

Том 64, Номер 9

ISSN 0044-4669

Сентябрь 2024



ФИЦ ИУ РАН

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ



НАУКА

— 1727 —

СОДЕРЖАНИЕ

Том 64, номер 9, 2024

К восьмидесятипятилетию со дня рождения академика Юрия Гавриловича Евтушенко 1573

ОБЩИЕ ЧИСЛЕННЫЕ МЕТОДЫ

Отыскание комплекснозначных решений уравнений Брента сведением к нелинейной задаче наименьших квадратов

И. Е. Капорин 1578

О мультиоператорных аппроксимациях 24-го порядка в схемах для уравнений с конвективными членами

А. И. Толстых 1589

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ

Методология быстрого автоматического дифференцирования и управление тепловыми динамическими системами

В. И. Зубов 1604

Синтез регулятора для линейно-квадратичной задачи оптимального управления

А. С. Антипин, Е. В. Хорошилова 1618

Синтез оптимального управления в модели рамсеевского типа

Н. В. Трусов, А. А. Шананин 1635

ОБЫКНОВЕННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

О мультипликативном свойстве определяющих полиномов

С. А. Абрамов 1661

О граничных свойствах конформных отображений

А. П. Солдатов 1667

УРАВНЕНИЯ В ЧАСТНЫХ ПРОИЗВОДНЫХ

О численном гашении колебаний струны с использованием нескольких стационарных актюаторов

И. Е. Михайлов 1680

МАТЕМАТИЧЕСКАЯ ФИЗИКА

Явно-неявная численная схема для задач динамики упруговязкопластических сред с разупрочнением <i>А. В. Шевченко, И. С. Никитин, В. И. Голубев, И. Б. Петров</i>	1689
Решение задач кинетического взаимодействия быстрых групп частиц с использованием аналитических и численных методов <i>В. В. Аристов, И. В. Воронич</i>	1699
Численный расчет влияния процесса миграции метана на результаты сейсмической разведки в зонах вечной мерзлоты <i>Е. К. Гусева, В. И. Голубев, И. Б. Петров</i>	1708
Спутник на эллиптической орбите: о численном обнаружении периодических движений и исследовании их устойчивости <i>А. А. Буров, В. И. Никонов</i>	1718
Численное исследование переходных режимов течения колмогорова в квадратной ячейке <i>А. О. Посудневская, С. В. ФОРТОВА, А. Н. Долуденко, И. В. Колоколов, В. В. Лебедев</i>	1727
Алгоритм адаптации расчетной сетки к полю течения, содержащему головной скачок уплотнения <i>И. В. Воронич, Н. С. Смирнова, В. А. Титарев</i>	1737
Численное исследование ударно-волновых процессов в материалах под воздействием сверхкоротких лазерных импульсов с использованием модели Баера–Нунциато <i>П. А. Чупров, С. В. ФОРТОВА, В. В. Шепелев</i>	1749
О точности решеточных уравнений Больцмана низкого и высокого порядков в приложениях к медленным изотермическим микротечениям <i>О. В. Ильин</i>	1760
Численное моделирование течения полимерного раствора для колмогоровского потока <i>В. В. Денисенко, С. В. ФОРТОВА, В. В. Лебедев, И. В. Колоколов</i>	1771



К ВОСЬМИДЕСЯТИПЯТИЛЕТИЮ СО ДНЯ РОЖДЕНИЯ АКАДЕМИКА ЮРИЯ ГАВРИЛОВИЧА ЕВТУШЕНКО

28 декабря 2023 года исполнилось 85 лет со дня рождения академика РАН Юрия Гавриловича Евтушенко.

Родился Юрий Гаврилович в г. Краснодаре. В начале Великой Отечественной войны его семья эвакуировалась в г. Фрунзе Киргизской ССР. В 1956 г. Юрий Гаврилович с отличием окончил среднюю школу № 6 в г. Фрунзе и в этом же году поступил в Московский физико-технический институт (МФТИ) на аэромеханический факультет. После успешного окончания института в 1962 г., получив диплом инженера-физика, он по 1965 г. обучался в аспирантуре МФТИ.

Трудовая деятельность Ю.Г. Евтушенко началась в Центральном аэрогидродинамическом институте им. профессора Н.Е. Жуковского (ЦАГИ), куда он был направлен по распределению после окончания аспирантуры и где проработал с 1965 г. по 1967 г. вначале в должности старшего инженера, а затем в должности старшего научного сотрудника. Здесь он занимался вопросами, связанными с движением ракет и спутников.

В 1967 г. Юрий Гаврилович перешел на работу в ВЦ АН СССР. Именно с этим институтом оказалась связанной вся его дальнейшая научная жизнь. Здесь, пройдя все научные должности от младшего научного сотрудника до директора института, он продолжает работать до сих пор. В настоящее время Ю.Г. Евтушенко — главный научный сотрудник, научный руководитель отделения 2 ФИЦ ИУ РАН.

В 1966 г. Ю.Г. Евтушенко в МФТИ успешно защитил диссертацию «Асимптотические методы расчета движения искусственного спутника Земли» на соискание ученой степени кандидата физико-математических наук. Научным руководителем работы был выдающийся ученый Н.Н. Моисеев, оказавший большое влияние на научное развитие Юрия Гавриловича. Под влиянием Н.Н. Моисеева Ю.Г. Евтушенко резко поменял тематику своих исследований. Он целиком переключился на изучение численных методов решения различных задач оптимизации, и в 1981 г. в ВЦ АН СССР защитил докторскую диссертацию на тему «Численные методы решения экстремальных задач и их применение в системах оптимизации» по специальности «Вычислительная математика».

Полученные в этой новой области результаты принесли ему мировую известность.

Выполненные Ю.Г. Евтушенко пионерские работы по отысканию глобального экстремума функций многих переменных — яркое тому подтверждение. Им предложен метод неравномерных покрытий, который успешно применяется многими специалистами для решения разнообразных задач в области исследования операций. Эти результаты позволили создать библиотеку программ для параллельной глобальной оптимизации, с успехом примененную, например, при проведении расчетов молекулярных кластеров на высокопроизводительных вычислительных комплексах.

Ю.Г. Евтушенко разработал процедуры построения эпсилон-сети паретовских множеств для задач многокритериальной оптимизации и построил алгоритмы нахождения гарантированных оценок в многошаговых задачах исследования операций. Он доказал новые теоремы об альтернативах, обобщающие классические исследования Фаркаша, Фредгольма, Гейла. На основе этих результатов создан ряд эффективных вычислительных

методов решения линейных систем равенств и неравенств, а также задач линейного программирования большой размерности.

К достижениям Ю.Г. Евтушенко следует отнести и создание ряда новых методов решения задач нелинейного программирования. Им разработан единый подход к систематизации и классификации численных методов нелинейного программирования, что позволило объединить разнообразные методы и создать универсальную вычислительную систему оптимизации (ДИСО). Эта система широко использовалась при решении различных задач народного хозяйства, в том числе задач САПР авиационной и космической техники. Им также разработана общая методология построения человеко-машинных интерактивных систем принятия решений.

Одним из важных направлений научной деятельности Ю.Г. Евтушенко является разработка методов решения задач оптимизации и оптимального управления. Им были разработаны новые методы решения таких задач, использующие редукцию к задачам нелинейного программирования. Ю.Г. Евтушенко сформулировал общий подход к дифференцированию сложных функций, возникающих в многоступенчатых процессах, заданных на графах (БАД-методология). Им получены соотношения, записанные в канонической форме и позволяющие автоматически строить вычислительные схемы нахождения точных производных сложных функций. Данная методология оказалась исключительно плодотворной при решении задач оптимального управления сложными динамическими системами, поведение которых описывается уравнениями с частными производными. Эта методология эффективно использовалась при решении многих разнообразных теоретически сложных прикладных задач, имеющих большое значение для народного хозяйства нашей страны. Так, она была применена при решении актуальных задач из области структурного материаловедения, при решении задачи моделирования и управления процессом сварки материалов, задачи управления процессом кристаллизации металла в литейном деле, при решении ряда обратных коэффициентных задач. Каждая из решенных задач представляет не только самостоятельный интерес, но также вносит вклад в общую теорию оптимального управления сложными системами. Можно твердо сказать, что в настоящее время БАД-методология является необходимым, неотъемлемым элементом численного решения сложных задач оптимального управления, так как для одних задач применение БАД-методологии позволяет повысить эффективность методов построения решения, для других получить решение удалось только с ее помощью. За цикл работ “Методы оптимизации и управления и их приложение к решению практически важных задач” в 2018 г. Ю.Г. Евтушенко (совместно А.Ф. Албу и В.И. Зубовым) был удостоен премии им. Н.Н. Моисеева РАН.

Интересными представляются исследования, проводимые Ю.Г. Евтушенко в последнее время (совместно А.А. Третьяковым) и связанные с развитием теории p -регулярности и аппарата p -фактор анализа. Результаты этих исследований нашли применение в различных областях математики и приложений. Так, исследован вопрос существования решения краевой задачи, непрерывно зависящего от граничных условий (случай задачи с вырождением). Доказано несколько теорем о неявной функции в случае вырождения, что является развитием теории p -регулярности. Рассматривались вопросы применения теории p -регулярности для вырожденных задач оптимизации с ограничениями типа неравенства. Показана возможность сведения этой задачи к задаче с ограничениями типа равенства. Были сформулированы условия оптимальности p -го порядка для вырожденной задачи вариационного исчисления и для задачи оптимального управления. Исследованы вырожденные динамические системы и предложены новые конструкции функций Ляпунова, так называемые p -фактор функции Ляпунова, позволяющие анализировать устойчивость решений данных нелинейных задач. Для решения различного рода нелинейных задач математической физики, таких как вырожденное уравнение Бюргерса, доказано существование решения в окрестности особой точки и зависящего от некоего параметра.

Под руководством Ю.Г. Евтушенко был решен класс практически важных оптимизационных задач, возникающих при проектировании авиационной и космической техники (оптимизация траекторий и конструкций). В 1981 году цикл этих работ был удостоен премии Совета Министров СССР в области науки и техники.

Результаты исследований, выполненных Ю.Г. Евтушенко в области оптимизации, высоко оценены научной общественностью, и в 1990 г. он был избран членом-корреспондентом АН СССР, а в 2006 г. стал действительным членом РАН по Отделению математических наук.

В настоящее время академик Ю.Г. Евтушенко известен мировой научной общественности как крупный ученый в области теории оптимизации, вычислительной математики и информатики. Он является автором ряда монографий и более 200 научных работ. Научные исследования Юрия Гавриловича характеризуются оригинальностью и высоким научным уровнем. Все работы Юрия Гавриловича отличаются сочетанием фундаментальности в подходе к решению задач и тщательностью проработки возникающих математических и вычислительных проблем.

За время своей активной работы Ю.Г. Евтушенко из учеников и единомышленников создал уникальный научный коллектив, признанный ведущей научной школой Российской Федерации по направлению “Методы решения сложных задач оптимизации”. Деятельность школы Ю.Г. Евтушенко поддерживается разными научными грантами: гранты Президента Российской Федерации, гранты Российского фонда фундаментальных

исследований, гранты Президиума и Отделений РАН, гранты Российского научного фонда.

Наряду с активной научной работой Ю.Г. Евтушенко много сил и времени отдает педагогической деятельности, участвуя в подготовке высококвалифицированных специалистов. За эту работу в 1992 г. ему присвоено звание профессора. В настоящее время Юрий Гаврилович является профессором МФТИ, где он несколько лет возглавлял кафедру “Нелинейные процессы и управление”, а в настоящее время руководит лабораторией «Численные методы прикладной структурной оптимизации» МФТИ, является профессором Национального исследовательского университета «Высшая школа экономики», является профессором МГУ им. М.В. Ломоносова, где заведует кафедрой “Исследование операций”. Ю.Г. Евтушенко — председатель диссертационного совета 24.1.224.02 при ФИЦ ИУ РАН. Его ученики работают во многих научных и учебных коллективах России и за рубежом (Австралия, Азербайджан, Бразилия, Великобритания, Иран, США, Швеция, Узбекистан). Хочется отметить, что Юрий Гаврилович — замечательный педагог, к нему тянутся ученики, и он отвечает им взаимностью.

Ю.Г. Евтушенко много внимания уделяет изданию математической литературы. В 1992 году он принял участие в организации международного научного журнала “Optimization Methods and Software”, который несколько лет возглавлял в качестве главного редактора. На протяжении многих лет Ю.Г. Евтушенко является членом редколлегии “Журнала вычислительной математики и математической физики”. Он — член редколлегии международного научного журнала “Informatics”, член редакционного комитета международного научного журнала “Problems of Nonlinear Analysis in Engineering Systems”.

Ю.Г. Евтушенко является членом Европейской академии наук, иностранным членом Национальной академии наук Украины, входит в состав Национального комитета российских математиков, является членом Национального комитета по индустриальной и прикладной математике, членом Международного общества “Математическое программирование”, членом Международного компьютерного общества IEEE, членом Международного фонда ученых и инженеров имени В.Ф. Уткина.

Большая научно-организационная работа, проводимая в течение многих лет Ю.Г. Евтушенко, высоко отмечена государством. Он награжден орденом “Знак Почета”, медалью ордена “За заслуги перед Отечеством II степени”, Благодарностью Президента Российской Федерации, отмечен нагрудным знаком «Почётный наставник» Минобрнауки России. Он является лауреатом премии Совета Министров СССР, лауреатом премии Правительства Российской Федерации, лауреатом премии им. Н.Н. Моисеева РАН.

Юрий Гаврилович — очень чуткий и отзывчивый человек, удивительно простой и доступный в общении, скромный, доброжелательный, объективный. Редколлегия и редакция желают ему крепкого здоровья, жизненной активности и ждут новых интересных научных результатов.

