.1, докажем справедливость следующего утверждения.

Теорема 4.2. Полностью дискретный метод (4.25) – (4.27), (4.28), (4.29), дополненный условиями (4.11) и (4.12), является безусловно градиентно устойчивым (диссипативным). А именно, при любом $\Delta t > 0$ выполнено следующее неравенство:

$$\frac{1}{\Delta t} (\Psi_h[\mathbf{p}^{n+1}] - \Psi_h[\mathbf{p}^n]) \leq -(M_{\alpha\beta} \delta_r^* \hat{\mu}_\beta, \delta_r^* \hat{\mu}_\alpha) \leq 0. \quad (4.30)$$

Доказательство. Поскольку Ψ_h^c и Ψ_h^e — выпуклые, верны следующие неравенства (легко получаемые из определений выпуклого функционала и вариационной производной):

$$\left(\frac{\delta \Psi^c}{\delta \rho_\alpha}[\mathbf{p}^{n+1}], (\rho_\alpha^{n+1} - \rho_\alpha^n) \right)_* \geq \Psi_h^c[\mathbf{p}^{n+1}] - \Psi_h^c[\mathbf{p}^n], \quad (4.31)$$

$$- \left(\frac{\delta \Psi^e}{\delta \rho_\alpha}[\mathbf{p}^n], (\rho_\alpha^{n+1} - \rho_\alpha^n) \right)_* \geq -\Psi_h^e[\mathbf{p}^{n+1}] + \Psi_h^e[\mathbf{p}^n], \quad (4.32)$$

напомним, что по повторяющимся индексам α подразумевается суммирование от 1 до 2.

Скалярно умножим левую часть (4.28) на $\hat{\mu}_\alpha$ и воспользуемся неравенствами (4.31) и (4.32):

$$\begin{aligned} \left(\frac{\rho_\alpha^{n+1} - \rho_\alpha^n}{\Delta t}, \hat{\mu}_\alpha \right)_* &= \frac{1}{\Delta t} \left(\rho_\alpha^{n+1} - \rho_\alpha^n, \left\{ \frac{\delta \Psi^c}{\delta \rho_\alpha}[\mathbf{p}^{n+1}] - \frac{\delta \Psi^e}{\delta \rho_\alpha}[\mathbf{p}^n] \right\} \right)_* \geq \\ &\geq \frac{1}{\Delta t} (\Psi_h[\mathbf{p}^{n+1}] - \Psi_h[\mathbf{p}^n]). \end{aligned} \quad (4.33)$$

Скалярно умножая правую часть (4.28) на $\hat{\mu}_\alpha$ и пользуясь (4.18) (см. выкладки из доказательства теоремы 4.1) и (4.33), получаем (4.30), что и завершает доказательство.

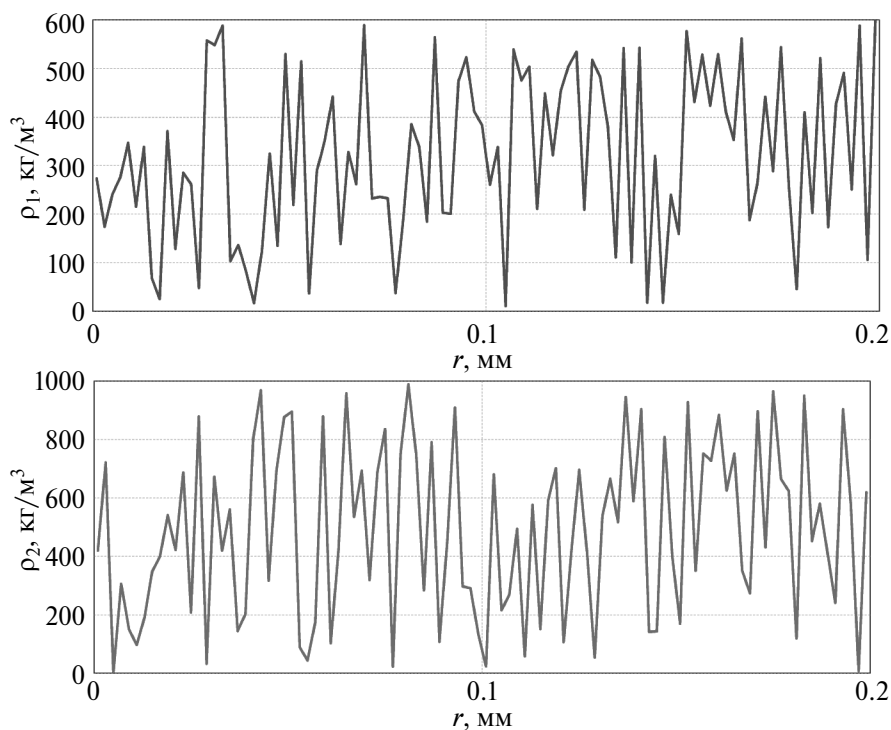
Метод (4.25) – (4.27), (4.28), (4.29) является нелинейным относительно ρ_α^{n+1} . Поэтому для решения соответствующей системы уравнений относительно ρ_α^{n+1} используется метод Ньютона. Для этого линеаризуем слагаемое $\partial_{\rho_\alpha} \Psi_0^c(\mathbf{p}^{n+1})$, разлагая его в ряд Тейлора в окрестности ρ_α^n . Добавим к верхнему индексу $n+1$ индекс s , обозначающий ньютоновские итерации. В результате получим

$$\begin{aligned} \partial_{\rho_\alpha} \Psi_0^c(\mathbf{p}^{n+1,s+1}) &= \partial_{\rho_\alpha} \Psi_0^c(\mathbf{p}^{n+1,s}) + \partial_{\rho_1} \partial_{\rho_\alpha} \Psi_0^c(\mathbf{p}^{n+1,s})(\rho_1^{n+1,s+1} - \rho_1^{n+1,s}) + \\ &+ \partial_{\rho_2} \partial_{\rho_\alpha} \Psi_0^c(\mathbf{p}^{n+1,s})(\rho_2^{n+1,s+1} - \rho_2^{n+1,s}). \end{aligned} \quad (4.34)$$

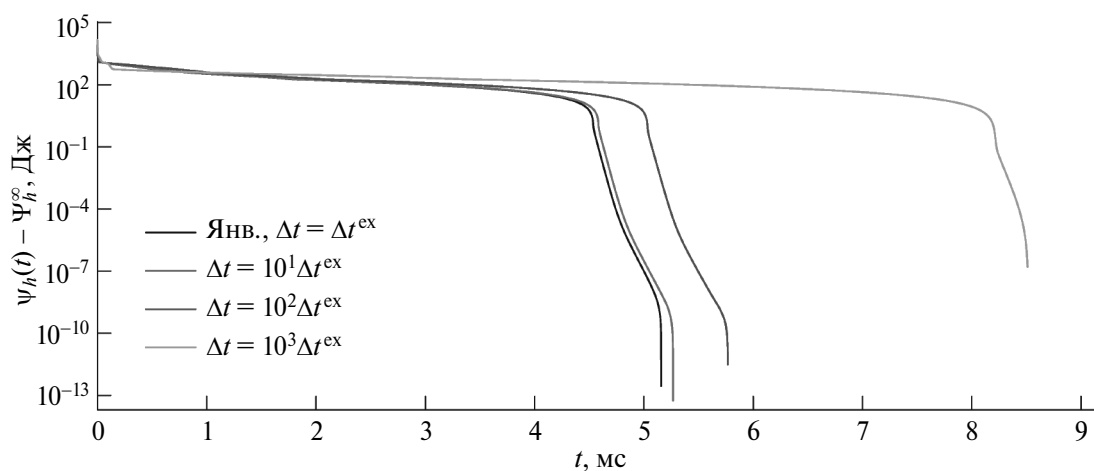
Заменяя $\mathbf{p}^{n+1}$ в (4.25) – (4.27), (4.28), (4.29) на $\mathbf{p}^{n+1,s+1}$ и используя (4.34), получаем линейную задачу. Критерий остановки зададим в виде

$$\max_{\alpha=1,2} \frac{\|\rho_\alpha^{n+1,s+1} - \rho_\alpha^{n+1,s}\|_2}{\|\rho_\alpha^{n+1,s}\|_2} < 10^{-10}.$$

При выполнении последнего неравенства полагаем, что решение для очередного временного слоя получено, т.е. $\mathbf{p}^{n+1} := \mathbf{p}^{n+1,s+1}$. Начальное приближение на каждом временном слое имеет вид $\mathbf{p}^{n+1,0} = \mathbf{p}^n$. Система линейных алгебраических уравнений на каждой итерации метода Ньютона решается методом Гаусса с использованием функции `gesv` из пакета LAPACK.



Фиг. 5. Начальные распределения плотностей компонентов, заданных согласно (5.1).



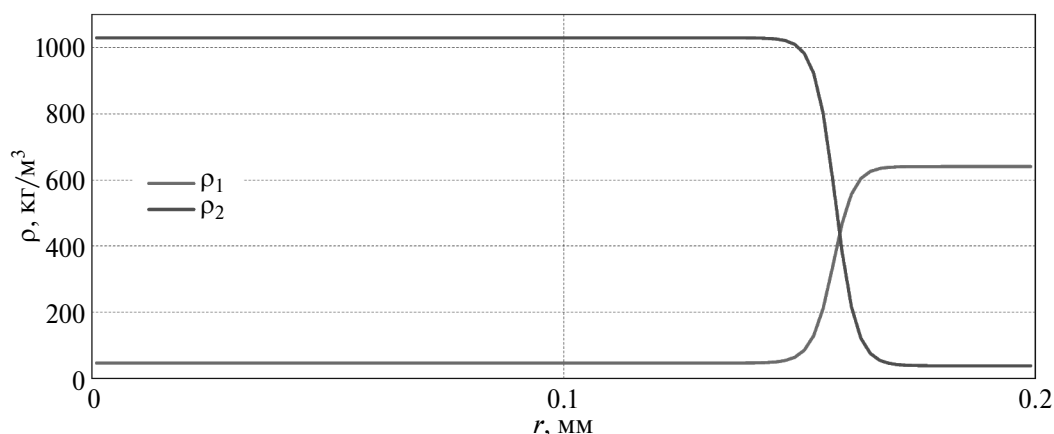
Фиг. 6. Зависимость полной энергии $\Psi_h(t)$ от времени, $\Psi_h^\infty := \Psi_h(t^\infty)$ — опорное (референсное) значение полной энергии, соответствующее равновесному состоянию.

5. ЧИСЛЕННОЕ ИССЛЕДОВАНИЕ ПОСТРОЕННЫХ МЕТОДОВ

В данном разделе представлено численное исследование разработанных методов на примере задачи о «расслоении» двухкомпонентной смеси: начальные распределения плотностей компонентов задаются случайным образом, а затем в смеси самопроизвольно постепенно начинают образовываться подобласти с практически однородным компонентным составом (фазы).

Таблица 1. Равновесные значения полной энергии Гельмгольца Ψ_h^∞ , полученные в задаче со случайными начальными данными для явной схемы и неявных при различных Δt

Δt	Ψ_h^∞	t^∞	n_t^∞
Явная, Δt^{ex}	452.4495827684581855	5.15799	1031598
$10^1 \Delta t^{\text{ex}}$	452.4495827386461428	5.26695	105339
$10^2 \Delta t^{\text{ex}}$	452.4495827430440045	5.768	11536
$10^3 \Delta t^{\text{ex}}$	452.4495637673383044	8.515	1703
$10^4 \Delta t^{\text{ex}}$	452.4495187180730227	45.05	901



Фиг. 7. Равновесные распределения плотностей компонентов ρ_a^∞ , полученные в расчете явным методом. Распределения компонентов, полученные с помощью явно-неявного метода при различных шагах по времени визуально неотличимы от представленного.

Зададим разностную сетку на области $\Omega = (R_0, R)$ с $R_0 = 0$, $R = Nh = 0.2$ мм с числом расчетных ячеек $N = 100$ и шагом по пространству $h = 0.002$ мм. Параметры смеси выберем следующими:

$$\begin{aligned} \rho_1^A &= 600 \text{ кг/м}^3, & \rho_2^A &= 5 \text{ кг/м}^3, \\ \rho_1^B &= 4 \text{ кг/м}^3, & \rho_2^B &= 1000 \text{ кг/м}^3, \\ \lambda_{11} &= \lambda_{22} = 2.8 \times 10^{-7} \text{ мм}^7/(\text{мс}^2 \cdot \text{мкг}), & \lambda_{12} &= \lambda_{21} = 0, \\ M_{11} &= M_{22} = M(\rho), & M_{12} &= M_{21} = 0, \end{aligned}$$

$A_\psi = 2.1 \times 10^{-8} \text{ мм}^5/(\text{мс}^2 \cdot \text{мкг})$. Зададим $M(\rho) = M_0 = 1 \text{ мкг} \cdot \text{мс}/\text{мм}^3$. Начальные условия имеют вид

$$\rho_{\alpha,i}^0 = \rho_\alpha^A + \zeta_{\alpha,i}(\rho_\alpha^B - \rho_\alpha^A), \quad (5.1)$$

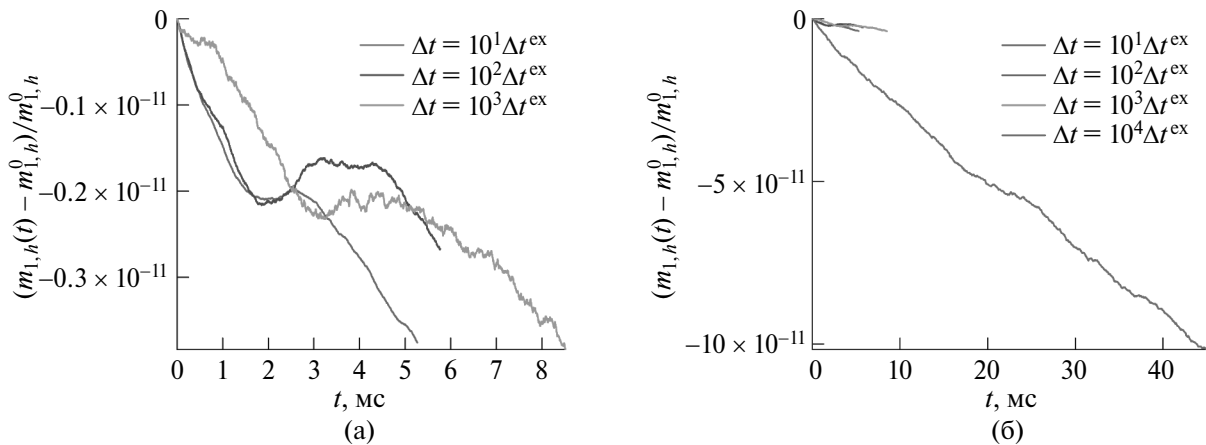
где $\zeta_{\alpha,i} \in H(\omega^*)$ — равномерно распределенная на интервале $(0, 1)$ случайная величина, $\alpha = 1, 2$. На фиг. 5 приведены соответствующие распределения плотностей компонентов для одной из реализаций $\zeta_{\alpha,i}$.

Эмпирически установлено, что при выбранных значениях параметров для явной схемы максимальный шаг по времени, при котором энергия Ψ_h является невозрастающей функцией (см. фиг. 6) равен $\Delta t = \Delta t^{\text{ex}} = 5 \times 10^{-6}$ мс (при $\Delta t = 6 \times 10^{-6}$ мс энергия Ψ_h возрастает). Момент времени $t^\infty = 1031600 \cdot \Delta t^{\text{ex}} = 5.158$ мс здесь принят за момент, при котором достигнуто равновесное состояние. Такой выбор максимального времени расчета обусловлен тем, что при $t > t^\infty$ энергия начинает расти в 11–12 знаках после запятой, причем рост энергии при $t > t^\infty$ связан не со свойствами схемы, а с накоплением ошибок округлений в арифметических операциях в ЭВМ (аналогичный эффект наблюдался, например, в [16]).

Использование явно-неявного метода, как было отмечено выше, позволяет использовать любое значение шага по времени Δt . На фиг. 6 приведено сравнение эволюции полной энергии для различных Δt (приведены значения относительно опорных, которые представлены в табл. 1). Во всех рассмотренных случаях полная энергия является невозрастающей функцией времени. В табл. 1 приведены равновесные значения полной энергии Гельмгольца и соответствующие момент времени t^∞ и номер временного слоя n_t^∞ (начальные распределения задаются на временном слое с номером 0). Видно, что при увеличении Δt финальное значение полной энергии практически не изменяется. Также проверено, что практически одинаковы и равновесные распреде-

Таблица 2. Наблюдаемые значения относительных дисбалансов Δm_α , соответствующие порядки чисел обусловленности в явно- неявном методе при различных Δt

Δt	$\kappa(A)$	Дисбаланс Δm_1	Дисбаланс Δm_2
$10^1 \Delta t^{\text{ex}}$	10^1	-3.76×10^{-12}	-1.90×10^{-11}
$10^2 \Delta t^{\text{ex}}$	10^2	-2.67×10^{-12}	-1.61×10^{-11}
$10^3 \Delta t^{\text{ex}}$	10^3	-3.83×10^{-12}	-2.26×10^{-11}
$10^4 \Delta t^{\text{ex}}$	10^4	-1.01×10^{-10}	-9.36×10^{-11}
$10^5 \Delta t^{\text{ex}}$	10^5	-2.31×10^{-10}	-5.87×10^{-10}
$10^6 \Delta t^{\text{ex}}$	10^6	-3.11×10^{-9}	-4.12×10^{-9}



Фиг. 8. Зависимость относительного отклонения наблюдаемой полной массы $m_{1,h}$ в расчетах с помощью явно- неявного метода.

ления ρ_α . На фиг. 7 представлены равновесные распределения плотностей компонентов, полученных с использованием явной схемы. Видно, что в результате осталась только одна капля (точнее один сферический слой) фазы B .

Приведенные результаты показывают, что время установления t^∞ увеличивается с увеличением Δt , что объясняется ошибкой аппроксимации по времени. Отметим, что из табл. 1 также видно, что

$$\frac{(t^\infty)_{10^4 \Delta t^{\text{ex}}} - (t^\infty)_{10^3 \Delta t^{\text{ex}}}}{(t^\infty)_{10^3 \Delta t^{\text{ex}}} - (t^\infty)_{10^2 \Delta t^{\text{ex}}}} = 11.2 \approx 10.$$

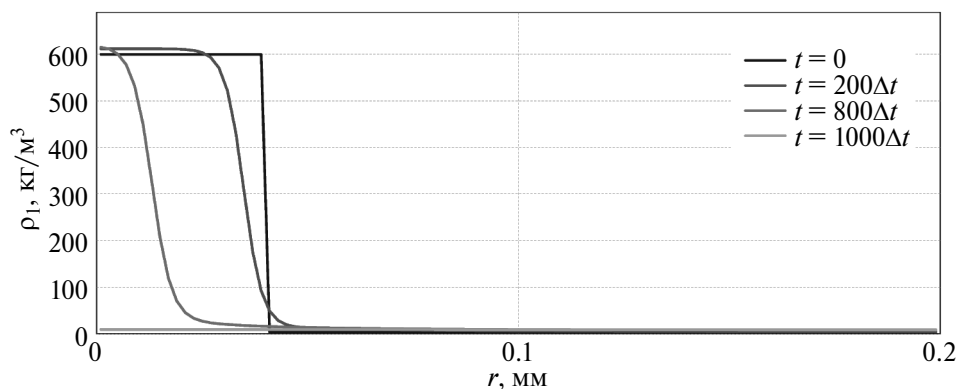
С учетом увеличения Δt в каждом последующем вычислительном эксперименте в 10 раз, это согласуется с первым порядком точности используемой аппроксимации по времени.

Отдельно обсудим консервативность по массе. Как для явного, так и для неявного методов это свойство выполняется теоретически (в смысле точной арифметики). В программной реализации явного метода это свойство выполняется с машинной точностью. Однако в программной реализации явно- неявного метода это свойство в некоторых случаях заметно нарушается при увеличении шага по времени. В табл. 2 для различных Δt приведены наблюдаемые в расчетах дисбалансы массы, рассчитанные по формуле

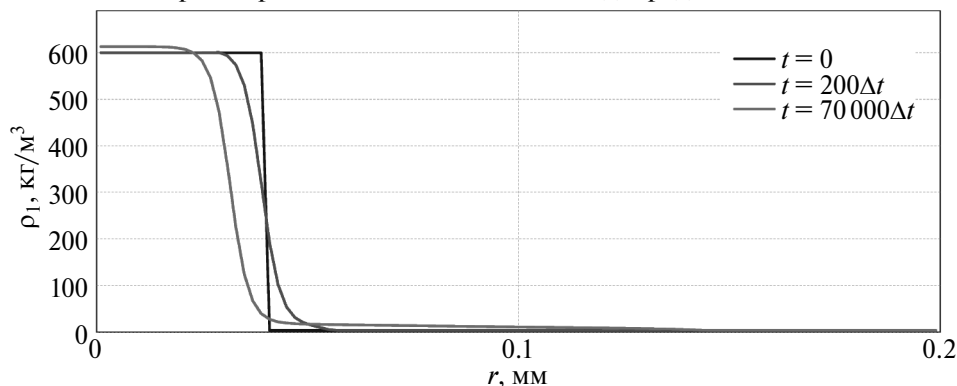
$$\Delta m_\alpha = \frac{m_{\alpha,h}^\infty - m_{\alpha,h}^0}{m_{\alpha,h}^0},$$

где $m_{\alpha,h}^0 := (\rho_\alpha^0, 1)_*$, $m_{\alpha,h}^\infty := (\rho_\alpha^\infty, 1)_*$ — полные массы компонентов с номером α на моменты времени $t = 0$ и $t = t^\infty$ соответственно (равновесное состояние).

Наличие дисбалансов связано с конечной точностью решения системы нелинейных уравнений методом Ньютона и ненулевой ошибкой решения системы линейных уравнений на каждой его итерации. Последняя может быть существенно больше машинного нуля при больших шагах по времени — как показывают представленные результаты, число обусловленности $\kappa(A) := \|A^{-1}\| \cdot \|A\|$ системы линейных алгебраических уравнений линейно растет по порядку величины при увеличении шага по времени. В результате можно наблюдать, что чем больше шаг по времени, тем больше среднее число обусловленности и тем сильнее дисбаланс по массе.

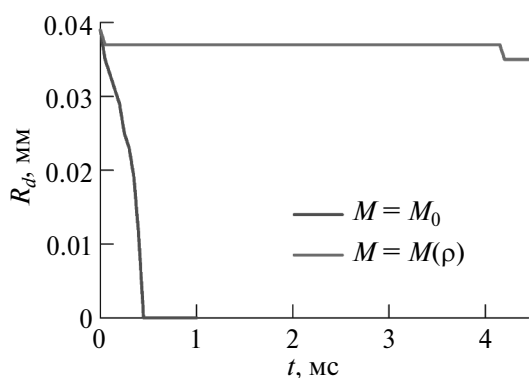


(а) Постоянная $M = M_0$. При стремлении смеси к равновесию капля фазы А полностью «растворяется» и смесь становится однородной.

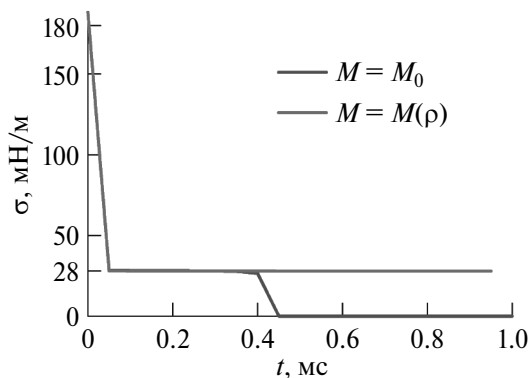


(б) $M = M(\rho)$, заданная выражением (6.1). Уменьшение радиуса капли происходит значительно медленнее. Капля существует при $t \gg t_s$.

Фиг. 9. Эволюция межфазной границы капли с $R_d^0 = 0.2R$ при различных способах задания диффузионной подвижности.



Фиг. 10. Радиус $R_d(t)$ при различных видах M .

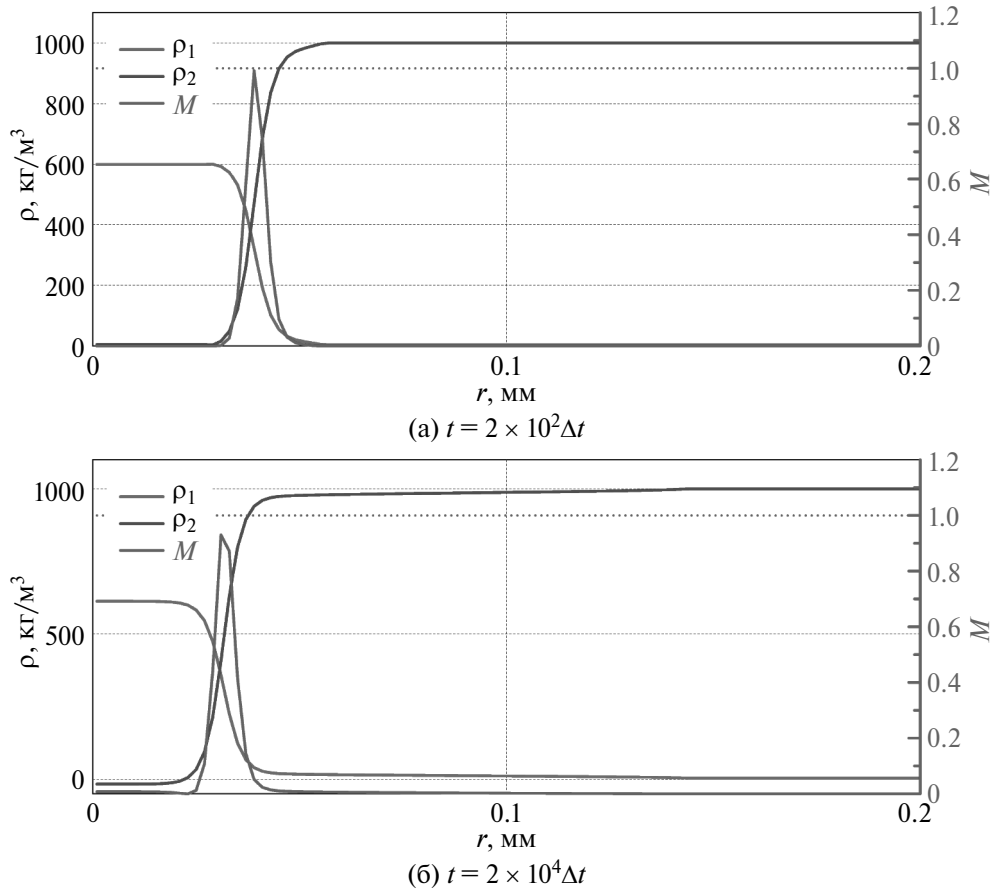


Фиг. 11. Поверхностное натяжение $\sigma(t)$ при различных видах M .

В дополнение на фиг. 8 представлена зависимость полной массы компонента $\alpha = 1$ от времени. Видно, что при небольших шагах относительное отклонение составляет величину порядка 10^{-12} (см. фиг. 8а), при шаге $\Delta t = 10^4 \Delta t^{\text{ex}}$ дисбаланс массы более значительный (см. фиг. 8б). Аналогичная ситуация имеет место и для компонента с номером $\alpha = 2$.

6. ЧИСЛЕННОЕ ИССЛЕДОВАНИЕ ВЛИЯНИЯ ВИДА ДИФФУЗИОННОЙ ПОДВИЖНОСТИ

В данном разделе приведен пример применения разработанного выше явно-неявного метода для численного исследования влияния способа задания диффузионной подвижности $M(\rho)$ на динамику и свойства межфазной границы. Параметры смеси возьмем из предыдущего раздела. Шаг по времени выберем как $\Delta t = 10^2 \Delta t^{\text{ex}}$.



Фиг. 12. Пространственные распределения ρ_α (левая ось) и $M(\rho)$ (правая ось) в различные моменты времени

В начальный момент времени зададим одну каплю фазы A , окруженную фазой B :

$$\rho_{\alpha,i} = \rho_\alpha^A + \theta(r_i) (\rho_\alpha^B - \rho_\alpha^A), \quad \theta(r_i) = \begin{cases} 1, & r_i < R_d^0, \\ 0, & r_i \geq R_d^0, \end{cases}$$

где R_d^0 — начальное значение радиуса капли. Далее везде положим $R_d^0 = 0.2R$.

На фиг. 9а приведены распределения ρ_1 , полученные при $M(\rho) = M_0 = 1 \text{ мкг} \cdot \text{мс}/\text{мм}^3$. При стремлении к равновесию капля фазы A полностью растворяется: в момент времени $t = t_s \approx 0.5 \text{ мс}$ смесь становится гомогенной (состоит из фазы B).

На фиг. 9б приведены распределения ρ_1 для различных моментов времени t , полученные с диффузионной подвижностью, заданной в виде (параметр l_{AB}^2 определен в (4.22))

$$M(\rho) = M_0 16 \frac{l_A^2(\rho) l_B^2(\rho)}{l_{AB}^4}, \quad M_0 = 1 \text{ мкг} \cdot \text{мс}/\text{мм}^3, \quad (6.1)$$

$$l_A^2(\rho) = (\rho_1 - \rho_1^A)^2 + (\rho_2 - \rho_2^A)^2,$$

$$l_B^2(\rho) = (\rho_1 - \rho_1^B)^2 + (\rho_2 - \rho_2^B)^2.$$

На фиг. 10 и 11 приведены зависимости наблюдаемых радиуса R_d и поверхностного натяжения σ от времени t . При этом $R_d = r_i$ для такого i , что $(\rho_{1,i} - \rho_{2,i})(\rho_{1,i+1} - \rho_{2,i+1}) < 0$; σ вычисляется с помощью дискретного аналога выражения (3.19).

Видно, что в случае диффузионной подвижности, заданной с помощью (6.1), капля фазы A продолжает существовать значительно дольше, чем для постоянной $M = M_0$.

Важно отметить, что поверхностное натяжение устанавливается в обоих случаях одинаково быстро (см. фиг. 11). При растворении капли поверхностное натяжение σ естественно обращается в нуль.

Дополнительно для наглядности на фиг. 12а и 12б приведены распределения диффузионной подвижности M вместе с распределением плотностей компонентов в различные моменты времени. Видно, что M велико

в межфазном слое и практически равно нулю вне его. При этом в «середине» межфазного слоя его значение близко к 1, которое и использовано в случае постоянной $M = M_0$ (поэтому значения σ устанавливаются одинаково быстро для рассмотренных видов M).

СПИСОК ЛИТЕРАТУРЫ

1. *Hirt C., Nichols B.* Volume of fluid (VOF) method for the dynamics of free boundaries // *J. Comput. Phys.* 1981. V. 39. № 1. P. 201–225.
2. Расчет газодинамических течений на основе метода концентраций / С.М. Бахрах, Ю.П. Глаголева, М.С. Самигулин, В.Д. Фролов, Н.Н. Яненко, Ю.В. Янилкин // *Докл. АН СССР*. 1981. Т. 257. № 3. С. 566–569.
3. *Gibou F., Fedkiw R., Osher S.* A review of level-set methods and some recent applications // *J. Comput. Phys.* 2018. V. 353. P. 82–109.
4. *Bellotti T., Graille B., Massot M.* Finite Difference formulation of any lattice Boltzmann scheme // *Numerische Mathematik*. 2022. V. 152. № 1. P. 1–40.
5. *Anderson D.M., McFadden G.B., Wheeler A.A.* Diffuse-interface methods in fluid mechanics // *Ann. Rev. Fluid Mech.* 1998. V. 30. № 1. P. 139–165.
6. *Cahn J.W., Hilliard J.E.* Free energy of a nonuniform system. I. Interfacial free energy // *J. Chem. Phys.* 1958. V. 28. № 2. P. 258–267.
7. Isogeometric analysis of the isothermal Navier–Stokes–Korteweg equations / H. Gomez, T.J. Hughes, X. Nogueira, V.M. Calo // *Comput. Meth. Appl. Mech. Engineer.* 2010. V. 199. № 25–28. P. 1828–1840.
8. *Aihara S., Takada N., Takaki T.* Highly conservative Allen–Cahn-type multi-phase-field model and evaluation of its accuracy // *Theoretic. and Comput. Fluid Dynamic*. 2023.
9. Modelling of the surface tension of binary and ternary mixtures with the gradient theory of fluid interfaces / C. Miqueu, B. Mendiboure, C. Graciaa, J. Lachaise // *Fluid Phase Equilibria*. 2004. V. 218. № 2. P. 189–203.
10. *Celny D., Vins V., Hruby J.* Modelling of planar and spherical phase interfaces for multicomponent systems using density gradient theory // *Fluid Phase Equilibria*. 2019. V. 483. P. 70–83.
11. *Rehner P., Gross J.* Predictive density gradient theory based on nonlocal density functional theory // *Phys. Rev. E*. 2018. V. 98. № 6. P. 063312.
12. Демьянов А., Динариев О., Евсеев Н. Основы метода функционала плотности в гидродинамике. Физматлит, 2009.
13. *Eyre D.J.* An unconditionally stable one-step scheme for gradient systems, 1997. preprint.
14. *Shen J., Xu J., Yang J.* A new class of efficient and robust energy stable schemes for gradient flows // *SIAM Rev.* 2019. V. 61. № 3. P. 474–506.
15. *Jamet D., Torres D., Brackbill J.* On the theory and computation of surface tension: the elimination of parasitic currents through energy conservation in the second-gradient method // *J. Comput. Phys.* 2002. V. 182. № 1. P. 262–276.
16. *Balashov V., Savenkov E.* Thermodynamically consistent spatial discretization of the one-dimensional regularized system of the Navier–Stokes–Cahn–Hilliard equations // *J. Comput. Appl. Math.* 2020. V. 372. P. 112743.
17. Балашов В.А., Савенков Е.Б. Регуляризованная изотермическая модель типа фазового поля двухкомпонентной двухфазной сжимаемой жидкости и ее одномерная пространственная дискретизация // *Дифференц. уравнения*. 2020. Т. 56. № 7. С. 887–900.
18. *Balashov V.* Dissipative spatial discretization of a phase field model of multiphase multicomponent isothermal fluid flow // *Comput. Math. Appl.* 2021. V. 90. P. 112–124.
19. *Yue P., Zhou C., Feng J.J.* Spontaneous shrinkage of drops and mass conservation in phase-field simulations // *J. Comput. Phys.* 2007. V. 223. № 1. P. 1–9.