СПИСОК ТРУДОВ Ю.Г. ЕВТУШЕНКО ЗА ПОСЛЕДНИЕ 5 ЛЕТ

1. *Evtushenko Y.G., Tret'yakov A.A., Tyrtysnikov E.E.* New Approach to Farkas' Theorem of the Alternative // *Doklady Mathematics*. 2019. Vol. 99. No 2. P. 208–210.
2. *Evtushenko Y.G., Tret'yakov A.A.* Constructive generalization of classical sufficient second-order optimality conditions // *Doklady Mathematics*. 2019. Vol. 100. No 1. P. 372–373.
3. *Голиков А.И., Евтушенко Ю.Г., Капорин И.Е.* Метод ньютоновского типа для решения систем линейных уравнений и неравенств // *Ж. вычисл. матем. и матем. физ.* 2019. Т. 59. № 12. С. 2086–2101.
4. *Албу А.Ф., Евтушенко Ю.Г., Zubov V.I.* Об одном подходе к определению вариации функционала с особенностями // *Ж. вычисл. матем. и матем. физ.* 2019. Т. 59. № 8. С. 1277–1295.
5. *Albu A.F., Evtushenko Y.G., Zubov V.I.* On Optimization Problem Arising in Computer Simulation of Crystal Structures // *Optimization and Applications. OPTIMA 2019. Communications in Computer and Information Science*. Springer, Cham. 2019. Vol. 1145. P. 115–126.
6. *Evtushenko Y., Golikov A.* Theorems of Alternative and Optimization // *Lecture Notes in Computer Science*. Vol. 12422. Springer, Cham. 2020. P. 86–96.
7. *Albu A.F., Evtushenko Y.G., Zubov V.I.* On Optimization Problem Arising in Computer Simulation of Crystal Structures // *Communications in Computer and Information Science*. Springer, Cham. 2020. Vol. 1145. P. 115–126.
8. *Албу А.Ф., Евтушенко Ю.Г., Zubov V.I.* Опыт использования методологии Быстрого Автоматического Дифференцирования для решения обратных коэффициентных задач // *Ж. вычисл. матем. и матем. физ.* 2020. Т. 60. № 1. С. 18–28.

9. Албу А.Ф., Евтушенко Ю.Г., Zubov B.I. О выборе разностных схем при решении обратных коэффициентных задач // Ж. вычисл. матем. и матем. физ. 2020. Т. 60. № 10. С. 1643–1655.
10. Евтушенко Ю.Г., Zubov B.I., Албу А.Ф. Применение методологии Быстрого Автоматического Дифференцирования к решению обратных коэффициентных задач // Прикладная матем. и фундаментальная информатика. 2020. Т. 7. № 1. С. 4–14.
11. Evtushenko Y.G., Zubov V.I., Albu A.F. On Fast Automatic Differentiation To Solving The Inverse Coefficient Problems // Proc. of the Workshop on Applied Mathematics and Fundamental Computer Science. CEUR Workshop Proc. 2020. V. 2642.
12. Evtushenko Y., Golikov A. Theorems of Alternative and Optimization // Lecture Notes in Computer Science (Olenev N. et. al. (eds)). Optimization and Applications (OPTIMA 2020). 2020. Vol. 12422. Springer, Cham. P. 86–96.
13. Евтушенко Ю.Г., Третьяков А.А. Новый взгляд на некоторые основополагающие результаты в оптимизации // Ж. вычисл.матем. и матем. физ. 2020. Т. 60. № 9. С. 1462–1471.
14. Евтушенко Ю.Г., Третьяков А.А. Локально полиномиальный метод решения систем линейных неравенств // Ж. вычисл.матем. и матем. физ. 2020. Т. 60. № 2. С. 216–220.
15. Brezhneva O., Evtushenko Y.G., Tret'yakov A.A. New perspective on some classical results in analysis and optimization // Optimization Methods and Software. 2020. P. 1–10.
16. Evtushenko Y., Malkova V., Tret'yakov A. P-Regularity Theory and Nonlinear Optimization Problems // Communications in Computer and Information Science (CCIS). 2020. Vol. 1145. P. 237–253. Springer, Cham.
17. Evtushenko Y., Malkova V., Tret'yakov A. P-Regularity Theory: Applications to Optimization // Lecture Notes in Computer Science (Olenev N. et. al. (eds)). Optimization and Applications (OPTIMA 2020). 2020. Vol. 12422. P. 97–109. Springer, Cham.
18. Албу А.Ф., Евтушенко Ю.Г., Zubov B.I. Об одном подходе к численному решению обратной коэффициентной задачи. // Докл. АН. Математика, информатика, процессы управления. 2021. Т. 499. С. 58–62.
19. Албу А.Ф., Евтушенко Ю.Г., Zubov B.I. О некоторых задачах оптимального управления динамическими системами // Сб. основных докладов XXVIII Моисеевских чтений – Международной научно-практической конференции «Моисеев Н.Н. о России в XXI веке: глобальные вызовы, риски и решения». Нижегородский государственный педагогический университет имени Козьмы Минина. 2021. С. 84–93.
20. Евтушенко Ю.Г., Zubov B.I., Албу А.Ф. Оптимальное управление тепловыми процессами с фазовыми переходами. Коллективная монография. М.: МАКС Пресс, 2021. С. 248.
21. Албу А.Ф., Евтушенко Ю.Г., Zubov B.I. Использование оптимизационных методов второго порядка для решения обратной коэффициентной задачи в трехмерной постановке // Труды Института математики и механики УрО РАН. 2021. Т. 21. № 4. С. 19–34.
22. Денисов Д.В., Евтушенко Ю.Г., Третьяков А.А. Некоторые свойства гладких выпуклых функций и метод Ньютона // Докл. АН. 2021. Т. 497. С. 12–17.
23. Евтушенко Ю.Г., Третьяков А.А. Новый класс функций Ляпунова для исследования на устойчивость вырожденных динамических систем. Элементы теории p-регулярности // Докл. АН. 2021. Т. 499. С. 8–12.
24. Evtushenko Y.G., Zubov V.I., Albu A.F. An Algorithm for Computing Gradients of Functionals in Structural Materials Science // Proc. of the Workshop on Applied Mathematics and Fundamental Computer Science. CEUR Workshop Proc. AMFCS 2021. 2021. Vol. 2928.
25. Evtushenko Y., Malkova V., Tret'yakov A. Exit from Singularity. New Optimization Methods and the p-Regularity Theory Applications // Lecture Notes in Computer Science (LNCS). 2021. Vol. 13078. P. 3–19. Springer, Cham.
26. Евтушенко Ю.Г., Zubov B.I., Албу А.Ф. Численное исследование устойчивости алгоритма идентификации коэффициента теплопроводности в трехмерном случае // Прикладная математика и фундаментальная информатика. 2022. Т. 9. № 1. С. 4–16.

27. *Evtushenko Yu.G., Zubov V.I., Albu A.F.* Numerical Study of Stability of an Algorithm for Identifying the Thermal Conductivity in the Three-Dimensional Case // *J. of Mathematical Sciences*. 2022. Vol. 267. No 4. P. 474–482.
28. *Евтушенко Ю.Г., Третьяков А.А.* О сходимости непрерывных аналогов численных методов решения вырожденных оптимизационных задач и систем нелинейных уравнений // *Ж. вычисл. матем. и матем. физ.* 2022. Т. 62. № 10. С. 1632–1638.
29. *Евтушенко Ю.Г., Беднарчук Е., Прусинска А., Третьяков А.А.* Об эквивалентности вырожденных и некорректных задач. Р-фактор метод регуляризации // *Докл. АН. Математика, информатика, процессы управления*. 2022. Т. 506. С. 41–44.
30. *Brezhneva O., Evtushenko Yu., Malkova V., Tret'yakov A.* Degenerate Equality Constrained Optimization Problems and P-Regularity Theory // *Lecture Notes in Computer Science (LNCS)*, N.N. Olenov et al. (Eds.): OPTIMA 2022. Springer, Cham. 2022. Vol. 13781. P. 16–31.
31. *Евтушенко Ю.Г., Медак Б., Третьяков А.А.* Теория р-регулярности и существование непрерывно зависящего от граничных условий решения краевой задачи // *Ж. вычисл.матем. и матем. физ.* 2023. Т. 63. № 6. С. 920–936.
32. *Евтушенко Ю.Г., Третьяков А.А.* Метод локализации фиктивных экстремумов в задаче глобальной оптимизации // *Докл. АН*. 2023. Т. 512. № 1. С. 78–80.
33. *Brezhneva O., Evtushenko Y., Malkova V., Tret'yakov A.* The pth-Order Karush–Kuhn–Tucker Type Optimality Conditions for Nonregular Inequality Constrained Optimization Problems // *Lecture Notes in Computer Science (LNCS) ser.*, 2023. Vol. 14395. P. 14–32. Springer Nature Switzerland AG 2023.

ОТЫСКАНИЕ КОМПЛЕКСНОЗНАЧНЫХ РЕШЕНИЙ УРАВНЕНИЙ БРЕНТА СВЕДЕНИЕМ К НЕЛИНЕЙНОЙ ЗАДАЧЕ НАИМЕНЬШИХ КВАДРАТОВ

© 2024 г. И. Е. Капорин^{1,*}¹ 119333 Москва, ул. Вавилова, 44, ФИЦ ИУ РАН, Россия

*e-mail: igorkaporin@mail.ru

Поступила в редакцию 12.03.2024 г.

Переработанный вариант 12.03.2024 г.

Принята к публикации 31.05.2024 г.

Отыскание нетривиальных решений трилинейных уравнений Брента соответствует построению асимптотически быстрых алгоритмов перемножения матриц является важной, но в общем случае весьма сложной вычислительной задачей. Предлагаются способы параметризации уравнений Брента, основанные на использовании симметрий тензора матричного произведения, которые позволяют многократно уменьшить размерность задачи. Численное решение полученных трилинейных или кубических систем нелинейных уравнений осуществляется посредством сведения к нелинейной задаче наименьших квадратов и применения к ней специально разработанного итерационного метода, не требующего вычисления производных. Найденные решения параметризованных уравнений Брента, как правило, имеют ранг не больший (а иногда и меньший) по сравнению с известными результатами. Так, получен алгоритм перемножения двух матриц 4-го порядка за 48 активных умножений. Библ. 16. Табл. 1.

Ключевые слова: уравнения Брента; быстрое умножение матриц; алгоритм Штрассена; нелинейная задача наименьших квадратов.

DOI: 10.31857/S0044466924090015, EDN: WLIAJA

1. ВВЕДЕНИЕ

Уравнения, введенные Р. Брентом [1] в связи с задачей построения быстрых алгоритмов перемножения матриц, являются трилинейными относительно неизвестных. Формально, эти уравнения представляют собой каноническое разложение ранга r трехмерного $n_3 n_1 \times n_1 n_2 \times n_2 n_3$ -тензора специального вида, см. ниже (1). Точное решение уравнений Брента обеспечивает существование $O(N^\omega)$ -алгоритма перемножения двух $N \times N$ -матриц, где $\omega = 3 \log r / \log(n_1 n_2 n_3)$. Соответствующую базовую конструкцию будем сокращенно обозначать как $(n_1, n_2, n_3; r)$ -алгоритм. В некоторых случаях, частные решения уравнений Брента могут быть получены «на кончике пера», например в случаях, соответствующих алгоритмам Штрассена [2] ($n_1 = n_2 = n_3 = 2$, $r = 7$) или Ладермана [3], ($n_1 = n_2 = n_3 = 3$, $r = 23$) а также связанных с результатами Пана [4] по перемножению двух или трех независимых пар матриц. Однако для отыскания потенциально лучших схем матричного умножения приходится прибегать к попыткам численного решения уравнений Брента увеличенных размеров. Одним из последних достижений в этом направлении является $(3, 3, 6; 40)$ -алгоритм Смирнова [5]. Важным с практической точки зрения является преобразование полученного алгоритма, например с использованием технологии смены базиса [6] с целью понижения числа операций сложения и умножения на константу.

В настоящей работе рассматриваются некоторые известные и предлагаются новые способы специализации структуры решения уравнений Брента, предположительно не повышающие минимальный ранг, при котором решение существует. Частный случай предлагаемой параметризации (см. ниже (23)) был впервые представлен в [7] для вещественных матриц 3-го порядка для использования в качестве тестовой задачи.

С помощью предлагаемого итерационного метода решения нелинейной задачи наименьших квадратов корректность предлагаемых конструкций проверена экспериментально для значительного количества задач малого размера. При этом получены новые улучшения для малых размеров матриц, в частности, представлены алгоритмы типа $(2, 4, 5; 32)$ и $(4, 4, 4; 48)$.

2. ОБЩИЕ УРАВНЕНИЯ БРЕНТА И ИХ РЕДУЦИРОВАННАЯ ФОРМА

Ниже рассмотрим общий случай, когда размеры матриц n_1, n_2, n_3 принимают произвольные значения.

2.1. Стандартные уравнения Брента

В работе Р. Брента [1] была предложена следующая система трilinearных уравнений, где в качестве неизвестных фигурируют коэффициенты $n_1 \times n_3$ -матриц X_t , $n_2 \times n_1$ -матриц Y_t и $n_3 \times n_2$ -матриц Z_t :

$$-\delta(i_2 - j_1)\delta(j_2 - k_1)\delta(k_2 - i_1) + \sum_{t=1}^r (X_t)_{i_2, i_1} (Y_t)_{j_2, j_1} (Z_t)_{k_2, k_1} = 0 \quad (1)$$

при всех

$$1 \leq i_2, j_1 \leq n_1, \quad 1 \leq j_2, k_1 \leq n_2, \quad 1 \leq k_2, i_1 \leq n_3.$$

Любому решению уравнения Брента при $r < n_1 n_2 n_3$ отвечает тройка нетривиальных билинейных алгоритмов типа $(n_1, n_2, n_3; r)$, позволяющих перемножать пары матриц размеров $n_1 \times n_2$ на $n_2 \times n_3$, или $n_2 \times n_3$ на $n_3 \times n_1$, или $n_3 \times n_1$ на $n_1 \times n_2$ с использованием r умножений в каждом из трех случаев. Чтобы в этом убедиться, достаточно домножить (1) на $(A)_{i_1, i_2} (B)_{j_1, j_2}$ и просуммировать по всем i_1, i_2, j_1, j_2 , что приводит к

$$(AB)_{k_2, k_1} = \sum_{t=1}^r \left(\sum_{i_1, i_2} (X_t)_{i_2, i_1} (A)_{i_1, i_2} \right) \left(\sum_{j_1, j_2} (Y_t)_{j_2, j_1} (B)_{j_1, j_2} \right) (Z_t)_{k_2, k_1},$$

и аналогично в остальных двух случаях. Транспонирование дает еще три алгоритма, например, для умножения матриц $n_3 \times n_2$ на $n_2 \times n_1$ и т.д.

При $n \gg 1$ рекурсивное использование соотношений (1) путем подходящего разбиения матриц на блоки приводит к алгоритму перемножения двух $n \times n$ -матриц за $O(n^\omega)$ операций, где

$$\omega = \frac{3 \log r}{\log(n_1 n_2 n_3)}. \quad (2)$$

Следует заметить, что отыскание решений уравнения Брента при значениях r , близких к минимально возможным, является крайне сложной задачей. Уравнение (1) появилось в работе [1] как инструмент построения алгоритмов матричного умножения, потенциально более эффективных, чем метод Штрассена [2]. Однако потребовалось более сорока лет, чтобы наконец в работе [5] появилось первое нетривиальное (в смысле выполнения условия $\omega = \log_{n_1 n_2 n_3} r^3 < \log_2 7$) решение уравнения Брента типа (3,3,6;40) для малых размеров задачи. Указанное решение обеспечивает построение $O(N^{2.7743})$ -алгоритма перемножения двух матриц размеров $N \times N$. Позже в [6] была показана возможность сократить почти на порядок число операций типа сложения и умножения на константу в этом алгоритме.

2.2. Эквивалентное трilinearное тождество

Как заметил В.Я. Пан (см. [4] и цитированные там источники), решение уравнений Брента эквивалентно построению трilinearного тождества вида

$$\text{tr}(ABC) = \sum_{t=1}^r \text{tr}(X_t A) \text{tr}(Y_t B) \text{tr}(Z_t C), \quad (3)$$

где X_t, Y_t, Z_t те же, что и ранее искомые матричные коэффициенты, а A, B, C представляют собой свободные параметры в виде матриц с размерами $n_3 \times n_1$, $n_1 \times n_2$ и $n_2 \times n_3$ соответственно.

Действительно, (3) легко получаются из (1) домножением на $(A)_{i_1, i_2} (B)_{j_1, j_2} (C)_{k_1, k_2}$ и суммированием по всем $i_1, i_2, j_1, j_2, k_1, k_2$. В обратную сторону, чтобы получить уравнения Брента, достаточно подставить в (3)

$$A = e_{i_1} e_{i_2}^T, \quad B = e_{j_1} e_{j_2}^T, \quad C = e_{k_1} e_{k_2}^T, \quad (4)$$

где e_i представляет собой i -й единичный столбцовый вектор соответствующей длины, и использовать соотношения типа $e_{i_2}^T e_{j_1} = \delta(i_2 - j_1)$ и $\text{tr}(X_t e_{i_1} e_{i_2}^T) = (X_t)_{i_2, i_1}$.

2.3. Параметризованные уравнения Брента

Трилинейная система (1) включает $M = (n_1 n_2 n_3)^2$ уравнений и $N = (n_3 n_1 + n_1 n_2 + n_2 n_3)r$ неизвестных. Решение этой задачи либо не существует (при слишком малых r), либо неединственно, причем изолированные решения отсутствуют (т.е. каждая ветвь решений представляет собой непрерывное многообразие). Кроме того, попытки применения численных методов осложняются наличием “приближенных решений”, для которых ранг разложения обычно меньше, чем для точных решений уравнений Брента. При этом для таких “решений” стремление невязки уравнений Брента к нулю сопровождается неограниченным ростом некоторых коэффициентов разложения. Попытки применения к уравнениям Брента общих оптимизационных методов типа описанных в [8] оказываются недостаточно эффективными. Здесь можно упомянуть метод попеременной квадратичной минимизации, на основе которого получен результат [5], однако там были использованы дополнительные эвристики, направленные на отыскание решений с рациональными коэффициентами. Таким образом, по крайней мере для $\min(n_1, n_2, n_3) \geq 3$, отыскание частных решений задачи (1) является исключительно трудной вычислительной проблемой. Поэтому представляет интерес сведение уравнений Брента к задаче с меньшим числом уравнений и неизвестных. Подход, обсуждавшийся в [13] (в более узком контексте $n_1 = n_2 = n_3$), основан на частичном переносе свойств инвариантности левой части (3) на правую часть (3). Это достигается за счет специальной параметризации матричных коэффициентов X_t, Y_t, Z_t .

Определяющим свойством инвариантности функции

$$\tau(A, B, C) = \tau(ABC) \quad (5)$$

относительно преобразования аргументов является тождество

$$\tau(A, B, C) \equiv \tau(M^{-1}AK, K^{-1}BL, \Lambda^{-1}CM), \quad (6)$$

где K, Λ , и M — произвольные невырожденные матрицы порядков n_1, n_2 и n_3 соответственно. Так, можно показать, что из (6) следует (5) с точностью до ненулевого скалярного множителя. Поэтому потребуем, чтобы свойство (6) соблюдалось также и для правой части (3), хотя бы и для ограниченного выбора матриц K, Λ, M :

$$\begin{aligned} \sum_{t=1}^r \operatorname{tr}(X_t A) \operatorname{tr}(Y_t B) \operatorname{tr}(Z_t C) &= \sum_{t=1}^r \operatorname{tr}(X_t M^{-1}AK) \operatorname{tr}(Y_t K^{-1}BL) \operatorname{tr}(Z_t \Lambda^{-1}CM) = \\ &= \sum_{t=1}^r \operatorname{tr}(KX_t M^{-1}A) \operatorname{tr}(\Lambda Y_t K^{-1}B) \operatorname{tr}(MZ_t \Lambda^{-1}C). \end{aligned}$$

Естественным способом удовлетворить это равенство для некоторых троек матриц K, Λ, M , представляются условия

$$KX_t M^{-1} = X_{\sigma(t)}, \quad \Lambda Y_t K^{-1} = Y_{\sigma(t)}, \quad MZ_t \Lambda^{-1} = Z_{\sigma(t)}, \quad t = 1, 2, \dots, r, \quad (7)$$

где $\sigma(\cdot)$ является перестановкой чисел $(1, 2, \dots, r)$. Чтобы согласованно замкнуть предлагаемые уравнения (7), положим

$$\sigma^p(t) = t, \quad t = 1, 2, \dots, r \quad (8)$$

(то есть $\sigma(\cdot)$ является корнем p -й степени из тождественной подстановки) и пусть

$$r = pq.$$

Тогда определим $\sigma(\cdot)$ как прямую сумму q циклических перестановок порядка p , то есть

$$\sigma(t) = t + 1 - p \delta(t \pmod{p}), \quad t = 1, \dots, r.$$

Например, если $r = 15, p = 3$ и $q = 5$, то $\sigma = (2, 3, 1, 5, 6, 4, 8, 9, 7, 11, 12, 10, 14, 15, 13)$. Наконец, потребуем также

$$K^p = I_{n_1}, \quad \Lambda^p = I_{n_2}, \quad M^p = I_{n_3}, \quad (9)$$

то есть выберем матрицы K, Λ, M как корни p -й степени из единичных матриц соответствующих размеров. Отсюда получаем формулы для матричных коэффициентов

$$X_t = K^s X_{t-s} M^{-s}, \quad Y_t = \Lambda^s Y_{t-s} K^{-s}, \quad Z_t = M^s Z_{t-s} \Lambda^{-s},$$

где

$$s = (t - 1) \pmod{p}, \quad t = 1, \dots, r.$$

Учитывая, что введенные выше индексы $t - s$ образуют неубывающую ступенчатую функцию вида $1, \dots, 1, p + 1, \dots, p + 1, \dots, r - p + 1, \dots, r - p + 1$, соответствующая замена обозначений матричных коэффициентов X_t, Y_t, Z_t в виде $X_t^{\text{new}} = X_{(t-1)p+1}^{\text{old}}$ приводит к формуле

$$\text{tr}(ABC) = \sum_{t=1}^q \sum_{s=0}^{p-1} \text{tr}(K^s X_t M^{-s} A) \text{tr}(\Lambda^s Y_t K^{-s} B) \text{tr}(M^s Z_t \Lambda^{-s} C). \quad (10)$$

Полученное при условии (9) тождество (10) дает (при переходе к покомпонентной записи) *параметризованное уравнение Брента*. Уже на этом этапе видно, что в (10) число неизвестных X_t, Y_t, Z_t сократилось в p раз по сравнению с (3).

2.4. Редуцированные уравнения Брента

Если конкретизировать выбор матриц K, Λ, M в виде

$$K = \text{Diag}_{1 \leq m \leq n_1}(\Omega^{m-1}), \quad \Lambda = \text{Diag}_{1 \leq m \leq n_2}(\Omega^{m-1}), \quad M = \text{Diag}_{1 \leq m \leq n_3}(\Omega^{m-1}), \quad (11)$$

где

$$\Omega = \exp(2\pi i/p), \quad i = \sqrt{-1}, \quad (12)$$

то можно сократить примерно в p раз также и число уравнений. Для этого достаточно подставить указанные формулы для K, Λ, M в (10) и перейти к покомпонентным соотношениям типа (1), подставляя (4):

$$\begin{aligned} \delta(i_2 - j_1)\delta(j_2 - k_1)\delta(k_2 - i_1) &= \sum_{t=1}^q \sum_{s=0}^{p-1} (e_{i_2}^T K^s X_t M^{-s} e_{i_1}) (e_{j_2}^T \Lambda^s Y_t K^{-s} e_{j_1}) (e_{k_2}^T M^s Z_t \Lambda^{-s} e_{k_1}) = \\ &= \sum_{t=1}^q \sum_{s=0}^{p-1} \left(\Omega^{(i_2-1)s} (X_t)_{i_2, i_1} \Omega^{(1-i_1)s} \right) \left(\Omega^{(j_2-1)s} (Y_t)_{j_2, j_1} \Omega^{(1-j_1)s} \right) \left(\Omega^{(k_2-1)s} (Z_t)_{k_2, k_1} \Omega^{(1-k_1)s} \right) = \\ &= \sum_{t=1}^q \left(\sum_{s=0}^{p-1} \Omega^{(i_2-i_1+j_2-j_1+k_2-k_1)s} \right) (X_t)_{i_2, i_1} (Y_t)_{j_2, j_1} (Z_t)_{k_2, k_1} = \\ &= p \delta((i_2 - i_1 + j_2 - j_1 + k_2 - k_1) \pmod{p}) \sum_{t=1}^q (X_t)_{i_2, i_1} (Y_t)_{j_2, j_1} (Z_t)_{k_2, k_1}. \end{aligned}$$

Учитывая, что в случае $\delta(i_2 - j_1)\delta(j_2 - k_1)\delta(k_2 - i_1) = 1$ также выполнено и равенство $\delta((i_2 - i_1 + j_2 - j_1 + k_2 - k_1) \pmod{p}) = 1$, получаем, умножая уравнения на p^{-1} , *редуцированные уравнения Брента*:

$$-\frac{1}{p} \delta(i_2 - j_1)\delta(j_2 - k_1)\delta(k_2 - i_1) + \sum_{t=1}^q (X_t)_{i_2, i_1} (Y_t)_{j_2, j_1} (Z_t)_{k_2, k_1} = 0, \quad (13a)$$

$$1 \leq i_2, j_1 \leq n_1, \quad 1 \leq j_2, k_1 \leq n_2, \quad 1 \leq k_2, i_1 \leq n_3, \quad (13б)$$

$$(i_2 - i_1 + j_2 - j_1 + k_2 - k_1) \pmod{p} = 0. \quad (13в)$$

Точному решению этих уравнений отвечает алгоритм перемножения двух $n \times n$ -матриц за $O(n^\omega)$ операций, где

$$\omega = \frac{3 \log pq}{\log(n_1 n_2 n_3)}. \quad (14)$$

Замечание 1. Понятно, что свойство инвариантности правой части (3) в нашем случае ограничено использованием $p - 1$ троек матриц вида

$$(K^m, \Lambda^m, M^m), \quad m = 1, \dots, p - 1.$$

Замечание 2. Построенная задача (13) не эквивалентна каноническому тензорному разложению, а является специальным случаем *частичного* разложения тензора в силу наличия дополнительного условия (13в).

Замечание 3. Можно рассмотреть несколько более общую конструкцию вида

$$K = \text{Diag}_{1 \leq j \leq n_1}(\Omega^{k(j)}), \quad \Lambda = \text{Diag}_{1 \leq j \leq n_2}(\Omega^{l(j)}), \quad M = \text{Diag}_{1 \leq j \leq n_3}(\Omega^{u(j)}),$$

где $\kappa(\cdot), \lambda(\cdot), \mu(\cdot)$ — произвольные целочисленные функции. В этом случае в редуцированных уравнениях Брента условие (13в) следует заменить на

$$(\kappa(i_2) - \mu(i_1) + \lambda(j_2) - \kappa(j_1) + \mu(k_2) - \lambda(k_1)) \pmod{p} = 0.$$

Неизвестно, можно ли на таком пути улучшить результаты, полученные при используемом в настоящей работе простейшем выборе $\kappa(j) = \lambda(j) = \mu(j) = j - 1$ для $j \geq 1$.

2.5. Обобщение редуцированных уравнений Брента

Сопоставление описанной выше конструкции (10) с результатами Штрассена (где $n_1 = n_2 = n_3 = 2$ и $r = 7$) и Ладермана (где $n_1 = n_2 = n_3 = 3$ и $r = 23$) приводит к необходимости ее обобщения на случай, не предполагающий наличия подходящего разложения r на множители.