20. Буслаев В.С. Вариационное исчисление. Изд-во Ленинградского ун-та, 1980. С. 288.
21. Михлин С.Г. Курс математической физики. М.: Наука, 1968. С. 576.
22. Гельфанд И.М., Фомин С.В. Вариационное исчисление. Гос. изд-во физ.-матем. лит-ры, 1961. С. 228.
23. Калиткин Н. Численные методы. БХВ-Петербург, 2011. С. 592.

DENSITY GRADIENT MODEL IN A SPHERICALLY SYMMETRIC FORMULATION AND ITS EXPLICIT-IMPLICIT DISSIPATIVE DISCRETIZATION FOR STUDYING INTERFACE DYNAMICS

V. A. Balashov^{a,*}, E. A. Pavlishina^{b,**}, E. B. Savenkov^{a,***}

^a*Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences, Miusskaya sq., 4, Moscow, 125047 Russia*

^b*National Research University Moscow Institute of Physics and Technology, Institutsky Pereulok, Dolgoprudny, 141701 Russia*

^{*}*e-mail: vladislav.balashov@gmail.com*

^{**}*e-mail: pavlishina.ea@phystech.edu*

^{***}*e-mail: savenkov@keldysh.ru*

Received 02 April, 2024

Revised 02 April, 2024

Accepted 02 May, 2024

Abstract. This work is dedicated to the development of an unconditionally gradient-stable (dissipative) numerical method for solving a conservative density gradient model in a spherically symmetric formulation. The algorithm is constructed using the Eyre method based on convex splitting of the system's free energy. The gradient stability of the algorithm is proven in both semi-discrete and fully discrete cases. Theoretical results are validated through several test calculations. The proposed numerical method is applied to analyze the impact of the specified diffusion mobility on the nature of interface evolution.

Keywords: density gradient theory, dissipative method, explicit-implicit approximation, convex splitting, spherically symmetric formulation.

СХЕМА РАСЧЕТА НЕСТАЦИОНАРНЫХ ТЕЧЕНИЙ ТЕПЛОПРОВОДНОГО ГАЗА В ТРЕХТЕМПЕРАТУРНОМ ПРИБЛИЖЕНИИ

© 2024 г. В.Т. Жуков^{1,*}, О.Б. Феодоритова^{1,**}

¹ 125047 Москва, Миусская пл., 4, Институт прикладной математики им. М.В. Келдыша РАН

*e-mail: vic.zhukov@mail.ru

**e-mail: feodor@kiam.ru

Поступила в редакцию 10.03.2024 г.

Переработанный вариант 10.03.2024 г.

Принята к публикации 05.05.2024 г.

Представлена методика численного моделирования нестационарных течений теплопроводного газа в трехтемпературном приближении. Методика построена с использованием основополагающих принципов С.К. Годунова. Для интегрирования по времени расчет каждого временного шага проводится с помощью расщепления определяющих уравнений на гиперболическую и параболическую подсистемы. Первая подсистема решается с помощью обобщения схемы Годунова, вторая — с помощью явно-итерационной чебышёвской схемы. Для дискретизации используются подвижные криволинейные адаптивные сетки, дискретная схема записывается в криволинейных координатах с сохранением симметрий дифференциальной задачи. Методика реализована в виде параллельного кода для многопроцессорных компьютеров. Основное назначение — обеспечение расчетных исследований по проблеме управляемого термоядерного синтеза, но возможно использование и в других приложениях вычислительной аэро-газодинамики. Библ. 28.

Ключевые слова: численное моделирование, высокотемпературная газовая динамика, многотемпературные модели, разностные схемы, схема Годунова, чебышёвские итерации.

DOI: 10.31857/S0044466924080153, EDN: XZXXLP

1. ВВЕДЕНИЕ

В данной статье изложены основные элементы методики численного моделирования нестационарных течений теплопроводного газа с детализацией алгоритмической реализации трехтемпературного приближения [1–6]. Методика построена с использованием основополагающих принципов С.К. Годунова и ведет свое происхождение от знаменитой схемы решения уравнений газодинамики, изобретенной им в 1954 г. и опубликованной спустя пять лет [7]. В этот период под руководством С.К. Годунова начата разработка двумерной газодинамической программы, первый вариант которой заработал в 1958 г. Участниками разработки являлись в основном сотрудники отдела вычислительной газовой динамики Института прикладной математики. Этим отделом руководил С.К. Годунов до переезда в 1969 г. в Новосибирск, а затем отдел возглавил А.В. Забродин, под началом которого методика развивалась вплоть до 2008 г.

Базовая многообластная методика, реализованная с использованием подвижных криволинейных адаптивных сеток, представлена в ч. 1 книги [8]; ее уместно называть методикой Годунова–Забродина–Проконова.

Атмосферу работы над методикой иллюстрирует эпизод из книги воспоминаний С.К. Годунова [9], посвященный поиску ответа на вопрос, важно ли решать задачу Римана с учетом соседства четырех ячеек, имеющих общую вершину, или можно ограничиться учетом соседства ячеек только по ребрам разностной сетки. После исследований и обсуждений с участием С.К. Годунова, И.М. Гельфанда, К.В. Брушлинского в двумерной методике реализована грубая модель с учетом соседства ячеек только по ребрам сетки. Печатных материалов, относящихся к этому вопросу, не сохранилось. Недавно в [10] доказано, что при замене точного решения многомерной задачи о распаде произвольного разрыва на решение одномерных задач с данными слева и справа от грани каждой ячейки (т.е. без учета неоднородного характера течения в окрестности вершин ячейки) вносимая погрешность имеет первый порядок малости по времени на гладких решениях. Тем самым получено обоснование применения классической схемы Годунова в многомерном случае для ячеек с плоскими гранями. В рассматриваемой здесь методике ребра ячеек сетки могут быть как прямолинейными отрезками, так и дугами окружностей. Последний случай теоретически не обоснован, но применяется в расчетных исследованиях.

К настоящему времени методика обеспечивает расчеты нестационарных двумерных уравнений газовой динамики на подвижных криволинейных сетках. Она до сих пор служит для решения прикладных задач, связанных с расчетными исследованиями аэрокосмических изделий, взрывных воздействий, процессов образования планетных систем и др.

Авторы данной статьи включились в работу в 1983 г. на этапе создания однотемпературного компьютерного кода для обеспечения исследований по проблеме инерциального тяжелоионного синтеза (ИТИС) [11–14]. Для расчета термоядерных микромишеней ИТИС в 1990 гг. был создан параллельный код для многопроцессорной вычислительной системы с распределенной памятью. Создание методики трехтемпературной газодинамики НЗТ (Hydrodynamics with Three Temperature) было завершено в начале 2000-х годов. В ее разработке участвовали Г.Б. Алалыкин, Е.А. Забродина, Г.Н. Новожилова, Л.А. Плинер, Г.П. Прокопов и др.

К настоящему времени работоспособность инерциального метода термоядерного синтеза можно считать доказанной. В декабре 2022 г. на установке NIF (США) впервые удалось осуществить реакцию самоподдерживающегося термоядерного синтеза; при облучении сферической дейтериево-тритиевой микромишени диаметром 2 мм выделилось больше энергии, чем использовалось на ее зажигание [15].

Разработка методики НЗТ проводилась на основе концепции, сформулированной С.К. Годуновым. Согласно этой концепции вычислительная модель является не просто разностной аппроксимацией дифференциальных уравнений, но представляет собой дискретный аналог физико-математической модели изучаемого явления. Это означает, что разностная схема должна отражать физические закономерности исследуемого процесса: законы сохранения, закон возрастания энтропии на ударной волне, сохранение симметрий исходной дифференциальной задачи.

Последнее свойство, в частности, инвариантность относительно группы вращений, обеспечено и в реализации трехтемпературного приближения. Несоблюдение этого свойства может приводить к эффектам, заметно искажающим физическую картину изучаемых процессов. Впервые инвариантная разностная схема для двумерных нестационарных уравнений газовой динамики в плоском и осесимметричном случаях построена именно в методике Годунова–Забродина–Прокопова, см. [8, гл. IV]. Для трехмерных нестационарных уравнений Эйлера такая схема построена А.С. Шведовым [16, 17]. Это направление, как и включение в методику алгоритмов повышенной точности на основе решения обобщенной задачи Римана (И.С. Меньшов [18]), мы рассматривать не будем.

Основное внимание в данной статье уделяется алгоритмической реализации трехтемпературного приближения, в котором сплошная среда описывается как газ с единой плотностью и общим вектором скорости частиц. Сохраняя прежние возможности методики, код НЗТ позволяет проводить моделирование физически содержательных течений сплошной среды в одно-, двух- и трех- температурных приближениях. Методика может быть адаптирована для применения в различных двухтемпературных моделях, используемых для описания лазерно-плазменных процессов, электрон-фононного взаимодействия в полупроводниках. Однако описание трехтемпературной методики, представленное в [2–6], не является достаточно полным. Данная статья восполняет указанный пробел.

2. ОСНОВНЫЕ УРАВНЕНИЯ ТРЕХТЕМПЕРАТУРНОЙ ГАЗОДИНАМИКИ

Для случая трехтемпературной модели сплошная среда, где протекают физические процессы, состоит из трех видов частиц: ионов, электронов и фотонов (это частицы, моделирующие перенос энергии излучения). При описании уравнений и параметров среды мы для определенности имеем в виду многооболочечные сферические или цилиндрические термоядерные микромишени ИТИС [2–6, 11–14].

Сплошная среда описывается как газ с единой плотностью частиц ρ и общим вектором скорости $\vec{u}$. Массой электронов пренебрегаем по сравнению с массой ионов, а фотоны считаются не имеющими массы. Температура T , давление p и удельная внутренняя энергия ϵ (приходящаяся на единицы массы среды) свои для каждого сорта частиц и являются функциями координат и времени. Переменные T , p , ϵ , а также связанные с ними коэффициенты уравнений, отмечены индексами i , e , f соответственно.

Система дифференциальных уравнений такой модели среды в переменных Эйлера имеет вид [2]:

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \rho \vec{u} = 0, \quad (1)$$

$$\frac{\partial (\rho \vec{u})}{\partial t} + \operatorname{div} (\rho \vec{u} \vec{u}) + \operatorname{grad} p = 0, \quad (2)$$

$$\frac{\partial (\rho \epsilon_i)}{\partial t} + \operatorname{div} (\rho \epsilon_i \vec{u}) + p_i \operatorname{div} \vec{u} = \operatorname{div} (K_i \operatorname{grad} T_i) + c_{ei}(T_e - T_i) + Q_i, \quad (3)$$

$$\frac{\partial (\rho \epsilon_e)}{\partial t} + \operatorname{div} (\rho \epsilon_e \vec{u}) + p_e \operatorname{div} \vec{u} = \operatorname{div} (K_e \operatorname{grad} T_e) + c_{ei}(T_i - T_e) + c_{ef}(T_f - T_e) + Q_e, \quad (4)$$

$$\frac{\partial(\rho \epsilon_f)}{\partial t} + \operatorname{div}(\rho \epsilon_f \vec{u}) + p_f \operatorname{div} \vec{u} = \operatorname{div}(K_f \operatorname{grad} T_f) + c_{ef}(T_e - T_f) + Q_f. \quad (5)$$

Уравнения рассматриваются в ограниченной области пространства при физически разумных краевых и начальных условиях. Левые части этих уравнений описывают процессы, изучаемые в газовой динамике. В правой части диффузионные слагаемые вида $\operatorname{div}(K \operatorname{grad} T)$ описывают перенос тепла ионами и электронами и процесс диффузии энергии излучения в приближении “серой материи”. Следующие слагаемые в правой части уравнений (3)–(5) моделируют релаксационные процессы обмена энергией между ионами и электронами, электронами и фотонами. Участвующие в их описании коэффициенты релаксации c_{ei}, c_{ef} , также как и коэффициенты теплопроводности K_i, K_e, K_f , могут зависеть от искомым функций, пространственных переменных и времени. Детальное описание этих коэффициентов приведено в [3]. Для представления о характере нелинейности коэффициентов, приведем только их качественную форму:

$$K_i = \kappa_{io} T_i^{5/2} / \Lambda_{ii}, \quad K_e = \kappa_{eo} T_e^{5/2} / \Lambda_{ei}, \quad K_f = \kappa_{fo} T_f^3 l, \\ c_{ei} = \kappa_{ei} \rho^2 \Lambda_{ei} / T_e^{3/2}, \quad c_{ef} = \kappa_{ef} \rho^2 (T_f^2 + T_e^2)(T_f + T_e) / T_e^{3.5}.$$

Коэффициенты $\kappa_{io}, \kappa_{eo}, \kappa_{fo}, \kappa_{ei}, \kappa_{ef}$ зависят от средних атомных весов и зарядов веществ, а также от различных констант (скорости света, постоянной Стефана–Больцмана и др.) и системы единиц, которая используется в расчетах. Кулоновские логарифмы $\Lambda_{ii}, \Lambda_{ei}$ и длина свободного пробега l зависят от искомым функций ρ, T_i, T_e, T_f .

Введем оператор L равенством

$$L \begin{pmatrix} T_i \\ T_e \\ T_f \end{pmatrix} = \begin{pmatrix} -\nabla \cdot K_i \nabla + c_{ei} I & -c_{ei} I & 0 \\ -c_{ei} I & -\nabla \cdot K_e \nabla + (c_{ei} + c_{ef}) I & -c_{ef} I \\ 0 & -c_{ef} I & -\nabla \cdot K_f \nabla + c_{ef} I \end{pmatrix} \begin{pmatrix} T_i \\ T_e \\ T_f \end{pmatrix}, \quad (6)$$

полагая здесь, что коэффициенты теплопроводности и релаксации определены по значениям температуры и плотности с нижнего слоя по времени. Этот оператор является линейным эллиптическим самосопряженным и неотрицательно определенным (при содержательных ограничениях на коэффициенты и краевые условия, которые учтены обычным образом в определении оператора и правых частях).

Члены $Q = \{Q_i, Q_e, Q_f\}$ в правых частях уравнений (3)–(5) учитывают энерговыделение в ионы, электроны и фотоны от различных источников (термоядерные реакции в топливе и нейтронно-ядерные реакции в делящемся веществе тяжелой оболочки мишени в случае гибридной мишени).

В систему (1)–(5) входят уравнение неразрывности (1), векторное уравнение импульса и подсистема трех уравнений теплопроводности. Величина p в этой системе представляет суммарное давление частиц: $p = p_i + p_e + p_f$. Уравнение (2) можно записать в инвариантной форме $\partial(\rho \vec{u}) / \partial t + \operatorname{div} \Pi = 0$ где Π — тензор плотности потока импульса; поток вектора импульса $\rho \vec{u}$ в направлении единичного вектора $\vec{n}$ есть $p \vec{n} + (\vec{u}, \vec{n}) \rho \vec{u}$. Уравнения (3), (4) описывают изменения внутренней энергии ионов и электронов соответственно, уравнение (5) — изменение энергии излучения в диффузионном приближении.

Согласно [13] уравнение (5) может быть записано для плотности энергии излучения $U = \rho \epsilon_f$; оно описывает безмассовый газ с показателем адиабаты $\gamma_f = 4/3$, уравнением состояния $p_f = (1/3) \rho \epsilon_f$, $U = \sigma T_f^4$, здесь σ — постоянная Стефана–Больцмана.

Если сложить уравнения (3)–(5) и ввести результирующие величины $\epsilon = \epsilon_i + \epsilon_e + \epsilon_f$, $Q = Q_i + Q_e + Q_f$, $-\vec{W} = K_i \operatorname{grad} T_i + K_e \operatorname{grad} T_e + K_f \operatorname{grad} T_f$ то получим уравнение изменения суммарной внутренней энергии ϵ :

$$\partial(\rho \epsilon) / \partial t + \operatorname{div}(\rho \epsilon \vec{u}) + p \operatorname{div} \vec{u} = -\operatorname{div} \vec{W} + Q. \quad (7)$$

Так же, как в обычной газодинамике, как следствие получается уравнение изменения кинетической энергии и дивергентная форма закона сохранения полной энергии [2]:

$$\frac{\partial}{\partial t} [\rho(\epsilon + 0.5|\vec{u}|^2)] + \operatorname{div}[(\rho(\epsilon + 0.5|\vec{u}|^2) + p)\vec{u}] = -\operatorname{div} \vec{W} + Q. \quad (8)$$

Величина $e = \rho(\epsilon + 0.5|\vec{u}|^2)$ представляет полную энергию среды единичного объема. Согласно принципам расчета разрывных решений на подвижных сетках, при определении закона движения поверхностей разрыва должен использоваться закон сохранения полной энергии (8) с целью обеспечения закона возрастания энтропии. Для реализации схемы Годуновского типа Г.П. Прокопов разработал модификации задачи о распаде разрыва для трехтемпературной среды, см. [2, 19].

Отметим, что разностная аппроксимация системы уравнений (1)–(5) записывается в криволинейной локальной подвижной системе координат. Такая запись обеспечивает, в частности, инвариантность разностных уравнений относительно группы вращений. Система для нахождения искомых величин на новый момент времени $t + \tau$ окажется переопределенной (семь уравнений для определения шести неизвестных). Поэтому уравнение для температуры ионов можно исключить, а энергию ионов определять из уравнения для полной энергии. Мы используем метод расщепления, основанный на интегрировании по времени гиперболической подсистемы с помощью обобщенной на трехтемпературной случай схемы Годунова и параболической подсистемы из трех уравнений теплопроводности с помощью явно-итерационной чебышёвской схемы [20]. Параболический этап трактуется как предварительный, нужный для получения тепловых потоков с целью их использования в завершающих процедурах расчета временного шага. В итоге результирующая схема на дискретном уровне обеспечивает выполнение законов сохранения массы, импульса и полной энергии.

Система уравнений (1)–(5) дополняется уравнениями состояния для каждой из трех компонент среды $ko = i, e, f$: $p_{ko} = p_{ko}(\rho, \varepsilon_{ko})$, $\varepsilon_{ko} = \varepsilon_{ko}(\rho, T_{ko})$, $p_{ko} = p_{ko}(\rho, T_{ko})$.

В качестве основного варианта используются уравнения состояния в “двучленной” форме, удобной для решения на интерфейсах ячеек элементарных задач Римана о распаде произвольного разрыва:

$$p_i = [(\gamma_i - 1)\varepsilon_i + c_{oi}^2]\rho - \rho_{oi}c_{oi}^2, \quad p_e = [(\gamma_e - 1)\varepsilon_e + c_{oe}^2]\rho - \rho_{oe}c_{oe}^2, \quad p_f = (\gamma_f - 1)\rho\varepsilon_f$$

с заданными константами $\rho_{oi}, \rho_{oe}, c_{oi}, c_{oe}$. Кроме того, для обеспечения правильной аппроксимации скорости звука в общем уравнении состояния среды полагается, что показатели адиабаты для ионов и электронов одинаковы, т.е. $\gamma_i = \gamma_e = \gamma$. Трехтемпературная модель в качестве общего уравнения состояния среды использует “двучленное” уравнение $p = [(\gamma - 1)\varepsilon + c_0^2]\rho - \gamma p^0$ с некоторыми эффективными параметрами $p^0 = \rho_0 c_0^2 / \gamma$ и c_0, ρ_0 . Адиабатическая скорость звука вычисляется по формуле $c_a = \sqrt{\gamma(p + p^0) / \rho}$, суммарное давление и внутренняя энергия есть $p = p_i + p_e + p_f$ и $\varepsilon = \varepsilon_i + \varepsilon_e + \varepsilon_f$ соответственно. Уравнения состояния электронов и ионов удобно задавать в следующем виде: $p_i = (\gamma - 1)\rho\varepsilon_i$, $\varepsilon_i = R \cdot T_i \cdot Z / (A \cdot (Z + 1))$, $p_e(\gamma - 1)\varepsilon_e$, $\varepsilon_e = R \cdot T_e \cdot Z / (A \cdot (Z + 1))$, где Z — порядковый номер элемента в периодической системе, A — его атомный вес.

Для описания геометрии задается начальное положение границ всех областей задачи. При t_0 задаются начальные данные для u, v, ρ, ε , где внутренняя энергия ε — векторная величина: $\varepsilon = (\varepsilon_i, \varepsilon_e, \varepsilon_f)$. На границах всех областей могут быть заданы граничные условия, зависящие от времени: давление и температура, давление и тепловой поток, скорость и температура, скорость и тепловой поток и др. Каждая граница может быть подвижна, и ее закон движения либо известен заранее, либо определяется во время расчета. Учет теплопереноса через внутреннюю границу предполагает выполнение равенства температур и тепловых потоков. Граничные условия на внутренних и внешних границах, включающие температуру и/или тепловой поток, задаются для каждого компонента среды.

3. ОСНОВНЫЕ ЭЛЕМЕНТЫ РАСЧЕТНОЙ МЕТОДИКИ

Общая схема расчета движения сплошной среды основана на декомпозиции (при необходимости) области решения Ω на подобласти. Эта операция включает три этапа: геометрическое разбиение исходной расчетной области на подобласти (введением дополнительных границ), движение границ на каждом шаге по времени, построение сетки в каждой подобласти. Разбиение исходной области на подобласти определяется в основном необходимостью ее представления набором “топологических четырехугольников”. Несколько подобластей могут объединяться в структуру, называемую ярусом, если при этом объединении сохраняется топология четырехугольника, а сами подобласти составляют в ярусе прямоугольную решетку. На этапе расчета теплопроводности несколько ярусов могут объединяться в счетные объекты, называемые кластерами, в которых сквозным образом осуществляется расчет теплопереноса. Заметим, что между ярусами могут быть пустоты, как например, в задаче соударения пластин при моделировании сварки взрывом. Эта задача входила в близкий круг интересов С.К. Годунова, и авторы данной статьи проводили при участии С.К. Годунова расчетные исследования процесса высокоскоростного соударения в газодинамическом приближении.

Каждая из подобластей является либо физической областью с заданными материальными параметрами, либо ее частью. Принадлежностями области являются ячейки, границы, физические параметры (параметры уравнения состояния, вид коэффициентов теплопроводности и т.д.).

В каждом ярусе строится четырехугольная криволинейная сетка, индуцированная двумя семействами координатных линий. Внутренние и внешние границы ярусов могут представлять выделенные разрывы в искомом решении. Как правило, выделяются контактные разрывы, основные ударные волны. Если ярус разделен на несколько областей, то именно границы областей служат опорными для построения координатных линий. Стороны четырехугольников аппроксимируются дугами окружностей или отрезками прямых [8]. Кривизна ребра

задается углом, равным половине центрального угла, стягивающего ребро. Объединение полученных четырехугольников, называемых ячейками сетки, составляет элементарную счетную область Ω_h . Сетка в каждом ярусе строится на каждом шаге по времени либо интерполяцией, либо минимизацией некоторого вариационного функционала [8].

Отметим, что интерес С.К. Годунова к проблеме построения класса вариационных квазиизометрических сеток, см. [21, 22], возник, в том числе, в связи с обсуждением проблемы моделирования микромишенной термоядерной синтеза.

Построенная в расчетной области разностная сетка называется геометрической или исходной, она задается декартовыми координатами своих узлов и метрическими параметрами ребер. Центр ячейки определяется как точка, в которой дуги окружностей, соединяющие середины противоположных ребер, пересекаясь, делятся пополам.

Расчетная сетка является подвижной. Сетку на фиксированный момент времени t_j будем называть t_j -слоем. Конструируемая система разностных уравнений связывает величины на двух слоях — на t_j (нижнем) и $t_{j+1} = t_j + \tau$ (верхнем). Шаг по времени τ переменный и определяется ограничениями, диктуемыми газодинамикой, см. [8]. Подсистема уравнений теплопроводности интегрируется с помощью явно-итерационной схемы ЛИ-М [20] с шагом по времени τ , величина которого ограничена газодинамическим условием устойчивости (см. [8]) гиперболической подсистемы, т.е. $\tau \approx \tau_{\text{conv}}$, обычно с некоторым запасом, например, $\tau = 0.5\tau_{\text{conv}}$. В рамках методики НЗТ интегрирование с таким шагом в задачах ИТИС является вполне допустимым в силу следующих соображений. Динамику микромишенной можно условно разделить на два этапа. На первом этапе происходит сжатие холодной мишени. На втором этапе в сжатом термоядерном топливе возникает интенсивное энерговыделение, и топливо нагревается до высокой температуры. На этом этапе становится существенной роль теплопроводных процессов и релаксации температур. Первый этап является длительным по времени (до 90% от всего времени существования микромишени). На этом этапе ограничение на шаг по времени, диктуемое параболическим условием устойчивости явной схемы, $\tau \approx \tau_{\text{diff}}$, является более слабым, чем конвективное, динамическое ограничение $\tau \approx \tau_{\text{conv}}$. Поэтому преимущество схемы ЛИ-М при расчете таких процессов проявляется очень ярко: интегрирование по времени ведется с шагом $\tau = 0.5\tau_{\text{conv}}$ и на первом этапе схема ЛИ-М автоматически превращается в явную схему. На втором этапе, где существенную роль играет теплопроводность, т.е. $\tau_{\text{diff}} \ll \tau$, схема ЛИ-М выполняет тоже автоматически (без подбора эмпирических параметров) внутренний цикл явных итераций, число q которых определяется верхней границей λ_{max} оператора теплопроводности и шагом τ по формуле $q \simeq \sqrt{\tau/\tau_{\text{diff}}} \simeq \sqrt{\tau\lambda_{\text{max}}}$.

Сетка на каждом слое строится по определенным правилам: сначала вычисляется положение всех границ счетной области на момент времени t_{j+1} , а затем строится сама сетка. На каждом слое по времени вычисляются сеточные функции: (u, v) — декартовы компоненты скорости, ρ — плотность, ε — удельная внутренняя энергия, T — температура. В рассматриваемой модели переменные ε и T являются вектор-функциями. Величины $(u, v, \rho, \varepsilon)$ считаются постоянными по ячейке и, как и температура T , относятся к центрам ячеек.

Формулы перехода на верхний слой $t = t_{j+1}$ строятся на основе дискретизации системы (1)–(5), записанной в форме законов сохранения в локальной криволинейной системе координат. Такая запись соответствует фундаментальным принципам, развитым С.К. Годуновым, о термодинамически согласованных уравнениях, образующих симметрические гиперболические системы, составленные из дивергенций — законов сохранения.

Схема строится для каждой ячейки геометрической сетки на интервале времени от t_j до $t_{j+1} = t_j + \tau$ с учетом движения этой сетки. Ячейка сетки топологически представляет собой куб в трехмерном пространстве x, y, t : его нижним и верхним основаниями являются, соответственно, ячейки на плоскостях $t = t_j$ и $t = t_{j+1}$, а четыре боковыми гранями — некоторые поверхности, образуемые в результате перемещения криволинейных ребер ячейки за время τ .

Для вычисления величин на момент $t_j + \tau$, т.е. на верхней грани каждой ячейки, кроме известных величин на нижней грани, необходимо знать газодинамические и тепловые потоки через боковые грани. Эти потоки на каждой из боковых граней предполагаются постоянными в течение всего временного шага τ . Газодинамические потоки на боковых гранях находятся по величинам, полученным из решения задачи о распаде разрыва. К таким величинам относятся нормальная и касательная к боковой грани компоненты вектора скорости $\vec{U} = (U, V)$, плотность вещества R , его давление P и внутренняя энергия E , тепловой поток. Значения первых четырех из перечисленных величин находятся из решения задачи об адиабатическом или изотермическом распаде разрыва [2, 3].

Возможны ситуации, когда не во всех, а лишь в некоторых областях, используется трехтемпературное приближение. В других же областях может использоваться двухтемпературная или однотемпературная модели. С точки зрения задачи о распаде разрыва это означает, что на границах между такими областями возникает необходимость проведения расчета, когда слева и справа от разрыва количество компонент для моделирования

среды может быть различным. Основной вклад в разработку теории и алгоритмов решения таких обобщенных задач о распаде разрыва внес Г.П. Прокопов.

Тепловые потоки находятся в процессе решения системы трех уравнений теплопроводности, см. разд. 4. В дальнейшем применяются следующие обозначения: $J, \hat{J}$ — площадь нижней и верхней граней пространственно-временной ячейки соответственно, J_{ij} — длина ребра ij ячейки (или площадь боковой грани, связанной с этим ребром) на слое $t + 0.5\tau$, U_{ij}^* — скорость этого ребра сетки в направлении внешней нормали, U_{ij}^n — нормальная компонента вектора скорости среды, α_{ij}, β_{ij} — компоненты единичной нормали n_{ij} в средней точке криволинейного отрезка между узлами i и j сетки на плоскости $t = t_j$. Величины на верхнем слое t_{j+1} отмечены крышкой, на среднем слое — верхней чертой, величины на нижнем слое t_j не имеют черты. Величины, принадлежащие центру верхней или нижней ячейки сетки, отмечены индексом “0”. Для трехтемпературного случая энергия, давление, температура представляют “наборы” из трех величин, отвечающих трем компонентам среды.

Суммарное давление на ребре ij и суммарная внутренняя энергия складываются из ионной, электронной, фотонной компонентов соответственно: $\bar{P}_{ij} = (\bar{P}_i + \bar{P}_e + \bar{P}_f)_{ij}$, $\bar{E}_{ij} = (\bar{\epsilon}_i + \bar{\epsilon}_e + \bar{\epsilon}_f)_{ij}$. Так же получается суммарный тепловой поток на ребре $\bar{\Delta}_{ij} = (\bar{\Delta}^i + \bar{\Delta}^e + \bar{\Delta}^f)_{ij}$. Дополнительно при переходе от t_j к $t_j + \tau$ может осуществляться покомпонентно вложение (или выделение) тепловой энергии в ионы, электроны, фотоны. Полное вложение энергии получается сложением вклада энергии отдельных компонент $Q = Q_i + Q_e + Q_f$.

Векторная форма записи дискретной системы уравнений имеет вид (для простоты записи ограничимся случаем “плоского” течения в декартовой системе координат):

$$\frac{\hat{J}\hat{f} - Jf}{\tau} = -A^{\text{flux}}, \quad (9)$$

где $f = (\rho, \rho u, \rho v, e, \rho \epsilon_e, \rho \epsilon_f)$, а поток A^{flux} через границу ячейки имеет следующие компоненты:

$$A^{\text{flux}} = (A^p, A^{\rho u}, A^{\rho v}, A^e, A^{\rho \epsilon_e}, A^{\rho \epsilon_f}). \quad (10)$$

Компоненты e и A^e , входящие в f и A^{flux} , это полная энергия среды единичного объема, $e = \rho(\epsilon + 0.5|\vec{u}|^2)$ и ее поток. Здесь переменная $\epsilon = \epsilon^i + \epsilon^e + \epsilon^f$. В простейшем случае (без учета источников и релаксационных членов) поток A^{flux} есть сумма газодинамических и тепловых потоков через четыре боковые грани ячейки. Обозначим этот поток через $A_{\text{surf}}^{\text{flux}}$. Его составляющие, потоки через каждую грань ячейки в направлении вектора внешней нормали n_{ij} в единицу времени есть векторные величины с компонентами

$$\begin{aligned} A_{ij}^p &= J_{ij} [\bar{R}(\bar{U}^n - \bar{U}^*)]_{ij}, \\ A_{ij}^{\rho u} &= J_{ij} [\bar{R}(\bar{U}^n - \bar{U}^*)\bar{U} + \alpha\bar{P}]_{ij}, \\ A_{ij}^{\rho v} &= J_{ij} [\bar{R}(\bar{U}^n - \bar{U}^*)\bar{V} + \beta\bar{P}]_{ij}, \\ A_{ij}^e &= J_{ij} [\bar{R}(\bar{U}^n - \bar{U}^*)(\bar{E} + (\bar{U}^2 + \bar{V}^2)/2) + \bar{P}\bar{U}^n]_{ij} + \bar{\Delta}_{ij}, \end{aligned} \quad (11)$$

где $\bar{E} = \bar{\epsilon}_i + \bar{\epsilon}_e + \bar{\epsilon}_f$, $\bar{\Delta}_{ij} = \bar{\Delta}_{ij}^i + \bar{\Delta}_{ij}^e + \bar{\Delta}_{ij}^f$. В качестве p_e^*, p_f^* используются величины не на боковых гранях, а в центре ячейки, взятые на нижнем временном слое t_j : $p_e^* = (p_e)_0 = p_e(\rho_0, \epsilon_{e0})$, $p_f^* = (p_f)_0 = p_f(\rho_0, \epsilon_{f0})$.

Формулы имеют ясный физический смысл и задают потоки массы, импульса и энергии через поверхность, определенную криволинейным отрезком ij при его перемещении во времени со скоростью U_{ij}^* . В общем случае в формулы для $A^{\rho \epsilon_i}, A^{\rho \epsilon_e}, A^{\rho \epsilon_f}$ из (10), (11), кроме потоков через боковые грани, входят также слагаемые $Q_{k_0} J$ и $R_{k_0}(T_i, T_e, T_f) J$, описывающие источники энергии и релаксационные члены, здесь $k_0 = i, e, f$. Заметим, что поток $A^{\rho \epsilon_i}$ входит в расчет потока A^e полной энергии. Поэтому окончательно имеем

$$\begin{aligned} A^{\rho \epsilon_i} &= A_{\text{surf}}^{\rho \epsilon_i} + R_i(T_i, T_e, T_f)\hat{J} + Q_i\hat{J}, \\ A^{\rho \epsilon_e} &= A_{\text{surf}}^{\rho \epsilon_e} + R_e(T_i, T_e, T_f)\hat{J} + Q_e\hat{J}, \\ A^{\rho \epsilon_f} &= A_{\text{surf}}^{\rho \epsilon_f} + R_f(T_i, T_e, T_f)\hat{J} + Q_f\hat{J}. \end{aligned} \quad (12)$$

Слагаемое $Q = Q_i + Q_e + Q_f$ — объемный интеграл от суммарной вкладываемой энергии, и оно входит также в выражение $A^e = A_{\text{surf}}^e + Q\hat{J}$, относящееся к потоку полной энергии. Детальная запись разностных уравнений в криволинейной системе координат является достаточно громоздкой и полностью приведена в [8].