Например, для $r = 7$ при $n_1 = n_2 = n_3 = 2$, выбор $p = 1$ тривиален, а при $p = 7$ — невозможен (детальное обсуждение нижних границ для q заслуживает отдельного рассмотрения). Однако в этом случае существует решение редуцированных уравнений Брента, отвечающих разложению $r = 8$ и $p = 2$. При этом полученное завышенное значение ранга можно понизить на единицу, если заметить, что одна из пар матричных коэффициентов, например X_1 и Y_1 , образована диагональными матрицами. В этом случае $KX_1M^{-1} = X_1$ и $\Lambda Y_1K^{-1} = Y_1$, так что первая сумма по s сворачивается в один член ранга 1:

$$\sum_{s=0}^1 \text{tr}(K^s X_1 M^{-s} A) \text{tr}(\Lambda^s Y_1 K^{-s} B) \text{tr}(M^s Z_1 \Lambda^{-s} C) = \text{tr}(X_1 A) \text{tr}(Y_1 B) \text{tr}\left(\left(\sum_{s=0}^1 M^s Z_1 \Lambda^{-s}\right) C\right).$$

В общем случае для понижения ранга редуцированных уравнений Брента добавим к уравнениям (13) дополнительные $3q$ линейных матричных уравнений

$$K^{d_t} X_t - X_t M^{d_t} = 0, \quad \Lambda^{d_t} Y_t - Y_t K^{d_t} = 0, \quad M^{d_t} Z_t - Z_t \Lambda^{d_t} = 0, \quad t = 1, \dots, q, \quad (15)$$

где числа d_t образуют неубывающую последовательность,

$$1 \leq d_1 \leq d_2 \leq \dots \leq d_q = p,$$

и каждое из них является делителем p . Тогда, с учетом представления $s = u + d_t v$, где $0 \leq u \leq d_t - 1$ и $0 \leq v \leq \frac{p}{d_t} - 1$, (10) преобразуется к виду

$$\text{tr}(ABC) = \sum_{t=1}^q \frac{p}{d_t} \sum_{u=0}^{d_t-1} \text{tr}(K^u X_t M^{-u} A) \text{tr}(\Lambda^u Y_t K^{-u} B) \text{tr}(M^u Z_t \Lambda^{-u} C). \quad (16)$$

Таким образом, ранг трилинейного разложения уменьшается до величины

$$r = \sum_{t=1}^q d_t. \quad (17)$$

Например, результату Штрассена соответствует разложение $7 = 1 + 2 + 2 + 2$ получаемое при $p = 2$ и $q = 4$, а результату Ладермана отвечает $23 = 1 + 2 + 4 + 4 + 4 + 4 + 4$ (при $p = 4$ и $q = 7$). Далее используется формула

$$r = (q - q_0)p + \sum_{t=1}^{q_0} d_t, \quad (18)$$

где обычно достаточно брать $q_0 \leq 2$.

Замечание 4. Одно из условий (15) является избыточным при используемом выборе диагональных матриц K, Λ, M , то есть выводится из остальных двух. Кроме того, можно сократить число неизвестных для каждого слагаемого при $d_t < p$, так как уравнения (15) фактически лишь определяют структуру разреженности матриц X_t, Y_t, Z_t согласно условиям

$$(X_t)_{i_2, i_1} = 0, \quad i_2 - i_1 \not\equiv 0 \pmod{p/d_t}, \quad t = 1, \dots, q, \quad (19a)$$

$$(Y_t)_{j_2, j_1} = 0, \quad j_2 - j_1 \not\equiv 0 \pmod{p/d_t}, \quad t = 1, \dots, q, \quad (19b)$$

$$(Z_t)_{k_2, k_1} = 0, \quad k_2 - k_1 \not\equiv 0 \pmod{p/d_t}, \quad t = 1, \dots, q. \quad (19v)$$

Отметим также, что для слагаемых с $d_t = p$ условия (15) или (19) выполняются при любых X_t, Y_t, Z_t в силу (9).

Замечание 5. Наложение условий (15) отвечает использованию общего вида перестановки $\sigma(\cdot)$, удовлетворяющей условию $\sigma^p(t) = t$.

3. РЕДУЦИРОВАННЫЕ УРАВНЕНИЯ БРЕНТА ДЛЯ КВАДРАТНЫХ МАТРИЦ

В случае

$$n_1 = n_2 = n_3 = n \quad (20)$$

можно еще в три раза сократить число неизвестных за счет замены матриц Y_t и Z_t на матрицы X_s с подходящими индексами s .

3.1. Использование инвариантности относительно циклической перестановки

Потребуем, чтобы правая часть тождества (3) обладала тем же свойством, что и левая, а именно, была бы инвариантна относительно циклической перестановки матричных аргументов:

$$\text{tr}(ABC) = \text{tr}(BCA) = \text{tr}(CAB). \quad (21)$$

Для этого произведем замену матриц вида $Y_t = X_{\alpha(t)}$ и $Z_t = X_{\alpha^2(t)}$:

$$\text{tr}(ABC) = \sum_{t=1}^r \text{tr}(X_k A) \text{tr}(X_{\alpha(k)} B) \text{tr}(X_{\alpha^2(k)} C), \quad (22)$$

где $\alpha(\cdot)$ является перестановкой $(1, 2, \dots, r)$ такой, что ее для ее третьей степени справедливо $\alpha^3(t) = t$, то есть α представляет собой корень 3-й степени из тождественной перестановки длины r . Эта конструкция была представлена в [13], а ее обобщение — в [7], где впервые использована матричная параметризация уравнений Брента, обсуждаемая в настоящей работе.

При условии (20) упрощается выбор диагональных матриц в виде $K = \Lambda = M$. Тогда параметризованная форма трилинейного тождества приобретает вид

$$\text{tr}(ABC) = \sum_{t=1}^q \sum_{s=0}^{p-1} \text{tr}(\Lambda^s X_t \Lambda^{-s} A) \text{tr}(\Lambda^s X_{\beta(t)} \Lambda^{-s} B) \text{tr}(\Lambda^s X_{\beta^2(t)} \Lambda^{-s} C), \quad (23)$$

где

$$\Lambda = \text{Diag}_{1 \leq j \leq n}(\Omega^{j-1}), \quad \Omega = \exp(2\pi i/p), \quad i = \sqrt{-1}. \quad (24)$$

и $\beta(\cdot)$ является перестановкой индексов $(1, 2, \dots, q)$ такой, что $\beta^3(t) = t$, то есть $\beta(\cdot)$ представляет собой корень 3-й степени из тождественной перестановки длины q .

3.2. Редуцированные уравнения Брента для квадратных матриц

Переход к покомпонентной записи приводит к соответствующим версиям уравнений Брента:

$$\delta(i_2 - j_1)\delta(j_2 - k_1)\delta(k_2 - i_1) = \sum_{t=1}^r (X_t)_{i_2, i_1} (X_{\alpha(t)})_{j_2, j_1} (X_{\alpha^2(t)})_{k_2, k_1} \quad (25)$$

и, для параметризованной версии,

$$\delta(i_2 - j_1)\delta(j_2 - k_1)\delta(k_2 - i_1) = \sum_{t=1}^q \sum_{s=0}^{p-1} (\Lambda^s X_t \Lambda^{-s})_{i_2, i_1} (\Lambda^s X_{\beta(t)} \Lambda^{-s})_{j_2, j_1} (\Lambda^s X_{\beta^2(t)} \Lambda^{-s})_{k_2, k_1}, \quad (26)$$

где $1 \leq i_1, i_2, j_1, j_2, k_1, k_2 \leq n$. Подставляя явное выражение для Λ из (24), получаем редуцированные уравнения Брента для квадратных матриц:

$$-\frac{1}{p}\delta(i_2 - j_1)\delta(j_2 - k_1)\delta(k_2 - i_1) + \sum_{t=1}^q (X_t)_{i_2, i_1} (X_{\beta(t)})_{j_2, j_1} (X_{\beta^2(t)})_{k_2, k_1} = 0, \quad (27a)$$

$$1 \leq i_1, i_2, j_1, j_2, k_1, k_2 \leq n, \quad (27b)$$

$$(i_2 - i_1 + j_2 - j_1 + k_2 - k_1) \pmod{p} = 0. \quad (27b)$$

Используя представление $q = q_0 + 3q_1$, без ограничения общности можно положить

$$\begin{aligned} t &= (1 \ 2 \ 3 \ \dots \ q_0 \ q_0 + 1 \ q_0 + 2 \ q_0 + 3 \ \dots \ q_0 + 3q_1), \\ \beta(t) &= (1 \ 2 \ 3 \ \dots \ q_0 \ q_0 + 2 \ q_0 + 3 \ q_0 + 1 \ \dots \ q_0 + 3q_1 - 2), \\ \beta^2(t) &= (1 \ 2 \ 3 \ \dots \ q_0 \ q_0 + 3 \ q_0 + 1 \ q_0 + 2 \ \dots \ q_0 + 3q_1 - 1). \end{aligned}$$

Так, для $q = 12$ при $q_0 = q_1 = 3$, такая конструкция имеет вид

$$\begin{aligned} t &= (1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12), \\ \beta(t) &= (1\ 2\ 3\ 5\ 6\ 4\ 8\ 9\ 7\ 11\ 12\ 10), \\ \beta^2(t) &= (1\ 2\ 3\ 6\ 4\ 5\ 9\ 7\ 8\ 12\ 10\ 11). \end{aligned}$$

Соответственно, параметризованная форма уравнений Брента для квадратных матриц принимает вид

$$\begin{aligned} & -\delta(i_2 - j_1)\delta(j_2 - k_1)\delta(k_2 - i_1) + \\ & + \sum_{t=1}^{q_0} \sum_{s=0}^{p-1} (\Lambda^s X_t \Lambda^{-s})_{i_2, i_1} (\Lambda^s X_t \Lambda^{-s})_{j_2, j_1} (\Lambda^s X_t \Lambda^{-s})_{k_2, k_1} + \\ & + \sum_{t=1}^{q_1} \sum_{s=0}^{p-1} ((\Lambda^s X_{q_0+3t-2} \Lambda^{-s})_{i_2, i_1} (\Lambda^s X_{q_0+3t-1} \Lambda^{-s})_{j_2, j_1} (\Lambda^s X_{q_0+3t} \Lambda^{-s})_{k_2, k_1} + \\ & + (\Lambda^s X_{q_0+3t-1} \Lambda^{-s})_{i_2, i_1} (\Lambda^s X_{q_0+3t} \Lambda^{-s})_{j_2, j_1} (\Lambda^s X_{q_0+3t-2} \Lambda^{-s})_{k_2, k_1} + \\ & + (\Lambda^s X_{q_0+3t} \Lambda^{-s})_{i_2, i_1} (\Lambda^s X_{q_0+3t-2} \Lambda^{-s})_{j_2, j_1} (\Lambda^s X_{q_0+3t-1} \Lambda^{-s})_{k_2, k_1}) = 0, \end{aligned} \quad (28)$$

и редуцированная форма переписывается в виде

$$\begin{aligned} & -\frac{1}{p}\delta(i_2 - j_1)\delta(j_2 - k_1)\delta(k_2 - i_1) + \\ & + \sum_{t=1}^{q_0} (X_t)_{i_2, i_1} (X_t)_{j_2, j_1} (X_t)_{k_2, k_1} + \\ & + \sum_{t=1}^{q_1} ((X_{q_0+3t-2})_{i_2, i_1} (X_{q_0+3t-1})_{j_2, j_1} (X_{q_0+3t})_{k_2, k_1} + \\ & + (X_{q_0+3t-1})_{i_2, i_1} (X_{q_0+3t})_{j_2, j_1} (X_{q_0+3t-2})_{k_2, k_1} + \\ & + (X_{q_0+3t})_{i_2, i_1} (X_{q_0+3t-2})_{j_2, j_1} (X_{q_0+3t-1})_{k_2, k_1}) = 0 \end{aligned} \quad (29)$$

для тех же $i_1, \dots, k_2$, что и в (27).