4. ТЕПЛОПРОВОДНОСТЬ

В трехтемпературном случае $(\Delta^i, \Delta^e, \Delta^f)_{ij}$ представляют тепловые потоки через грани ячеек для трех компонентов среды. Ниже излагается реализация схемы (9) в части нахождения температур и тепловых потоков и используются обозначения:

$$T = (T_i, T_e, T_f), \quad \varepsilon = (\varepsilon_i, \varepsilon_e, \varepsilon_f), \quad \Delta = \Delta^i + \Delta^e + \Delta^f, \quad Q = Q_i + Q_e + Q_f.$$

Напомним, что нижние индексы (i, e, f) относятся к величинам в центрах ячеек, верхние — к потокам через боковую поверхность.

Разделим систему на две подсистемы, принимая некоторые упрощения, в виде

$$\frac{\hat{J}f^g - Jf}{\tau} = -A^g + Q\hat{J}, \quad (13)$$

$$\frac{\hat{J}\hat{f} - \hat{J}f^g}{\tau} = -\Delta(\bar{T}) + R(\bar{T})\hat{J}. \quad (14)$$

Подсистема (13) есть результат дискретизации уравнений газодинамики без учета переноса тепла теплопроводностью. Здесь $f^g = (\hat{\rho}, \hat{\rho}\hat{u}, \hat{\rho}\hat{v}, \hat{\rho}\varepsilon_i^g, \hat{\rho}\varepsilon_e^g, \hat{\rho}\varepsilon_f^g)$ — результат решения системы (13), не содержащей членов, описывающих тепловые потоки и релаксационные процессы. Подсистема (14), наоборот, содержит опущенные члены. Этот прием — разностный аналог расщепления по процессам: в сумме уравнения (13), (14) дают исходную систему (9). В (13) $A^g = (A^p, A^{pu}, A^{pv}, A^{gi}, A^{ge}, A^{gf})$ — газодинамические потоки. Энергию Q можно распределить между подсистемами, но для упрощения описания мы относим ее к гиперболической подсистеме (13). В формулу (14) входит некоторая температура $\bar{T}$. Если взять эту температуру с нижнего слоя по времени, то получим явную схему с обременительным ограничением на величину шага по времени $\tau \simeq h^2$, где h — параметр, характеризующий сетку, например, минимальный диаметр пространственной ячейки. Ниже мы покажем, как делается переход на верхний слой по времени с помощью явно-итерационной чебышёвской схемы ЛИ-М.

Уравнения подсистемы (14) можно записать для сеточных вектор-функций $\varepsilon = (\varepsilon_i, \varepsilon_e, \varepsilon_f)$ и $T = (T_i, T_e, T_f)$ в виде

$$\frac{\hat{J}\hat{\rho}(\hat{\varepsilon} - \varepsilon^g)}{\tau} = -\Delta(\bar{T}) + R(\bar{T})\hat{J}. \quad (15)$$

Так как удельная энергия ε связана с T уравнением состояния $\varepsilon = E(\rho, T)$, то (15) можно рассматривать как нелинейную неявную разностную схему относительно векторной сеточной функции $\bar{T}$. Заметим, что эти три уравнения связаны между собой, так как релаксационный член $R(\bar{T})$ связывает сразу все компоненты температуры T или их попарные комбинации. В [2] приведен алгоритм решения нелинейной системы алгебраических уравнений (15), основанный на дальнейшем расщеплении системы (15): сначала с помощью схемы ЛИ-М на шаге по времени решаются три уравнения теплопроводности в отсутствии процессов релаксации температур, т.е. при $R(\bar{T})\hat{J} \equiv 0$, затем решается нелинейная система в отсутствии тепловых потоков, т.е. при $\Delta(\bar{T}) \equiv 0$, независимо для каждого узла сетки. В процессе расчетов выяснилось, что такое расщепление приводит к вычислительным трудностям, так как процедура релаксации приводит в силу своей локальности (и неочевидной сходимости ньютоновских итераций) к существенно немоному распределению температуры.

Поэтому выбран алгоритм совместного расчета процессов теплопроводности и релаксации, т.е. решение дискретной параболической подсистемы, отвечающей дискретизации оператора (6) и обеспечивающей выполнение трех соотношений (15). На шаге по времени эта подсистема возникает при аппроксимации редуцированных дифференциальных уравнений трехтемпературной теплопроводности. В операторном виде эту систему можно записать в виде

$$\rho \frac{d\varepsilon}{dt} = L_h T, \quad (16)$$

где оператор L_h есть результат дискретизации дифференциального оператора (6). В (16) производная по времени берется вдоль траектории пространственной точки, в алгоритме интегрирования по времени оператор L_h в (16) считается линейным: коэффициенты теплопроводности и релаксации, зависящие от температуры и плотности, определяются этими переменными с нижнего слоя.

Разностная аппроксимация потока тепла Δ_{ij} через внутреннее ребро ij основана на аппроксимации поля температуры в окрестности ребра многочленом (линейным, квадратичным или квадратично-кубическим). Аппроксимация выполняется в локальной криволинейной системе координат, связанной с ребром сетки, и обеспечивает сохранение симметрий. Детальное описание различных способов построения такой аппроксимации дано в [23, 24]. Разработанные дискретные аппроксимации L_h являются консервативными и сохраняют свойство самосопряженности в сеточном аналоге скалярного произведения $L_2(\Omega_h)$. Для сохранения неотрицательной определенности дискретного оператора L_h достаточно обеспечить нестрогое диагональное преобладание.

При сильных искажениях сетки это удастся выполнить с помощью перехода на аппроксимации, приводящие к пятиточечным разностным шаблонам.

Переходим к описанию многошаговой процедуры решения уравнения (16), схеме ЛИ-М. Для согласования с (15) перепишем уравнение (16) в виде

$$\frac{d\varepsilon}{dt} = D_h T + R_h T, \quad (17)$$

где

$$D_h = \frac{1}{\hat{J}\hat{\rho}} \begin{pmatrix} \nabla \cdot K_i \nabla & 0 & 0 \\ 0 & \nabla \cdot K_e \nabla & 0 \\ 0 & 0 & \nabla \cdot K_f \nabla \end{pmatrix}, \quad R_h = \frac{1}{\hat{\rho}} \begin{pmatrix} c_{ei} I & -c_{ei} I & 0 \\ -c_{ei} I & c_{ei} I + c_{ef} I & -c_{ef} I \\ 0 & -c_{ef} I & -c_{ef} I \end{pmatrix}. \quad (18)$$

Запишем схему ЛИ-М для вектор-функций ε и T в виде:

$$\varepsilon^{(k)} = \frac{1}{1 + \tau b_k} \left\{ \varepsilon^g + \tau b_k^{(k-1)} \varepsilon + \tau D_h^{(k-1)} T + \tau R_h^{(k-1)} T \right\}, \quad T^{(k)} = F \left(\varepsilon^{(k)} \right), \quad k = 1, \dots, q, \quad (19)$$

где ε^g — внутренняя энергия после расчета гиперболического этапа (13), $\varepsilon^{(k)}$, $T^{(k)}$ — итерационные приближения, k — номер внутреннего шага (итерации), F — функция, определяемая уравнением состояния $\varepsilon = E(\rho, T)$; начальное приближение $T^{(0)}$ — температура с нижнего слоя, операторы D_h , R_h определены в (18).

Для нахождения $\varepsilon^{(k)}$ по найденным значениям $T^{(k)}$ в (19) используется линеаризация уравнения состояния для каждой температурной функции. А именно, используется аппроксимация уравнения состояния вида $E(\rho, T) \cong E(\rho_0, T_0) + \partial E / \partial T \cdot (T - T_0)$. Здесь ρ_0 , T_0 — известные значения сеточных функций — плотность и температура на нижнем слое t_j в текущей ячейке; $\partial E / \partial T$ вычисляется по значениям ρ_0 , T_0 в той же ячейке.

Пусть $AK = \partial E / \partial T$ и $BK = E(\rho_0, T_0) - T_0 \partial E / \partial T$. Тогда температура $T^{(k)}$ легко находится из соотношения $\varepsilon^{(k)} = AK \cdot T^{(k)} + BK$.

Число итераций (шагов) q в схеме (19) определяется величиной шага по времени и коэффициентами разностной аппроксимации по известному правилу. С помощью этих величин определяется набор параметров $\{b_k\}$ в итерациях (19). Среди них по построению есть нулевой; при выборе $b_q = 0$ легко видеть, что последняя итерация эквивалентна шагу явной схемы. Это означает, что схема ЛИ-М может быть представлена в виде процедуры “предиктор-корректор” — предиктор вычисляет $\bar{T}$ по формулам (19), в которых отсутствует итерация с $b_q = 0$, а корректором является явная схема (15). Для энергии ионов окончательная коррекция делается после реализации явной схемы, отвечающей закону сохранения полной энергии, в котором тепловой поток в правой части получен суммированием тепловых потоков, полученных в схеме (19) на последней итерации.

Обобщение (19) схемы ЛИ-М на рассматриваемый векторный случай оказывается работоспособным. Можно попытаться внести в (19) элемент неясности заменой $R_h^{(k-1)} T^{(k-1)}$ на $R_h^{(k)} T^{(k)}$ и находить $T^{(k)}$ из решения линейной системы размерности 3×3 относительно трех неизвестных функций: $T_i^{(k)}$, $T_e^{(k)}$, $T_f^{(k)}$. Проведенные численные расчеты для модельного простейшего уравнения $u_t = u_{xx} - a \cdot u$, $a > 0$ показали, что такой прием ухудшает точность. Поэтому расчетная схема реализована в каноническом варианте (19) со следующими дополнениями. По формуле (19) вычисляется промежуточная температура $\bar{T}$ и тепловые потоки на ребрах ячеек $\Delta(\bar{T})$ — это этап “предиктор”. Найденные значения подставляются в систему уравнений (15) и находятся все три компонента внутренней энергии на верхнем слое, отвечающие аппроксимации уравнения сохранения внутренней энергии (7). Но навязывание сохранения внутренней энергии эквивалентно закону сохранения энтропии [2]. На сильных разрывах это противоречит принципу возрастания энтропии, закон движения разрывов должен базироваться на законе сохранения полной энергии. Поэтому с помощью введенных в разд. 3 аппроксимаций и определенных в этом разделе тепловых потоков, мы реализуем расчет уравнения полной энергии (8). Из решения этого уравнения дополнительно получаем новые значения внутренней энергии ионов. По внутренней энергии частиц каждого сорта находим температуру с помощью уравнения состояния. Если уравнение состояния является нелинейным, то температура находится ньютоновскими итерациями.

Обсудим схему ЛИ-М (19) более подробно. Ее конструкция диктуется исключительно требованием аппроксимации и устойчивости, что принципиально отличает эту схему от чебышёвского ускорения сходимости итерационных процессов, применяемых для решения эллиптических уравнений или для решения неявных схем для параболических уравнений, см. [20]. Для построения нужного многочлена Чебышёва $F_p(L_h)$ степени p необходимо задать верхнюю границу $\lambda_{\max}$ спектра сеточного оператора L_h . Эта граница находится по теореме Гершгорина о кругах спектра, т.е. на основе расчета сумм модулей шаблонных коэффициентов, отвечающих

оператору L_h . Зависимость $\lambda_{\max}$ от h известна: $\lambda_{\max} \simeq h^{-2}$. Многочлен $F_p(\lambda)$ строится на отрезке $[0; \lambda_{\max}]$, исходя из условий устойчивости и аппроксимации точного оператора послойного перехода $\exp(-\tau\lambda)$. В частности, выполнено $F_p(0) = 1$, $|F_p(\lambda)| \leq 1$. С помощью многочлена F_p оператор послойного перехода S схемы LINS можно записать в виде рациональной функции оператора L_h [20]:

$$S(L_h) = (I - F_p^2(L_h)) \cdot (I + \tau L_h)^{-1}.$$

Оператор S переводит решение с нижнего слоя по времени на следующий слой и реализуется циклом, состоящим из $q = 2p - 1$ элементарных шагов. Каждый из них по трудоемкости эквивалентен одному шагу явной схемы. Множитель $(I + \tau L_h)^{-1}$ в S есть оператор послойного перехода чисто неявной схемы. На спектре оператора L_h выполнено важное неравенство $0 \leq S(\lambda) \leq (1 + \tau\lambda)^{-1}$, что следует из свойств многочлена F_p . Это неравенство означает, что вычислительная устойчивость схемы ЛИ-М не хуже, чем у неявной схемы. Напоминим, что для устойчивости достаточно выполнение более слабого неравенства $|S(\lambda)| \leq 1$.

В [20] приведено строгое изложение и обоснование конструкции чебышёвского многочлена F_p . Важно, что степень многочлена определяется без участия эмпирических параметров формулой $p = \arg \min [\operatorname{ctg}^2(0.25\pi/p) - \tau \cdot \lambda_{\max}]$, $p \in \mathbb{Z}^+$. Этой точной формулой следует пользоваться для аккуратного перехода схемы ЛИ-М в явную схему при $\tau\lambda_{\max} \approx 1$. Обозначив $[x] = \min\{n \in \mathbb{Z} : n \geq x\}$, можно записать удобную для практического использования приближенную формулу

$$p = \left\lceil 0.25\pi \sqrt{\tau\lambda_{\max} + 1} \right\rceil.$$

Параметры b_k в итерациях (19) определены как $\{b_1, \dots, b_q\} = \{a_1, \dots, a_{p-1}, a_1, \dots, a_p\}$ где $a_m = \lambda_{\max}(z_1 - \beta_m)/(1 + z_1)$, $z_1 = \cos(0.5\pi/p)$, $m = 1, \dots, p$, выражаются через корни β_m многочлена Чебышёва I рода степени p из множества $K_p = \{\cos((i - 0.5)\pi/p), i = 1, 2, \dots, p\}$. Это множество K_p упорядочено для устойчивости [24], причем так, что $a_p = 0$.

Строгий результат, важный для практики параллельных вычислений, состоит в следующем. Вместо ограничения $\tau \simeq O(h^2)$, диктуемого явной схемой, эффективный размер временного шага схемы ЛИ-М при $\tau \sim h$ составляет в терминах вычислительных затрат величину $O(h^{3/2})$. Формально величина шага τ не ограничена, так как схема ЛИ-М сохраняет устойчивость при любом значении $\tau > 0$. Однако в нестационарных высокотемпературных задачах интегрирование с шагом $\tau \sim \text{const} \cdot h$ часто является физически содержательным ограничением. Схема ЛИ-М прошла тщательную проверку на тестовых и практических задачах, подтвердив высокую точность и монотонность.

Отдельного упоминания заслуживает процедура расчета коэффициентов теплопроводности на ребрах (интерфейсах) ячеек. Рассмотрим случай, когда любой из коэффициентов теплопроводности можно записать в виде произведения двух функций $f(\rho, r)$ и $S(T)$: $\kappa(T, \rho, r) = f(\rho, r) \cdot S(T)$, причем функция f может быть разрывной по пространственной координате r ; искомая плотность $\rho(t, r)$ тоже может быть разрывной (контактный разрыв). Функция $S(T)$ является непрерывной. На решениях типа тепловых волн она является линейной в окрестности фронта.

Пусть $\kappa^m = \kappa(T^m, \rho^m, r^m)$ — значения коэффициента теплопроводности в центре ячейке с одной стороны ребра (−) и с другой стороны (+). Обозначим через h^m расстояния (длины дуг) от центров ячеек до середины ребра, где мы намерены вычислить коэффициент теплопроводности, и обозначим суммарную длину как $h = h^- + h^+$. Введем интерполяционные приближения в виде средних арифметических значений

$$K^- = \frac{h^+ \kappa(T^-, \rho^-, r^-) + h^- \kappa(T^+, \rho^-, r^-)}{h}, \quad K^+ = \frac{h^- \kappa(T^+, \rho^+, r^+) + h^+ \kappa(T^-, \rho^+, r^+)}{h},$$

а в качестве коэффициента теплопроводности K на ребре берем среднее гармоническое значение этих приближений

$$K = \frac{K^- K^+ h}{K^- h^+ + K^+ h^-}.$$

В [25] Р.П. Федоренко показал, что при расчете тепловых волн именно такая комбинированная аппроксимация приводит к корректному описанию — не препятствует правильному движению тепловой волны (как это происходит в случае среднего гармонического усреднения) и не приводит к погрешности на контактных разрывах (как в случае среднего арифметического усреднения). Для случая представления $\kappa(T, \rho, r) = f(\rho, r)S(T)$ и равномерной сетки имеем

$$K = \frac{2f^- f^+}{f^- + f^+} \frac{S(T^-) + S(T^+)}{2}.$$

Код НЗТ обеспечивает расчеты реакций горения термоядерного топлива. При описании термоядерного горения микромишеней из смеси дейтерия и трития учитываются 4 реакции с участием дейтерия, трития и гелия в предположениях, что мишень прозрачна для нейтронов и энергия выделяется локально в точке реакции. Система уравнений кинетики термоядерного горения детально представлена в [13, 14]. Модель и расчеты с учетом нейтронно-ядерного горения представлены в [27]. Отметим, что методика НЗТ может быть приспособлена для расчетов задач, алгоритм решения которых связан с движением основных разрывов и динамической перестройкой блочно-неструктурированной сетки, см. [28].

5. ЗАКЛЮЧЕНИЕ

В работе рассмотрена история создания методики численного моделирования двумерных нестационарных течений теплопроводного газа в трехтемпературном приближении и основные алгоритмы с детализацией схемы решения системы уравнений теплопроводности с учетом релаксации температур. Показано, что в трехтемпературной методике сохраняется идейная и алгоритмическая основа газодинамической методики, у истоков которой стоял С.К. Годунов. Методика реализована в виде параллельного кода для многопроцессорных компьютеров. Численная схема использует подвижные криволинейные адаптивные сетки, обеспечивает выполнение законов сохранения и симметрий дифференциальной задачи, что повышает эффективность методики при моделировании сложных высокотемпературных динамических процессов. Основное назначение — обеспечение расчетных исследований по проблеме управляемого термоядерного синтеза, но возможно использование и в других приложениях вычислительной аэро-газодинамики.

СПИСОК ЛИТЕРАТУРЫ

1. Жуков В.Т., Забродин А.В., Феодоритова О.Б. Метод решения двумерных уравнений динамики теплопроводного газа в областях сложной формы // Ж. вычисл. матем. и матем. физ. 1993. Т.33. № 8. С. 1240–1250.
2. Забродин А.В., Прокопов Г.П. Методика численного моделирования двумерных нестационарных течений теплопроводного газа в трехтемпературном приближении // Вопр. атомной науки и техн. Сер. Матем. моделирование физ. процессов. 1998. Вып. 3. С. 3–16.
3. Алалыкин Г.Б., Жуков В.Т., Забродин А.В., Забродина Е.А., Новожилова Г.Н., Плинер Л.А., Прокопов Г.П., Феодоритова О.Б. Методика численного моделирования двумерных нестационарных течений теплопроводного газа в трехтемпературном приближении в областях сложной формы с подвижными границами (НЗТ). М.: ИПМ им. М.В. Келдыша РАН, отчет НИР 8-1-04, 2004, 244 с.
4. Гиззаткулов Н.М. Разностная схема для двумерных нестационарных уравнений газовой динамики в трехтемпературном приближении // Матем. моделирование. 2004. Т. 16. № 10. С. 107–119.
5. Гиззаткулов Н.М. Численное моделирование двумерной нестационарной газовой динамики в трехтемпературном приближении с учетом нейтронного и термоядерного горения // Дисс. ... канд. физ.-матем. наук. М.: ИПМ им. М.В. Келдыша РАН. 2005.
6. Жуков В.Т. Развитие вычислительных моделей динамики мишеней термоядерного синтеза. // Дисс. докт. физ.-матем. наук. М.: ИПМ им. М.В. Келдыша РАН. 2010.
7. Годунов С.К. Разностный метод численного расчета разрывных решений газодинамики // Матем. сборник. 1959. Т. 47(89). № 3. С. 271–306.
8. Численное решение многомерных задач газовой динамики. Под ред. С. К. Годунова / С. К. Годунов, А. В. Забродин, М. Я. Иванов [и др.]. М.: Наука, 1976. С. 400.
9. Годунов С.К. Воспоминания о разностных схемах. Новосибирск: Науч. Книга. 1997. С. 40.
10. Тишкин В.Ф., Жуков В.Т., Мышецкая Е.Е. К обоснованию схемы Годунова в многомерном случае // Матем. моделирование. 2016. Т. 28. № 2. С. 86–96.
11. Разработка энергетической установки синтеза и деления на основе микромишеней прямого действия и мощного тяжелоионного драйвера / Н.Н. Алексеев, М.М. Баско, Е.А. Забродина [и др.] // Атомная энергия. 2004. Т. 97. № 3. С. 200–209.
12. Долголева Г.В., Забродин А.В. Кумуляция энергии в слоистых системах и реализация безударного сжатия. М.: Физматлит, 2004. 70 с.

13. Баско М.М. Уравнения одномерной радиационной гидродинамики с теплопереносом и гидродинамикой горения. М.: Препринт ИТЭФ. 1985. № 145. С. 58.
14. Баско М.М. Тяжелоионные мишени инерциального термоядерного синтеза // Дисс.... докт. физ.-матем. наук. М.: 1995.
15. Abu-Shawareb H. et al. Achievement of Target Gain Larger than Unity in an Inertial Fusion Experiment // Phys. Rev. Lett. 132. 065102. 2024.
16. Шведов А.С. Инвариантные разностные схемы для уравнений газовой динамики // Докл.АН СССР. 1987. Т. 292. № 1. С. 46–50.
17. Шведов А.С. Разностная схема для уравнений газовой динамики, сохраняющая групповые свойства решений // Матем. заметки. 1990. Т. 4. Вып. 4. С. 140–151.
18. Меньшов И.С. Повышение порядка аппроксимации схемы Годунова на основе решения обобщенной задачи Римана // Ж. вычисл. матем. и матем. физ. 1987. Т. 27. № 12. С. 1853–1860.
19. Прокопов Г.П. Задача о распаде разрыва в трехтемпературной газовой динамике // Препринты ИПМ им. М.В. Келдыша. 2004. № 66. С. 28.
20. Жуков В.Т. О явных методах численного интегрирования для параболических уравнений // Матем. моделирование. 2010. Т. 22. № 10. С. 127–158.
21. Годунов С.К., Жуков В.Т., Феодоритова О.Б. Численный метод квазиизометрической параметризации для двумерных криволинейных областей // Ж. вычисл. матем. и матем. физ. 2020. Т. 60. № 4. С. 578–589.
22. Godunov S.K., Feodoritova O.B., Zhukov V.T. On one class of quasi-isometric grids // Advances in Grid Generation. New York: Nova Science Publishers. 2007. P. 53–69.
23. Жуков В.Т., Феодоритова О.Б. Разностные схемы для уравнения теплопроводности на основе локальных среднеквадратичных приближений // Препринты ИПМ им. М.В. Келдыша. 1989. № 97. 19 с.
24. Жуков В.Т., Феодоритова О.Б. Разностные схемы решения нестационарного двумерного уравнения теплопроводности на криволинейных сетках и их реализация на параллельной вычислительной системе // Вопр. атомной науки и техн. Сер. Матем. моделирование физ. процессов. 1992. Вып. 3. С. 66–71.
25. Лебедев В.И., Финогенов С.А. О порядке выбора итерационных параметров в чебышёвском циклическом методе // Ж. вычисл. матем. и матем. физ. 1971. Т. 11. № 2. С. 425–438.
26. Федоренко Р.П. Введение в вычислительную физику. Долгопрудный: Изд. дом Интеллект, 2008. 503 с.
27. Жуков В.Т., Забродина Е.А., Имшенник В.С., Масленников М.В., Николаева О.В. Моделирование гибридной микромишени инерциального тяжелоионного синтеза с учетом нейтронно-ядерных реакций // Вопр. атомной науки и техн. Сер. Матем. моделирование физ. процессов. 2014. Вып. 2. С. 45–58.
28. Веселова Е.А., Дерюгин Ю.Н., Зеленский Д.К. Методика ЛОГОС-ВОЛНА расчета двумерных задач газовой динамики с учетом теплопроводности на подвижных неструктурированных сетках // Вопр. атомной науки и техн. Сер. Матем. моделирование физ. процессов. 2021. Вып. 4. С. 50–66.

SCHEME FOR CALCULATING UNSTEADY FLOWS OF HEAT-CONDUCTING GAS IN A THREE-TEMPERATURE APPROXIMATION

V. T. Zhukov*, O. B. Feodoritova**

Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences, Miusskaya sq., 4, Moscow, 125047 Russia

**e-mail: vic.zhukov@mail.ru,*

***e-mail: feodor@kiam.ru*

Received 10 March, 2024

Revised 10 March, 2024

Accepted 05 May, 2024

Abstract. A numerical modeling technique for unsteady flows of heat-conducting gas in a three-temperature approximation is presented. The methodology is based on the foundational principles of S.K. Godunov. For time integration, each time step is computed by splitting the governing equations into hyperbolic and parabolic subsystems. The first subsystem is solved using a generalized Godunov scheme, while the second uses an explicit-iterative Chebyshev scheme. Adaptive, curvilinear moving grids are used for discretization, and the discrete scheme is formulated in curvilinear coordinates, preserving the symmetries of the differential problem. The methodology is implemented as a parallel code for multiprocessor computers. Its primary purpose is to support computational studies in controlled thermonuclear fusion, though it can also be applied to other areas of computational aerogas dynamics.

Keywords: numerical modeling, high-temperature gas dynamics, multi-temperature models, finite difference schemes, Godunov scheme, Chebyshev iterations.

РАСЧЕТ НЕСТАЦИОНАРНЫХ ЗАКРУЧЕННЫХ ТЕЧЕНИЙ В СОПЛАХ И ТРУБАХ С ПОМОЩЬЮ НОВОЙ ЛОКАЛЬНО НЕЯВНОЙ СХЕМЫ ГОДУНОВСКОГО ТИПА

© 2024 г. Н. А. Зайцев^{1,*}, Ю. Г. Рыков^{1,**}

¹125047 Москва, Миусская пл., 4, ИПМ им. М.В.Келдыша, ИПМ РАН, Россия

*e-mail: nikolai_zaitsev@mail.ru

**e-mail: yu-rykov@yandex.ru

Поступила в редакцию 15.03.2024 г.

Переработанный вариант 15.03.2024 г.

Принята к публикации 05.05.2024 г.

Представлена численная схема нового класса для расчета нестационарных закрученных течений в соплах и трубах на основе уравнений сжимаемого невязкого газа. Основным достоинством таких схем является способность эффективно проводить расчеты нестационарных задач, содержащих различные масштабы. Конструкция схем такого типа основана на хорошо известном подходе Годунова, заключающемся в расчете потоков на гранях расчетных ячеек (объемов) из решения вспомогательных одномерных задач в окрестности каждой грани и аппроксимации законов сохранения. На основе анализа текущего решения в окрестности грани происходит переключение с явной схемы для расчета потока на неявную. Схема абсолютно устойчива и не порождает паразитных осцилляций. Работоспособность данной схемы продемонстрирована на задаче расчета нестационарных закрученных течений в соплах и трубах. Исследованы особенности постановки задач указанного типа, предложены варианты корректной постановки задач. Исследованы свойства решений задачи о течении закрученных потоков при наличии центрального тела, покрывающего не всю ось симметрии в расчетной области. Библ. 7. Фиг. 11.

Ключевые слова: консервативная локально неявная схема с пересчетом, схема годуновского типа, течение закрученного газа в сопле, сверхзвуковое течение, ударные волны, разномасштабные задачи.

DOI: 10.31857/S0044466924080164, EDN: XZUOOM

ВВЕДЕНИЕ

При создании новых численных методов для решения системы уравнений Эйлера большие усилия тратятся на преодоление двух основных трудностей: расчету разрывных решений и минимизацией вычислительной работы. Среди огромного количества численных методов расчета разрывных решений особое место принадлежит методу Годунова [1, 2]. Центральным пунктом годуновских схем является алгоритм вычисления потоков консервативных переменных через границы разностной сетки. В оригинальном варианте метода эти потоки вычислялись из точного решения задачи о распаде разрыва между состояниями в двух соседних ячейках. Однако решение задачи о распаде произвольного разрыва требует организации итерационного процесса и больших затрат вычислительных мощностей. По этой причине рядом авторов были предложены методы приближенного решения задачи Римана о распаде произвольного разрыва (см., например, [3], [4]). Строгая консервативность схем обеспечивается дивергентным замыканием — явной аппроксимацией интегральных законов сохранения для каждой расчетной ячейки (объема).

Вторая трудность, с которой приходится бороться, это минимизация вычислительной работы. Прежде всего за счет увеличения шага по времени. Если этот шаг диктуется не сложностью решения, а условиями устойчивости схемы, то совершается лишняя вычислительная работа. Эта проблема решается отчасти переходом к неявным схемам, но решить систему разностных уравнений для многомерной неявной схемы, это задача того же порядка сложности. Этим недостатком не обладают неявные схемы, которые вычисляют потоки на гранях расчетной ячейки на основе решения одномерных задач, которые эффективно решаются методом прогонки. К этому направлению принадлежат схемы независимых потоков [5, 6]. Но неявные схемы больше размазывают имеющиеся в решении разрывы, хуже отслеживают скорость и направление распространения информации и зачастую обладают более плохими осцилляционными свойствами, чем явные схемы. Поэтому качество

численного решения будет лучше, если использовать неявную схему только там, где это диктуется условием устойчивости и использовать явную схему везде, где это возможно.

В данной работе представлена консервативная локально-неявная схема годовновского типа, предназначенная для быстрого расчета нестационарных краевых разномасштабных газодинамических задач с особенностями в решении, такими как газодинамические разрывы, большие градиенты, области с низким давлением и т. п. Эта схема является развитием идеи локально неявной консервативной схемы независимых потоков с пересчетом, разработанной Ю. Б. Радвогиным в [6] для линейных гиперболических систем уравнений. В настоящей работе мы обобщили схему Радвогина для нелинейной начально-краевой задачи о закрученном потоке невязкого сжимаемого газа в канале с центральным телом, описываемом системой уравнений Эйлера.

Мы называем эту схему локально неявной консервативной с пересчетом. Основная особенность этого метода — способность эффективно решать задачи разного масштаба. А именно, хотя в некоторых областях (или по некоторым направлениям и т. д.) течение достаточно регулярное, можно использовать большой шаг по пространству и использовать явную схему (что предпочтительнее из-за более высокого разрешения, лучших свойств устойчивости и т. д.), в других областях течение может проявлять нерегулярное поведение со сложной тонкой структурой и требует мелкого шага по пространству. Использование явной схемы в таких случаях обременительно из-за жесткого ограничения на шаг по времени. В этом случае в нашей схеме используется неявный вариант схемы и сохраняется достаточно большой шаг по времени. Такая методология позволяет эффективно рассчитывать задачи разного масштаба в целом.

Предлагаемая нами схема для уравнений Эйлера, вслед за схемами а [6] построена на следующих принципах:

- Схема является схемой годовновского типа, т.е. решение на следующем временном слое вычисляется на основе дивергентного замыкания, аппроксимирующего систему законов сохранения. Все разнообразие схем такого типа определяется различными способами расчета потоков на гранях расчетных ячеек (объемов).
- Потоки на гранях расчетных ячеек рассчитываются независимо для каждого направления на основе вспомогательных одномерных задач.
- Везде, где это возможно используется локально явная схема.
- Если число Куранта для данного направления данной характеристики в данной точке превышает единицу, то используется локально неявная схема.
- Переключение между различными типами локальной схемы должно быть непрерывным.
- Используется принцип разностей против потока, т.е. используются характеристические свойства уравнений.

Ниже подробно описана реализация предложенной методики расчета нестационарных процессов в закрученных потоках в соплах нетривиальной геометрии. В качестве базовой модели мы используем уравнения Эйлера в цилиндрических координатах. Закрутка в сочетании с эффектами продольного течения и геометрии порождает сложные нестационарные режимы с почти периодическим поведением, возможными бифуркациями и перестройкой течения. Кроме того, такая задача, очевидно, имеет как минимум два масштаба, унаследованных от геометрии: продольное направление и радиальное направление. В радиальном направлении нам нужна гораздо более мелкая сетка, чтобы уловить особенности потока. Таким образом, задача о закрученном потоке в трубках/соплах подходит для проверки локально неявной методики.

Наш опыт показывает, что дозвуковой режим для закрученных потоков, вообще говоря, проще сверхзвукового. Поэтому мы будем проводить демонстрационные расчеты для случаев, когда по продольной координате возникает локально сверхзвуковое течение. Расчеты выявили колеблющуюся ударную волну и развивающиеся вихри в диффузоре. Эти вихри могут расти, толкать ударную волну и разрушать ее. Затем ударная волна снова возникает возле горловины сопла, приближается к диффузору и процесс повторяется.

Наша работа организована следующим образом. В разд. 1 мы строго формулируем задачу, а в разд. 2 описываем основные принципы численной схемы. Разд. 3 посвящен численному методу для уравнений Эйлера с учетом граничных условий, а в разд. 4 мы сосредоточимся на модификации метода для случая осесимметричных течений с закруткой в канале с центральным телом. Наконец, в разд. 5 мы представляем результаты расчетов и соответствующее обсуждение, включая возможность расчета уравнений Навье—Стокса с помощью этого метода.

1. ФОРМУЛИРОВКА ЗАДАЧИ

Рассмотрим нестационарное осесимметричное течение закрученного невязкого газа в сопле с закручивающим устройством, см. фиг. 1.

Соответствующую систему дифференциальных уравнений в цилиндрических координатах (с учетом того, что производные по углу обращаются в нуль) можно представить в следующих двух основных формах.

1. Дивергентная форма:

$$\frac{\partial U}{\partial t} + \frac{\partial \hat{F}}{\partial z} + \frac{\partial \hat{G}}{\partial r} = \hat{Q}, \quad (1)$$

где

$$U = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ e \end{pmatrix}; \quad \hat{F} = \begin{pmatrix} \rho u \\ P + \rho u^2 \\ \rho uv \\ \rho uw \\ Hu \end{pmatrix}; \quad \hat{G} = \begin{pmatrix} \rho v \\ \rho uv \\ P + \rho v^2 \\ \rho vw \\ Hv \end{pmatrix}; \quad \hat{Q} = \begin{pmatrix} -\rho v/r \\ -\rho uv/r \\ \rho(w^2 - v^2)/r \\ -2\rho vw/r \\ -Hv/r \end{pmatrix}.$$

Здесь ρ обозначает плотность; P — давление; u, v и w — аксиальная, радиальная и тангенциальная компоненты скорости соответственно; c — скорость звука; $e = E + \frac{\rho}{2}(u^2 + v^2 + w^2)$, $H = W + \frac{\rho}{2}(u^2 + v^2 + w^2)$; E и $W = E + P$ — внутренняя энергия и энтальпия единицы объема.