3.3. Обобщенная форма редуцированных уравнений Брента

Добавляя к уравнениям (29) дополнительные условия типа (15)

$$\Lambda^{d_t^{(0)}} X_t - X_t \Lambda^{d_t^{(0)}} = 0, \quad t = 1, \dots, q_0, \quad (30a)$$

$$\Lambda^{d_t^{(1)}} X_u - X_u \Lambda^{d_t^{(1)}} = 0, \quad u - q_0 \in \{3t - 2, 3t - 1, 3t\}, \quad t = 1, \dots, q_1, \quad (30b)$$

(здесь, как и ранее, предполагаем, что d_t^0 и d_t^1 являются делителями p), получаем обобщенную форму редуцированных уравнений Брента. Аналогично уравнениям (19), условия (30) можно заменить на требование разрезности вида

$$(X_t)_{i_2, i_1} = 0, \quad i_2 - i_1 \not\equiv 0 \pmod{\frac{p}{d_t^{(0)}}}, \quad 1 \leq t \leq q_0, \quad (31a)$$

$$(X_u)_{i_2, i_1} = 0, \quad i_2 - i_1 \not\equiv 0 \pmod{\frac{p}{d_t^{(1)}}}, \quad u - q_0 \in \{3t - 2, 3t - 1, 3t\}, \quad 1 \leq t \leq q_1. \quad (31b)$$

Соответствующие параметризованные уравнения Брента принимают вид

$$\begin{aligned} & -\delta(i_2 - j_1)\delta(j_2 - k_1)\delta(k_2 - i_1) + \\ & + \sum_{t=1}^{q_0} \frac{p}{d_t^{(0)}} \sum_{s=0}^{d_t^{(0)}-1} (\Lambda^s X_t \Lambda^{-s})_{i_2, i_1} (\Lambda^s X_t \Lambda^{-s})_{j_2, j_1} (\Lambda^s X_t \Lambda^{-s})_{k_2, k_1} + \\ & + \sum_{t=1}^{q_1} \frac{p}{d_t^{(1)}} \sum_{s=0}^{d_t^{(1)}-1} ((\Lambda^s X_{q_0+3t-2} \Lambda^{-s})_{i_2, i_1} (\Lambda^s X_{q_0+3t-1} \Lambda^{-s})_{j_2, j_1} (\Lambda^s X_{q_0+3t} \Lambda^{-s})_{k_2, k_1} + \\ & + (\Lambda^s X_{q_0+3t-1} \Lambda^{-s})_{i_2, i_1} (\Lambda^s X_{q_0+3t} \Lambda^{-s})_{j_2, j_1} (\Lambda^s X_{q_0+3t-2} \Lambda^{-s})_{k_2, k_1} + \\ & + (\Lambda^s X_{q_0+3t} \Lambda^{-s})_{i_2, i_1} (\Lambda^s X_{q_0+3t-2} \Lambda^{-s})_{j_2, j_1} (\Lambda^s X_{q_0+3t-1} \Lambda^{-s})_{k_2, k_1}) = 0, \end{aligned} \quad (32)$$

и ранг такого разложения равен

$$r = \sum_{t=1}^{q_0} d_t^{(0)} + 3 \sum_{t=1}^{q_1} d_t^{(1)}. \quad (33)$$

4. НЕЛИНЕЙНЫЙ МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

Ниже рассмотрим нелинейный метод наименьших квадратов, применимый в случае комплекснозначной функции невязки $f: \mathbb{C}^n \rightarrow \mathbb{C}^m$ и не требующий вычисления ее производных. Для достаточно гладкой функции невязки f рассмотрим нелинейную задачу наименьших квадратов

$$x = \arg \min_{x \in \mathbb{C}^n} \frac{1}{2} \|f(x)\|^2, \quad (34)$$

где $\|f(x)\|^2 = f^H(x)f(x)$ и $f^H = (\bar{f})^T$ обозначает операцию транспонирования с сопряжением; также, будем обозначать $f = f(x)$. Предлагается, аналогично [11], использовать шаги минимизации нормы невязки вдоль направлений, выбираемых в подпространствах, образованных столбцами псевдослучайных прямоугольных $n \times d$ -матриц U , где $d \leq n$. При этом предпочтительна (хотя бы приближенная) ортонормированность столбцов u_j и нормализованность строк каждой матрицы $U = [u_1 | \dots | u_d]$. Если x является текущим приближением к решению, то следующее приближение строится в виде

$$x_+ = x + \alpha U s, \quad 0 < \alpha < 2,$$

где параметр длины шага α получается из условия типа Армихо [12] (см. ниже (38)). В качестве вектора $s \in \mathbb{C}^d$ берется решение регуляризованной линейной задачи наименьших квадратов

$$s = \arg \min_{s \in \mathbb{C}^d} (\|f + V s\|^2 + \xi \|s\|^2), \quad 0 < \xi \ll \text{tr}(V^H V), \quad (35)$$

где ξ является параметром регуляризации, а $n \times d$ -матрица V строится в виде

$$V = \left[\frac{f(x + \xi u_1) - f(x)}{\xi} \mid \dots \mid \frac{f(x + \xi u_d) - f(x)}{\xi} \right] \in \mathbb{C}^{n \times d}. \quad (36)$$

Отсюда получаем

$$s = -(V^H V + \xi I)^{-1} V^H f.$$

Теперь, определяя величины

$$\theta^2 = \frac{-s^H V^H f}{f^H f}, \quad \gamma^2 = \frac{s^H s}{f^H f},$$

можно показать, что при выполнении условия сходимости

$$\left\| \frac{f(x + \alpha U s) - f}{\alpha} - V s \right\| \leq \left(\left(1 - \frac{\alpha}{2}\right) \theta^2 + \frac{\alpha}{2} \xi \gamma^2 \right) \|f\|, \quad (37)$$

справедлива следующая оценка убывания квадрата нормы невязки:

$$\frac{\|f(x + \alpha U s)\|^2}{\|f\|^2} \leq 1 - \left(\frac{\alpha(2 - \alpha)\theta^2 + \alpha^2 \xi \gamma^2}{2} \right)^2. \quad (38)$$

Таким образом, (37) представляет достаточное условие для убывания нормы невязки. Заметим, что левая часть (37) представляет собой норму разности двух различных аппроксимаций одной и той же величины $J(x)Us$, где $J(x) = \partial f / \partial x$ обозначает якобиан функции $f(x)$. Аналогичный подход был использован в [9, 10] и адаптирован к отказу от вычисления производных в [7].

Оценка (38) непосредственно используется для выбора подходящего значения параметра шага $\alpha \in \{1, 1/2, 1/4, \dots\}$ аналогично [12]. Так, в случае липшицева якобиана J можно убедиться в справедливости (37) для всех достаточно малых α , что обеспечивает корректность предлагаемого способа выбора параметра шага. Естественным способом остановки итераций является (помимо задания предела числа итераций) появление предельно малых значений величины $\alpha(2 - \alpha)\theta^2 + \alpha^2 \xi \gamma^2$.

Замечание 6. В качестве псевдослучайной последовательности использовались точки на единичном круге в комплексной плоскости, получаемые по рекуррентной формуле

$$\eta_{k+1} = \eta_k / \bar{\eta}_k, \quad k = 1, 2, \dots,$$

с начальным условием, например, $\eta_0 = 0.6 + 0.8i$. Элементы матрицы $U \in \mathbb{C}^{n \times d}$ задавались в виде $(U_k)_{i,j} = n^{-1/2} \eta_{i+(j-1)n+k}$, $1 \leq i \leq n$, $1 \leq j \leq d$, где через k обозначен номер итерации. Начальный вектор x также задавался псевдослучайным образом, например, $(x^{(m)})_j = 0.25 \eta_{mj}$, $m = 1, 2, \dots, m_{\max}$, где $m_{\max}$ обозначает число повторных применений итерационного метода для вычисления решения уравнения $f(x) = 0$.

Замечание 7. Для расчетов в двойной точности, в (35) и (36) выбирались значения $\xi = 10^{-13}$ и $\zeta = 10^{-8}$ соответственно. Максимальное число итераций и рестартов ограничивалось несколькими сотнями.

5. РЕЗУЛЬТАТЫ ЧИСЛЕННОГО ПОИСКА МАТРИЧНЫХ КОЭФФИЦИЕНТОВ

Расчеты были выполнены на настольном компьютере Pentium® Dual-Core CPU E6600@3.06 GHz (объем оперативной памяти 3.25 Gbytes). Использовался транслятор G95 Fortran 95 (<http://www.g95.org>).

Таблица 1. Найденные решения редуцированных уравнений Брента с наибольшими значениями p . В большинстве случаев удалось воспроизвести известные значения ранга, и даже улучшить два результата. Однако в трех случаях ранг увеличился.

n_1	n_2	n_3	r	ω	p	q	$\mathcal{D}$	$r(p, \mathcal{D})$	$\omega(p, \mathcal{D})$
2	2	2	7	2.807...	3	3	1	7	2.807...
2	2	3	11	2.894...	2	6	1	11	2.894...
2	2	4	14	2.855...	2	7	$\emptyset$	14	2.855...
2	2	5	18	2.894...	3	6	$\emptyset$	18	2.894...
2	2	6	21	2.873...	6	4	3	21	2.873...
2	3	3	15	2.810...	3	5	$\emptyset$	15	2.810...
2	3	4	20	2.827...	4	5	$\emptyset$	20	2.827...
2	3	5	25	2.839...	5	5	$\emptyset$	25	2.839...
2	3	6	30	2.847...	6	5	$\emptyset$	30	2.847...
2	4	4	26	2.820...	4	7	2	26	2.820...
2	4	5	33	2.843...	4	8	$\emptyset$	32	2.818...
2	4	6	39	2.839...	6	7	3	39	2.839...
2	5	5	40	2.828...	5	8	$\emptyset$	40	2.828...
2	5	6	48	2.836...	6	8	$\emptyset$	48	2.836...
2	6	6	57	2.836...	6	10	3	48	2.836...
3	3	3	23	2.854...	4	7	1;2	23	2.854...
3	3	4	29	2.818...	4	8	2	30	2.847...
3	3	5	36	2.824...	3	12	$\emptyset$	36	2.824...
3	3	6	40	2.774...	3	14	$\emptyset$	42	2.810...
3	4	5	47	2.821...	4	12	$\emptyset$	48	2.836...
2	2	2	7	2.807...	2	$4 = 1 + 3 \cdot 1$	1	7	2.807...
3	3	3	23	2.854...	4	$7 = 4 + 3 \cdot 1$	1;2	23	2.854...
4	4	4	49	2.807...	4	$12 = 3 + 3 \cdot 3$	$\emptyset$	48	2.792...

Вектор x численного решения считался допустимым, если его невязка была близка к пределу машинной точности, а максимум модуля компонент был меньше единицы:

$$\|f(x)\| < 10^{-15}, \quad \|x\|_{\infty} < 1. \quad (39)$$

Выполнение этих требований позволяет предполагать, что известные эквивалентные преобразования решения уравнений Брента (см., например, [14]) позволят получить явные выражения коэффициентов X_t , Y_t и Z_t через квадратичные и/или кубические иррациональности. Если получался вектор x с хорошей невязкой, но с нормой больше единицы, то он умножался на 0.9 и использовался как начальное приближение для рестарта, возможно, несколько раз. В итоге удавалось получить вектор x , удовлетворяющий (39).

Результаты расчетов приведены в таблице 1. В левой половине приводятся известные результаты, опубликованные в [2, 3, 5, 15, 16], а в правой — настройки редуцированных уравнений Брента и соответствующие значения ранга. Через $\mathcal{D}$ обозначается множество $\{d_1, d_2, \dots\}$ нетривиальных делителей p , использованных для построения обобщенной формы редуцированных уравнений Брента (13), (19) или, для квадратных матриц, (29), (31). В нижней части таблицы представлены результаты для квадратных матриц с указанием параметров перестановки $\beta(\cdot)$, использованных в (27), в виде $q = q_0 + 3q_1$.

В случае несовпадения рангов, меньшее значение выделено жирным шрифтом. Из таблицы видно улучшение до $r = 32$, достигнутое для задачи перемножения прямоугольных матриц с $n_1 = 2$, $n_2 = 4$ и $n_3 = 5$, а также улучшение ранга до $r = 48$ для алгоритма перемножения двух квадратных матриц 4-го порядка.

Автор выражает благодарность Е. Е. Тыртышникову за полезные обсуждения тематики данной работы.

СПИСОК ЛИТЕРАТУРЫ

1. *Brent R. P.* Algorithms for matrix multiplication. (Report No. STAN-CS-70-157). Stanford Univ. CA Dept. of Computer Science, 1970, 58 p.
2. *Strassen V.* Gaussian elimination is not optimal // Numer. Math. **13**, 354–356 (1969).
3. *Laderman J. D.* A noncommutative algorithm for multiplying 3x3 matrices using 23 multiplications. Bull. Amer. Math. Soc. **82**, 126–128 (1976).
4. *Pan V.* How we can speed up matrix multiplication? // SIAM Review, **26**(3), 393–415 (1984).
5. *Smirnov A. V.* The bilinear complexity and practical algorithms for matrix multiplication // Comp. Math. Math. Phys. **53**, 1781–1785 (2013).
6. *Karstadt E., Schwartz O.* Matrix multiplication, a little faster // J. of the ACM (JACM). 2020, **67**(1), 1–31.
7. *Kaporin I.* A Derivative-Free Nonlinear Least Squares Solver. In: Olenov N.N., Evtushenko Y.G., Jacimovic M., Khachay M., Malkova V. (eds.) Optimization and Applications. OPTIMA 2021. Lecture Notes in Computer Science, V. 13078. P. 217–230. Springer, Cham. (2021). https://doi.org/10.1007/978-3-030-91059-4_16
8. *Oseledets I. V., Savostyanov D. V.* Minimization methods for approximating tensors and their comparison // Computational Mathematics and Mathematical Physics, 2006, **46**(10), 1641–1650
9. *Kaporin I. E., Axelsson O.* On a class of nonlinear equation solvers based on the residual norm reduction over a sequence of affine subspaces // SIAM J. Sci. Comput. **16**(1), 228–249 (1994)
10. *Kaporin I.* Preconditioned Subspace Descent Method for Nonlinear Systems of Equations // Open Computer Science, 2020, **10**(1), 71–81
11. *Kozak D., Molinari C., Rosasco L., Tenorio L., Villa S.* Zeroth order optimization with orthogonal random directions. // Mathematical Programming. 2022. V. 199. P. 1–41.
12. *Armijo L.* Minimization of functions having Lipschitz continuous first partial derivatives // Pacific Journal of mathematics **16**(1), 1–3 (1966).
13. *Ballard G., Ikenmeyer C., Landsberg J. M., Ryder N.* The geometry of rank decompositions of matrix multiplication II: 3x3 matrices. // Journal of Pure and Applied Algebra, 2019, **223**(8), 3205–3224.
14. *Berger G. O., Absil P. A., De Lathauwer L., Jungers R. M., Van Barel M.* Equivalent polyadic decompositions of matrix multiplication tensors // J. of Computational and Applied Mathematics, 2022, 406, 113941. <https://doi.org/10.1016/j.cam.2021.113941>.

15. *Hopcroft J. E., Kerr L. R.* On minimizing the number of multiplications necessary for matrix multiplication // SIAM Journal on Applied Mathematics, 1971, **20**(1):30–36.
16. *Fawzi A. et al.* Discovering faster matrix multiplication algorithms with reinforcement learning. Nature, 2022, **610**(7930): 47–53.