Будем считать, что газ совершенен, т.е. $E = \frac{1}{\gamma-1}P$, $c^2 = \frac{\gamma P}{\rho}$, $\gamma = \text{const.}$

2. Недивергентная форма:

$$\frac{\partial X}{\partial t} + \hat{A} \frac{\partial X}{\partial z} + \hat{B} \frac{\partial X}{\partial r} = \hat{D}, \quad (2)$$

где

$$X = \begin{pmatrix} \rho \\ u \\ v \\ w \\ P \end{pmatrix}, \quad \hat{D} = \begin{pmatrix} \rho v/r \\ 0 \\ w^2/r \\ -vw/r \\ \rho c^2 v/r \end{pmatrix}, \quad \hat{A} = \begin{pmatrix} u & \rho & 0 & 0 & 0 \\ 0 & u & 0 & 0 & 1/\rho \\ 0 & 0 & u & 0 & 0 \\ 0 & 0 & 0 & u & 0 \\ 0 & \rho c^2 & 0 & 0 & u \end{pmatrix}, \quad \hat{B} = \begin{pmatrix} v & 0 & \rho & 0 & 0 \\ 0 & v & 0 & 0 & 0 \\ 0 & 0 & v & 0 & 1/\rho \\ 0 & 0 & 0 & v & 0 \\ 0 & 0 & \rho c^2 & 0 & v \end{pmatrix}.$$

Геометрия задачи (см. фиг. 1) определяется следующими параметрами: $z = 0$ — входное сечение, $z = L$ — выходное сечение; $r = R_w(z)$ и $r = R_b(z)$ — зависимость радиуса стенки сопла и радиуса стенки центрального тела от z . Заметим, что центральное тело может простирается как вдоль всего сопла ($R_b(z) \neq 0$), так и только вдоль его передней части ($R_b(z) \equiv 0$ при $z > z^*$), а может отсутствовать ($R_b(z) \equiv 0$).

Для простоты постановки входных граничных условий полагаем, что в некоторой окрестности входного сечения обе границы представляют собой цилиндрические поверхности, т.е. $\frac{dR_w}{dz} = \frac{dR_b}{dz} \equiv 0$ на интервале $0 \leq z \leq z_1$.

На границах расчетной области задаются следующие граничные условия.

1. Во входном сечении:

$$v = 0, \quad w = k(r)u. \quad (3)$$

Первое условие обусловлено цилиндрической геометрией входной части сопла, а второе условие имитирует закрутку, создаваемую закручивающим устройством лопатного типа. А именно, полагаем, что поток после прохождения лопаток поворачивается на угол α_{in} : $\text{tg}(\alpha_{in}) = k(r)$. Зависимость α_{in} от r моделирует крутку лопаток, т.е. зависимость угла поворота лопатки от r .



Фиг. 1. Общая схема расчетной области: 1 — входная область, 2 — форкамера, 3 — горловина, 4 — диффузор, 5 — рабочая камера, 6 — конфузор, 7 — выходная область, 8 — центральное тело.

Мы будем предполагать, что вход в сопло дозвуковой. В дозвуковом входном сечении необходимы четыре граничных условия. Таким образом, кроме двух кинематических условий (3) необходимо указать еще два условия. Такими условиями может быть, например, определение любой пары следующих функций: $P(r)$, $\rho(r)$, $T(r)$, где T — температура. Однако зависимость этих функций от r неизвестна и предположение о их постоянстве относительно r вряд ли верно. Поэтому более естественным путем является задание состояния газа, покоящегося в газгольдере — резервуаре большого объема, из которого газ поступает к лопаткам закручивающего устройства. Такое состояние можно определить через два параметра покоя: либо (P_0, T_0) , либо (P_0, ρ_0) , либо (ρ_0, T_0) . Поскольку процесс истечения газа из газгольдера к лопаткам можно рассматривать как стационарный и адиабатический, будем считать, что на входе в сопло энтропия $S = P\rho^{-\gamma}$ и интеграл Бернулли $B = \frac{1}{\rho}H$ равны своим значениям в критической точке, т.е. равны известным значениям S_0 и B_0 . Следовательно, к двум линейным граничным условиям (3) необходимо добавить два нелинейных условия:

$$P\rho^{-\gamma} = S_0, \quad (4)$$

$$\frac{c^2}{\gamma - 1} + \frac{u^2[1 + k^2(r)]}{2} = B_0 \quad (5)$$

где $S_0 = P_0\rho_0^{-\gamma}$, $B_0 = \frac{c_0^2}{\gamma - 1}$ (условие (3) учитывается в граничном условии (5)).

2. В выходном сечении возможны два случая. Первый случай — выходная секция сверхзвуковая. Тогда граница не требует никаких граничных условий. Второй случай — дозвуковое выходное сечение. Тогда для корректной постановки задачи необходимо наложить одно граничное условие (предполагается положительность u_{out}). Обычно в качестве такого граничного условия задается постоянное давление:

$$P = P_{\text{out}}. \quad (6)$$

3. На стенке сопла и на поверхности центрального тела ставятся условия непротекания:

$$\frac{v_w}{u_w} = R'_w(z), \quad \frac{v_b}{u_b} = R'_b(z). \quad (7)$$

Здесь и далее штрих обозначает дифференцирование по z .

Заметим, что на оси z (где нет центрального тела) никаких дополнительных граничных условий не требуется (точнее, требуется только ограниченность решения). Из-за осевой симметрии имеем для $r = 0$

$$\frac{\partial P}{\partial r} = \frac{\partial \rho}{\partial r} = \frac{\partial u}{\partial r} = v = w = 0.$$

Эти соотношения можно использовать в численном алгоритме, чтобы избежать потери аппроксимации, связанной с особенностью координат.

Для проведения расчетов используется расчетная система координат. Геометрия задачи позволяет легко перейти к простой расчетной области, введением нормирующей координаты ξ :

$$r = \xi R_w(z) + (1 - \xi) R_b(z).$$

Теперь область расчета представляет собой прямоугольник $\Omega : \{0 \leq \xi \leq 1, 0 \leq z \leq L\}$.

В координатах (z, ξ) система (1) принимает форму

$$\frac{\partial U}{\partial t} + \frac{\partial F}{\partial z} + \frac{\partial G}{\partial \xi} = Q, \quad (8)$$

где

$$F = \hat{F}, \quad G = c_B \hat{G} - c_A \hat{F}, \quad Q = \hat{Q} - \frac{\partial c_A}{\partial \xi} \hat{F},$$

$$c_B = 1/(R_w - R_b), \quad c_A = c_B [\xi R'_w + (1 - \xi) R'_b], \quad \frac{\partial c_A}{\partial \xi} = c_B [R'_w - R'_b].$$

Для недивергентной формы (2) имеем в новых координатах:

$$\frac{\partial X}{\partial t} + A \frac{\partial X}{\partial z} + B \frac{\partial X}{\partial \xi} = D, \quad (9)$$

где

$$A = \hat{A}, \quad B = c_B \hat{B} - c_A \hat{A}, \quad D = \hat{D}. \quad (10)$$

Итак, мы рассматриваем решение начально-краевой задачи для системы (8) (или (9)) с граничными условиями (3)–(7).

2. ОСНОВНЫЕ ПРИНЦИПЫ РАЗНОСТНОЙ СХЕМЫ

Особенностью рассматриваемой задачи является вытянутая вдоль оси z геометрия сопла, что приводит к различию геометрических масштабов вдоль оси z и вдоль оси r , что в свою очередь приводит к необходимости использовать по радиусу более мелкую сетку. К тому же требованию приводит наличие закрутки потока газа: при приближении линий тока к оси z возникают большие градиенты в потоке из-за резкого роста тангенциальной скорости. Таким образом, естественно использовать для решения этой задачи неявную схему по r (по ξ) и явную по z . Последняя дешевле и лучше разрешает движущиеся ударные волны. Практика численных расчетов показывает, что неявные схемы имеют меньшую разрешающую способность, чем явные, и поэтому желательно минимизировать объем их использования.

Приемлемый компромисс может быть достигнут при использовании следующего принципа: использовать неявную схему в тех и только тех ситуациях, когда модуль соответствующего числа Куранта превышает единицу. Схемы, удовлетворяющие этому условию, можно назвать *локально-неявными*. Понятно, что такие схемы нелинейны и их структура зависит от локальных свойств решений. Что касается продольной координаты z , структура течения диктует выбор явной схемы.

Таким образом, эффективная для этой задачи схема должна обладать следующими свойствами:

- 1) явная по направлению координаты z ;
- 2) локально-неявная по направлению координаты ξ , т.е. в зависимости от значения числа Куранта становится либо явной, либо неявной;
- 3) переключение между явной и неявной схемами должно быть непрерывным.

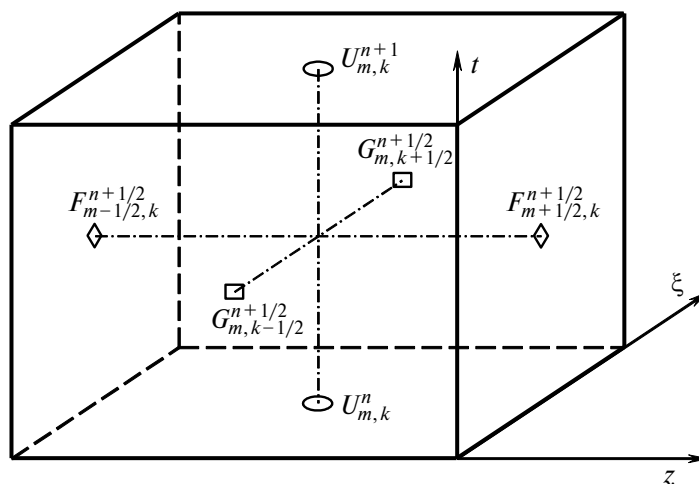
Опыт использования различных локально-неявных схем подтвердил их эффективность [7], [6].

Мы будем пользоваться технологией конструирования локально-неявных консервативных схем с пересчетом, предложенной в [6] для системы линейных двумерных гиперболических уравнений. Там же доказана абсолютная устойчивость этой схемы. Если использование неявного варианта схем ограничить только одним направлением (в нашем случае поперек оси z , то схема будет устойчива при числе Куранта по явному направлению меньше 1. Поскольку устойчивость разностной схемы в признаке устойчивости Неймана (спектральный признак) применяется к линеаризованным уравнениям с замороженными коэффициентами, доказательство устойчивости приведенное в [6] применимо и к предлагаемой здесь схеме для уравнений Эйлера.

3. РАЗНОСТНАЯ СХЕМА ДЛЯ УРАВНЕНИЙ ЭЙЛЕРА

Следуя методологии понятия конечных объемов, мы разбиваем расчетную область в пространстве (t, z, ξ) на объемы координатными поверхностями $t = t^n \equiv \tau n$, $n = 0, 1, 2, \dots$; $z = z_{m+1/2} \equiv h_z m$, $m = 0, 1, 2, \dots, M$, $h_z = L/M$; $\xi = \xi_{k+1/2} \equiv h_\xi k$, $k = 0, 1, 2, \dots, K$, $h_\xi = 1/K$. Значения сеточных функций $U_{m,k}^n$ на верхнем и нижнем слоях по времени будем относить к центрам этих пространственных ячеек $z_m = h_z(m - 1/2)$, $m = 0, 1, 2, \dots, M$, $\xi_k = h_\xi(k - 1/2)$, $k = 0, 1, 2, \dots, K$ т.е. $U_{m,k}^n = U(t^n, z_m, \xi_k)$.

Предлагаемый алгоритм использует обе формы основной системы: недивергентную (9) и дивергентную (8). Основным элементом алгоритма является расчет потоков F и G на боковых сторонах расчетной ячейки, т.е. $F_{m+1/2,k}$, $G_{m,k+1/2}$ (см. фиг. 2).



Фиг. 2. Расчетная ячейка.

Начнем с z -направления, для которого потоки можно вычислить по явной схеме. При построении этой схемы используются левые собственные векторы ψ^A матрицы Якоби $\frac{\partial F}{\partial U}$ и ее собственные значения λ^A . Пусть φ^A — левый собственный вектор матрицы A : $\varphi^A A = \lambda^A \varphi^A$. Поскольку $\psi^A \frac{\partial F}{\partial U} = \lambda^A \psi^A$ и $\frac{\partial F}{\partial U} = \frac{\partial U}{\partial X} A \left(\frac{\partial U}{\partial X} \right)^{-1}$, то $\psi^A = \varphi^A \left(\frac{\partial U}{\partial X} \right)^{-1}$.

Матрица A имеет следующие собственные значения и векторы:

$$\lambda_1^A = \lambda_2^A = \lambda_3^A = u, \quad \lambda_4^A = u + c, \quad \lambda_5^A = u - c,$$

$$\phi_1^A = (c^2, 0, 0, 0, -1), \quad \phi_2^A = (0, 0, 1, 0, 0), \quad \phi_3^A = (0, 0, 0, 1, 0), \quad \phi_4^A = (0, \rho c, 0, 0, 1), \quad \phi_5^A = (0, \rho c, 0, 0, -1).$$

Матрица $\frac{\partial U}{\partial X}$ имеет вид:

$$\frac{\partial U}{\partial X} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ u & \rho & 0 & 0 & 0 \\ v & 0 & \rho & 0 & 0 \\ w & 0 & 0 & \rho & 0 \\ \frac{V^2}{2} & \rho u & \rho v & \rho w & \frac{1}{\gamma - 1} \end{pmatrix},$$

где $V^2 = u^2 + v^2 + w^2$. Тогда в качестве ψ_i^A имеем следующие векторы:

$$\begin{aligned} \psi_1^A &= \left(\frac{c^2}{\gamma - 1} - \frac{V^2}{2}, u, v, w, -1 \right); \\ \psi_2^A &= (-v, 0, 1, 0, 0); \\ \psi_3^A &= (-w, 0, 0, 1, 0); \\ \psi_4^A &= \left(-\frac{uc}{\gamma - 1} + \frac{V^2}{2}, \frac{c}{\gamma - 1} - u, -v, -w, 1 \right); \\ \psi_5^A &= \left(-\frac{uc}{\gamma - 1} - \frac{V^2}{2}, \frac{c}{\gamma - 1} + u, v, w, -1 \right). \end{aligned}$$

Определим неизвестные потоки F из следующей одномерной системы:

$$\frac{\partial F}{\partial t} + \frac{\partial F}{\partial U} \frac{\partial F}{\partial z} = 0. \quad (11)$$

Заморозим матричные коэффициенты и после преобразования (11) к диагональному виду получим для инвариантов $I_i^A = \psi_i^A F$ следующую систему:

$$\frac{\partial I_i^A}{\partial t} + \lambda_i^A \frac{\partial I_i^A}{\partial z} = 0. \quad (12)$$

Затем на основании вышеизложенного выводится алгоритм расчета $F_{m+1/2,k}$.

- 1) Расчет $X_{m+1/2,k}$ на боковой стороне с помощью простого усреднения по точкам (m, k) и $(m + 1, k)$.
- 2) Расчет $(\psi_i^A)_{m+1/2,k}$ и $(\lambda_i^A)_{m+1/2,k}$ по $X_{m+1/2,k}$ для всех $i = 1, \dots, 5$.
- 3) Определение $F_{m+1/2,k}$ из системы линейных уравнений, $i = 1, \dots, 5$:

$$(\psi_i^A)_{m+1/2,k} F_{m+1/2,k} = \begin{cases} (\psi_i^A F)_{m,k}, & (\lambda_i^A)_{m+1/2,k} > 0 \\ (\psi_i^A F)_{m+1,k}, & (\lambda_i^A)_{m+1/2,k} < 0 \end{cases}. \quad (13)$$

Чтобы изменение знака λ_i^A не приводило к нефизическим решениям, в алгоритм включен анализ знаков λ_i^A в точках, соседних с данной гранью (по m). В случае $(\lambda_i^A)_{m,k} (\lambda_i^A)_{m+1,k} < 0$ соответствующее уравнение в системе (13) необходимо изменить следующим образом:

$$(\psi_i^A)_{m+1/2,k} F_{m+1/2,k} = (\psi_i^A)_{m+1/2,k} \frac{F_{m,k} + F_{m+1,k}}{2}.$$

Для входного сечения четыре уравнения системы (13), соответствующие положительным λ_i^A , заменяются граничными условиями. Для дозвукового выхода последнее уравнение из (13) заменяется граничным условием

$P = P_{\text{out}}$. В случае сверхзвукового выхода правая часть в (13) полностью определена (ввиду положительности всех λ_i^A на такой границе).

Собственные векторы ψ_i^A , необходимые для расчета граничных потоков, вычисляются не путем усреднения по соседним точкам (это невозможно), а путем простой экстраполяции или даже переноса значений ψ_i^A из соседней внутренней точки.

Теперь давайте обратимся к направлению ξ . Как и в случае направления z потоки G на гранях расчетной ячейки определим из одномерной системы. Но поскольку в рассматриваемой задаче поперечные этому направлению скачки отсутствуют, расчет соответствующих потоков G можно провести, используя одномерную часть недивергентной формы основной системы (9):

$$\frac{\partial X}{\partial t} + B \frac{\partial X}{\partial \xi} = 0. \quad (14)$$

Это заметно сокращает время расчета, так как упрощает формулы. Согласно (10) матрица имеет вид:

$$B = \begin{pmatrix} c_B u - c_A v & -c_A \rho & c_B v & 0 & 0 \\ 0 & c_B u - c_A v & 0 & 0 & -\frac{c_A}{\rho} \\ 0 & 0 & c_B u - c_A v & 0 & \frac{c_B}{\rho} \\ 0 & 0 & 0 & c_B u - c_A v & 0 \\ 0 & -c_A \rho c^2 & c_B \rho c^2 & 0 & c_B u - c_A v \end{pmatrix}.$$

Собственные значения матрицы B имеют вид:

$$\lambda_1^B = \lambda_2^B = \lambda_3^B = -c_A u + c_B v, \quad \lambda_4^B = \lambda_1^B + d, \quad \lambda_5^B = \lambda_1^B - d,$$

где $d = \sqrt{c_A^2 + c_B^2}$.

Им соответствуют следующие собственные векторы φ_i^B :

$$\begin{aligned} \varphi_1^B &= (c^2, 0, 0, 0, -1), \\ \varphi_2^B &= (0, c_B, -c_A, 0, 0), \\ \varphi_3^B &= (0, 0, 0, 1, 0), \\ \varphi_4^B &= (0, \alpha c, \beta c, 0, 1), \\ \varphi_5^B &= (0, \alpha c, \beta c, 0, -1), \end{aligned} \quad (15)$$

где $\alpha = -c_A/d$, $\beta = c_B/d$.

Поскольку теперь искомым вектор не U , а X (см. (9)), аналогом (12) является

$$\frac{\partial I_i^B}{\partial t} + \lambda_i^B \frac{\partial I_i^B}{\partial \xi} = 0, \quad i = 1, \dots, 5, \quad (16)$$

где $I_i^B = \phi_i^B X$, т. е. инварианты не уравнения для потоков (11), а уравнения для неконсервативных величин (14).

Локально-неявная по ξ схема строится в виде разностной системы вида:

$$A_{k+1/2} I_{k-1/2}^B + B_{k+1/2} I_{k+1/2}^B + C_{k+1/2} I_{k+3/2}^B = D_{k+1/2}, \quad k = 0, \dots, K, \quad (17)$$

где I^B — вектор-столбец с компонентами I_i^B , A, B, C являются (5×5) матрицами.

Каждое уравнение (i -я строка) этой системы является разностной аппроксимацией соответствующего уравнения системы (16) (индекс m опущен для упрощения написания):

$$(a_i)_{k+1/2} (I_i^B)_{k-1/2} + (b_i)_{k+1/2} (I_i^B)_{k+1/2} + (c_i)_{k+1/2} (I_i^B)_{k+3/2} = (d_i)_{k+1/2}, \quad k = 0, \dots, K,$$

То есть аппроксимация уравнения (14) приводит к диагональным матричным коэффициентам A, B, C . Коэффициенты a_i, b_i, c_i стоят на диагонали в i -й строке матриц A, B, C , а коэффициент d_i стоит в i -й строке вектора правых частей D . Эти коэффициенты вычисляются по следующему правилу в соответствии с направлением

распространения (знаком собственного числа) соответствующего инварианта и величиной соответствующего числа Куранта:

$$\begin{aligned} (q_i^B)_{k+1/2} \geq 1 : \quad & a_i = 1 - (q_i^B)_k, \quad b_i = (q_i^B)_k, \quad c_i = 0, \quad d_i = (I_i^B)_k; \\ 0 < (q_i^B)_{k+1/2} \leq 1 : \quad & a_i = 0, \quad b_i = 1, \quad c_i = 0, \quad d_i = (I_i^B)_k; \\ -1 \leq (q_i^B)_{k+1/2} < 0 : \quad & a_i = 0, \quad b_i = 1, \quad c_i = 0, \quad d_i = (I_i^B)_{k+1}; \\ (q_i^B)_{k+1/2} \leq -1 : \quad & a_i = 0, \quad b_i = -(q_i^B)_{k+1}, \quad c_i = 1 + (q_i^B)_{k+1}, \quad d_i = (I_i^B)_{k+1}. \end{aligned} \quad (18)$$

Числа Куранта $q_i^B = \lambda_i^B \kappa_\xi$ и инварианты I_i^B вычисляются на нижнем слое $t = t^n$. Все значения газодинамических величин, относимые к боковым сторонам расчетной ячейки, т.е. имеющие дробный индекс по отношению к k , рассчитываются по усредненным значениям базовых величин $X_{k+1/2}^n = \frac{1}{2}(X_k^n + X_{k+1}^n)$.

Система (17) замыкается граничными условиями непротекания при $\xi = 0$ и $\xi = 1$:

$$-c_A u + c_B v = 0,$$

которое можно выразить через инварианты I_4^B и I_5^B как

$$I_4^B + I_5^B = 0. \quad (19)$$

Напомним, что $I_{4,5}^B = (-c_A u + c_B v) \rho c / d \pm P$. Следовательно, на обеих этих границах $\lambda_{1,2,3}^B = q_{1,2,3}^B = 0$, $I_4^B = P$, $I_5^B = -P$, и

$$G = \begin{pmatrix} 0 \\ -c_A P \\ c_B P \\ 0 \\ 0 \end{pmatrix}.$$

Видно, что для вычисления G на границе необходимо иметь ровно одно скалярное условие. На границе $\xi = 0$ этим условием является значение приходящего из расчетной области инварианта с I_5^B , на границе $\xi = 1$ — инварианта с I_4^B . Поэтому

$$G|_{\xi=0} = \begin{pmatrix} 0 \\ c_A I_5^B \\ -c_B I_5^B \\ 0 \\ 0 \end{pmatrix}, \quad G|_{\xi=1} = \begin{pmatrix} 0 \\ -c_A I_4^B \\ c_B I_4^B \\ 0 \\ 0 \end{pmatrix}.$$

Таким образом, для вычисления G на границе нет необходимости вычислять там же инварианты I_1^B, I_2^B, I_3^B , и можно назначить им любые значения, например, 0. А для того, чтобы эти величины не возникали в разностных уравнениях на соседних с границей гранях мы наложим ограничения на шаг по времени τ :

$$|q_{1+1/2}^1| \leq 1, \quad |q_{K-1/2}^1| \leq 1.$$

Это условие не является стеснительным, поскольку на соседних гранях (на границах области) $|q_{1-1/2}^1| = 0$, $|q_{K+1/2}^1| = 0$ и является естественным, так как сигнализирует о недостаточном количестве точек. Кроме того, мы потребуем, чтобы на противоположных гранях одной ячейки числа q^k с одинаковыми номерами инвариантов k не лежали по разные стороны от отрезка $[-1, 1]$, что также сигнализирует о недостаточном количестве точек.

В результате уравнения для первых трех инвариантов отщепляются от общей системы уравнений (17) в виде трех скалярных трехдиагональных разностных уравнений, решение которых можно найти с помощью экономичного метода скалярной прогонки.

Что касается оставшихся инвариантов I_4^B, I_5^B , то эти инварианты можно найти из соответствующей подсистемы (17) для 2-векторов $I^{(2)} = (I_4^B, I_5^B)$. В этой подсистеме на каждой границе разностное уравнение, соответствующее входящей в расчетную область характеристике (т.е. для $\xi = 0$ — четвертое уравнение, а для $\xi = 1$ — пятое уравнение), должно быть заменено граничным условием (13). В результате на нижней границе матричный коэффициент A подсистемы для 2-вектора $I^{(2)}$ гарантированно обращается в нуль, а на верхней границе то же происходит с матричным коэффициентом C . Таким образом, полученная подсистема для I_4^B и I_5^B имеет стандартную трехдиагональную структуру и для ее решения можно использовать метод векторной прогонки для

двумерных векторов, которая гораздо эффективнее, чем прогонка для пятимерных векторов. Дополнительное ускорение расчета можно получить, если учесть, что на всех гранях расчетных ячеек, кроме крайних, матрицы A, B, C являются диагональными.

Решение обеих подсистем дает на всех гранях полный набор инвариантов $I_1^B, \dots, I_5^B$. После этого компоненты вектора X на тех же гранях находятся по формулам (см. (15) и (16)):

$$\begin{aligned} P &= (I_4^B - I_5^B)/2, \\ \rho &= (I_1^B + P)/(c^2)_f, \\ u &= \frac{(\beta \rho c)_f I_2^B - (c_A)_f (I_4^B + I_5^B)/2}{(\beta \rho c)_f (c_B)_f - (\alpha \rho c)_f (c_A)_f}, \\ v &= \frac{(\alpha \rho c)_f I_2^B - (c_B)_f (I_4^B + I_5^B)/2}{(\alpha \rho c)_f (c_A)_f - (\beta \rho c)_f (c_B)_f}, \\ w &= I_3^B, \end{aligned}$$

где индекс f означает, что это величины, вычисленные на соответствующей грани путем усреднения по значениям в соседних целочисленных точках нижнего слоя (аналогично расчету $(q_i^B)_{k+1/2}$ в (18)).

Далее по X вычисляется U , а по нему G . На этом процедура расчета потоков закончена.

Описанный выше метод расчета потоков F (по явной схеме) и G (по локально неявной схеме) составляет основу алгоритма перехода на следующий слой. Как и в скалярном случае (см. разд. 2), эта процедура выполняется в два этапа, которые можно назвать предиктором и корректором.

На этапе предиктора вычисляются потоки $\tilde{F}$ и $\tilde{G}$ по описанному выше алгоритму и значениям консервативных величин на нижнем временном слое:

$$\tilde{F} = F(U^n), \quad \tilde{G} = G(U^n),$$

и с помощью дивергентного замыкания вычисляются предварительные значения консервативных величин на верхнем слое:

$$\tilde{U}_{m,k}^{n+1} = U_{m,k}^n - \kappa_z (\tilde{F}_{m+1/2,k} - \tilde{F}_{m-1/2,k}) - \kappa_\xi (\tilde{G}_{m,k+1/2} - \tilde{G}_{m,k-1/2}) + \tau Q_{m,k}^n, \quad m = 1, \dots, M; \quad k = 1, \dots, K;$$

и, в свою очередь, $\tilde{X}_{m,k}$:

$$\rho = U_1, \quad u = U_2/U_1, \quad v = U_3/U_1, \quad w = U_4/U_1, \quad P = (\gamma - 1) \left[U_5 - \rho(u^2 + v^2 + w^2)/2 \right].$$

Этап корректора отличается от первого только заменой вектора U^n , участвующего в вычислении F и G , на взвешенное значение

$$U^v = v\tilde{U} + (1 - v)U^n.$$

По этим значениям вычисляются уточненные значения потоков

$$F = F(U^v), \quad G = G(U^v),$$

и уточненные значения консервативных величин на верхнем слое:

$$U_{m,k}^{n+1} = U_{m,k}^n - \kappa_z (F_{m+1/2,k} - F_{m-1/2,k}) - \kappa_\xi (G_{m,k+1/2} - G_{m,k-1/2}) + \tau Q_{m,k}^n, \quad m = 1, \dots, M; \quad k = 1, \dots, K;$$

Нужно брать величину $v = 0.25$, так как только это значение обеспечивает устойчивость при любых числах Куранта q_i^B .

Подчеркнем, что в этом методе этап коррекции делается не для повышения точности схемы, а для повышения области ее устойчивости: в результате этапа корректора схема становится устойчивой при всех числах Куранта по оси z , не превосходящих по модулю 1, и при всех числах Куранта по оси ξ , т. е. поперек канала.

4. ОСОБЕННОСТИ В ЗАКРУЧЕННЫХ ОСЕСИММЕТРИЧНЫХ ПОТОКАХ

Поскольку рассматриваемое течение не зависит от угловой координаты, четвертое уравнение системы (1) дает

$$\frac{d\Omega}{dt} = 0, \quad \frac{d}{dt} = \frac{\partial}{\partial t} + u \frac{\partial}{\partial z} + v \frac{\partial}{\partial r}, \quad (20)$$

где $\Omega = wr$. Это уравнение выражает сохранение циркуляции вдоль траектории жидкой частицы потока. Под траекторией мы понимаем ее след на плоскости (z, r) . (Настоящая траектория — спиральная.)

Рассмотрим некоторые следствия закона сохранения циркуляции и начнем со стационарного течения в сопле с центральным телом, см. фиг. 1. Предположим, что $w \neq 0$ во входном сечении в точке C , расположенной на поверхности (образующей) центрального тела. При движении вперед вдоль поверхности тела частица опускается к оси z , т.е. ее r -координата уменьшается. Но $\Omega = \text{const}$ вдоль линии тока. Поэтому w вдоль этой линии увеличивается. Таким образом, эта линия тока не может достичь оси z , так как при $r \rightarrow 0$ $w \rightarrow \infty$, и должна существовать точка, в которой линия тока отрывается от тела и идет вниз по каналу, оставаясь на конечном расстоянии от оси симметрии. В этом случае под такой линией должна возникнуть зона застоя.

В нестационарном течении ситуация та же, поскольку при приближении жидкой частицы к оси $r = 0$ азимутальная скорость будет неограниченно расти, что противоречит ограниченности скорости потока. Вообще, закон сохранения циркуляции запрещает любые линии тока с $\Omega \neq 0$, простирающиеся до оси z . Этот факт фактически делает некорректной задачу о невязком течении закрученного газа в сопле с центральным телом конечной длины. Местоположение точки отрыва и структура течения, строго говоря, существенно зависят от исходных данных.

Эту трудность можно преодолеть двумя способами. Во-первых, можно задать такие условия на входе, чтобы закрутка на центральном теле равнялась нулю. На практике такое требование может быть выполнено путем изменения угла поворота закручивающих лопаток по радиусу так, чтобы на центральном теле закрутка не возникала. Во-вторых, можно ввести тонкую трубку, например цилиндрическую, радиусом ε , не позволяющую линиям тока доходить до оси $r = 0$. Разумеется, толщина этой трубки должна быть достаточно малой, чтобы не оказывать существенного влияния на основную часть потока.

Для оценки толщины области отрыва в окрестности оси симметрии или, что практически то же самое, радиус искусственной трубки, продлевающей центральное тело на весь диапазон изменения z , рассмотрим следующую модельную задачу. Предположим, что течение изоэнтропическое, изоэнергетическое, стационарное и не зависит от продольной координаты z . Тогда $v = 0$ и все остальные функции, зависящие только от r , удовлетворяют следующей системе уравнений:

$$P = S_0 \rho^\gamma, \quad (21)$$

$$\frac{\gamma}{\gamma - 1} \frac{P}{\rho} + \frac{u^2 + w^2}{2} = H_0, \quad (22)$$

$$r \frac{dP}{dr} = \rho w^2, \quad (23)$$

где $S_0 = \frac{P_0}{\rho_0^\gamma}$, $H_0 = \frac{\gamma}{\gamma - 1} \frac{P_0}{\rho_0}$.

Вводя скорость звука $c^2 = \gamma P / \rho$, перепишем (22) в следующей форме:

$$c^2 + \frac{\gamma - 1}{2} (u^2 + v^2) = (\gamma - 1) H_0. \quad (24)$$

Из (21) и (23) следует, что

$$r \frac{dc^2}{dr} = (\gamma - 1) w^2. \quad (25)$$

Исключение w^2 из (24) и (25) дает

$$\frac{r}{2} \frac{dc^2}{dr} + c^2 = f(r), \quad (26)$$

где

$$f(r) = (\gamma - 1) \left[H_0 - \frac{u^2(r)}{2} \right],$$

решением (26) является

$$c^2(r) = \frac{2}{r^2} \int_{\varepsilon}^r s f(s) ds,$$

где ε — произвольная постоянная.

Предположим, что $u(r) = \text{const} = u_0$. Обозначая через $f_0 > 0$ соответствующее значение f , получим

$$c^2(r) = f_0 \left(1 - \frac{\varepsilon^2}{r^2} \right).$$

Тогда

$$\begin{aligned}w(r) &= \sqrt{\frac{2f_0}{\gamma-1}} \frac{\varepsilon}{r}, \\ \rho(r) &= \left[\frac{f_0}{\gamma S_0} \left(1 - \frac{\varepsilon^2}{r^2} \right) \right]^{\frac{1}{\gamma-1}}, \\ P(r) &= S_0 \left[\frac{f_0}{\gamma S_0} \left(1 - \frac{\varepsilon^2}{r^2} \right) \right]^{\frac{1}{\gamma-1}}.\end{aligned}$$

Для продольного числа Маха M_z получим

$$M_z = u_0 \left[f_0 \left(1 - \frac{\varepsilon^2}{r^2} \right) \right]^{\frac{1}{2}}.$$

И окончательно

$$\Omega = wr = \varepsilon \sqrt{\frac{2}{\gamma-1}} f_0 = \text{const.} \quad (27)$$

Следовательно, рассматриваемое течение существует только за пределами ε -окрестности оси z , т.е. только для $r > \varepsilon$. При $r \rightarrow \varepsilon + 0$, значения P и ρ стремятся к нулю, а значения w и M_z — к бесконечности.

Из (27) следует, что

$$\varepsilon = \Omega(2H_0 - u_0^2)^{-1/2}, \quad (28)$$

что и является искомой оценкой ε . Переменность $u(r)$ не приводит к принципиальному изменению картины течения, а лишь усложняет формулы.

5. РЕЗУЛЬТАТЫ РАСЧЕТОВ И ОБСУЖДЕНИЕ

Расчеты проводились для различных форм канала и различных параметров потока, но сценарий расчета для всех случаев был общим: вначале рассчитывалось стационарное течение без закрутки ($k(r) = 0$ в формуле (3)), а затем мгновенно или за короткий промежуток времени закручивающие лопасти поворачивались на заданный угол. При этом геометрия канала и параметры потока были таковы, что в минимальном сечении в стационарном течении достигалась звуковая скорость потока и ниже его течение было сверхзвуковым (по крайней мере на некотором участке). Это позволяет разделить до некоторой степени исследование течения в форкамере (область от конца центрального тела до критического сечения, или горловины) и в области диффузора (расширение канала за горловиной).

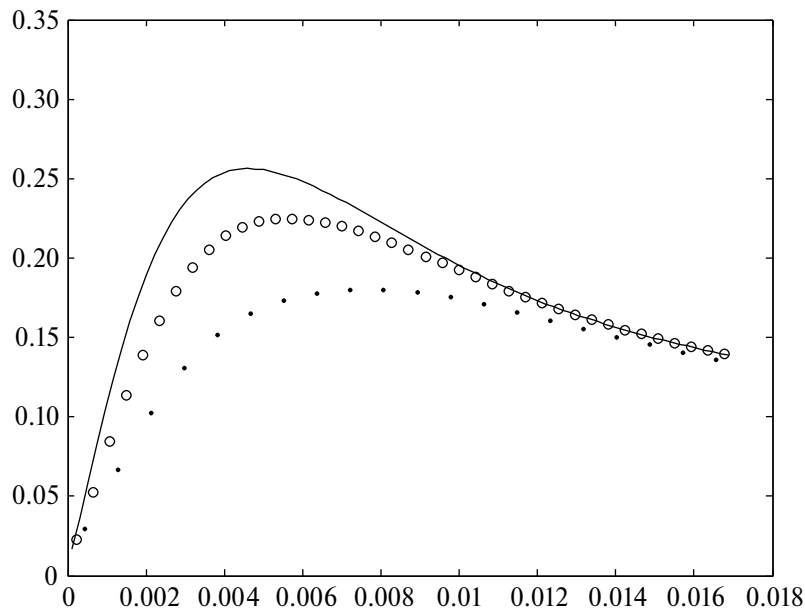
Коснемся еще раз вопроса о роли центрального тела. Ряд численных экспериментов, выполненных с помощью представленного алгоритма, позволяет утверждать, что стационарное течение в соплах без центрального тела даже при наличии закрутки существует и без специальных ухищрений может быть получено установлением по времени. Требуемое время установления соответствует примерно 3–5-кратному прохождению возмущения через форкамеру вниз и вверх по потоку. То же самое справедливо и для незакрученных течений в соплах с центральным телом.

Однако любая закрученность течения на входе в сопло на поверхности центрального тела принципиально меняет картину течения. В частности, например, стационарное течение становится невозможным. Действительно, величина $\Omega = rw$ переносится вдоль линии тока (см. уравнение (20)). Таким образом, из ограниченности скорости следует, что линия тока при движении вниз по течению обязана соскочить с центрального тела. Более того, по той же причине эта линия раздела не может подойти к оси z .

Если предположить, что вблизи оси z образуется застойная зона со скоростью близкой нулю и с постоянным давлением, то это противоречит структуре течения в окрестности критического сечения.

Из сказанного следует, что через определенное время после включения закрутки на входе в сопло должен образоваться отрыв в некоторой точке центрального тела и зона слабого возвратного течения. Дальнейшая эволюция течения в окрестности оси симметрии полностью определяется свойствами численной схемы и, в частности, схемной вязкостью.

Уже первые расчеты течения в форкамере подтвердили теоретические выводы о том, что структура потока существенно зависит от характера закрутки. В случае задания входной закрученности в (3) в виде $k(r) = k_w(r - R_b)/(R_w - R_b)$, т.е. линейно меняющейся от нуля на центральном теле до максимального значения на стенке канала), особая точка на поверхности центрального тела не возникает, и течение быстро стабилизируется.



Фиг. 3. Сравнение профиля $w(r)$ в сечении $z = 140$ мм для трех расчетов: 175×20 (пунктирная линия), 350×40 (кружки) и 350×80 (сплошная линия).

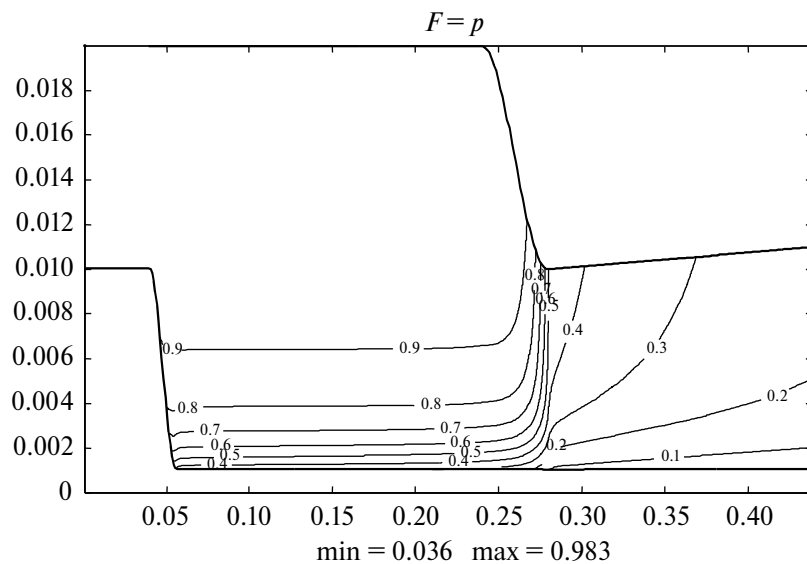
Принципиально другая ситуация возникает, если на входе $k(r) = k_w$, т. е. угол поворота закручивающих лопаток постоянен по радиусу. При таком способе задания закрученности течение визуально довольно быстро стабилизируется, после чего начинается медленный квазистационарный процесс, в котором чрезвычайно медленно вымывается тонкая осевая струя незакрученного течения.

В форкамере азимутальная скорость непрерывно растет с увеличением r от нуля на оси симметрии до своего максимума в некоторой точке r_* , а потом убывает.

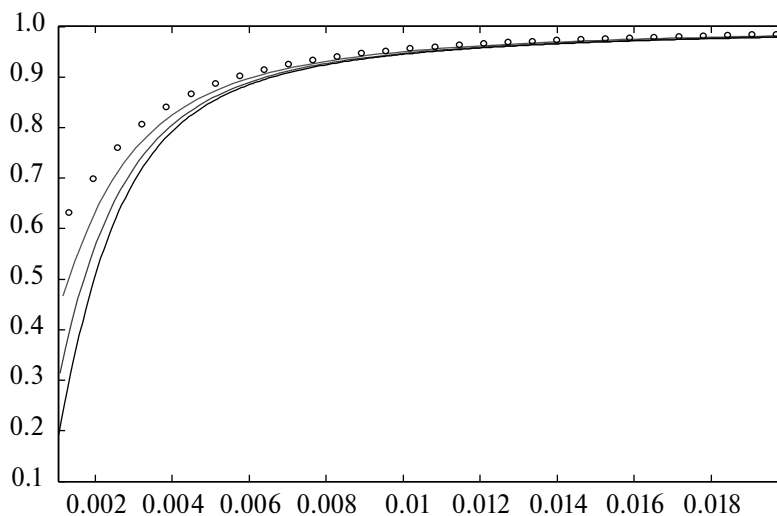
На нисходящей поверхности центрального тела возникает особая точка, разбивающая течение в форкамере на внешнюю и внутреннюю области, отличающиеся распределением скорости w по радиусу. Во внутренней области азимутальная скорость растет от нуля на оси симметрии до своего максимума на границе областей, а во внешней области убывает (см. фиг. 3). Из этого следует, что, схемная вязкость позволяет проникать закрутке в область под контактной поверхностью, разделяющей эти области. Увеличение числа расчетных точек приводит к уменьшению схемной вязкости, уменьшению толщины внутренней области и к увеличению максимальной величины скорости w . В дальнейшем при достаточной подробности сетки эта конфигурация разваливается: либо давление становится отрицательным, либо развиваются осцилляции, которые в конце концов тоже приводят к отрицательному давлению и/или плотности.

Поэтому все остальные расчеты проводились в модифицированной области, как предлагалось в предыдущем разделе: в область течения помещалась тоненькая трубочка (иголка), которая отделяла область течения от оси $r = 0$. Радиус иголки подбирался двумя способами. Либо проводился предварительный расчет без иголки, определялась точка отрыва линии тока на поверхности центрального тела и затем делался основной расчет с трубочкой, содержащей эту точку отрыва. Либо радиус оценивался по формуле (28). Оба способа дают близкие результаты. Величина вносимого иглой возмущения зависит от параметров задачи — чем интенсивнее закрутка, тем толще должна быть игла. Если величина возмущения оказывается неприемлемой, то, очевидно, нужно менять постановку задачи (учитывать вязкость, допускать обратное течение на выходной границе и т.д.). В наших расчетах площадь иголки составляла около 1% от площади минимального (критического) сечения.

Для оценки точности метода была проведена серия расчетов для канала с центральным телом, геометрия которого показана на фиг. 4, с углом поворота закручивающих лопаток 45° и толщиной иглы 0.001 (радиус критического сечения равен 0.01). На этом же рисунке приведены изолинии давления установившегося течения, отнесенного к давлению торможения p_0 . Видно резкое падение давления к оси сопла. На фиг. 5, 6 и 7 приведены профиль $P(r)/P_0$, и профили $u(r)$ и $w(r)$, отнесенные к местной скорости звука, (числа Маха) в сечении $z = 0.180$, полученные в расчетах этой конфигурации на четырех сетках: 55×30 , 110×60 , 220×120 и 440×240 точек. Видна неравномерная сходимость решения по числу точек: решение практически не меняется около наружной стенки сопла и достаточно медленно сходится около оси. Этот факт объясняется, по-видимому, резким изменением величин в окрестности оси (иголки).



Фиг. 4. Изолинии давления при угле поворота лопаток 45° (толщина иглы 0.001).

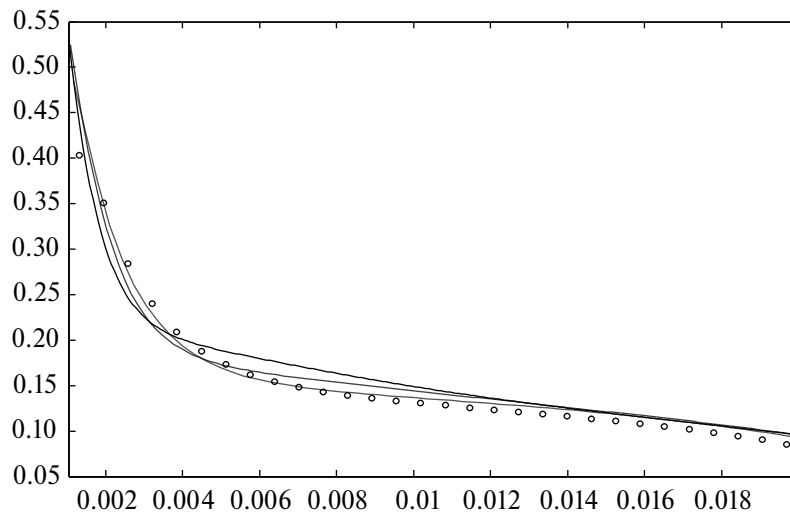


Фиг. 5. Профили $P(r)$ при $z = 180$: на четырех сетках: кружки — 55×30 точек, штрих-пунктир — 110×60 , штриховая линия — 220×120 , сплошная линия — 440×240 .

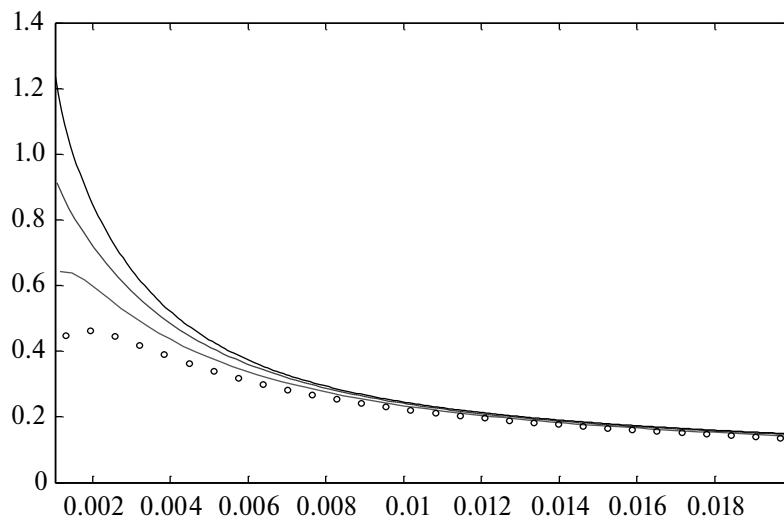
Обращают на себя внимание еще два факта. Во-первых, около оси сопла заметно наличие достаточно мощной скоростной струи, в которой скорость более чем в 2 раза превосходит скорость в основном потоке. Это объясняется, по-видимому, тем, что около оси давление существенно ниже, чем в основном потоке, и перепад давления по отношению к давлению на входе около оси больше. И, во-вторых, любопытно, что азимутальная скорость становится сверхзвуковой в некоторой окрестности оси еще до прохождения потоком критического сечения. В диффузоре (канал после горловины) отмеченные особенности также имеют место.

Цель следующего численного эксперимента состояла в оценке возможностей алгоритма при расчете течения с внутренними, в том числе висячими ударными волнами. В качестве исследуемого течения был выбран сверхзвуковой поток в трубе переменного сечения. Геометрия трубы представлена на фиг. 8. (Масштаб искажен: реальные поперечные размеры следует уменьшить в 20 раз.) На входе в трубу слева сверхзвуковое течение с $M = 1.1$. Горизонтальные и наклонные стенки сопряжены гладким образом. Центральное тело в расчетной области отсутствует (поэтому отсутствует и искусственная игла).

Расчеты проводились на последовательности разностных сеток; представленные расчеты получены для сетки 960×640 . Такая подробная сетка потребовалась, чтобы получить известный теоретический результат для осесимметричных (незакрученных) потоков, что регулярное Маховское отражение ударных волн от оси невозможно, обязательно образуется лямбда-скачок (ударная волна подходит к оси по нормали). На этом же рисунке



Фиг. 6. Профили $u(r)$, отнесенные к местной скорости звука, при $z = 180$: на четырех сетках: кружки — 55×30 точек, штрихпунктир — 110×60 , штриховая линия — 220×120 , сплошная линия — 440×240 .

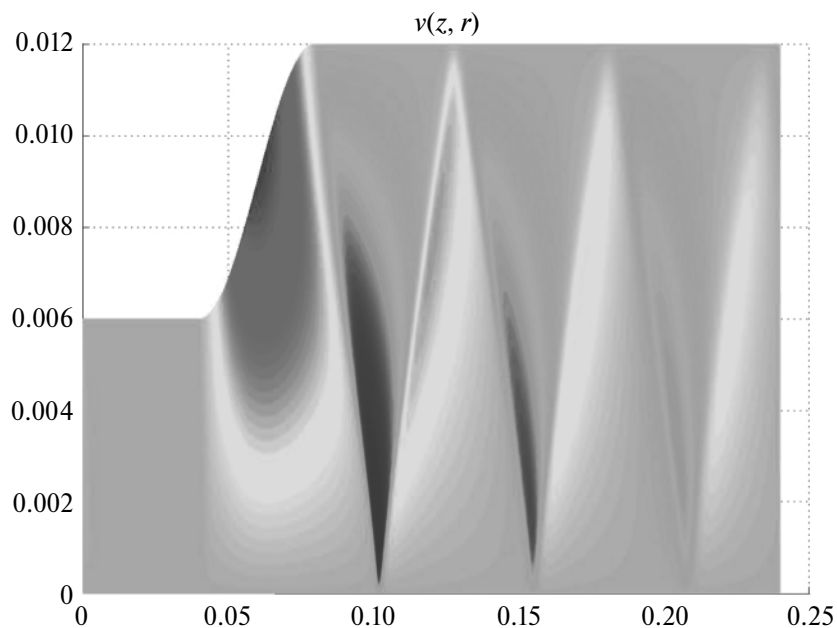


Фиг. 7. Профили $w(r)$, отнесенные к местной скорости звука, при $z = 180$: на четырех сетках: кружки — 55×30 точек, штрихпунктир — 110×60 , штриховая линия — 220×120 , сплошная линия — 440×240 .

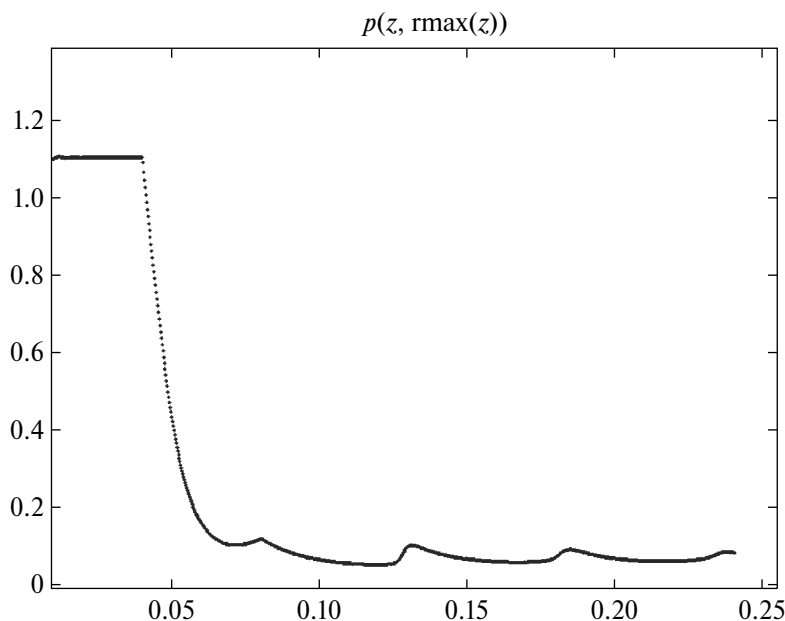
приведено распределение радиальной скорости v , на котором хорошо видна структура скачков. Отчетливо видна система косых ударных волн, причем первая волна — висячая. Хорошо видно, что при прохождении через очередную ударную волну v меняет знак, и что в решении отсутствуют паразитные осцилляции.

На фиг. 9 представлено распределения давления P на стенке трубы. Отчетливо видно, что первая ударная волна висячая. Причем точка формального зарождения скачка располагается довольно близко к стенке, что порождает на поверхности трубы излом производной по z (не скачок). Но уже вторичная волна, пришедшая на стенку в точке $z \approx 0.127$ на фиг. 9 отчетливо видна как скачок, размытый в силу разностных эффектов. Следующий скачок на стенке также хорошо виден, хотя его размытие больше.

Следующий расчет посвящен расчету закрученного течения в канале с дозвуковым выходом. Геометрия сопла показана на фиг. 10 и 11. Трудность постановки задачи при дозвуковом выходе заключается в том, что в закрученных потоках давление сильно меняется поперек канала, поэтому задавать на выходе из канала постоянное давление, как это обычно делается, неестественно и приводит к неудовлетворительным последствиям: в лучшем случае течение в окрестности выхода очень сильно искажается, и это искажение возмущает всю область течения. Но, как правило, этим дело не ограничивается: на выходе в приосевой области возникает втекание, что требует других граничных условий и другой постановки задачи. Чтобы избежать этих трудностей мы поместили выход из расчетной области достаточно далеко от оси. Так, чтобы закрутка уже была незначительной, и

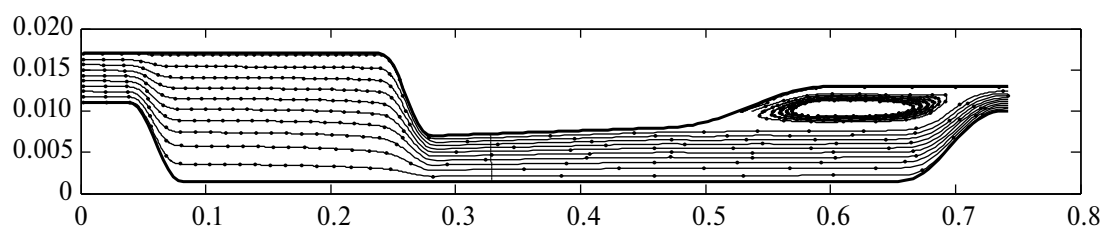


Фиг. 8. Распределение компоненты скорости v .

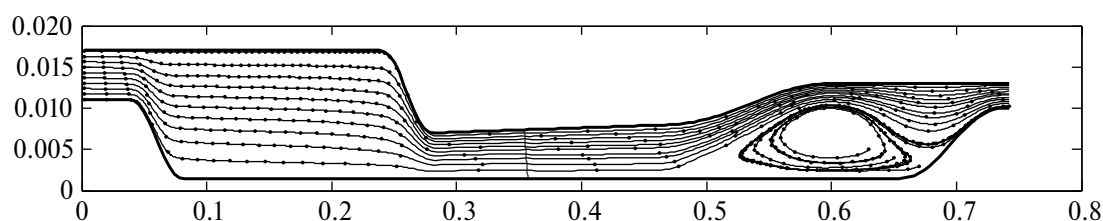


Фиг. 9. Профиль давления вдоль стенки трубы.

чтобы можно было задавать граничное условие в виде $P = \text{const}$. Но и в этом случае получить стационарное течение достаточно сложно. В большинстве случаев течение выходит на квазипериодический режим с большой амплитудой перемещения ударной волны вдоль канала. При этом в области диффузора (расширение канала после горловины) возникает крупномасштабный, бурно эволюционирующий, вихрь. Он возникает либо в верхней части сопла, либо в нижней, но в некоторых расчетах время от времени происходит перескакивание вихря снизу на верх и наоборот. Найти какую-нибудь закономерность в этом процессе не удалось. На фиг. 10 и 11 изображены линии тока для двух моментов времени одного из расчетов — векторные линии поля (u, v) , полученные интегрированием по времени системы уравнений $\{dz/dt = u, dr/dt = v\}$ с фиктивным временем t при разных начальных данных $(z(0), r(0))$. На полученных линиях тока через равные промежутки фиктивного времени t нанесены точки, что позволяет видеть скорость жидкости вдоль линии тока (по длине отрезка между точками). На тех же рисунках штриховой линией нанесено положение ударной волны (точнее, линия



Фиг. 10. Линии тока и ударная волна при верхнем расположении вихря.



Фиг. 11. Линии тока и ударная волна при нижнем расположении вихря.

$M = 1$). Видно заметное перемещение ударной волны. В других расчетах оно может быть еще больше (вплоть до разрушения ударной волны и временный переход на чисто дозвуковое течение). Эволюция ударной волны не связана непосредственно с размером и положением вихря — течение носит существенно нестационарный характер (это видно также из незамкнутости линий тока).

ВЫВОДЫ

В целом, представленный численный метод позволяет эффективно проводить расчеты как стационарных, так и нестационарных закрученных потоков в канале с центральным телом (и без него). Такие течения могут иметь достаточно сложную структуру: приосевая струя, резкие перепады давления, сверхзвуковая закрутка уже в камере предварительного смешения, система скачков в диффузоре, нетривиальная и неоднородная эволюция положения ударной волны в центральной части канала. и т. д. На типичной сетке отношение шагов по радиусу и вдоль оси канала достигало $1/25$ и не было ограничительным для алгоритма. В результате проведенных численных расчетов продемонстрирована специфика постановки задач для закрученных потоков в каналах сложной геометрии: закрутка на входе определяется моделированием угла лопаток; запуск сопла осуществляется постепенно; центральное тело, если оно есть, не должно заканчиваться, если около него есть закрутка, оно должно быть продолжено вспомогательной трубкой подходящего радиуса; если выход дозвуковой, то он должен находиться достаточно далеко от оси и т. д. Также можно отметить особенности закрученных потоков в каналах с центральным телом: стационарное решение, если оно есть, как правило, не является притягивающим; в течении возникают крупномасштабные движущиеся вихри; ударная волна может распространяться на большие расстояния, если для этого достаточно места и т. д.

Таким образом, предложенная методология расчета достаточно хорошо работает для показанного сложного многомасштабного потока. Примененный здесь метод независимых потоков, рассчитываемых на основе локально неявных схем с пересчетом, позволяет конструировать алгоритмы второго порядка точности при сохранении абсолютной устойчивости алгоритма. Эта же техника позволяет строить абсолютно устойчивые схемы и для параболических уравнений и систем (см. [5]). На основании этого можно предположить, что аналогичный подход окажется эффективным и для задач, связанных с уравнениями Навье—Стокса.

СПИСОК ЛИТЕРАТУРЫ

1. Годунов С. К. Разностный метод численного расчета разрывных решений уравнений гидродинамики // Матем. сборник. 1959. №3. С. 271—306.
2. Годунов С. К., Забродин А. В., Иванов М. Я., Крайко А. Н., Прокопов Г. П. Численное решение многомерных задач газовой динамики. М.: Наука, 1975. 400 с.
3. Roe P. L. Approximate Riemann solvers, parameter vectors, and difference schemes // J. of Comp. Phys. 1981. V. 43. P. 357—372.

4. Osher S., Solomon F. Upwind difference schemes for hyperbolic systems of conservation laws // Math. Comput. 1982. V. 38. P. 339–348.
5. Гейн С. В., Зайцев Н. А., Посвянский В. С., Радвогин Ю. Б. Метод независимых потоков для численного решения многомерного уравнения теплопроводности : Препринты ИПМ им. М. В. Келдыша. 2003. № 53. 23 с.
6. Радвогин Ю. Б. Экономичные безусловно устойчивые локально-неявные разностные схемы решения двумерных гиперболических систем : Препринты ИПМ им. М. В. Келдыша, 2003. № 35. 32 с.
7. Radvogin Yu. B., Zaitsev N. A. A locally implicit second order accurate difference schemes for solving 2D time-dependent hyperbolic systems and Euler equations // Proc. of Intern. Conf. on Spectral and High order Methods. Herzliya. 1998. in Appl. Num. Math. M. 2000. V. 33. P. 525–532.

COMPUTATION OF UNSTEADY SWIRLING FLOWS IN NOZZLES AND PIPES BY APPLYING A NEW LOCALLY IMPLICIT GODUNOV-TYPE SCHEME

N. A. Zaitsev*, Yu. G. Rykov**

Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences, Miusskaya sq., 4, Moscow, 125047 Russia

*e-mail: nikolai_zaitsev@mail.ru

**e-mail: yu-rykov@yandex.ru

Received 15 March, 2024

Revised 15 March, 2024

Accepted 05 May, 2024

Abstract. A new class of numerical schemes for calculating unsteady swirling flows in nozzles and pipes based on compressible inviscid gas equations is presented. The main advantage of these schemes is their ability to efficiently handle unsteady problems with multiple scales. The construction of such schemes is based on the well-known Godunov approach, which involves calculating fluxes at the faces of computational cells (volumes) by solving auxiliary one-dimensional problems in the vicinity of each face and approximating conservation laws. The scheme switches from an explicit method to an implicit one for flux calculation based on an analysis of the current solution near the cell face. The scheme is absolutely stable and does not generate spurious oscillations. Its effectiveness is demonstrated in the calculation of unsteady swirling flows in nozzles and pipes. Specifics of setting up problems of this type are investigated, and options for proper problem formulation are proposed. Properties of the solution for swirling flows with a central body covering only part of the symmetry axis in the computational domain are also studied.

Keywords: conservative locally implicit scheme with recalculation, Godunov-type scheme, swirling gas flow in a nozzle, supersonic flow, shock waves, multi-scale problems.

МЕТОД СВОБОДНОЙ ГРАНИЦЫ ДЛЯ СОПРЯЖЕННЫХ ЗАДАЧ ДИНАМИКИ ГАЗА И ТВЕРДОГО ТЕЛА

© 2024 г. И. С. Меньшов^{1,*}

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

*e-mail: menshov@kiam.ru

Поступила в редакцию 02.04.2024 г.

Переработанный вариант 02.04.2024 г.

Принята к публикации 05.05.2024 г.

В настоящей работе предлагается новый подход к численному моделированию течения газа около стационарных и движущихся твердых тел, который позволяет использовать эйлеровы сетки, не связанные с геометрией тела. Тела предполагаются абсолютно жесткими и недеформируемыми, их упругие свойства не учитываются. Газ является невязким и нетеплопроводным и описывается в рамках уравнений сжимаемой жидкости. Предлагаемый подход основывается на осреднении уравнений исходной модели по малому пространственному фильтру. В результате получается система осредненных уравнений, включающих дополнительную величину — параметр объемной доли твердого тела, пространственное распределение которого дает цифровое представление геометрии тела (аналог функции порядка). Эта система уравнений действует во всем пространстве. При таком подходе стандартная краевая задача в части пространства (занятого газом) сводится фактически к задаче Коши во всем пространстве. Для одномерной модели рассматривается численное решение осредненных уравнений методом Годунова. При этом в пересекаемых ячейках вводится разрывное восполнение решения, что приводит к рассмотрению составной задачи Римана, описывающей распад начального разрыва при наличии ограничивающей стенки. Доказывается, что аппроксимация численного потока на решении составной задачи Римана обеспечивает перенос функции порядка без численной диссипации. Библ. 25. Фиг. 6.

Ключевые слова: газовая динамика около движущейся поверхности твердого тела, осредненные уравнения Эйлера, составная задача Римана, численный метод Годунова.

DOI: 10.31857/S0044466924080171, EDN: XZTYNI

1. ВВЕДЕНИЕ

Численное моделирование взаимодействия течения сжимаемой жидкости со стационарными или движущимися твердыми телами является важной задачей, имеющей большое прикладное значение и в то же время остающейся весьма актуальной в научном плане. Стандартным классическим подходом к решению этой задачи является применение лагранжево-эйлеровых методов, когда геометрия поверхности тела отслеживается явным образом движущимися вместе с телом узлами сетки. Граничные условия при этом естественным образом выставляются на поверхности тела [1]. Применительно к сложной меняющейся во времени геометрии подходы, основанные на согласованных с геометрией сетках, становятся неэффективными, поскольку требуют многократной генерации или перестройки сетки. Последнее является затратной вычислительной процедурой, что в ряде случаев может приводить к практически невыполнимым задачам.

Альтернативой лагранжево-эйлеровых подходов могут служить развиваемые в последнее время чисто эйлеровы методы, когда геометрия тела рассматривается как свободная, не связанная с пространственной дискретизацией расчетной области. Сетка при таких подходах никак не связана с телом и, как правило, не меняется в процессе счета. Можно выделить две группы эйлеровых методов. К первой относятся методы, которые основаны на явной реконструкции геометрии тела на заданной несогласованной сетке. При этом используется цифровое представление геометрии посредством параметра доли объема сеточного элемента, занимаемого газом, т.н. VOF (volume of fluid)-методы [2,3], или функции расстояния от узла сетки до поверхности тела, LS (level-set)-методы [4,5]. Для выполнения граничных условий на реконструированной поверхности обычно применяется метод окаймляющих границу тела фиктивных ячеек (ghost cells), в которых параметры течения доопределяются надлежащим образом [6,7].

Вторую группу представляют методы диффузной границы. Эти методы изначально строились для численного моделирования многофазных сред, состоящих из различных материалов, разделенных межфазными границами [8]. Пространственное распределение каждого материала (фазы) здесь задается характеристической

функцией — функцией Хевисайда соответствующей области, которая является решением уравнения переноса в поле скорости среды. Таким образом, в отличие от методов первой группы, в методе диффузной границы реконструкция геометрии в процессе расчета не производится. Вся информация о положении границы содержится в распределении функции порядка. При численном интегрировании в силу численной диссипации характеристическая функция теряет разрывный тип и размазывается по пространству так, что граница представляется узкой переходной областью, где значения характеристической функции плавно меняются от 0 до 1. Недостатком такого подхода является переходная (смесевая) область, которая требует специального моделирования.

В работах [9–11] метод диффузной границы применяется для численного моделирования взаимодействия течения газа с упругопластической средой при больших деформациях. В этих работах рассматривается специальная модель переходной зоны вблизи границы, включающей как газовую, так и твердую компоненту. Вывод этой модели проводится на основе вариационного принципа и принципа термодинамической согласованности.

В [12] модель диффузной границы используется для численного моделирования невязких сжимаемых течений около твердых тел произвольной формы на простых равномерных декартовых сетках. Геометрия тела представляется в этой работе скалярным полем, которое задает распределение объемной доли газа в ячейках сетки. Уравнение переноса на объемную долю численно решается вместе с модифицированной системой уравнений газовой динамики, учитывающей присутствие твердой компоненты. Модифицированная система представляет собой фактически редуцированную модель Байера–Нунциато [8]. Система определяющих уравнений решается сквозным образом во всей области, включая твердое тело. Несмотря на высокоточные пространственные и временные расчетные схемы — DG и ADER, граница тела, тем не менее, обязательно размазывается в пределах нескольких ячеек из-за численной вязкости. Важные для практических приложений данные, такие как, например, распределения параметров течения по поверхности тела, при этом теряются. Это является одним из недостатков метода. Другой состоит в том, что система определяющих уравнений, хотя и гиперболическая, становится неконсервативной из-за членов с градиентом объемной доли. Ее решение, вообще говоря, не единственно и определяется конкретным видом соотношений на сильном разрыве, которые зависят от выбора пути в фазовом пространстве [13].

В вычислительной аэродинамике следует отметить также семейство эйлеровых методов, которые также не требуют сеточной привязки к расчетной области, но используют явным образом ее геометрическое представление для постановки граничных условий. В литературе они называются методами погруженных границ (Immersed Boundary Method) [6, 14]. Граничные условия либо рассчитываются непосредственно путем рассмотрения усеченных ячеек сетки (cut-cell method), либо моделируются посредством добавления в правую часть определяющих уравнений специальных источников членов. К ним относятся метод штрафных функций [15, 16], метод компенсационных потоков [17, 18] и ряд других.