FINDING COMPLEX-VALUED SOLUTIONS TO THE BRENT EQUATIONS BY REDUCING THEM TO A NONLINEAR LEAST SQUARES PROBLEM

I. E. Kaporin*

*119333 Moscow, Vavilov Str., 44, Federal Research Center Computer Science and Control
of the Russian Academy of Sciences, Russia*

**e-mail: igorkaporin@mail.ru*

Received: 12.03.2024

Revised: 12.03.2024

Accepted: 05.06.2024

Abstract. Finding nontrivial solutions to the trilinear Brent equations corresponds to the construction of asymptotically fast matrix multiplication algorithms is an important, but in general a very difficult computational task. Methods of parameterization of the Brent equations based on the use of symmetries of the matrix product tensor are proposed, which make it possible to repeatedly reduce the dimension of the problem. The numerical solution of the obtained trilinear or cubic systems of nonlinear equations is carried out by reducing to a nonlinear least squares problem and applying to it a specially developed iterative method that does not require calculation of derivatives. The found solutions of the parameterized Brent equations, as a rule, have a rank no higher (and sometimes even lower) than the known results. Thus, an algorithm for multiplying two 4th-order matrices in 48 active multiplications is obtained.

Keywords: Brent equations; fast matrix multiplication; Strassen algorithm; nonlinear least squares problem.

О МУЛЬТИОПЕРАТОРНЫХ АППРОКСИМАЦИЯХ 24-ГО ПОРЯДКА В СХЕМАХ ДЛЯ УРАВНЕНИЙ С КОНВЕКТИВНЫМИ ЧЛЕНАМИ

© 2024 г. А. И. Толстых^{1,*}

¹ 119333 Москва, ул. Вавилова, 44, ФИЦ ИУ РАН, Россия

*e-mail: tol@ccas.ru

Поступила в редакцию 19.01.2024 г.

Переработанный вариант 04.04.2024 г.

Принята к публикации 31.05.2024 г.

В рамках исследования мультиоператорных аппроксимаций и схем, использующих экономично обрабатываемые двухточечные операторы, рассмотрены аппроксимации 24-го порядка первых производных в задачах с конвективными членами. Основное внимание уделено спектральным свойствам, характеризующим их высокую точность и разрешающую способность. Для иллюстрации этих свойств приведены примеры решения модельных задач. Рассмотрены возможности использования таких мультиоператорных схем в случае разрывных решений. Библ. 16. Фиг. 7. Табл. 2.

Ключевые слова: мультиоператоры, аппроксимация 24-го порядка производной, уравнения Эйлера, задачи с разрывными решениями.

DOI: 10.31857/S0044466924090024, EDN: WLDQMD

ВВЕДЕНИЕ

Построение схем повышенных порядков аппроксимации, эффективно реализуемых на высокопроизводительных вычислительных системах, является важным направлением развития современных численных методов. Существует множество задач, для которых применение таких методов играет существенную роль. Это связано не только с возможностью значительного сокращения затрат времени вычислений при желаемой точности решений, но и с возможностью осуществлять более точное численное моделирование ряда физических процессов при использовании разумного числа узлов разностных сеток.

На современном уровне при решении многих задач оказываются востребованными численные методы, обладающие не только высокими порядками аппроксимаций. Во многих случаях желательно, чтобы реализующие их коды допускали бы эффективное использование высокопроизводительных вычислительных систем. Это означает возможность их распараллеливания с высокой параллельной эффективностью.

Перечисленными свойствами могут обладать схемы с аппроксимациями производных, использующими мультиоператоры. Концепция мультиоператоров (линейных комбинаций базисных операторов) была введена в [1] в качестве нестандартного подхода к увеличению порядков аппроксимаций формул численного анализа. В традиционных численных методах это увеличение достигается путем увеличения числа базисных функций в лагранжевых полиномах, на которых основаны используемые приближенные формулы. В широко применяемых сеточных методах возрастание числа базисных функций проявляется в расширении шаблонов сеточных операторов. Хорошо известно, что при этом несмотря на формальное повышение порядков аппроксимаций, могут возникать нежелательные эффекты, ограничивающие порядки реально применяемых схем.

В мультиоператорных методах при конструировании базисных операторов используются шаблоны с двумя или тремя узлами. При этом порядки аппроксимации мультиоператоров линейно возрастают с числом базисных операторов и их рост может быть весьма значительным.

На ранних этапах развития мультиоператорных схем в качестве базисных операторов использовались однопараметрические нецентральные (несимметричные) компактные аппроксимации третьего или пятого порядков, основанные на трехточечных операторах [2]. Это позволяло аппроксимировать производные в уравнениях механики жидкости и газа с более высокими порядками и получать достаточно точные решения двумерных задач при сравнительно небольшом числе узлов сетки. Это вполне соответствовало уровню вычислительной

техники в виде однопроцессорных ПК. Не останавливаясь на обзоре решенных задач, отметим успешное применение аппроксимаций 9-го порядка для описания так называемого скрич-эффекта при истечении из сопл сверхзвуковых струй [3].

Хотя порядки аппроксимации мультиоператоров такого типа превосходили порядки стандартных методов, их применение при решении пространственных задач с использованием суперкомпьютерных вычислений не выглядело перспективным.

Новым этапом развития мультиоператорной технологии явились семейства мультиоператоров с базисными операторами в виде компактных аппроксимаций с обрабатываемыми двухточечными операторами. Можно выделить два свойства этих семейств.

Во-первых, при построении мультиоператоров значительно улучшается обусловленность линейных систем, определяющие коэффициенты при базисных операторах. Это позволяет, увеличивая их число, получать очень высокие порядки аппроксимации.

Во-вторых, использование бегущего счета при обращении двухдиагональных матриц вместо трехточечных прогонок в прежних мультиоператорных схемах не только сокращает число арифметических операций, приходящихся на узел сетки, но и допускает высоко эффективное распараллеливание. Хотя первые идеи этого подхода были обозначены в [4], описание семейств с доказательствами существования и единственности были представлены позднее в [5].

Одно из семейств мультиоператоров указанного выше типа отличается особой простотой. В нем базисные операторы являются компактными аппроксимациями первых производных, определяемыми произведением двух сеточных операторов вида $A^{-1}(c)B$, где оба оператора являются двухточечными, а первый из них линейно зависит от параметра c . При вычислении действий таких операторов на сеточные функции число арифметических операций, отнесенное к числу узлов сетки, равно трем. Первый опыт их применения связан с численными решениями двумерных и пространственных нестационарных уравнений Эйлера и Навье—Стокса при использовании схемы 16-го порядка. В трехмерных случаях численные решения были получены в результате суперкомпьютерных вычислений, реализующих распараллеленные коды [6], [7]. В первом случае осуществлялось моделирование ламинарно-турбулентного перехода, а во втором изучалась неустойчивость истекающих их сопел дозвуковых струй. В обоих случаях оценки показали параллельную эффективность, достаточно близкую к идеальной.

Успешный опыт применения мультиоператорных аппроксимаций 16-го порядка из данного семейства делает естественным дальнейшие исследования представителей этого семейства, обладающих более высокими порядками.

Вместе с тем следует подчеркнуть, что высокие порядки любых схем повышенной точности достигаются при наличии достаточной гладкости всех функций, входящих в исходные уравнения и, в частности, якобианов в случае криволинейных систем координат. Насколько степени гладкости этих функций влияют на точность получаемых решений зависит от конкретных задач и могут допускать те или иные оценки. При этом мультиоператорные схемы не пригодны в случае неструктурированных сеток, а в случае равномерных сеток в криволинейной системе координат предполагается гладкое (например, аналитическое) преобразование декартовых координат. Таким образом, мультиоператорные схемы следует рассматривать не как универсальный инструмент, а как дополнение к традиционным методам (в том числе невысокого порядка), повышающий возможности численного моделирования для подробного описания ряда сложных физических процессов.

В данной работе рассматривается аппроксимация 24-го порядка в контексте схем для уравнений с конвективными членами и, в частности для уравнений Эйлера. К этой категории можно отнести также уравнения Навье—Стокса, в которых члены с вязкостью могут аппроксимироваться независимо с теми или иными порядками.

1. КЛАСС СХЕМ И ОБРАЗУЮЩИЕ ИХ МУЛЬТИОПЕРАТОРЫ

1.1. Структура схем

Рассматриваемые в данной работе мультиоператоры предназначены для аппроксимации пространственных производных в гиперболических системах вида

$$\frac{\partial \mathbf{u}}{\partial t} + \sum_{m=1}^3 \frac{\partial \mathbf{f}_m(\mathbf{u})}{\partial x_m} = 0, \quad (1)$$

где $\mathbf{u}$ и $\mathbf{f}_m$ — векторные функции. При этом имеется в виду их применение в задачах для уравнений Эйлера и Навье—Стокса.

Рассматриваемый класс схем для уравнений (1) имеет вид

$$\frac{\partial \mathbf{u}}{\partial t} + \sum_{m=1}^3 L_M^{(m)} \mathbf{f}_m(\mathbf{u}) + \sum_{m=1}^3 C_M D_M^{(m)} \mathbf{u} = 0, \quad C_M \geq 0, \quad (2)$$

где $L_M^{(m)}$ и $D_M^{(m)}$ – мультиоператоры (т.е. линейные комбинации базисных операторов [2]), соответственно, аппроксимирующие производные по переменной x_m и образующие диссипативную часть погрешности аппроксимации схемы. В данной работе предполагается, что формулы, определяющие эти мультиоператоры, одинаковы для всех пространственных переменных.

По построению $L_M^{(m)}$ и $D_M^{(m)}$ являются, соответственно, кососимметричными и положительными самосопряженными операторами в гильбертовом пространстве сеточных функций со скалярным произведением в виде суммирования их значений по узлам сетки. При этом действия $D_M^{(m)} \mathbf{u}$ операторов $D_M^{(m)}$ на сеточные функции $\mathbf{u}$, являющиеся проекциями на сетки достаточно гладких функций, входят в погрешность аппроксимации схемы. Неотрицательные постоянные C_M , рассматриваемые как свободные параметры, позволяют в случае необходимости регулировать уровни диссипации, демпфирующей паразитические осцилляции численных решений.

Полудискретные схемы (2) в приближении “замороженных коэффициентов” являются схемами с положительными операторами и поэтому устойчивы в среднеквадратичной норме при $C > 0$; при $C = 0$ они нейтрально устойчивы. После той или иной дискретизации производной по t устойчивость схем в этом приближении может быть исследована традиционными методами. Опыт применения мультиоператорных схем в настоящее время ограничен явными схемами для нестационарных задач. При использовании метода Рунге–Кутты для интегрирования по переменной t схемы в этом случае оказываются явными условно-устойчивыми схемами. Выбор его порядка и шагов по t легко определяется результатами методических расчетов для конкретных задач.

В случае уравнений с диссипативными членами и, в частности, уравнений Навье–Стокса, к аппроксимациям (2) добавляются те или иные аппроксимации членов с вязкостью. В частности, последние также могут использовать мультиоператоры того или иного порядка (см. [2]). Однако в случае больших чисел Рейнольдса оценки, численные эксперименты, а также опыт применения схем высокого порядка, отраженный в различных публикациях, показывают, что для поддержания высокой точности решений нет необходимости аппроксимировать члены с вязкостью с тем же высоким порядком, что и остальные члены уравнений.

В данной статье рассматриваются мультиоператоры $L_M^{(m)}$, аппроксимирующие первые производные с 24-м порядком, а также мультиоператоры $D_M^{(m)}$, действия которых на проекции гладких функций стремятся к нулю с 23-м порядком относительно шагов сетки. Для описания структуры мультиоператоров в схеме (2) достаточно ограничиться случаем одной пространственной переменной $x = x_1$.

1.2. Аппроксимация производных

На сетке $\omega_h : x_j = jh, j = 0, \pm 1, \pm 2, \dots, h = \text{const}$, мультиоператорная аппроксимация производной по x имеет вид

$$\frac{\partial}{\partial x} \approx L_M(c_1, c_2, \dots, c_M) = \sum_{i=1}^M \gamma_i l(c_i), \quad \sum_{i=1}^M \gamma_i = 1, \quad (3)$$

где $c_1, c_2, \dots, c_M$ – набор M не совпадающих свободных параметров, а $l(c_i)$ – базисные операторы. Определяющее их семейство операторов $l(c)$ имеет вид

$$l(c) = (L_l(c) + L_r(c))/2,$$

где L_l и L_r являются “левыми” и “правыми” компактными аппроксимациями производных вида

$$L_l(c) = \frac{1}{h} R_l(c)^{-1} \Delta_-, \quad L_r(c) = \frac{1}{h} R_r(c)^{-1} \Delta_+, \quad R_l(c) = I + c \Delta_-, \quad R_r(c) = I - c \Delta_+.$$

Здесь Δ_- и Δ_+ – операторы левых и правых двухточечных разностей ($\Delta_- u_j = u_j - u_{j-1}$, $\Delta_+ u_j = u_{j+1} - u_j$). Из этих формул следует, что вычисление при заданных значениях параметров c_i действия мультиоператора на сеточную функцию u_j сводится к обращениям двухдиагональных матриц методом бегущего счета. Например, вычисление действия $v_j = L_l(c_i) u_j$ оператора $L_l(c_i)$ осуществляется согласно формуле

$$v_j = \alpha v_{j-1} + \beta \Delta_- u_j, \quad \alpha = c_i / (1 + c_i), \quad \beta = 1 / (1 + c_i). \quad (4)$$

Аналогично осуществляется вычисление действия $w_j = L_r(c)u_j$ оператора $L_r(c_i)$, при котором значение сеточной функции с индексом j определяется по ее значению с индексом $j + 1$. Эти процессы очевидно устойчивы при $c_i > -1/2$. В дальнейшем будем предполагать, что это условие выполнено. Результатом являются значения $l(c_i)u_j = (v_j + w_j)/2$, которые затем умножаются на коэффициенты γ_i и суммируются по i . Эти коэффициенты определяются из условия обращения в ноль $M - 1$ низших членов разложения в ряд Тейлора погрешности аппроксимации L_M при предположении о достаточной гладкости функции $u(x)$. $M - 1$ таких условий вместе с равенством единице суммы коэффициенты γ_i в (3) образуют линейную систему с матрицей $M \times M$, зависящей только от c_i (см. [2]). Можно легко убедиться в том, что коэффициенты разложения в ряд по степеням h действия $l(c)u_j$ оператора $l(c_i)$ являются полиномами только с четными степенями h . Из этого следует, что решение системы, определяющее γ_i , $i = 1, 2, \dots, M$, обеспечивает погрешность $O(h^{2M})$ мультиоператорной аппроксимации производной $L_M(c_1, c_2, \dots, c_M)$.