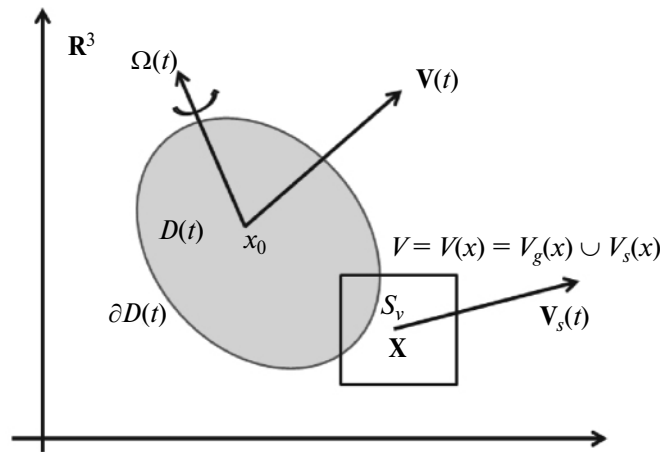
В настоящей работе предлагается эйлерова методика для задач взаимодействия течения газа с поверхностью твердого тела, которая объединяет методы диффузной границы и VOF. Математическая модель строится по аналогии с моделями многофазных сред. Система газ–твердое тело рассматривается как двухфазная среда. Определяющая система уравнений выводится путем осреднения исходных газодинамических уравнений по малому пространственному фильтру. Для численного интегрирования получающейся в результате системы мы используем метод Годунова [19]. Чтобы исключить численную диффузию границы тела, аппроксимация численного потока для пересекаемых ячеек выполняется на решении составной задачи Римана [20]. В этой задаче распад произвольного разрыва развивается в ограниченном границей тела полупространстве. С этой целью на каждом временном шаге мы переходим в пересекаемых ячейках от диффузного к геометрическому представлению границы. Численная реализация предложенной методики демонстрируется в настоящей работе пока только на одномерной модели. Показывается, что в этом случае получается бездиссипативный перенос геометрии по эйлеровой сетке. Обобщение на многомерный случай будет предметом последующих публикаций.

2. МОДЕЛЬ ОСРЕДНЕННЫХ УРАВНЕНИЙ

Рассматривается течение газа около замкнутой области $D(t)$, занятой твердым телом. Движение тела характеризуется в общем случае вектором поступательной скорости $\mathbf{V}_0(t)$ и вектором мгновенной угловой скорости $\boldsymbol{\Omega}(t)$. Поле скорости, индуцированное движением тела, определяется во всем пространстве как $\mathbf{V}_s = \mathbf{V}_0(t) + [\boldsymbol{\Omega}(t) \times (\mathbf{x} - \mathbf{x}_0)]$, где $\mathbf{x}$ — радиус-вектор, $\mathbf{x}_0$ — радиус-вектор центра масс тела (см. фиг. 1). Заметим, что это поле скорости бездивергентно, $(\nabla, \mathbf{V}_s) = 0$. Для дальнейшего поле скорости считается известным.

Введем характеристическую функцию геометрии тела:

$$\chi = \chi(t, \mathbf{x}) = \begin{cases} 0, & \text{если } \mathbf{x} \in D(t), \\ 1, & \text{если } \mathbf{x} \notin D(t). \end{cases} \quad (1)$$



Фиг. 1. Поле скорости движения твердого тела и пространственный фильтр для вывода осредненных уравнений.

Эта функция будет удовлетворять уравнению переноса

$$\frac{\partial \chi}{\partial t} + \mathbf{V}_s \nabla \chi = 0, \quad (2)$$

которое может также быть записано в дивергентной форме,

$$\frac{\partial \chi}{\partial t} + \operatorname{div}(\mathbf{V}_s \chi) = 0. \quad (3)$$

Будем предполагать, что течение газа описывается системой уравнений Эйлера (невязкий и нетеплопроводный газ), которая записывается в векторной консервативной форме как

$$\frac{\partial \mathbf{q}}{\partial t} + \frac{\partial \mathbf{f}_m}{\partial x_m} = 0, \quad (4)$$

где $\mathbf{q} = [\rho, \rho u_k, \rho E]^T$ — вектор консервативных переменных, $\mathbf{f}_m = [\rho u_m, \rho u_m u_k + p \delta_{mk}, \rho u_m H]^T$ — потоковый вектор, $k = 1, 2, 3$. Решение этой системы ищется в области $x \in R^3 \setminus D(t)$ с граничными условиями непротекания на границе тела, $u_m n_m = V_{sm} n_m$ для $\mathbf{x} \in \partial D(t)$. Здесь $E = \frac{1}{2} u_m u_m + \frac{p}{\rho(\gamma-1)}$ — удельная полная энергия, $H = E + \frac{p}{\rho}$ — удельная полная энтальпия, $\mathbf{n} = [n_k]$ — вектор единичной внешней к $D(t)$ нормали (т.е., направленной в сторону газа), остальные обозначения являются общепринятыми.

Для выведения модели диффузной границы для рассматриваемого течения газа введем фильтр — малый элемент пространства $V(\mathbf{x})$ с опорной точкой $\mathbf{x}$. В общем случае он состоит из подобласти, занимаемой газом, и подобласти, занимаемой телом, $V(\mathbf{x}) = V_g(\mathbf{x}) + V_s(\mathbf{x})$ (фиг. 1). Определим средние для выбранного фильтра:

$$\begin{aligned} V_g &= \int_V \chi dV, \\ \bar{\varphi} &= \frac{1}{V_g} \int_V \chi \varphi dV, \\ \alpha &= \frac{V_g}{V} = \frac{1}{V} \int_V \chi dV, \end{aligned} \quad (5)$$

где $\varphi = \varphi(t, \mathbf{x})$ — некоторый параметр течения. Очевидно, V_g имеет смысл объема подобласти $V_g(\mathbf{x})$, занимаемого газом, а α — объемной доли газа. Здесь мы используем одно обозначение V для области $V(\mathbf{x})$ и ее объема. Предполагается, что эта область геометрически фиксирована, так что объем V есть константа, не зависящая от $\mathbf{x}$. При этом объемы V_g и $V_s = V - V_g$ переменные и зависят от опорной точки $\mathbf{x}$. Отметим, что введенные в (5) средние значения определены только в точках, для которых $\alpha > 0$, т.е., когда фильтр целиком находится в газовой подобласти или пересекается границей тела.

Для дальнейшего рассмотрения нам понадобятся следующие утверждения.

Лемма 1. Для произвольной пробной функции $\varphi = \varphi(t, \mathbf{x})$ справедливы следующие соотношения:

$$\begin{aligned} \int_V \varphi \frac{\partial \chi}{\partial t} dV &= - \int_{S_V} \varphi (\mathbf{V}_s, \mathbf{n}) dS, \\ \int_V \varphi \nabla \chi dV &= \int_{S_V} \varphi \mathbf{n} dS, \end{aligned} \quad (6)$$

где $S_V = V \cap \partial D(t)$ — часть поверхности твердого тела $\partial D(t)$ внутри фильтра V , $\mathbf{n}$ — единичная внешняя нормаль к S_V , направленная в сторону газа.

Доказательство. Пусть граница тела $\partial D(t)$ определяется уравнением $\Phi(t, \mathbf{x}) = 0$. Путем надлежащей нормировки функцию Φ можно выбрать так, чтобы выполнялись условия $\Phi(t, \mathbf{x}) < 0$ при $\mathbf{x} \in D(t)$, $\Phi(t, \mathbf{x}) > 0$ при $\mathbf{x} \notin D(t)$, $|\nabla \Phi(t, \mathbf{x})| = 1$. Тогда характеристическая функция $\chi(t, \mathbf{x}) = H[\Phi(t, \mathbf{x})]$, где $H(\cdot)$ — функция Хевисайда, а соответствующие производные вычисляются через δ -функцию Дирака:

$$\frac{\partial \chi}{\partial t} = \delta[\Phi(t, \mathbf{x})] \frac{\partial \Phi}{\partial t}, \quad \nabla \chi = \delta[\Phi(t, \mathbf{x})] \nabla \Phi = \delta[\Phi(t, \mathbf{x})] \mathbf{n}. \quad (7)$$

Учитывая, что производные $\Phi(t, \mathbf{x})$ связаны со скоростью $\mathbf{V}_s$ уравнением

$$\frac{\partial \Phi}{\partial t} + (\mathbf{V}_s, \nabla \Phi) = 0, \quad (8)$$

приходим к искомым соотношениям (6).

Лемма 2. Для произвольной пробной функции $\varphi = \varphi(t, \mathbf{x})$ справедливы следующие соотношения:

$$\begin{aligned} \frac{\partial}{\partial t} \int_V \varphi dV &= \int_V \frac{\partial \varphi}{\partial t} dV, \\ \nabla \int_V \varphi dV &= \int_V \nabla \varphi dV. \end{aligned} \quad (9)$$

Доказательство. Первое соотношение очевидно в силу того, что фильтр $V = V(\mathbf{x}) = \text{const}$. Второе справедливо, так как

$$\frac{\partial}{\partial x_k} \int_V \varphi dV = \int_{\partial V} \varphi n_k dS$$

по определению производной, и

$$\int_V \frac{\partial \varphi}{\partial x_k} dV = \int_{\partial V} \varphi n_k dS$$

в соответствии с теоремой Остроградского—Гаусса. Здесь n_k обозначают компоненты единичной внешней нормали на границе фильтра.

Поделим уравнение (2) на V и проинтегрируем по фильтру $V(\mathbf{x})$. Тогда с учетом леммы 1 и 2 получим

$$\frac{\partial \alpha}{\partial t} + \frac{1}{V} \int_{S_V} (\mathbf{V}_s, \mathbf{n}) ds = 0. \quad (10)$$

Аналогичным образом можно получить осредненные уравнения Эйлера, умножая (4) на характеристическую функцию χ и интегрируя по объему $V(\mathbf{x})$. Результатом будет следующая система уравнений:

$$\frac{\partial \alpha \bar{\mathbf{q}}}{\partial t} + \frac{\partial \alpha \bar{\mathbf{f}}_m}{\partial x_m} - \frac{1}{V} \int_{S_V} [n_k \mathbf{f}_k - \mathbf{q}(\mathbf{V}_s, \mathbf{n})] ds = 0, \quad (11)$$

которая с учетом граничных условий на поверхности тела, запишется в виде

$$\begin{aligned} \frac{\partial \alpha \bar{\mathbf{q}}}{\partial t} + \frac{\partial \alpha \bar{\mathbf{f}}_m}{\partial x_m} - \frac{1}{V} \int_{S_V} \mathbf{G} ds &= 0, \\ \mathbf{G} &= \begin{pmatrix} 0 \\ pn_k \\ p\mathbf{V}_s, \mathbf{n} \end{pmatrix}. \end{aligned} \quad (12)$$

Вектор $\mathbf{G}$ фактически является компенсационным потоком, который вводится в методе внутренней границы, описанном в работе [17, 18]. Заметим, что давление в векторе $\mathbf{G}$ соответствует давлению на поверхности тела; оно не совпадает со средним давлением в векторах $\bar{\mathbf{q}}$ и $\bar{\mathbf{f}}_m$ и служит для обеспечения граничного условия на теле.

Лемма 3. *Справедливо следующее соотношение между объемной долей и внешней нормалью границы тела:*

$$\frac{\partial \alpha}{\partial x_k} = \frac{1}{V} \int_{S_V} n_k ds. \quad (13)$$

Доказательство. По определению и леммам 1 и 2 имеем

$$\frac{\partial \alpha}{\partial x_k} = \frac{\partial}{\partial x_k} \left(\frac{1}{V} \int_V \chi dV \right) = \frac{1}{V} \int_V \frac{\partial \chi}{\partial x_k} dV = \frac{1}{V} \int_{S_V} n_k dS. \quad (14)$$

В силу этой леммы мы будем рассматривать вместо уравнения (10) для объемной доли его дифференциальный аналог —

$$\frac{\partial \alpha}{\partial t} + \mathbf{V}_s, \nabla \alpha = 0.$$

Это уравнение удобнее рассматривать в консервативном виде, так как поле $\mathbf{V}_s$ бездивергентное:

$$\frac{\partial \alpha}{\partial t} + \nabla(\alpha \mathbf{V}_s) = 0.$$

Система уравнений (10), (11) справедлива, в принципе, для любого фильтра. Уменьшая фильтр и стягивая его фактически к точке, можно получить предельное выражение для интеграла по поверхности тела:

$$\lim_{V \rightarrow 0} \frac{1}{V} \int_{S_V} \varphi \mathbf{n} dS = \varphi \nabla \alpha. \quad (15)$$

Тогда система (10)–(12) запишется в следующем виде:

$$\begin{aligned} \frac{\partial \alpha}{\partial t} + \mathbf{V}_s, \nabla \alpha &= 0, \\ \frac{\partial \alpha \mathbf{q}}{\partial t} + \frac{\partial \alpha \mathbf{f}_m}{\partial x_m} - \hat{\mathbf{G}} &= 0, \\ \hat{\mathbf{G}} &= \begin{pmatrix} 0 \\ p \frac{\partial \alpha}{\partial x_k} \\ p \mathbf{V}_s, \nabla \alpha \end{pmatrix}. \end{aligned} \quad (16)$$

Здесь мы опускаем черту над вектором консервативных переменных и вектором потоков и предполагаем, что компоненты этих векторов и уравнение состояния такие же, как в исходной задаче (4).

Система уравнений (16) является эквивалентной постановкой задачи об обтекании твердого тела газовой сжимаемой средой. Она совпадает с системой уравнений, которая рассматривалась в работе [12] для описания альтернативной модели течения сжимаемого газа около движущегося твердого тела. В этой работе авторы берут в качестве основы своих рассуждений неравновесную модель Байера–Нунзиато для двухфазного сжимаемого течения и затем переходят к предельному случаю, когда одна из фаз становится твердым телом.

В области, где $\alpha = 1$ система уравнений (16) переходит в классическую систему уравнений Эйлера. В области твердого тела $\alpha = 0$ она вырождается. Вектор $\hat{\mathbf{G}}$, включающий неконсервативные члены, играет роль штрафной функции в методе погруженной границы [15, 16]. Фактически он представляет собой компенсационный поток в методе свободной границы [17, 18], который определяет взаимодействие газа с твердой поверхностью обтекаемого тела и обеспечивает выполнение граничных условий на поверхности тела. Ниже мы будем придерживаться этой терминологии и называть неконсервативные члены компенсационными потоками. Отметим, что в [12] доказывается, что для решений (16) нормальные скорости тела и газа совпадают, $u_k n_k = \mathbf{V}_s \mathbf{n}$, в точках, где α скачком изменяется от 1 до 0.

Модель диффузной границы не связана с реальной областью течения газа и может рассматриваться во всем пространстве, т.е. не требует постановки граничных условий. Это позволяет свести краевую задачу в ограниченной области к начальной задаче Коши во всем пространстве. Движение тела при таком подходе полностью определяется пространственным распределением одного скалярного параметра α .

При численном решении это распределение задается дискретным образом на достаточно простой декартовой сетке. Для более точного представления может быть применена технология AMR локальной сеточной адаптации. При таком подходе можно полностью исключить задачу генерации сетки в области течения, которая в стандартных сеточных численных методах является достаточно затратной для задач со сложной пространственной и меняющейся во времени геометрией. Вся информация о геометрии тела полностью содержится в скалярной функции α , которая определяется решением уравнения переноса (16). Решение искомой задачи обтекания сводится к решению системы уравнений (16) во всем пространстве. Для исключения вырождения системы уравнений предполагается, что в твердом теле объемная доля газа не обращается точно в ноль, а имеет малое ненулевое значение $\alpha \approx 10^{-2} - 10^{-3}$ (см. [12]).

При численном решении (16) разрывное распределение объемной доли α становится сглаженным из-за численной вязкости, и граница представляется узкой переходной зоной. Точные данные о геометрии поверхности тела теряются. Это является основным недостатком подхода, так как в приложениях требуются не только интегральные данные аэродинамических характеристик, но и локальные, представляющие распределение газодинамических параметров по поверхности исследуемого объекта. Тем не менее, метод диффузной границы показывает хорошую точность описания внешнего течения. В частности, он успешно применяется для решения FSI задач (сопряженные задачи газовой динамики и динамики деформируемого твердого тела) [9–11].

В отличие от упомянутых выше работ построение численного метода проводится в настоящей работе на основе осредненных уравнений (10) и (12). Мы применяем метод конечного объема для дискретизации этих уравнений, что позволяет на подсеточном уровне учесть компенсационные потоки на поверхности твердого тела S_V в пересекаемых ячейках. Для этого на каждом временном шаге делается подсеточная реконструкция геометрии интерфейса (поверхности тела) на основе дискретного распределения объемной доли. Тогда при вычислении потоков через грани ячейки можно явно учесть эффект твердой стенки вблизи грани, а также должным образом аппроксимировать компенсационные потоки. Для этого используется метод Годунова [19] и решение составной задачи Римана, решение которой описывает распад произвольного разрыва в газе при наличии дополнительного контактного разрыва [20–22]. Это позволяет фактически убрать численную диссипацию при численном интегрировании уравнения для объемной доли. Ниже рассмотрим предлагаемую методику на примере одномерных уравнений.

3. ДИСКРЕТНАЯ МОДЕЛЬ

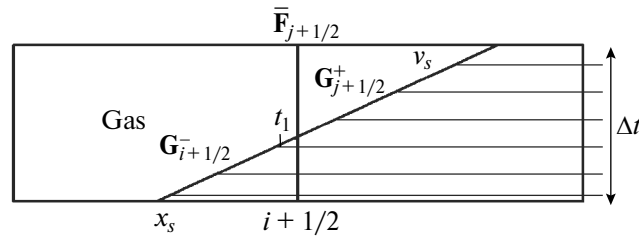
Основные элементы дискретной модели и численного метода рассмотрим на простой одномерной модели. Обобщение на пространственный случай будет предметом последующих публикаций. В одномерной модели граница тела представляется точкой x_s , расчетная область — интервалом, на котором введена сетка. Будем предполагать ее равномерной с шагом h .

Одномерный аналог системы уравнений (11) получается, если считать все параметры зависящими только от времени и одной пространственной координаты $x = x_1$, компоненты скорости тела $\mathbf{V}_s = (v_s, 0, 0)$, скорость газа $\mathbf{U} = (u, 0, 0)$, $V = h$, $S_v = x_s$. Интегрируя при этих предположениях (11) по времени на интервале $[t_n, t_{n+1}]$, $t_{n+1} = t_n + \Delta t$, и по ячейке сетки $[x_{i-1/2}, x_{i+1/2}]$, $x_{i+1/2} = x_{i-1/2} + h$ получаем дискретные уравнения следующего вида:

$$\mathbf{Q}_i^{n+1} = \mathbf{Q}_i^n - \frac{1}{h} \left[\int_{t^n}^{t^{n+1}} \mathbf{F}_{i+1/2}(t) dt - \int_{t^n}^{t^{n+1}} \mathbf{F}_{i-1/2}(t) dt - \int_{t^n}^{t^{n+1}} \mathbf{G}(x_s(t)) dt \right], \quad (17)$$

где $\mathbf{Q} = (\alpha, \alpha \mathbf{q})^T$ — расширенный вектор состояния, $\mathbf{F} = (\alpha v_s, \alpha \mathbf{f})^T$ — расширенный вектор потока, $\mathbf{G} = (0, \mathbf{f} - v_s \mathbf{q})^T$ — компенсационный поток на границе. Здесь $\mathbf{q}$ и $\mathbf{f}$ — стандартные векторы консервативных переменных и потока одномерной газодинамики. Вектор $\mathbf{Q} = 0$ и соответствующий вектор потока $\mathbf{F} = 0$ отвечают телу ($\alpha = 0$).

Таким образом, изменение вектора решения определяется вектором потока через грань ячейки $\bar{\mathbf{F}}_{i+1/2} = \int_{t^n}^{t^{n+1}} \mathbf{F}_{i+1/2}(t) dt$ и вкладом компенсационных потоков. При этом надо иметь в виду, что на временном шаге граница тела может пересечь грань ячейки и переместиться в соседнюю ячейку (фиг. 2). Поэтому вклад компенсационного потока должен быть перераспределен между двумя соседними ячейками. Если t_1 — момент времени, когда граница тела пересекает грань ячейки, то часть потока, отвечающая интервалу $[0, t_1]$, представляет собой вклад в i -ю ячейку, а другая часть $([t_1, \Delta t])$ — вклад в $i + 1$.



Фиг. 2. Численный и компенсационный поток на грани ячейки: t_1 — момент времени, когда граница тела пересекает грань ячейки.

Чтобы учесть эффект компенсационного потока, введем на каждой грани два вектора —

$$\mathbf{G}_{i+1/2}^- = \int_0^{t_1} \mathbf{G}(x_s(t)) dt, \quad \mathbf{G}_{i+1/2}^+ = \int_{t_1}^{\Delta t} \mathbf{G}(x_s(t)) dt. \quad (18)$$

Тогда разностная схема может быть записана в следующем виде:

$$\mathbf{Q}_i^{n+1} = \mathbf{Q}_i^n - \frac{1}{h} \left[(\bar{\mathbf{F}}_{i+1/2} - \mathbf{G}_{i+1/2}^-) - (\bar{\mathbf{F}}_{i-1/2} + \mathbf{G}_{i-1/2}^+) \right]. \quad (19)$$

Компенсационные потоки $\mathbf{G}_{i+1/2}^\pm$ возникают на гранях, разделяющих ячейки, одна из которых является пересекаемой. Они обеспечивают перенос массы, импульса и энергии при переходе границы тела через грань ячейки. В остальных случаях эти потоки равны нулю, и схема редуцируется в базовую газодинамическую схему. Необходимо также отметить, что компенсационные потоки должны учитывать направление нормали. Так, для случая на фиг. 2 потоки должны браться со знаком минус.

Аппроксимация численного потока $\bar{\mathbf{F}}_{i+1/2}$ и компенсационных потоков $\mathbf{G}_{i+1/2}^\pm$ выполняется методом С.К. Годунова на основе решения возникающей на грани составной (пристеночной) задачи Римана (СЗР). Эта задача в общем случае описывает распад произвольного разрыва вблизи твердой стенки. Подробно ее постановка, решение и аппроксимация численного и компенсационного потоков рассматриваются в разд. 4.

Возможны 4 типа СЗР в зависимости от того, находится грань в газе или теле, а также от паттерна — взаимного расположения газа и тела. Будем называть *левосторонним паттерном* расположение, когда газ находится слева от тела, и *правосторонним*, когда справа. Введем следующие условные обозначения для типов СЗР: СЗР_g для задачи, когда грань в газе, СЗР_s, когда грань в теле.

Решение первой задачи зависит от паттерна (*), координаты границы твердого тела (x_s), ее скорости (v_s) и начальных параметров газа слева и справа от разрыва ($\mathbf{Q}_l, \mathbf{Q}_r$): СЗР_g = СЗР_g(*, $x_s, v_s, \mathbf{Q}_l, \mathbf{Q}_r$), а решение второй — от паттерна, координаты границы, ее скорости и начальных параметров газа ($\mathbf{Q}$): СЗР_s = СЗР_s(*, $x_s, v_s, \mathbf{Q}$). Здесь $\mathbf{Q}$ обозначает расширенный вектор решения, $\mathbf{Q} = (\alpha, \alpha \mathbf{q})$, * = l для левостороннего паттерна, * = r для правостороннего.

Численный поток на грани σ строится либо на решении СЗР, либо на решении стандартной ЗР:

$$\begin{aligned} \alpha_l = 0, 0 < \alpha_r \leq 1 & \quad - \quad \mathbf{F}_\sigma = \mathbf{F}[\text{СЗР}_s(r, x_s, v_s, \mathbf{Q}_r)], \\ \alpha_r = 0, 0 < \alpha_l \leq 1 & \quad - \quad \mathbf{F}_\sigma = \mathbf{F}[\text{СЗР}_s(l, x_s, v_s, \mathbf{Q}_l)], \\ \alpha_l = 1, 0 \leq \alpha_r < 1 & \quad - \quad \mathbf{F}_\sigma = \mathbf{F}[\text{СЗР}_g(l, x_s, v_s, \mathbf{Q}_l, \mathbf{Q}_r)], \\ \alpha_r = 1, 0 \leq \alpha_l < 1 & \quad - \quad \mathbf{F}_\sigma = \mathbf{F}[\text{СЗР}_g(r, x_s, v_s, \mathbf{Q}_l, \mathbf{Q}_r)], \\ \alpha_r = 1, \alpha_l = 1 & \quad - \quad \mathbf{F}_\sigma = \mathbf{F}[\text{ЗР}(\mathbf{Q}_l, \mathbf{Q}_r)], \\ \alpha_r = 0, \alpha_l = 0 & \quad - \quad \mathbf{F}_\sigma = 0. \end{aligned} \quad (20)$$

Здесь через ЗР($\mathbf{Q}_1, \mathbf{Q}_2$) обозначается стандартное решение задачи Римана в газе, $\mathbf{F}[\cdot]$ — численный поток на решении задачи Римана.

4. СОСТАВНАЯ (ПРИСТЕНОЧНАЯ) ЗАДАЧА РИМАНА

Аппроксимация численного потока строится методом С.К. Годунова на основе решения составной (пристеночной) задачи Римана. При этом, как было сказано выше, необходимо рассмотреть четыре типа задач —

$C3P_s(l)$, $C3P_s(r)$, $C3P_g(l)$ и $C3P_g(r)$, которые различаются взаимным расположением газа, твердого тела и точки начального разрыва ($x = 0$). Фактически, достаточно рассмотреть один тип паттерна, например, левосторонний. Решение для правостороннего паттерна тогда можно получить, сведя задачу к левостороннему заменой знака координаты и скорости на противоположный.

Математическая постановка $C3P_g(l, x_s, v_s, Q_l, Q_r)$ и $C3P_s(l, x_s, v_s, Q_l)$ сводится к решению задачи Коши для расширенной системы уравнений газовой динамики

$$\frac{\partial Q}{\partial t} + \frac{\partial F}{\partial x} = 0, \quad (21)$$

где $Q = [\alpha, \alpha r, \alpha u, \alpha E]^T$ и $F = [\alpha v_s, \alpha u, \alpha u^2 + \alpha p, \alpha r H]^T$ с начальными данными при $t = 0$

$$\text{для } C3P_g - \quad \alpha = \begin{cases} 1, & x \leq x_s, \\ 0, & x > x_s, \end{cases} \quad Q = \begin{cases} Q_l, & x \leq 0, \\ Q_r, & 0 < x \leq x_s, \\ Q_s, & x > x_s, \end{cases} \quad (22)$$

$$\text{для } C3P_s - \quad \alpha = \begin{cases} 1, & x \leq x_s, \\ 0, & x > x_s, \end{cases} \quad Q = \begin{cases} Q_l, & x \leq x_s, \\ Q_s, & x > x_s. \end{cases} \quad (23)$$

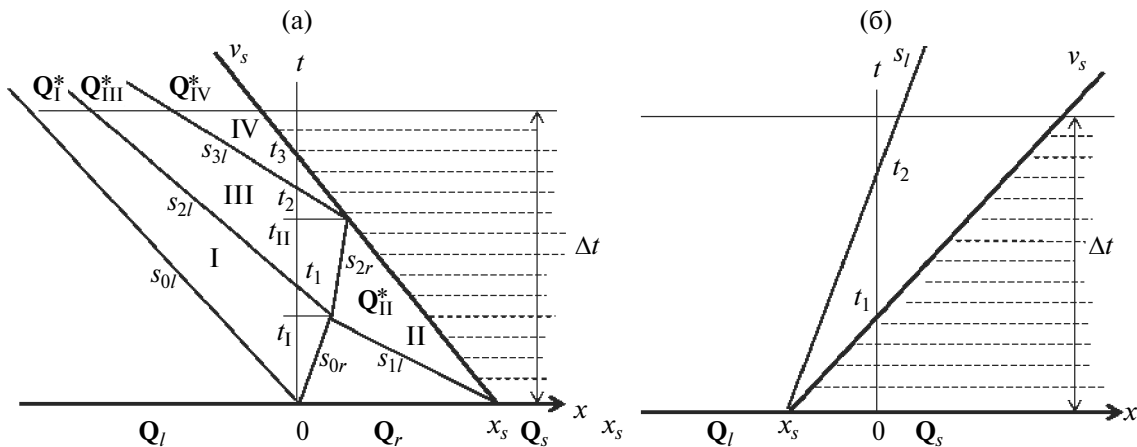
Здесь $Q_s = 0$ — вектор решения в области тела. Граница твердого тела движется с постоянной скоростью v_s . Мы также полагаем, что $F(Q_s) = 0$, и примитивный вектор для определенности принимает некоторые референсные значения, например, $\rho_s = 1$, $u_s = v_s$, $p_s = 1$, при $\alpha = 0$.

Первая задача, (21), (22), не является автомодельной. Ее решение описывает множественное взаимодействие разрывов, являющихся результатом распада начального разрыва ($x = 0$) и их последующих взаимодействий с поверхностью твердого тела. Характерная структура волн в этой задаче показана на фиг. 3а. Она определяется взаимодействием двух возмущенных зон, образующихся в результате распада начального разрыва ($x = 0$) с начальными параметрами Q_l , Q_r и воздействия потока Q_s на движущуюся со скоростью v_s поверхность тела.

Первая возмущенная зона (I) ограничена волнами s_{0l} и s_{0r} , вторая (II) — поверхностью тела v_s и волной s_{1l} . Взаимодействие первичных возмущенных зон начинается в момент времени t_I , когда пересекаются траектории волн s_{0r} и s_{1l} . Результатом является формирование вторичной возмущенной зоны (III), ограниченной волнами s_{2l} и s_{2r} . В момент времени t_{II} правосторонняя волна s_{2r} сталкивается с поверхностью тела, формируя третичную возмущенную зону (IV), ограниченную телом v_s и волной s_{3l} . Вектор состояния в возмущенных зонах обозначается верхним индексом * на фиг. 3.

В последующие моменты времени волны сталкиваются друг с другом и с телом, образуя новые волны и возмущенные зоны. Этот каскадный волновой процесс описать точно в общем случае не представляется возможным. Поэтому для аппроксимации численного потока мы воспользуемся приближенным решением, которое строится на основе методов HLL и HLLC [23]. Решение HLL используется для описания возмущенных областей I и III, HLLC — для пристеночных областей II и IV. Ниже кратко приведем эти решения для случая формирования возмущенной зоны I (HLL) и II (HLLC).

Начальными данными для HLL-приближения являются значения вектора состояния и вектора потока слева и справа от разрыва, Q_l , Q_r и F_l , F_r . Заметим, что в общем случае необязательно, что $F_l = F(Q_l)$ и $F_r = F(Q_r)$.



Фиг. 3. Волновая конфигурация в составной (пристеночной) задаче Римана: а — для случая $C3P_g$, б — для случая $C3P_s$.

Решение HLL определяется скоростями левой и правой волны, s_l и s_r , которые ограничивают возмущенную область, вектором состояния $\mathbf{Q}^*$ и вектором потока $\mathbf{F}^*$ возмущенной зоны. Эти векторы, при условии, что скорости волн известны, определяются из соотношений Рэнкина–Гюгонио на разрывах s_l и s_r :

$$\mathbf{Q}^* = \frac{\mathbf{F}_l - \mathbf{F}_r + s_r \mathbf{Q}_r - s_l \mathbf{Q}_l}{s_r - s_l}, \quad \mathbf{F}^* = \frac{s_r \mathbf{F}_l - s_l \mathbf{F}_r + s_l s_r (\mathbf{Q}_r - \mathbf{Q}_l)}{s_r - s_l}. \quad (24)$$

Скорости волн аппроксимируются в соответствии с [24] следующим образом:

$$s_l = \min [0, \min(u_l - a_l, \bar{u} - \bar{a})], \quad s_r = \max [0, \max(u_r + a_r, \bar{u} + \bar{a})], \quad (25)$$

где u — скорость газа, a — скорость звука, верхняя черта — среднее арифметическое значений слева и справа от разрыва.

Начальными данными для HLLC-приближения, описывающего возмущенную зону при взаимодействии газового потока с поверхностью движущегося тела, являются скорость тела v_l , вектор состояния потока $\mathbf{Q}_l$ и вектор потока $\mathbf{F}_l$. Опять же, вектор потока может быть в общем случае аппроксимацией и не совпадать с вектором потока $\mathbf{F}(\mathbf{Q}_l)$.

HLLC-приближение дает значение консервативного вектора $\mathbf{Q}^*$ и потокового вектора $\mathbf{F}^*$ в возмущенной зоне, ограниченной поверхностью тела v_s и волной s_l . Скорость волны определяется так же, как в HLL-решении:

$$s_l = \min(u_l - a_l, v_s - a_l). \quad (26)$$

Имея скорость волны, разделяющей возмущенную и невозмущенную области, определяем сначала вектор $\mathbf{Q}^*$ из законов сохранения массы, импульса и энергии —

$$\mathbf{F}(\mathbf{Q}^*, p^*) - \mathbf{F}(\mathbf{Q}_l) = s_l(\mathbf{Q}^* - \mathbf{Q}_l) \quad (27)$$

и условия $u^* = v_s$:

$$\mathbf{Q}^* = \begin{bmatrix} 1 \\ \frac{u_l - s_l}{v_s - s_l} \begin{pmatrix} \rho_l \\ \rho_l v_s \\ \rho_l E_l + (u_l - v_s) \left(\frac{p_l}{u_l - s_l} - \rho_l v_s \right) \end{pmatrix} \end{bmatrix}, \quad (28)$$

$$p^* = p_l + \rho_l(u_l - s_l)(u_l - v_s).$$

Аппроксимация потока в возмущенной области $\mathbf{F}^*$ определяется после этого из условия Рэнкина–Гюгонио:

$$\mathbf{F}^* = \mathbf{F}_l + s_l(\mathbf{Q}^* - \mathbf{Q}_l). \quad (29)$$

На основе HLL- и HLLC-солверов решение СЗР_g может быть построено в возмущенной области следующим образом:

$$\begin{aligned} \text{область I : } & \text{HLL}(\mathbf{Q}_l, \mathbf{F}(\mathbf{Q}_l), \mathbf{Q}_r, \mathbf{F}(\mathbf{Q}_r)) \Rightarrow (s_{0l}, s_{0r}, \mathbf{Q}_I^*, \mathbf{F}_I^*), \\ \text{область II : } & \text{HLLC}(\mathbf{Q}_r, \mathbf{F}(\mathbf{Q}_r), v_s) \Rightarrow (s_{1l}, \mathbf{Q}_{II}^*, \mathbf{F}_{II}^*), \\ \text{область III : } & \text{HLL}(\mathbf{Q}_I^*, \mathbf{F}_I^*, \mathbf{Q}_{II}^*, \mathbf{F}_{II}^*) \Rightarrow (s_{2l}, s_{2r}, \mathbf{Q}_{III}^*, \mathbf{F}_{III}^*), \\ \text{область IV : } & \text{HLLC}(\mathbf{Q}_{III}^*, \mathbf{F}_{III}^*, v_s) \Rightarrow (s_{3l}, \mathbf{Q}_{IV}^*, \mathbf{F}_{IV}^*). \end{aligned} \quad (30)$$

Что касается СЗР_s, хотя она является автомодельной, и ее решение может быть построено точно, для нее также используется приближенный солвер: $\text{HLLC}(\mathbf{Q}_l, \mathbf{F}(\mathbf{Q}_l), v_s) \Rightarrow (s_l, \mathbf{Q}^*, \mathbf{F}^*)$.

Рассмотренное выше приближенное решение СЗР_g справедливо, вообще говоря, при условии, что траектории левосторонних волн s_{0l}, s_{2l}, s_{3l} не пересекаются на временном шаге Δt . Если пересечение имеет место, то возникают дополнительные волны, которые меняют рассмотренное выше решение. Одним из решений этой проблемы может быть корректировка скоростей волн, например, так: $s_{0l}, s_{2l}, s_{3l} \leftarrow \min(s_{0l}, s_{2l}, s_{3l})$ [21]. Следует отметить, что ситуация с указанным выше пересечением, хотя теоретически и возможна, практически не возникает, как показывают многочисленные расчеты.

Численный поток, определяющий изменение консервативного вектора решения в счетной ячейке на шаге Δt , вычисляется на решении рассмотренной выше составной (пристеночной) задачи Римана. В случае СЗР_g с решением возмущенной зоны (30) поток аппроксимируется интегралом от потока $\mathbf{F}^*$ по шагу Δt :

$$\mathbf{F}_\sigma = [\tau_1 \mathbf{F}_I^* + (\tau_2 - \tau_1) \mathbf{F}_{III}^* + (\tau_3 - \tau_2) \mathbf{F}_{IV}^*] \Delta t, \quad (31)$$

где $\tau_k = t_k/\Delta t$, $k = 1, 2, 3$, а t_k — моменты времени, в которые волны пересекают грань ячейки (прямую $x = 0$ на фиг. 3а).

Эти моменты времени определяются моментами времени t_I первичного взаимодействия возмущенных областей и времени t_{II} вторичного взаимодействия потока с поверхностью тела. Они вычисляются по координате x_s и скоростям соответствующих волн:

$$t_I = \frac{x_s}{s_{0r} - s_{1l}}, \quad t_{II} = \frac{x_s(s_{2r} - s_{1l})}{(s_{0r} - s_{1l})(s_{0r} - v_s)}, \quad (32)$$

а времена t_k в (31) —

$$t_1 = \frac{s_{2l} - s_{0r}}{s_{2l}} t_I, \quad t_2 = \frac{(s_{2r} - s_{0r})t_I + (s_{3l} - s_{2r})t_{II}}{s_{3l}}, \quad t_3 = \frac{(s_{2r} - s_{0r})t_I + (v_s - s_{2r})t_{II}}{v_s}. \quad (33)$$

Выражение для потока (31) верно для волновой конфигурации, изображенной на фиг. 3а, когда все интересные нас взаимодействия волн происходят на шаге Δt . Если это не так, и какие-то волновые взаимодействия выходят за пределы временного интервала $[0, \Delta t]$, формула (31) уже не работает. Чтобы сделать ее универсальной, справедливой для любых волновых конфигураций, можно скорректировать параметры τ_k следующим образом:

$$\tau_k \leftarrow \bar{\tau}_k = \begin{cases} \tau_k, & \text{если } 0 \leq \tau_k \leq 1, \\ 1, & \text{если иначе.} \end{cases} \quad (34)$$

Для СЗР_s (фиг. 3б) мы используем HLLC-решение для пристеночной возмущенной области (28), (29). Численный поток в этом случае выражается следующим образом:

$$\mathbf{F}_\sigma = [(\bar{\tau}_2 - \bar{\tau}_1)\mathbf{F}^* + (1 - \bar{\tau}_2)\mathbf{F}_l] \Delta t, \quad (35)$$

где $(\bar{\tau}_k = \tau_k, \text{ если } \tau_k \in [0, 1], \text{ и } (\bar{\tau}_k = 1, \text{ в противном случае, } \tau_k = t_k/\Delta t, k = 1, 2, \text{ а } t_1 = -x_s/v_s, t_2 = -x_s/s_l \text{ — моменты времени, в которые поверхность тела и волна пересекают грань ячейки (прямую } x = 0 \text{ на фиг. 3б), соответственно.}$

Векторы $\mathbf{G}_\sigma^\pm$, определяющие вклад компенсационного потока на шаге Δt в ячейке слева (знак “−”) и справа (знак “+”) от грани в уравнении (19), вычисляются на соответствующей грани по решению СЗР_g:

$$\begin{aligned} \mathbf{G}_\sigma^+ &= [(\mathbf{F}_{II}^* - v_s \mathbf{Q}_{II}^*)\bar{\tau}_{II} + (\mathbf{F}_{IV}^* - v_s \mathbf{Q}_{IV}^*)(\bar{\tau}_3 - \bar{\tau}_{II})] \Delta t, \\ \mathbf{G}_\sigma^- &= (\mathbf{F}_{IV}^* - v_s \mathbf{Q}_{IV}^*)(1 - \bar{\tau}_3) \Delta t. \end{aligned} \quad (36)$$

Если грань приходится на тело, то потоки определяются по решению СЗР_s:

$$\begin{aligned} \mathbf{G}_\sigma^- &= (\mathbf{F}^* - v_s \mathbf{Q}^*)\bar{\tau}_1 \Delta t - \mathbf{G}_\sigma = -(\mathbf{F}^* - v_s \mathbf{Q}^*)(1 - \bar{\tau}_1) \Delta t, \\ \mathbf{G}_\sigma^+ &= (\mathbf{F}^* - v_s \mathbf{Q}^*)(1 - \bar{\tau}_1) \Delta t. \end{aligned} \quad (37)$$

Вычитание полного потока $\mathbf{G}_\sigma$ из выражения для левостороннего потока $\mathbf{G}_\sigma^-$ в (37) исключает двойной учет компенсационного потока, поскольку этот поток считается на обеих гранях пересекаемой ячейки.

Решение правосторонней задачи пристеночной задачи Римана (тело слева, газ справа) сводится к решению рассмотренной выше левосторонней задачи заменой знака у координаты и скорости на противоположный. Практически это сводится к следующему алгоритму.

Пусть надо определить потоки $\mathbf{F}_\sigma$ и $\mathbf{G}_\sigma^\pm$ для случая правостороннего паттерна, т.е., на решении СЗР_g($r, x_s, v_s, \mathbf{Q}_l, \mathbf{Q}_r$). Для этого мы рассмотрим вспомогательную задачу СЗР_g($l, -x_s, -v_s, \bar{\mathbf{Q}}_l, \bar{\mathbf{Q}}_r$), где $\bar{\mathbf{Q}}_l = \mathbf{I}^* \mathbf{Q}_r$, $\bar{\mathbf{Q}}_r = \mathbf{I}^* \mathbf{Q}_l$, $\mathbf{I}^* = \text{diag}(1, 1, -1, 1)$ — диагональная матрица, и вычислим для нее соответствующие потоки $\bar{\mathbf{F}}_\sigma$ и $\bar{\mathbf{G}}_\sigma^\pm$. После этого искомые потоки находятся простым обратным преобразованием: $\mathbf{F}_\sigma = -\mathbf{I}^* \bar{\mathbf{F}}_\sigma$, $\mathbf{G}_\sigma^\pm = -\mathbf{I}^* \bar{\mathbf{G}}_\sigma^\pm$.

5. ЧИСЛЕННЫЕ РЕЗУЛЬТАТЫ

Рассмотренный в предыдущих разделах метод численного решения сопряженной задачи динамики газа и твердого тела на несогласованной сетке позволяет удерживать разрешение границы тела в одномерной модели с точностью до одной расчетной ячейки при надлежащем выборе временного шага. Последний определяется стандартным курантовским условием устойчивости. Этот шаг гарантирует также, что граница тела не проходит за шаг по времени расстояние более одной ячейки.

Все расчеты, представленные ниже, проводились на стационарной равномерной сетке сквозным образом по схеме (19). Геометрия тела задавалась начальным распределением объемной доли газа в ячейках сетки. Скорость тела являлась заданной величиной при форсированном движении. При свободном движении она определялась из решения уравнения движения тела под действием давления со стороны газа, которое определяется при вычислении компенсационных потоков. Число Куранта равнялось 0.9.

Тест 1. Постановка первой задачи аналогична задаче Сода [25] с той лишь разницей, что на месте разрыва мы ставим узкий поршень, который форсированно движется с постоянной заданной скоростью, равной скорости контактного разрыва в оригинальной постановке, $v_s = 0.92745$. Начальные значения приведены в табл. 1.

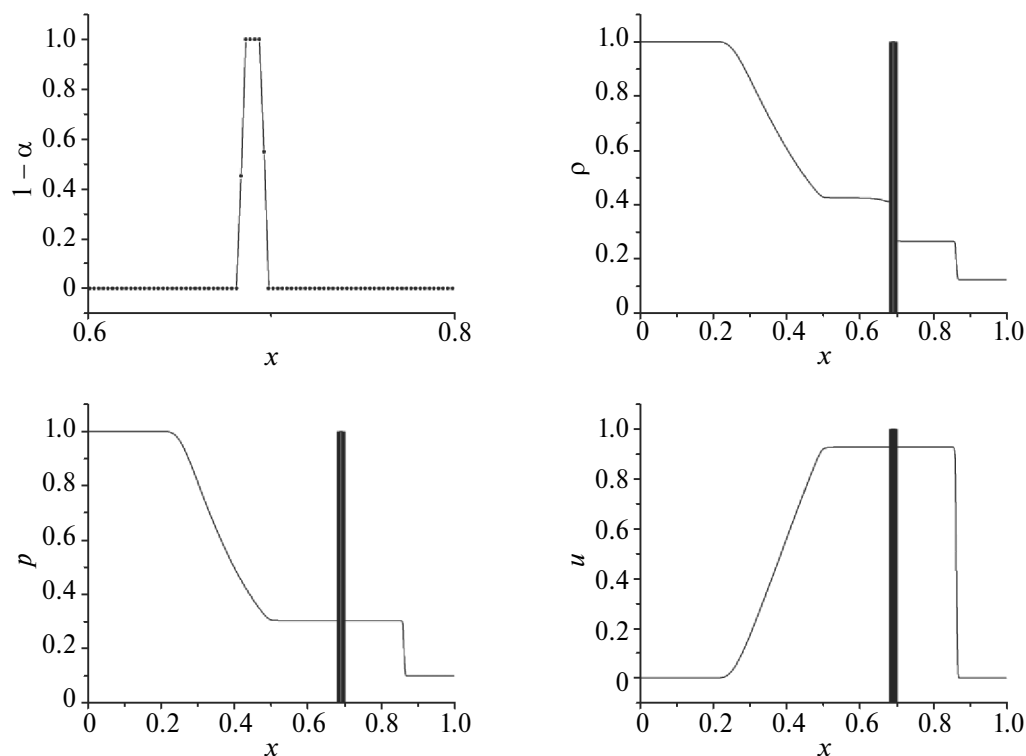
Таблица 1. Начальные значения для теста 1.

	$0 < x < 0.5$	$0.5 < x < 0.505$	$0.505 < x < 1$
α	1.0	0.0	1.0
ρ	1.0	0.0	0.125
u	0.0	0.0	0.0
p	1.0	0.0	0.1

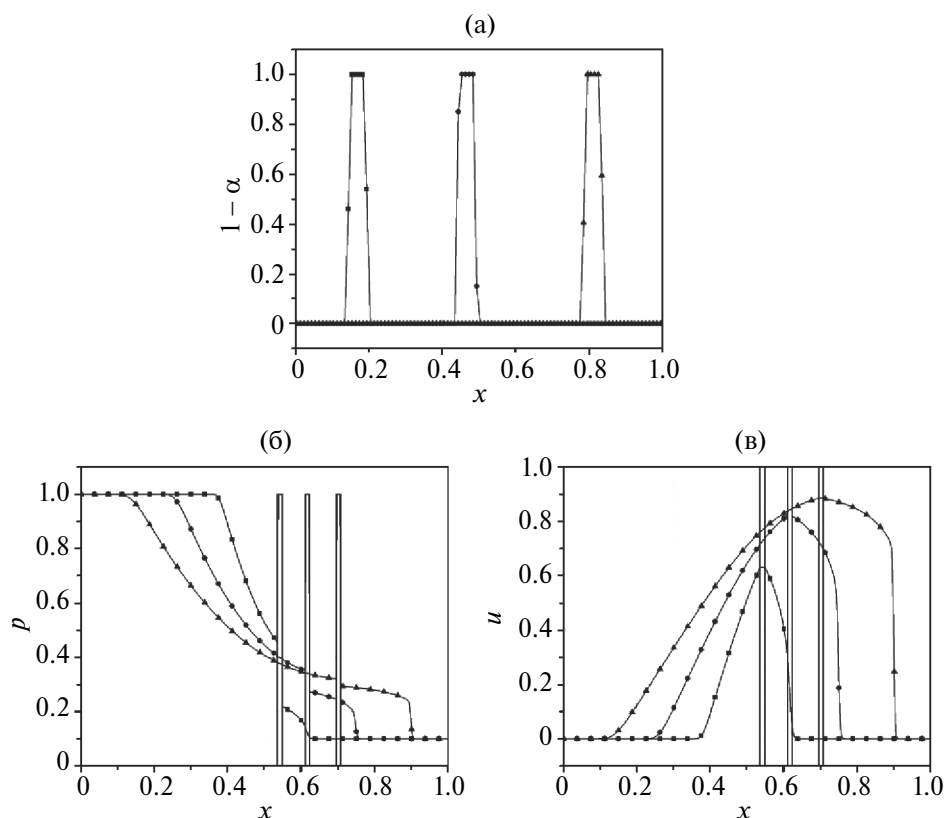
Численные решения приведены на графиках, изображенных на фиг. 4. Здесь тело представляет собой тонкий плоский поршень. Его положение показано на графиках двумя вертикальными линиями. Распределение объемной доли тела демонстрирует бездиссипативное разрешение границы; левая и правая границы поршня разрешаются с точностью до одной ячейки. При этом распределение газодинамических параметров, положение и амплитуды волны разрежения и ударной волны, возникающих при движении поршня, в точности совпадают с соответствующими решениями оригинальной задачи Сода.

Тест 2. В этом тесте рассматривается свободное движение тела под действием силы давления. Постановка задачи такая же, как в предыдущем тесте (Таблица 1). Масса поршня равна 1.5, его начальная скорость равна нулю.

Под действием перепада давления поршень начинает перемещаться справа налево. При этом в газе образуются нарастающие по времени волны сжатия и разрежения. Этот процесс показан на фиг. 5, где приведены



Фиг. 4. Численное решение для задачи Сода, форсированное движение: распределения объемной доли и плотности (верхний ряд), давления и скорости (нижний ряд).



Фиг. 5. Численное решение для задачи Сода, свободное движение: а — для распределения объемной доли и плотности, б, в — для давления и скорости на три последовательных момента времени.

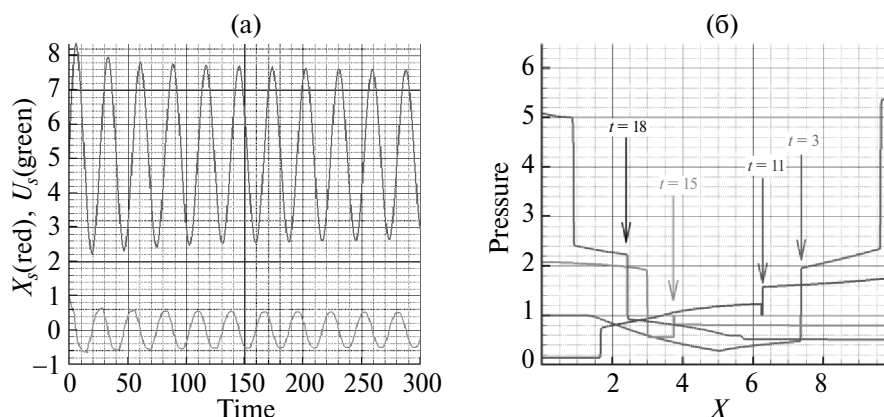
численные распределения объемной доли тела, давления и скорости на три последовательных момента времени. Граница поршня показана вертикальными линиями. Со временем волна сжатия опрокидывается, формируя фронт ударной волны. Скорость поршня увеличивается, амплитуда ударной волны растет, перепад давления на поршне уменьшается. Асимптотически этот процесс выходит на автомодельные распределения предыдущей задачи Сода. Здесь также отметим бездиссипативное разрешение границы тела (фиг. 5).

Тест 3. Этот тест представляет собой газодинамический осциллятор. В нем массивный поршень совершает свободное колебательное движение за счет меняющегося со временем перепада давления на его границах. Начальное распределение параметров в газе постоянное: $\rho = 1$, $u = 0$, $p = 1$. Газ занимает область $0 \leq x \leq 10$, начальное положение поршня — $5 \leq x \leq 5.1$. Масса поршня равна 1.5, начальная скорость $v_s = 1$. Левая и правая границы области ($x = 0$ и $x = 1$) соответствуют жестким стенкам с граничным условием $u = 0$.

В этом тесте при движении поршня также происходит формирование волн сжатия и разрежения, которые теперь при достижении границ области отражаются от стенок и повторно взаимодействуют с поршнем, замедляя его скорость до нуля и разворачивая направление движения на противоположное. Этот процесс иллюстрируется на фиг. 6б, где показаны распределения давления на 4 последовательных момента времени. В результате со временем возникает колебательное движение поршня с амплитудой примерно 2.5. Этот процесс показан на фиг. 6а графиками центра поршня (красный) и скорости (зеленый) в зависимости от времени. Отметим, что рассматриваемый процесс становится строго периодическим и не должен затухать со временем, так как какие-либо диссипативные процессы в рассматриваемой модели отсутствуют. Этот факт хорошо подтверждается приведенными численными результатами, благодаря бездиссипативному разрешению границ тела.

6. ЗАКЛЮЧЕНИЕ

Предложен новый подход к решению задач течения сжимаемой жидкости около стационарных и движущихся тел на основе цифрового представления геометрии и осредненных уравнений газовой динамики, действующих во всем пространстве, включая область тела. Методика использует математические модели многофазных сред, где распределение фаз задается функцией Хевисайда, а межфазный интерфейс моделируется в рамках метода диффузной границы. Это позволило продолжить область решения на все пространство и фактически



Фиг. 6. Газодинамический осциллятор: а — для координаты центра поршня (красный) и скорости (зеленый) в зависимости от времени, б — для распределения давления в газе на четыре последовательных момента времени.

свести краевую задачу в ограниченной области к начальной задаче Коши во всем пространстве. Движение тела при этом полностью определяется решением уравнения переноса для скалярного поля функции порядка. Вычислительная модель строится на простейших декартовых сетках. Для решения осредненных уравнений в рамках одномерной модели рассмотрен численный метод Годунова. При этом в пересекаемых ячейках вводится разрывное восполнение решения, что приводит к рассмотрению составной задачи Римана, описывающей распад начального разрыва при наличии ограничивающей стенки. Построено приближенное решение составной задачи Римана на основе гибридного HLL-HLLC-метода. Показано, что аппроксимация численного потока на решении составной задачи Римана обеспечивает, во-первых, перенос функции порядка без численной диссипации, а во-вторых, устойчивость разностной схемы при стандартном курантовском ограничении на шаг по времени.

СПИСОК ЛИТЕРАТУРЫ

1. Годунов С. К., Забродин А. В., Иванов М. Я., Крайко А. Н., Прокопов Г. П. Численное решение многомерных задач газовой динамики. М.: Издательство «Наука», 1976. 400 с.
2. Rider W. J., Kothar D. B. Reconstructing volume tracking // J. Comput. Phys. 1998. V. 141. P. 112–152.
3. Hirt C. W., Nichols B. D. Volume of fluid (VOF) method for dynamics of free boundaries // J. Comput. Phys. 1981. V. 39. P. 201–225.
4. Osher S., Sethian J. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton–Jacobi formulations // J. Comput. Phys. 1988. V. 79. P. 12–49.
5. Mulder W., Osher S., Sethian J. Computing interface motion in compressible gas dynamics // J. Comput. Phys. 1992. V. 100. P. 209–228.
6. Mittal R., Iaccarino G. Immersed boundary methods // Annual Review of Fluid Mechanics. 2005. V. 37. P. 239–261.
7. Sambasivan S. K., Udaykumar H. S. Ghost fluid method for strong shock interactions. Part 2: Immersed solid boundaries // AIAA J. 2009. V. 47. P. 2923–2937.
8. Baer M., Nunziato J. A two-phase mixture theory for the deflagration-to-detonation transition (DDT) in reactive granular materials // J. Multiph Flow. 1986. V. 12. P. 861–89.
9. Favrie N., Gavriluk S., Saurel R. Solid-fluid diffuse interface model in cases of extreme deformations // J. Comput. Phys. 2009. V. 228. P. 6037–6077.
10. Ndanou S., Favrie N., Gavriluk S. Multi-solid and multi-fluid diffuse interface model: applications to dynamic fracture and fragmentation // J. Comput. Phys. 2015. V. 295. P. 523–555.
11. Favrie N., Gavriluk S. Diffuse interface model for compressible fluid—compressible elastic-plastic solid interaction // J. Comput. Phys. 2012. V. 231. P. 2695–2723.

12. *Kemm F., Gaburro E., Thein F., Dumbser M.* A simple diffuse interface approach for compressible flows around moving solids of arbitrary shape based on a reduced Baer–Nunziato model. *Computers and Fluids*. 2020. V. 204. P. 104–536
13. *Dal Maso G., LeFloch P., Murat F.* Definition and weak stability of nonconservative products // *J. Math. Pures Appl.* 1995. V. 74. P. 483–548.
14. *Peskin C.* Flow patterns around heart valves: a numerical method // *J. Comput. Phys.* 1972. V. 10. P. 252–271.
15. *Angot P., Bruneau C. H., Fabrie P.* A penalization method to take into account obstacles in incompressible viscous flows // *Numer. Math.* 1999. V. 81. P. 497–520.
16. *Абалакин И. В., Васильев О. В., Жданова Н. С., Козубская Т. К.* Метод характеристических штрафных функций для численного моделирования сжимаемых течений на неструктурированных сетках // *Ж. вычисл. матем. и матем. физ.* 2021. Т. 61. № 8. С. 1336–1352.
17. *Меньшов И. С., Корнев М. А.* Метод свободной границы для численного решения уравнений газовой динамики в областях с изменяющейся геометрией // *Матем. моделирование*. 2014. Т. 26. № 5. С. 99–112.
18. *Меньшов И. С., Павлухин П. В.* Эффективный параллельный метод сквозного счета задач аэродинамики на несвязных декартовых сетках // *Ж. вычисл. матем. и матем. физ.* 2016. Т. 56. № 9. С. 1677–1691.
19. *Годунов С. К.* Разностный метод численного расчета разрывных решений уравнений гидродинамики // *Матем. сборник*. 1957. 47:3. С. 271–306.
20. *Igor Menshov, Pavel Zakharov.* On the composite Riemann problem for multi-material fluid flows // *Internat. Journal for Numerical Methods in Fluids*. 2014. V. 76(2). P. 109–127.
21. *Menshov I. S., Serezhkin I. S.* Numerical Model of Multiphase Flows Based on Sub-Cell Resolution of Fluid Interfaces // *Comput. Math. and Math. Phys.* 2022. V. 62. No. 10. P. 1723–1742.
22. *Chao Zhang, Igor Menshov.* Using the composite Riemann problem solution for capturing interfaces in compressible two-phase flows // *Appl. Math. and Comput.* 2019. V. 363.
23. Toro E. *Riemann solvers and numerical methods for fluid dynamics*. Springer. 2009. P. 719.
24. *Batten P., Clarke N., Lambert C., Causon D. M.* On the choice of wavespeeds for the HLLC Riemann solver. 1997. V. 18. No 6. P. 1553–1570.
25. *Sod G. A.* A Survey of Several Finite Difference Methods for Systems of Nonlinear Hyperbolic Conservation Laws // *J. Comput. Phys.* 1978 V. 27 (1): P. 1–31.

FREE BOUNDARY METHOD FOR COUPLED PROBLEMS OF GAS AND SOLID DYNAMICS

I.S. Menshov*

Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences, Miusskaya sq., 4, Moscow, 125047 Russia

**e-mail: menshov@kiam.ru*

Received 02 April, 2024

Revised 02 April, 2024

Accepted 05 May, 2024

Abstract. This paper presents a new approach to numerical modeling of gas flow around stationary and moving rigid bodies, allowing for the use of Eulerian grids that are not tied to the geometry of the body. The bodies are assumed to be absolutely rigid and undeformable, with their elastic properties disregarded. The gas is inviscid and non-heat-conducting, described by compressible fluid equations. The proposed approach is based on averaging the equations of the original model over a small spatial filter. This results in a system of averaged equations that includes an additional quantity — the solid volume fraction parameter — whose spatial distribution digitally represents the geometry of the body (analogous to an order function). This system of equations operates across the entire space. Under this approach, the standard boundary-value problem within the gas region is effectively reduced to a Cauchy problem over the entire space. For a one-dimensional model, the numerical solution of the averaged equations is considered using Godunov's method. In intersected cells, a discontinuous solution is introduced, leading to a compound Riemann problem that describes the decay of the initial discontinuity in the presence of a confining wall. It is shown that the approximation of the numerical flux for the compound Riemann problem solution ensures transport of the order function without numerical dissipation.

Keywords: gas dynamics near a moving solid surface, averaged Euler equations, compound Riemann problem, Godunov's numerical method.

УДК 519.72

О СПЕКТРАЛЬНЫХ ПОРТРЕТАХ МАТРИЦ ИНЦИДЕНТНОСТИ ГРАФОВ БЛИЖАЙШИХ СОСЕДЕЙ¹⁾

© 2024 г. А. А. Кислицын^{1,*}¹125047 Москва, Миусская пл., 4, ИПМ им. Келдыша РАН, Россия

*e-mail: alexey.kislitsyn@gmail.com

Поступила в редакцию 02.04.2024 г.

Переработанный вариант 02.04.2024 г.

Принята к публикации 05.05.2024 г.

В работе исследуется возможность применения метода С.К. Годунова построения спектрального портрета матрицы для оценки ранга матриц специального вида, которые появляются в ряде приложений, таких, как анализ структур графов ближайших соседей, теория конечных автоматов и оценка спектра разреженных матриц. Описан вычислительный алгоритм генерации ансамбля матриц случайных расстояний и связанных с ним графов ближайших соседей. На основе вычислительного эксперимента оценены параметры распределения степеней вершин случайных графов ближайших соседей. Оценки удается получить вследствие того, что указанное распределение не зависит от функции распределения случайных расстояний и является многомерным нормальным распределением. Доказано, что ранг матрицы инцидентности графа ближайших соседей равен суммарному числу вершин со степенями 0 и 1 по входящим ребрам, и получено распределение ранга такой матрицы. Показано, что в данной задаче определения ранга матрицы весьма эффективным является также метод, основанный на анализе распределения степеней вершин. Библ. 17. Фиг. 4.

Ключевые слова: граф ближайших соседей, распределение степеней вершин, спектральный портрет, распределение ранга матрицы.

DOI: 10.31857/S0044466924080189, EDN: XZTUET

1. ВВЕДЕНИЕ

Во многих задачах вычислительной линейной алгебры необходимо иметь оценки того, в каких пределах лежит возмущение спектра матрицы при возмущении ее элементов. Это связано с тем, что на практике при построении разностных схем для решения уравнений математической физики появляются матрицы коэффициентов, характеризующих свойства системы в вычислительной ячейке. Ими могут быть, например, коэффициенты теплопроводности или вязкости, получаемые в ходе физического эксперимента, и потому известные с некоторой ошибкой. В результате может оказаться, что свойства численного решения неожиданно сильно зависят от точности входных данных, которые можно трактовать как параметры со случайным возмущением. Возникающие при этом проблемы были описаны в известных монографиях [1, 2].

В дальнейшем выяснилось, что исследование устойчивости спектра матрицы при случайном возмущении ее элементов требуется проводить также в ряде статистических задач для оценки достаточности длины выборки, корректности модели распознавания образов, оценки ранга случайной матрицы. Примеры такого использования спектральных портретов стохастических матриц содержатся в книге [3].

Современное состояние исследований представлено в монографии С.К. Годунова [4], которой мы и будем следовать в части основных определений и методологии.

Представляет интерес использование спектральных портретов матриц, возникающих при анализе случайных разреженных матриц больших размеров. Такие матрицы возникают, в частности, как матрицы инцидентности графов ближайших соседей при моделировании структурных связей тех или иных множеств.

Анализ графов ближайших соседей (далее ГБС) представляет собой один из широко используемых методов кластеризации данных [5]. Графом *ближайших соседей* называется ориентированный граф, определенный для множества E точек в метрическом пространстве. Каждая точка данного множества является вершиной графа. Направленное ребро из точки A в точку B проводится для той точки B множества, расстояние до которой от

¹⁾ Работа выполнена при финансовой поддержке РНФ (код проекта 23-71-10055).

точки A минимально. Обозначим через $\rho_A(X)$ расстояние от данной точки $A \in E$ до некоторой другой точки $X \in E$. Тогда точка B , к которой проводится ребро из точки A , определяется условием

$$B = \arg \min_{\substack{X \in E \\ X \neq A}} \rho_A(X).$$

В большинстве случаев целью анализа является установление структурных связей между объектами, которым были поставлены в соответствие наборы чисел и введена мера близости в некотором подходящем пространстве. Тогда практическую важность приобретает вопрос о вариабельности установленных эмпирических закономерностей, поскольку разные выборки однотипных объектов будут порождать, вообще говоря, разные графы связей. В этом смысле граф связей является случайным, поскольку случаен сам отбор объектов в анализируемую группу.

Теория случайных графов развита очень подробно в части случайных графов, которые обычно трактуются в смысле Эрдёша–Реньи [6], когда вершины соединяются между собой по заданному вероятностному закону. Такие графы активно изучались в контексте решения многих прикладных задач (см., например, [7–9]). Асимптотические свойства случайных графов также анализировались в рамках конкретных вероятностных моделей применительно к различным онтологиям, социальным и транспортным сетям и аналогичным объектам (см. работы [10–13]). Однако в области статистического анализа графов конструктивных моделей и результатов существенно меньше. Это связано с тем, что выборка графов сама по себе не имеет места, поскольку граф визуализирует свойство изучаемой группы объектов. Поэтому понятие выборки адресуется именно группе объектов, которые должны быть однородными по изучаемому свойству в статистическом смысле.

Основная проблема состоит в том, что анализ одного конкретного графа не позволяет оценить вероятность его реализации как элемента из определенного множества случайных графов, поскольку нужное множество не формализовано. Исключением является граф ближайших соседей, который строится на основе матрицы попарных расстояний между точками. В этом случае функция распределения расстояний известна, что позволяет выяснить, какие графы ей отвечают. В работах [14–16] были получены результаты о статистике структур графов ближайших соседей. В основе вычислительного метода лежит теорема о независимости распределения этих структур от функции распределения расстояний между вершинами графа. Этот факт позволил получить в результате численных экспериментов табличные значения вероятностей, характеризующих ансамбль ГБС: распределение графов по числу несвязных фрагментов, фрагментов по числу вершин и вершин по степеням входящих ребер. Используя эти результаты, можно построить распределение ранга матрицы инцидентности и получить новый критерий независимости выборочных данных.

2. СПЕКТРАЛЬНЫЙ ПОРТРЕТ МАТРИЦЫ

Приведем здесь определения основных понятий, требующихся далее, следуя монографии С.К. Годунова [4].

Число λ принадлежит ε -спектру $\Lambda_\varepsilon(A)$ матрицы A , если существует такая возмущающая ее матрица Δ , что $\|\Delta\| \leq \varepsilon \|A\|$ и $\det(\lambda I - A - \Delta) = 0$, где I — единичная матрица.

Сингулярными числами σ_i матрицы A называются собственные значения эрмитовой матрицы $A^+ A$, где знак плюс сверху означает эрмитово сопряжение. Все эти числа неотрицательны. Занумеруем их в порядке возрастания, так что для матрицы порядка N минимальное сингулярное число есть σ_1 , а максимальное σ_N .

Числом обусловленности невырожденной матрицы A называется величина

$$\mu(A) = \frac{\sigma_N(A)}{\sigma_1(A)} = \|A\| \cdot \|A^{-1}\|. \quad (2.1)$$

Имеет место оценка для числа обусловленности возмущенной матрицы:

$$|\mu(A + \Delta) - \mu(A)| \leq \frac{2\mu(A)}{1 - \mu(A) \frac{\|\Delta\|}{\|A\|}} \cdot \frac{\|\Delta\|}{\|A\|}. \quad (2.2)$$

Эта оценка означает, что при малом возмущении, когда $\|\Delta\| \ll \|A\|/\mu(A)$, число обусловленности возмущенной матрицы также меняется мало.

Резольвентой матрицы A называется матрица

$$R(\lambda) = (\lambda I - A)^{-1}. \quad (2.3)$$

В терминах резольвенты ε -спектр определяется следующим образом: число λ принадлежит ε -спектру $\Lambda_\varepsilon(A)$ матрицы A , если

$$\|R(\lambda)\| \geq \frac{1}{\varepsilon \|A\|}. \quad (2.4)$$

На этой формуле основан практически применяемый численный алгоритм определения областей в комплексной плоскости параметра λ , отвечающих спектральным пятнам в зависимости от величины возмущения ε .