В гильбертовом пространстве H сеточных функций со скалярным произведением $(u, v) = h \sum_{-\infty}^{\infty} u_j v_j$, где суммирование значений сеточных функций производится по узлам сетки ω_h , базисные операторы и, соответственно, мультиоператоры L_M являются кососимметричными операторами, не создающими диссипацию.

Для доказательства этого удобно записать левые и правые двухточечные разности, используя кососимметричные и самосопряженные операторы, определенные, соответственно, формулами

$$\Delta_0 u_j = u_{j+1} - u_{j-1}, \quad \Delta_2 u_j = u_{j+1} - 2u_j + u_{j-1}.$$

При этом двухточечные разности принимают вид

$$\Delta_- = (\Delta_0 - \Delta_2)/2, \quad \Delta_+ = (\Delta_0 + \Delta_2)/2.$$

С учетом этих представлений использование алгебры перестановочных операторов (арифметические действия с операторами аналогичны операциям с комплексными числами, при которых самосопряженные и кососимметричные операторы рассматриваются как действительные и мнимые переменные) приводит после простейших выкладок к следующему выражению для семейства $l(c)$:

$$l(c) = \frac{1}{2}(I - c\Delta_2 - c\Delta_2^2)^{-1}\Delta_0, \quad (5)$$

где I - единичный оператор. При выводе этого выражения использовалось легко проверяемое тождество

$$\Delta_0^2 = 4\Delta_2 + \Delta_2^2.$$

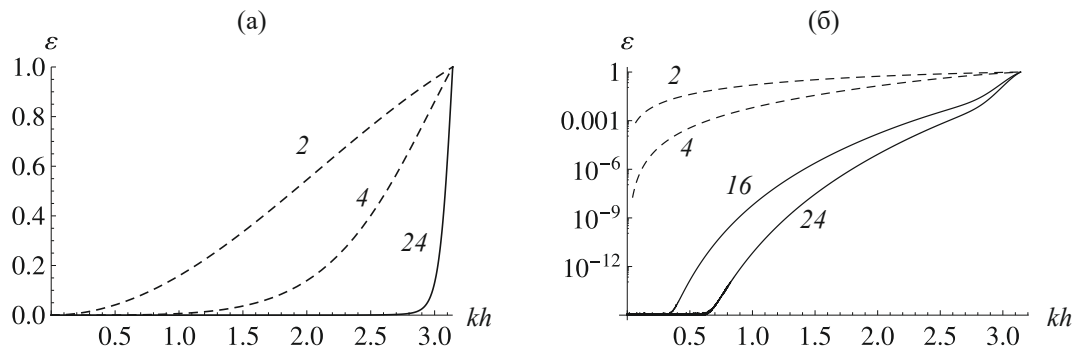
Из (5) следует, что базисные операторы $l(c_i)$ кососимметричны как произведение самосопряженного и кососимметричного операторов. Соответственно таковыми являются и их линейные комбинации, определяющие мультиоператоры.

1.3. Выбор параметров c_i

Из описанной выше процедуры построения мультиоператора следует, что любой выбор различающихся параметров $c_i > -1/2$ в случае разрешимости системы обеспечивает погрешность аппроксимации $O(h^{24})$ при $M = 12$. Однако более информативной оценкой погрешности является ее оценка в пространстве Фурье. Ее асимптотика имеет вид $O((kh)^{24})$, где k - переменная Фурье. В случае гармоник с большими значениями k (т.е. коротковолновых гармоник) для достижения малости этой погрешности для некоторых наборов параметров могут потребоваться достаточно малые значения шага h . В случае, когда желательно правильно воспроизводить мелкомасштабные детали точных решений при разумных шагах сетки, требуется выбор параметров, при котором в пространстве Фурье образ мультиоператора $h^{-1}\mu(\alpha)$, $\alpha = kh$, был бы близок к этому образу производной в как можно более широком диапазоне kh гармоник, поддерживаемых сеткой с шагом h . Как известно, предельное значение kh определяется равенством $\alpha = \pi$. Определим погрешность аппроксимации в пространстве Фурье как отличие от единицы отношения этих образов. Тогда можно показать, что оптимальный набор c_i можно рассматривать как тот, при котором величина

$$\varepsilon(\alpha; c_1, c_2, \dots, c_M) = |\mu(\alpha; c_1, c_2, \dots, c_M)/\alpha - 1|$$

была бы малой в как можно более широком диапазоне волновых чисел $\alpha < \pi$. Для упрощения поиска требуемых значений c_i будем рассматривать мультиоператоры, у которых c_i равномерно распределены между своими минимальными и максимальными значениями $c_{\min}, c_{\max}$, $c_{\min} > -1/2$. Здесь можно поставить задачу о минимизации некоторого функционала. Однако мы ограничимся в рассматриваемом случае $M = 12$ поиском этих пар, для которых графики функции $\varepsilon(\alpha)$ визуальнo совпадали бы с осью абсцисс в как можно более широком



Фиг. 1. Дисперсионные погрешности аппроксимаций производных в обычном (а) и логарифмическом представлении; цифры около кривых отмечают порядки аппроксимаций.

диапазоне $[0, \alpha]$. Примером такого выбора ($c_{\min} = -0.46$, $c_{\max} = 0.45$), определяющего рассматриваемый в дальнейшем конкретный мультиоператор, являются графики функции $\varepsilon(\alpha)$, изображенные на фиг. 1 в обычном и логарифмическом масштабах. Как следует из фиг. 1а, погрешность $\varepsilon(\alpha)$ визуально равна нулю приблизительно для гармоник с волновыми числами $kh < 2.75$. Отсюда следует, что погрешности приблизительно равны нулю, если длины волн гармоник $\lambda = 2\pi/k$ удовлетворяют неравенству $\lambda \gtrsim 2.3h$. Рассматривая такую оценку как вполне удовлетворительную, в дальнейшем будем считать, что для аппроксимации производных используется именно таким образом выбранный мультиоператор. Для сравнения на фиг. 1 приведены в виде штриховых линий графики погрешностей для трехточечной центральной разности второго порядка аппроксимации и компактной аппроксимации четвертого порядка. Как видно, эти погрешности значительно превосходят погрешности мультиоператора и имеют более узкие интервалы kh , где они визуально малы. Известны способы некоторого увеличения этих интервалов путем расширения шаблонов без увеличения порядков аппроксимации.

Более подробная оценка погрешности для построенного мультиоператора следует из графика фиг. 1б, где те же функции представлены с использованием логарифмического масштаба для их значений, превосходящих 10^{-15} . Из такого представления можно заключить, что визуально нулевая погрешность на первом графике при $kh = 2.75$ на самом деле приблизительно равна 0.004, а при $kh < 0.6$ применение мультиоператора с точностью до 10^{-15} эквивалентно дифференцированию. Кроме того, на фиг. 1 показана функция $\varepsilon(kh)$ для мультиоператорной аппроксимации 16-го порядка, использованной ранее в [8]. Как видно, она уступает аппроксимации 24-го порядка вследствие больших значений ε на всем интервале $[0, \pi]$ и более узкого интервала kh , на котором численное дифференцирование можно считать точным.

Рассмотрим простейшие тесты, иллюстрирующие точность и разрешающую способность схемы с построенной мультиоператорной аппроксимацией производной. Эти тесты основаны на вариантах задач для уравнений конвективного переноса вида

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial t} = 0. \quad (6)$$

Для получения численных решений будем использовать схему с описанной выше мультиоператорной аппроксимацией 24-го порядка без диссипативной добавки и стандартный метод Рунге–Кутты четвертого порядка.

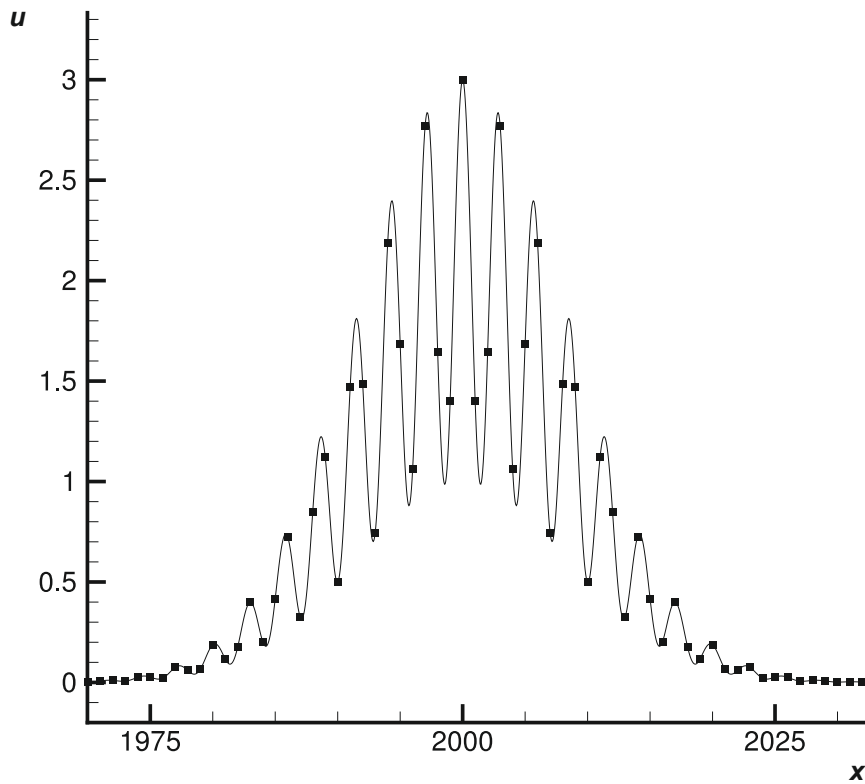
Тест 1. Сохранение дисперсионных свойств на больших интервалах времени.

При решении многих нестационарных задач аэрогидродинамики (в частности, задач аэроакустики, задач о развитии неустойчивости течений) возникает необходимость осуществлять вычисления на достаточно больших интервалах времени. Эти интервалы могут включать в себя отрезок времени до выхода на статистически стационарный режим и время, необходимое для вычисления статистических характеристик решений.

В [9] был предложен тест, выясняющий в контексте задач аэроакустики способность схем правильно описывать распространение коротковолновых гармоник на заданном интервале времени. Задача формулировалась следующим образом. Для уравнения (6) при $f(u) = u$ и начальным условием

$$u(x, 0) = [2 + \cos(\beta x)] [\exp(-\ln 2 (x/10)^2)]$$

требуется получить для значения $\beta = 1.7$ численное решение при $t = 800$, используя тестируемую схему с шагом сетки $h = 1$. Начальные данные описывают волновой пакет с достаточно коротковолновым наполнением,



Фиг. 2. Решения задачи о распространении волнового пакета при $t = 2000$; сплошная кривая представляет точное решение, а маркеры — численное решение в узлах сетки.

поскольку безразмерное волновое число kh гармоник, наложенных на экспоненту, для единичного шага сетки равно 1.7. Согласно точному решению этой задачи, волновой пакет, задаваемый начальными данными, должен перемещаться при возрастании времени t вдоль оси x в неизменном виде с единичной скоростью. Однако численные решения при $t = 800$ и заявленном шаге сетки даже в случае схем повышенной точности могут заметно отличаться от точного решения в силу накопления с течением времени указанных выше погрешностей. Кроме того, дополнительные ошибки могут образовываться вследствие диссипативности тестируемых схем. На фиг. 2 показано точное решение в виде сплошной линии, на которой маркерами отмечено численное решение для рассматриваемой схемы при $C = 0$, т.е. без диссипативной добавки. Оно получено при более жестких условиях $t = 2000$ и $\beta = 2.2$. Как видно, эти решения визуально не различимы.

Рассмотрим более подробно оценки интервалов времени, на которых численные решения для различных схем, не содержащих диссипативный механизм, сохраняют некоторую точность. В силу простоты уравнения его решение с использованием преобразования Фурье, может быть записано в виде

$$u(x, t) = \int_{-\infty}^{\infty} \hat{u}(k) e^{ik(x-t)} dk,$$

где $\hat{u}(k)$ — преобразование Фурье функции $u(x, 0)$. После замены производной по x некоторым разностным оператором решение разностной задачи можно получить аналогичным образом, считая что сеточные функции являются проекциями на сетку функций непрерывного аргумента, а интегрирование по времени осуществляется абсолютно точно. Это решение имеет вид

$$u_1(x, t) = \int_{-\infty}^{\infty} \hat{u}(k) e^{ik(x-t(1+\varepsilon(k)))} dk. \quad (7)$$

На основании анализа выражения для фурье-образа оператора можно заключить, что в случае $h = 1$ функция погрешности аппроксимации $\varepsilon(k)$, входящая в (7), является четной ограниченной функцией, определенной на интервале $k \in [-\infty, \infty]$. Элементарные выкладки для разности $\delta u = u_1(x, t) - u(x, t)$, использующие четность и нечетность функций из подынтегральных выражений, приводит к равенству

$$\delta u = 2 \int_{-\infty}^{\infty} \hat{u}(k) \cos k(x-t) \sin^2 \frac{k\varepsilon(k)t}{2} dk. \quad (8)$$

Правая часть этого равенства допускает приближенную оценку, поскольку функция

$$\hat{u}(k) = a(\exp(-b(k - \beta)^2) + \exp(-b(k + \beta)^2) + 4 \exp(-bk^2)),$$

где $a \approx 3.6$, $b \approx 36$, пренебрежимо мала везде за исключением очень узких окрестностей точек $k = 0, \pm\beta$. Это позволяет для приближенной оценки заменить в (8) функцию $\sin^2 \frac{k\epsilon(k)t}{2}$ на постоянное значение $\sin^2 \frac{\beta\epsilon(\beta)t}{2}$. В результате для относительной погрешности решения возникает приближенное равенство

$$\frac{\delta u(x, t)}{u(x, t)} \approx 2 \sin^2 \frac{\beta\epsilon(\beta)t}{2}. \quad (9)$$

При использовании функций $\epsilon(kh)$ для различных схем оно позволяет оценить без проведения расчетов времени, на которых эти схемы способны сохранять достаточную точность решений при заданных значениях волновых чисел гармоник. В частности, для схем 2-го, 4-го и 24-го порядков значения $\epsilon(1.7)$ равны соответственно 0.42, 0.06 и $4 \cdot 10^{-7}$. Если довольствоваться относительной погрешностью 0.01, то максимальные значения времени оказываются, согласно (9), равными 0.2, 1.4 и 208146. Если же ужесточить тест, положив $\beta = 2.2$, то в случае мультиоператорной схемы расчет можно вести до $t = 1821$. Эта оценка приблизительно согласуется с результатом расчета для $t = 2000$.