При исследовании расположения точек спектра представляют интерес замкнутые гладкие кривые γ_ε , представляющие изолинии ε -спектра. Контур γ_ε разбивает весь ε -спектр $\Lambda_\varepsilon(A)$ на две части, лежащие внутри и вне его. В этом смысле контур γ_ε осуществляет дихотомию ε -спектра матрицы. Качество дихотомии $\kappa_\gamma(A)$ оценивается нормой квадрата резольвенты (2.3) на данной кривой:

$$\kappa_\gamma(A) = \frac{\|A\|^2}{l_\gamma} \oint_\gamma \|R(\lambda)\|^2 d\lambda. \quad (2.5)$$

Здесь l_γ есть длина контура γ . Величина $\kappa_\gamma(A)$ выбрана как индикатор точности разделения спектра потому, что если на некоторой кривой γ нет точек спектра $\lambda(A)$, то норма резольвенты на такой кривой конечна: $\|R(\lambda)\|_\gamma < \infty$, как и интеграл от нее по этой кривой.

Если внутри области, ограниченной кривой γ_ε , оказалось несколько собственных значений, то с указанной точностью ε их естественно считать совпадающими. Тогда подпространство с базисом из собственного и присоединенных векторов для такого кратного собственного значения будет инвариантным подпространством для оператора A . Проектор Π на это инвариантное подпространство определяется формулой

$$\Pi_\gamma = \frac{1}{2\pi i} \oint_\gamma R(\lambda) d\lambda. \quad (2.6)$$

В частности, если контур γ охватывает все собственные значения, то всякая аналитическая функция от матрицы, будучи определенной на ее спектре, имеет интегральное представление

$$f(A) = \frac{1}{2\pi i} \oint_\gamma f(\lambda) R(\lambda) d\lambda. \quad (2.7)$$

Удобно рассматривать радиальную дихотомию, т.е. дихотомию, задаваемую кривой $\lambda = re^{i\varphi}$ при фиксированном значении r . Тогда параметр дихотомии $\kappa_r(A)$ является нормой эрмитовой матрицы $H_r(A)$, имеющей интегральное представление

$$H_r(A) = \frac{1}{2\pi} \int_0^{2\pi} (A^+ - re^{-i\varphi} I)^{-1} (A - re^{i\varphi} I)^{-1} d\varphi, \quad \kappa_r(A) = \|A\|^2 \cdot \|H_r(A)\|. \quad (2.8)$$

Интеграл в (2.8) сходится только в том случае, если на окружности $\lambda = re^{i\varphi}$ нет собственных значений матрицы A . Это представление используется при численном нахождении ε -спектра матрицы.

3. ОСНОВНЫЕ СВОЙСТВА ГРАФОВ БЛИЖАЙШИХ СОСЕДЕЙ

Мы будем предполагать единственность ближайшего соседа для каждой точки множества. Тогда из определения ГБС следуют его основные структурные свойства: с каждой вершиной связано ровно одно выходящее ребро, число вершин графа равно числу ребер, каждый связный ГБС содержит ровно один цикл. Для дальнейшего существенны следующие утверждения.

Пусть $n_p(N)$ есть число вершин степени p в графе ближайших соседей из N вершин. Тогда

$$\sum_{p=0}^{N-1} n_p(N) = N, \quad \sum_{p=0}^{N-1} p n_p(N) = N. \quad (3.1)$$

Отсюда следует, что число вершин нулевой степени равно

$$n_0 = \sum_{p=2}^{N-1} (p-1) n_p, \quad (3.2)$$

а сумма числа вершин нулевой и первой степени равна

$$n_0 + n_1 = N - \sum_{p=2}^{N-1} n_p. \quad (3.3)$$

Справедливы также следующие утверждения.

Теорема 1 (см. [14]). *Вероятность реализации структуры ГБС не зависит от распределения расстояний между точками.*

Утверждение следует из того, что монотонное преобразование одного стационарного распределения вероятностей в другое не меняет упорядоченности между элементами выборки, т.е. не влияет на структуру ГБС.

Теорема 2 (см. [16]). *Статистика результата генерации ансамбля графов ближайших соседей не зависит от выполнения неравенства треугольника для элементов матрицы, которые могут трактоваться как расстояния, а могут — как независимые случайные числа.*

Доказательство основано на том, что предъявляется алгоритм построения матрицы именно расстояний в некотором метрическом пространстве, после чего, используя теорему 1, делается преобразование получающегося распределения расстояний к равномерному распределению, при котором упорядоченность между элементами матрицы сохраняется, но неравенство треугольника уже не выполняется.

Новый результат, обсуждению которого посвящена настоящая работа, состоит в следующем.

Теорема 3. *Сумма числа нулевых и первых степеней вершин графа ближайших соседей равна рангу соответствующей матрицы инцидентности.*

Доказательство. Рассмотрим ГБС, порожденный некоторой матрицей расстояний $N \times N$. Сопоставим этой матрице расстояний матрицу инцидентности для соседа порядка k . Эта матрица того же порядка $N \times N$, что и исходная матрица случайных расстояний, в которой все элементы строки i , кроме одного, отвечающего соседству вершины в строке i и вершины в столбце j , заменены нулями. Отмеченный же элемент 1_{ij} положен равным единице. Поскольку в силу единственности соседа данного порядка в каждой строке матрицы есть ровно один ненулевой элемент и сумма всех таких элементов равна порядку матрицы N , то необходимым и достаточным условием невырожденности матрицы является в данном случае отсутствие нулевого столбца. Необходимость условия очевидна (для невырожденности необходимо отсутствие нулевого столбца), а достаточность следует из разложения определителя по i -й строке, что сводит вычисление определителя матрицы $N \times N$ к вычислению определителя матрицы $(N-1) \times (N-1)$ с точностью до знака $(-1)^{i+j}$. Совпадающих ненулевых столбцов в матрице инцидентности нет, так как иначе это означало бы, что у некоторой точки более одного ближайшего соседа. Следовательно, ранг матрицы инцидентности равен ее размеру за вычетом числа нулевых столбцов. Нулевые столбцы могут появиться только в том случае, если в каких-то других столбцах число единиц больше одной. Наличие, например двух единиц в столбце j означает, что вершина j является ближайшей для некоторых двух вершин. Следовательно, нулевых столбцов в матрице инцидентности столько, сколько имеется вершин степени 2, 3, и т.д. Поскольку же $\sum_{p=0}^{N-1} n_p = N$, то число ненулевых столбцов равно $N - \sum_{p=2}^{N-1} n_p = n_0 + n_1$. Это и есть порядок максимального отличного от нуля минора. Теорема 3 доказана.

Следствием этого утверждения является то, что распределение ранга матрицы инцидентности равно распределению суммы двух случайных величин — числа вершин со степенями 0 и 1 по входящим ребрам.

4. АЛГОРИТМ ГЕНЕРАЦИИ МАТРИЦ СЛУЧАЙНЫХ РАССТОЯНИЙ

Для численного анализа статистик структур ГБС был разработан алгоритм, основанный на теоремах 1 и 2. При сборе статистик графов первых ближайших соседей были использованы следующие особенности именно ГБС структур.

Поскольку ближайший сосед ровно один, то вместо матрицы смежности достаточно хранить всего одну строку, в k -й ячейке которой записан номер $j(k)$ вершины j , которая является ближайшей для вершины k .

Пересчет числа фрагментов начинается с первой ячейки. В первый фрагмент попадают вершины, связанные цепочкой: $1 \rightarrow j_1(1) \rightarrow j_2(j_1(1)) \rightarrow j_3(j_2(j_1(1))) \rightarrow \dots$. Цепочка продолжается до тех пор, пока в ней не появляется номер вершины, уже отмеченный ранее. После этого найденная последовательность номеров вершин записывается как фрагмент 1. Эти номера вычеркиваются из общей строки. Затем начинается аналогичный процесс для ячейки, номер которой следует за 1 (с учетом вычеркивания). Если на каком-то шаге появляется номер $j(k)$, найденный ранее, то вся эта цепочка присоединяется к соответствующему фрагменту. В результате собираются данные о количестве вершин в каждом фрагменте, что позволяет построить по достаточно большому числу таких экспериментов распределение фрагментов по числу вершин.

Распределение вершин по степеням как для отдельного фрагмента, так и для графа в целом строится по частотам встречаемости номеров вершин, являющихся чьими-то соседями: это количество номеров, встретившихся в строке $j(k)$ 0 раз, 1 раз, 2 раза и т.д. В итоге после реализации K экспериментов по генерации матрицы $N \times N$ случайных равномерно распределенных чисел получаем набор $(n_0^1, \dots, n_0^K)$ вариантов для числа вершин с нулевой степенью, что позволяет построить эмпирическое распределение вероятности количества

таких вершин $f_0(n)$, так что $\sum_{n=0}^N f_0(n) = 1$. Из тех же экспериментов получаем набор $(n_1^1, \dots, n_1^K)$ вариантов для числа вершин с первой степенью и соответствующее распределение $f_1(n)$ и т.д.

С учетом однотипности производимых вычислительных процедур целесообразно

перейти к параллельной обработке данных. Алгоритм для параллельной реализации выглядит следующим образом. На первом шаге задаются инициализирующие параметры. Например: количество реализаций, размер матрицы графа, количество потоков. Далее происходит настройка главного процесса, который будет раздавать задачи всем остальным. Инициализируется набор процессов, на которых непосредственно будут происходить вычисления, и на них запускается реализация основного вычислительного блока. В основном вычислительном блоке происходят следующие операции:

- генерируется стохастическая симметрическая матрица заданных размеров;
- в каждом столбце проводится упорядочение значений по близости к первому элементу столбца, что позволяет для каждого первого элемента столбца получить вектор его соседей, отсортированный по удаленности;
- формируется вектор соседей заданной удаленности;
- по сформированным векторам соседей выстраивается ориентированный граф с использованием библиотеки networkx;
- собирается статистика по количеству связанных компонент и количеству вершин;
- рассчитывается распределение вершин графа по степеням.

На последнем шаге проводится проверка контрольных сумм, общего числа сгенерированных вершин и количества графов.

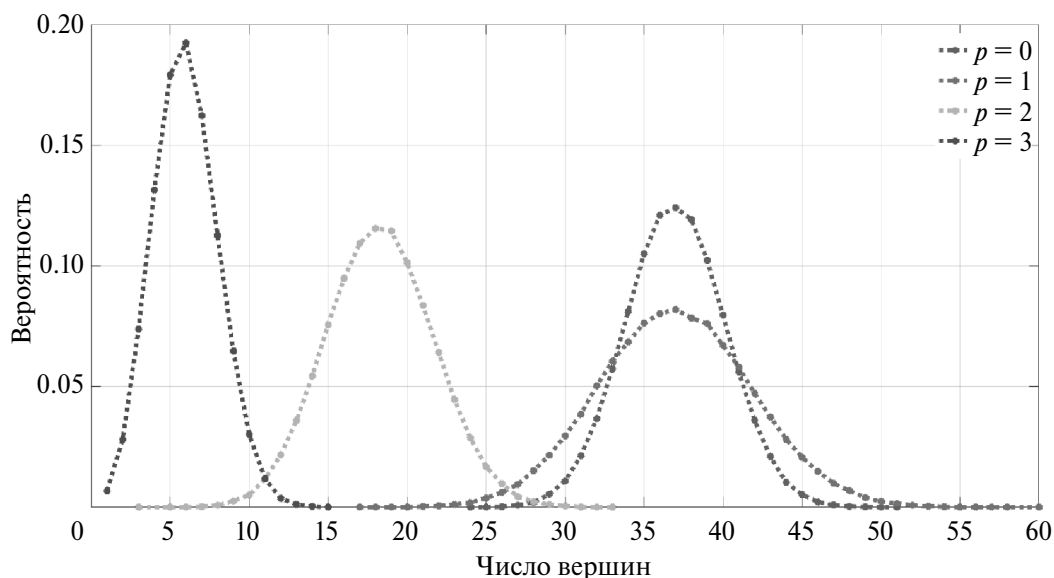
Подчеркнем, что в силу теоремы 2 нет необходимости генерировать расстояния между случайными точками в евклидовом n -мерном пространстве, можно сгенерировать непосредственно матрицу случайных чисел. Это заметно сокращает вычисления, а также позволяет проводить анализ ближайших соседей в различных квазинормах, таких, например, как расстояние Кульбака–Лэйблера.

5. РАСПРЕДЕЛЕНИЕ РАНГА МАТРИЦЫ ИНЦИДЕНТНОСТИ ГРАФОВ БЛИЖАЙШИХ СОСЕДЕЙ

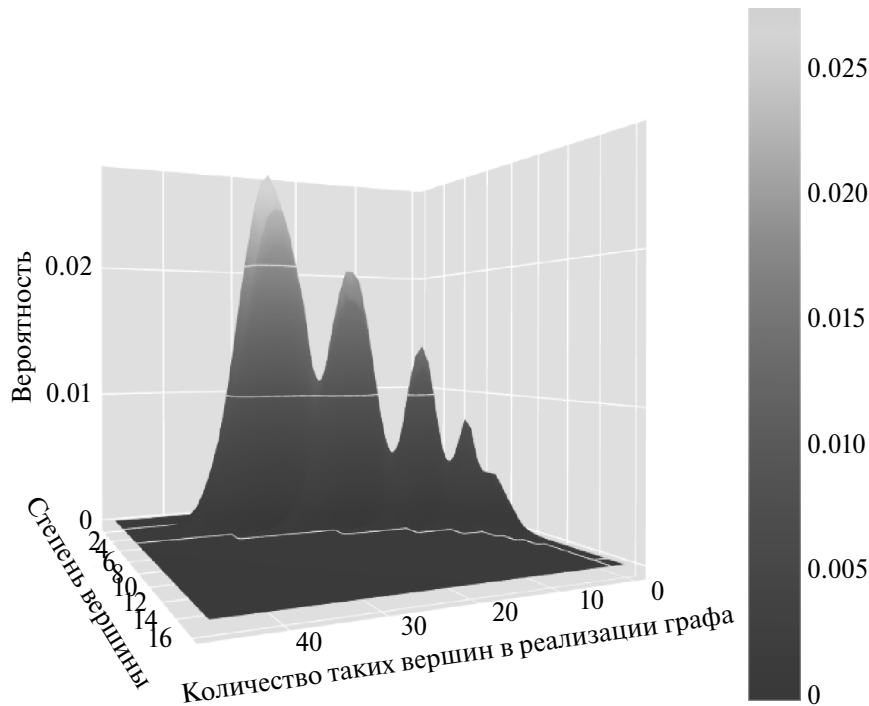
Важная информация о структуре графов ближайших соседей содержится в распределении вершин по степеням. В результате численного моделирования выяснилось, что зависимость этого распределения от числа вершин обладает определенным самоподобием, что позволяет численно найти характеризующие его константы.

На фиг. 1 приведено распределение вершин ближайших соседей по степеням для графа с числом вершин $N = 100$ по результатам 1 млн экспериментов. Показаны вероятности $f_P(n)$ того, что число вершин со степенью P равно n .

Если расположить эти распределения на одной плоскости в виде поверхности $F(p, n)$, которая есть совокупность распределений $f_p(n)$, то получим фазовое пространство возможных реализаций графа. При этом воз-



Фиг. 1. Распределение вершин по степеням для $N = 100$.



Фиг. 2. Распределение числа вершин графа по степеням для $N = 100$.

можно не любая траектория $(p, n(p))$, а лишь такая, для которой выполняются условия (3.1). Соответствующая поверхность $F(p, n)$ приведена на фиг. 2 для $N = 100$.

Совместное распределение степеней вершин ГБС асимптотически нормально. Это следует из того, что при случайном размещении n частиц по N ячейкам вероятность $P(\xi_1 = k_1, \dots, \xi_N = k_N | \xi_1 + \dots + \xi_N = n)$ совместного распределения заполнения ячеек есть полиномиальное распределение схемы размещения (n, N) :

$$P(\xi_1 = k_1, \dots, \xi_N = k_N | \xi_1 + \dots + \xi_N = n) = \frac{n!}{N^n k_1! \dots k_N!}, \quad (5.1)$$

которое при больших N стремится к многомерному нормальному распределению. Важно подчеркнуть, что теоретически параметры этого распределения ранее определены не были. По итогам вычислительного эксперимента выяснилось, что эмпирическое совместное распределение числа вершин по степеням $\{0, 1, \dots, p\}$ может быть приближенно описано распределением вида

$$f(n_0, n_1, \dots, n_p) = \frac{1}{\sqrt{\det C}} \left(\frac{N}{\pi} \right)^{(p+1)/2} \exp \left(-N \sum_{i,j=0}^p C_{ij}^{-1} (n_i/N - \mu_i)(n_j/N - \mu_j) \right), \quad (5.2)$$

где $\det C$ есть определитель ковариационной матрицы, C_{ij}^{-1} — элементы обратной ковариационной матрицы, n_j — число вершин со степенью j , μ_j — средняя доля вершин, имеющих степень j . Поскольку показатель экспоненты имеет порядок N , то в силу быстрого стремления плотности распределения к нулю при малых значениях n_j и при значениях порядка N приближение нормировки по всему p -мерному пространству от $-\infty$ до $+\infty$ по каждой координате допустимо, хотя формально аргументы распределения в (5.2) неотрицательны.

По результатам 1 млн генераций графов были получены следующие значения μ_j , не зависящие от N :

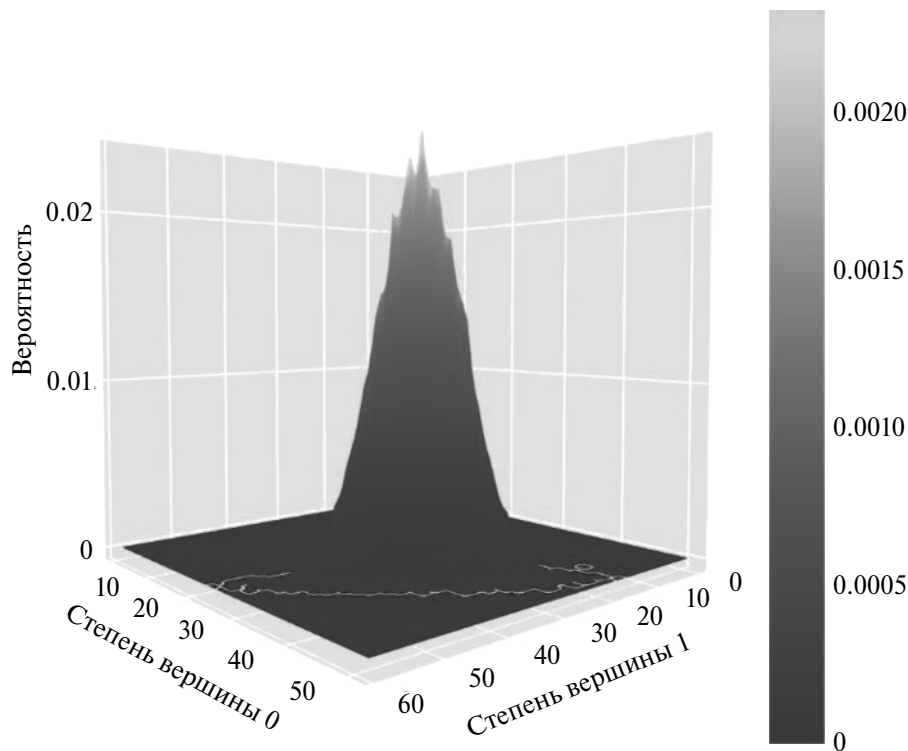
$$\mu_0 = 0.367; \quad \mu_1 = 0.367; \quad \mu_2 = 0.184; \quad \mu_3 = 0.061; \quad \mu_4 = 0.015; \quad \mu_5 = 0.003; \quad \mu_6 = 0.0005. \quad (5.3)$$

Статистическая достоверность предъявленных результатов выше 0.99.

Из совместного распределения (5.2) можно получить практические аппроксимации для плотностей распределения числа вершин по степеням. Например, для нулевой, первой и второй степеней в графе из N вершин имеем формулы:

$$f_p(n) = \sqrt{\frac{A_p}{\pi N}} \exp \left(-\frac{A_p}{N} (n - \mu_p N)^2 \right), \quad (5.4)$$

$$A_0 = 4.9, \mu_0 = 0.37; \quad A_1 = 2.1, \mu_1 = 0.37; \quad A_2 = 4.5, \mu_2 = 0.18.$$



Фиг. 3. Совместное распределение числа вершин со степенями 0 и 1 для $N = 100$.

В силу условия (3.1) распределения числа вершин по степеням коррелированы. Совместное распределение числа вершин степеней 0 и 1 для $N = 100$ показано на фиг. 3.

Этот важный аспект находит свое приложение в анализе матрицы инцидентности. Такие матрицы рассматриваются не только в теории графов, но и в теории конечных автоматов (см. [17]). Обычно сами матрицы для описания конечных автоматов постоянны, а случайным является подаваемый на них сигнал, так что по анализу выходного сигнала делается попытка определения указанной матрицы. В нашей задаче в некотором смысле обратная ситуация — случайной является матрица. Также распределение числа вершин с нулевой и первой степенями по входящим ребрам имеет отношение к спектральной теории случайных графов.

Несмотря на конструктивную простоту доказательства теоремы 3, в качестве следствия из нее ниже получен результат, который следует из (5.4). Приближенными в этой формуле являются коэффициенты A_p и μ_p , найденные в результате численного моделирования. Сама же формула нормального распределения является асимптотически точной по N для распределения степеней вершин.

Носитель совместного распределения для двух случайных величин (числа вершин со степенями 0 и 1) имеет вид эллипса, что соответствует коррелированному нормальному распределению вида:

$$f(n_0, n_1) = \frac{1}{2\pi N \sigma_0 \sigma_1 \sqrt{1 - \rho^2}} \exp \left(\frac{-1}{2N(1 - \rho^2)} \left(\frac{(\Delta n_0)^2}{\sigma_0^2} + \frac{(\Delta n_1)^2}{\sigma_1^2} - 2\rho \frac{\Delta n_0 \Delta n_1}{\sigma_0 \sigma_1} \right) \right), \quad (5.5)$$

где $\Delta n_j = n_j - \mu_j N$, n_j — число вершин со степенью j , μ_j — средняя доля вершин, имеющих степень j , ρ — коэффициент корреляции, а σ_j^2 — дисперсии распределений, показанных на фиг. 3. Эти дисперсии оценены в формулах (5.4). Получим отсюда выражение для распределения $F(m)$ суммы случайных величин $m = n_0 + n_1$. Поскольку совместное распределение (5.5) степеней вершин известно, то имеем

$$F_m = \int_{-\infty}^{\infty} f(x, m - x) dx. \quad (5.6)$$

В силу быстрого убывания распределений (5.4) пределы интегрирования удобно распространить на всю числовую ось. В результате стандартного интегрирования плотностей гауссового типа получаем из (5.5), (5.6):

$$F(m) = \frac{1}{\sigma_0 \sigma_1 A \sqrt{2\pi N}} \exp \left(\frac{-(m - (\mu_0 + \mu_1)N)^2}{2A^2 N \sigma_0^2 \sigma_1^2} \right), \quad A^2 = \frac{1}{\sigma_0^2} + \frac{1}{\sigma_1^2} + \frac{2\rho}{\sigma_0 \sigma_1}. \quad (5.7)$$

Если случайные величины n_0 и n_1 независимы, то дисперсия распределения (5.7) максимальна и равна $\sigma_{\max}^2 = \sigma_0^2 + \sigma_1^2$. Если же эти величины коррелированы, то дисперсия равна

$$\sigma_m^2 = A^2 \sigma_0^2 \sigma_1^2 = \sigma_0^2 + \sigma_1^2 + 2\rho\sigma_0\sigma_1. \quad (5.8)$$

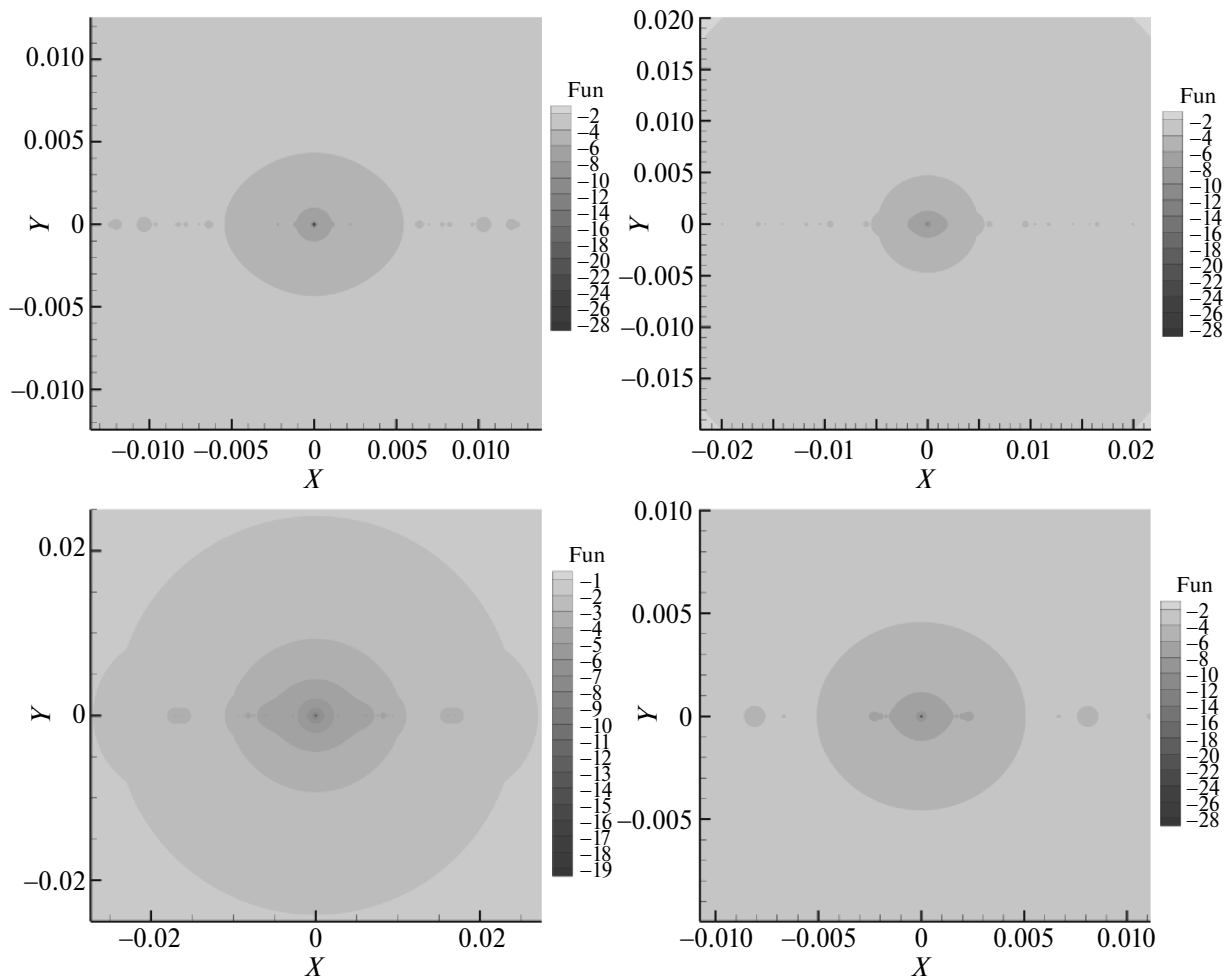
Минимальное значение дисперсии достигается при коэффициенте корреляции, равном $\rho = -1$, оно равно $\sigma_{\min}^2 = (\sigma_0 - \sigma_1)^2$. При равенстве дисперсий и линейной зависимости это распределение переходит в дельта-функцию.

По итогам вычислительного эксперимента были получены следующие числовые значения для параметров распределения (5.7):

$$\sigma_0^2 \approx 0.10; \quad \sigma_1^2 \approx 0.24; \quad \rho \approx -0.92. \quad (5.9)$$

Отсюда и из (5.8) находим, что дисперсия распределения (5.7) оказывается равной величине $\sigma_m^2 \approx 0.05$, что меньше, чем дисперсия распределения нулевых степеней. Следовательно, распределение суммы в данном случае имеет весьма малую дисперсию и может рассматриваться как приближенный закон сохранения средней суммы степеней.

Применим теперь к матрице инцидентности метод построения спектральных портретов, чтобы понять, где лежат точки спектра и с какой точностью требуется вычислять собственные значения. На фиг. 4 приведены фрагменты портретов матриц для случая 100×100 вблизи нулевых собственных значений. Точные ранги этих матриц равны соответственно 69, 64, 62 и 60. Границы областей, имеющих одну цветовую окраску, являются линиями уровня для L_2 -нормы резольвенты $\|(\lambda I - P)^{-1}\|$. Здесь P — матрица C_{ij}^{-1} из (5.2), элементы которой в данном случае определяются из (5.5). На легендах рисунков указаны показатели степени десятки, т.е. точности, с которой проведены расчеты спектра для данной цветовой окраски.



Фиг. 4. Фрагменты спектральных портретов четырех матриц инцидентности ГБС для $N = 100$.

Из вида ϵ -спектра этих матриц можно сделать следующие выводы. Во-первых, и это весьма важно, для более-менее корректного позиционирования нулевого собственного значения нужны вычисления с точностью на уровне $\epsilon = 10^{-10}$ и выше (в смысле величины отрицательной степени). С точностью же 10^{-4} спектр занимает некоторую область около нуля и отдельные точки не различимы. Во-вторых, с увеличением ранга (первые два портрета) спектр распознается более детально. Поэтому, с одной стороны, при численном анализе плохо обусловленных матриц, элементы которых известны точно, требуется достаточно высокая точность вычислений. В то же время эта точность, как показывают спектральные портреты, в практически важном случае не является недостижимой, а находится на вполне реализуемом уровне.

Таким образом, что касается ранга матриц рассматриваемого специального вида, то он может быть определен точно исходя из свойств ближайших соседей. Важно, что спектр матриц инцидентности, определяемый по методу Годунова, дает те же результаты: начиная с определенного уровня точности все ненулевые собственные значения идентифицируются как области спектрального портрета. Это означает, что вычислительные процедуры, проводимые с такими вырожденными матрицами, могут осуществляться корректно при надлежащей точности вычислений.

6. ЗАКЛЮЧЕНИЕ

В представленном исследовании приведено применение метода спектральных портретов плохо обусловленных матриц для статистического анализа структур графов на примере матриц инцидентности графов ближайших соседей. Показано, что точно решаемый пример с определением ранга матрицы инцидентности может быть верно решен и численно с использованием метода Годунова анализа ϵ -спектра матрицы. При этом корректно определяется не только ранг матрицы, что может быть определено и более простыми вычислительными процедурами, но и достаточно точное расположение собственных значений такой матрицы, что представляет самостоятельную ценность.

Результаты работы могут быть использованы для построения критерия независимости выборочных данных, поскольку найдено распределение суммы степеней вершин порядка 0 и 1 найдено с высокой статистической точностью, причем это распределение не зависит от распределения элементов в матрице расстояний, на основе которой проводится дальнейшее исследование.

Также следует указать о возможности применения данного метода, как и подхода с использованием спектральных портретов, к задачам в области случайных конечных автоматов.

СПИСОК ЛИТЕРАТУРЫ

1. Уилкинсон Дж.Х. Алгебраическая проблема собственных значений. М.: Наука, 1970. 564 с.
2. Голуб Дж., Ван Лоун Ч. Матричные вычисления. М.: Мир, 1999. 546 с.
3. Орлов Ю.Н., Осминин К.П. Методы статистического анализа литературных текстов. М.: Эдиториал УРСС/Книжный дом "ЛИБРОКОМ", 2012. 312 с.
4. Годунов С.К. Современные аспекты линейной алгебры. Новосибирск: Научн. книга, 1997. 388 с.
5. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning. Springer, 2001. 764 p.
6. Erdos P., Renyi A. On random graphs I // Publ. Math. Debrecen, 1959. Vol. 6. P. 290–297.
7. Колчин В.Ф. Случайные графы. М.: Физматлит, 2004. 256 с.
8. Bollobas B. Random Graphs. London: Cambridge University Press, 2001. 520 p.
9. Janson S., Luczak T., Rucinski A. Random graphs. New York: Wiley, 2000. 333 p.
10. Райгородский А.М. Модели случайных графов и их применения // Тр. МФТИ, 2010. Т. 2. № 4. С. 130–140.
11. Райгородский А.М. Модели Интернета. Долгопрудный: Издательский Дом Интеллект, 2013. 64 с.
12. Barabasi L.-A., Albert R. Emergence of scaling in random networks // Science. 1999. V. 286. P. 509–512.
13. Leskovec J., Chakrabarti D., Kleinberg J., Faloutsos C., Gharamani Z. Kronecker graphs: an approach to modeling networks // J. Machine Learning Research. 2010. V. 11. P. 985–1042.
14. Кислицын А.А. Исследование статистик графов ближайших соседей // Матем. моделирование. 2022. Т. 34. № 8. С. 110–126.

15. *Kislitsyn A.A., Orlov Yu.N.* Discussion about Properties of First Neighbor Graphs // Lobachevskii Journal of Mathematics 2022. V. 43. № 12. P. 109–118.
16. *Кислицын А.А.* Анализ каталога землетрясений // Матем. моделирование 2023. Т. 35. № 7. С. 63–82.
17. *Мельников С.Ю.* Идентификация конечных автоматов на основе метода многогранников. М. — Ижевск: Институт компьютерных технологий, 2013. 136 с.

ON SPECTRAL PORTRAITS OF NEAREST NEIGHBOR GRAPH INCIDENCE MATRICES

A.A. Kislitsyn*

Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences, Miusskaya sq., 4, Moscow, 125047 Russia

**e-mail: alexey.kislitsyn@gmail.com*

Received 02 April, 2024

Revised 02 April, 2024

Accepted 05 May, 2024

Abstract. This paper explores the possibility of applying S.K. Godunov's method of constructing spectral portraits of matrices to estimate the rank of special types of matrices that appear in various applications, such as nearest neighbor graph structure analysis, finite automata theory, and sparse matrix spectrum estimation. A computational algorithm for generating an ensemble of random distance matrices and the associated nearest neighbor graphs is described. Based on computational experiments, parameters of the vertex degree distribution of random nearest neighbor graphs are evaluated. These estimates are feasible because the distribution is independent of the random distance function and follows a multivariate normal distribution. It is proven that the rank of the incidence matrix of a nearest neighbor graph equals the total number of vertices with in-degree 0 and 1, and the rank distribution of such a matrix is derived. It is also shown that, in this context, a method based on analyzing the vertex degree distribution is highly effective for determining the matrix rank.

Keywords: nearest neighbor graph, vertex degree distribution, spectral portrait, matrix rank distribution.