Тест 2. Сравнение точности решений при уменьшении шагов сетки.

Рассмотрим периодическую задачу с начальными данными для уравнения Хопфа ($f(u) = u^2$ в (6)), используемую в ряде публикаций для оценки точности решений:

$$\begin{aligned} \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \frac{u^2}{2} &= 0, \quad -1 \leq x \leq 1, \\ u(0, x) &= 1 + 0.5 \sin(\pi x) \end{aligned}$$

Ее решение до некоторого момента времени t является гладким. При фиксированном значении t оно может быть получено с машинной точностью с помощью итерационной процедуры, описанной в [10]. В табл. 1 представлены C -нормы погрешностей численных решений (т.е. их отличий от “точного”), полученных на последовательности сеток с шагами $h = 2/N$, $N = 16, 32, 64, 128, 256$, а также локальные порядки сходимости. В таблице кроме рассматриваемой схемы (обозначенной как L_{24}) представлены для сравнения решения, полученные при использовании компактной аппроксимации 4-го порядка (L_4), а также мультиоператорной аппроксимации 9-го порядка [2] (L_{59}). Во всех случаях применялся метод Рунге–Кутты 4-го порядка с контролем шага по t . Как видно, при $h = 1/64$, $N = 128$ точность решений в случае рассматриваемой схемы имеет порядок точности машинной арифметики. Однако точность вычислений “точного” решения согласно алгоритму из [10] имеет тот же порядок. Это делает проблематичным сравнение с “точным” решением при этом значении h и тем более при его уменьшении. Возможно, что это влияет также на локальные показатели k_c скорости сходимости, вычисленные как отношение логарифмов текущей и предыдущей погрешности. Они не достигают порядка аппроксимации мультиоператорной схемы на данном интервале шагов сетки, что, однако, не представляется существенным с практической точки зрения в виду очень малых значений погрешностей. Чтобы не связывать оценку точности рассматриваемой схемы с точностью “точного” решения, сформулируем следующий вопрос: каким должен приблизительно быть шаг сетки, чтобы в процессе счета до рассматриваемого значения t осуществлялось бы с погрешностью 10^{-15} точное дифференцирование по x ? Чтобы ответить на него, вернемся к фиг. 1. Из приведенного там графика следует, что это происходит приблизительно при $kh < \approx 0.5$. С другой стороны, дискретные преобразования Фурье решений при $h = 1/64$ (т.е. $N = 128$) и $t \leq 0.2$

Таблица 1. Сеточные нормы C погрешностей решений и локальные скорости сходимости k_c

N	16	32	64	128	256
L_4	1.80e-2	2.486e-3	2.052e-4	1.40e-5	9.30e-7
k_c	—	2.6	3.6	3.9	3.9
L_{59}	1.86e-3	2.16e-5	5.02e-8	4.7e-11	9.96e-14
k_c	—	6.4	8.7	10.0	9.2
L_{24}	1.3e-3	7.5e-6	1.4e-9	1.2e-14	—
k_c	—	7.7	9.1	11.1	—

показывают, что их спектры визуально отличны от нуля при $k < 20$. Следовательно, приближенно можно считать, что шаги h должны быть такими, чтобы $20h < 0.5$. Это дает оценку $h < 1/40$.

1.4. Мультиоператор диссипации

Мультиоператор D_M для координаты x имеет вид

$$D_M(\bar{c}_1, \bar{c}_2, \dots, \bar{c}_M) = \sum_{i=1}^M \bar{\gamma}_i d(\bar{c}_i), \quad \sum_{i=1}^M \bar{\gamma}_i = 1.$$

Базисные операторы $d(\bar{c}_i)$ образованы подстановками наборов M значений параметра c в однопараметрическое семейство операторов, записанных в виде разности левых и правых операторов $L_l(c)$ и $L_r(c)$ в (3):

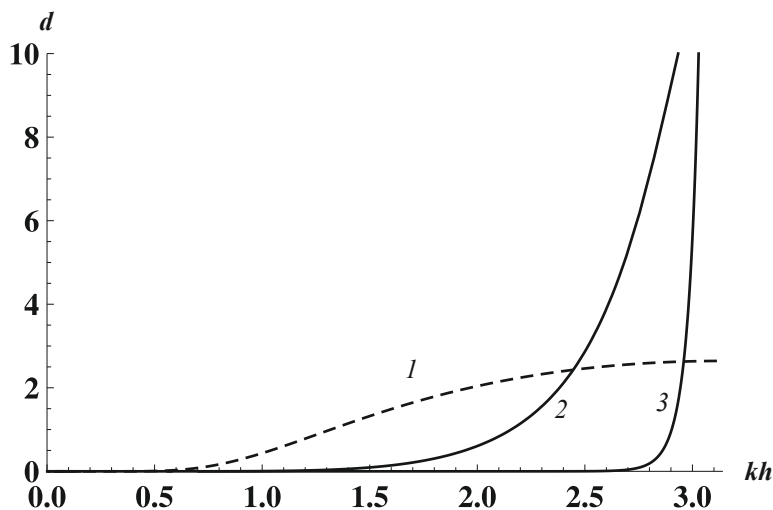
$$d(c) = L_l(c) - L_r(c).$$

Используя алгебраические манипуляции, аналогичные описанным выше для оператора $l(c)$, можно показать, что оператор $d(c)$ является самосопряженным. В отличие от L_M , коэффициенты $\bar{\gamma}_i$ определяются из условий обращения в ноль низших $M - 1$ членов в ряде Тейлора для функции $u(x)$, а не для погрешности аппроксимации производных. Таким образом, действие мультиоператора на гладкую функцию аппроксимирует ее производную порядка $2M$, умноженную на h^{2M-1} .

Поскольку базисные операторы являются самосопряженными, самосопряженным является и оператор $D_M(\bar{c}_{\min}, \bar{c}_{\max})$. Однако он необязательно является положительным, как этого требует устойчивость схемы. Поэтому возникает ограничение на выбор значений $\bar{c}_{\min}, \bar{c}_{\max}$, основанное на требовании положительности фурье-образа $\hat{D}_M(kh; \bar{c}_{\min}, \bar{c}_{\max})$ оператора D_M .

На фиг. 3 приведены варианты функций $d(kh) = \hat{D}_M(kh; \bar{c}_{\min}, \bar{c}_{\max})$, соответствующие различным значениям $\bar{c}_{\min}, \bar{c}_{\max}$. Все они имеют одну и ту же асимптотику $d = O(\alpha^{23})$, $\alpha = kh$, но разные интервалы kh , на которых графики этих функций визуально близки к оси абсцисс. При расширении этих интервалов значения функций монотонно возрастают до максимальных значений при $kh = \pi$; последние увеличиваются при этом расширении. В случае кривых 1, 2, 3 действие диссипативного мультиоператора на фурье-гармоники оказывается весьма малым, соответственно, при $kh < \approx 0.6$, $kh < \approx 1.5$, $kh < \approx 2.7$. Для краткости в дальнейшем будем называть эти спектры оператора D_M диссипациями 1, 2, 3.

Рассмотрим вопрос об использовании перечисленных или иных диссипаций. Поскольку действие мультиоператора является частью погрешности аппроксимации и имеет порядок $O(h^{23})$ при любом выборе параметров, значения $\bar{c}_i$ могут быть достаточно произвольными по крайней мере в случае умеренных градиентов точных решений. В этом случае роль мультиоператора сводится в силу его положительности к увеличению запаса устойчивости полудискретной схемы в приближении замороженных коэффициентов. Выбор постоянной C , характеризующей интенсивность диссипации также может быть достаточно произвольным при условии, что ее значение не является чрезмерно большим. Например, она вообще может быть равной нулю, как в случае



Фиг. 3. Фурье-образы диссипативного оператора трех типов.

приведенных выше примеров. В иных случаях достаточно универсальным выбором может быть $C = 1$. Таким образом, можно было бы рассматривать схемы с фиксированными параметрами не только мультиоператора производной, но и оператора диссипации. Однако, в отличие от L_M , параметры диссипации полезно не фиксировать в результате той или иной оптимизации, а рассматривать как свободные параметры, контролирующие ее спектральные свойства. Для упрощения анализа этих свойств будем предполагать, что $\bar{c}_i$ выбираются равномерно распределенными между минимальными и максимальными значениями $\bar{c}_{\min}$, $\bar{c}_{\max}$, обеспечивающими положительность мультиоператора. Такую свободу выбора можно использовать для улучшения монотонных свойств численных решений в случае больших градиентов или разрывов точных решений. В приводимых ниже расчетах использовались пары $\bar{c}_{\min} = 0.5$, $\bar{c}_{\max} = 5$, $\bar{c}_{\min} = -0.3$, $\bar{c}_{\max} = 0.1$ и $\bar{c}_{\min} = -0.3$, $\bar{c}_{\max} = 0.1$, соответствующие кривым 1, 2 и 3 на фиг. 3.

1.5. О граничных условиях

При вычислении действия мультиоператора на известную сеточную функцию согласно процедуре бегущего счета типа (4) для начала счета для каждого значения параметра требуется стартовое значение на границе, например значение v_0 в случае границы $x = x_0$. Являясь аппроксимацией производной на границе, оно является дополнительным граничным условием, необязательно входящим в математическую формулировку исходной задачи. Пусть таковой является задача численного дифференцирования в виде приближенного определения производной в узлах сетки $x_j = jh$, $j = 0, 1, \dots$, по значениям в этих узлах сеточной функции. Тогда значение на границе v_0 , аппроксимирующее производную с тем же порядком, что и внутри области, может быть найдено согласно методики, описанной в [4], [2]. Она состоит в использовании мультиоператоров экстраполяции желаемого порядка значений v_j во внутренних узлах на границу $x = x_0$. Такие мультиоператоры описаны в [2]. Они используют свои базисные операторы и свои параметры.

При случае численных решений краевых задач нет необходимости использовать этот подход. Практически достаточно использовать специфику этих задач. Например, в случае периодических задач проблема задания стартовых значений для бегущего счета отпадает (условия периодичности использовались в тесте 2). В случае удаленных границ эти значения, играющие роль производных на границе, могут задаваться либо из асимптотики точных решений, либо просто полагаться равными нулю (последнее условие использовалось в тесте 1). В других случаях может оказаться приемлемой обычная экстраполяция вместо описанной выше. Во всех случаях влияние начальных значений на решения во внутренних узлах убывает с увеличением расстояния от границы со скоростью убывания членов геометрической прогрессии. Более подробно вопросы граничных условий обсуждаются в [2].

2. О ПРИМЕНЕНИИ МУЛЬТИОПЕРАТОРОВ В СЛУЧАЕ РАЗРЫВНЫХ РЕШЕНИЙ

Описанные выше мультиоператоры являются линейными разностными операторами с постоянными коэффициентами. Это позволяет представить их действие в узле сетки x_j в виде разности “потоков” через границы ячейки $[x_j - h/2, x_j + h/2]$. Например, действие оператора $l(c_i)u_j$ можно записать в виде

$$\begin{aligned} l(c_i)u_j &= (q_{j+1/2} - q_{j-1/2})/h, \\ q_{j+1/2} &= (R_l(c_i)^{-1}u_j + R_r(c_i)^{-1}u_{j+1})/2, \\ q_{j-1/2} &= (R_l(c_i)^{-1}u_{j-1} + R_r(c_i)^{-1}u_j)/2. \end{aligned} \quad (10)$$

Суммирование по всем ячейкам приводит к взаимному уничтожению потоков через их внутренние грани. Суммирование по i с коэффициентами c_i потоков $q_{j\pm 1/2}$ приводит к аналогичному представлению действия мультиоператора L_M . Это свойство позволяет использовать мультиоператорные схемы для аппроксимаций уравнений, записанных в виде законов сохранения. Соответственно, возникает возможность проведения расчетов в случае разрывных решений.

В этом случае основной смысл применения схем повышенных порядков состоит в получении решений, которые при одних и тех же сетках по сравнению с монотонными схемами первого порядка, во-первых, допускают значительно меньшее “размазывание” разрывов, и, во-вторых, обеспечивают большую точность решений в областях их гладкости на удалении от разрывов. Основной проблемой при этом является необходимость предотвращения немонотонности численных решений, следующей из классической теоремы С.К. Годунова.

При применении мультиоператорных схем рассматриваемого класса уменьшение до приемлемых уровней амплитуд возникающих паразитических осцилляций численных решений по крайней мере в случае разрывов сравнительно небольшой интенсивности может быть достигнуто в результате действия оператора диссипации D_M . При этом может оказаться полезным управление спектральными свойствами этого оператора, создающее возможность учета специфики решений конкретной задачи. Ограниченный опыт применения мульт-

тиоператорных схем меньших порядков при решении модельных задач показывает, что численные решения могут либо мало, либо заметно чувствительны к выбору типов диссипации из фиг. 2. При решении уравнений Навье—Стокса с использованием мультиоператоров 9-го порядка с одним из этих типов в случае чисел Маха, не превосходящих два, заметных проявлений немонотонности схемы обнаружено не было [3].

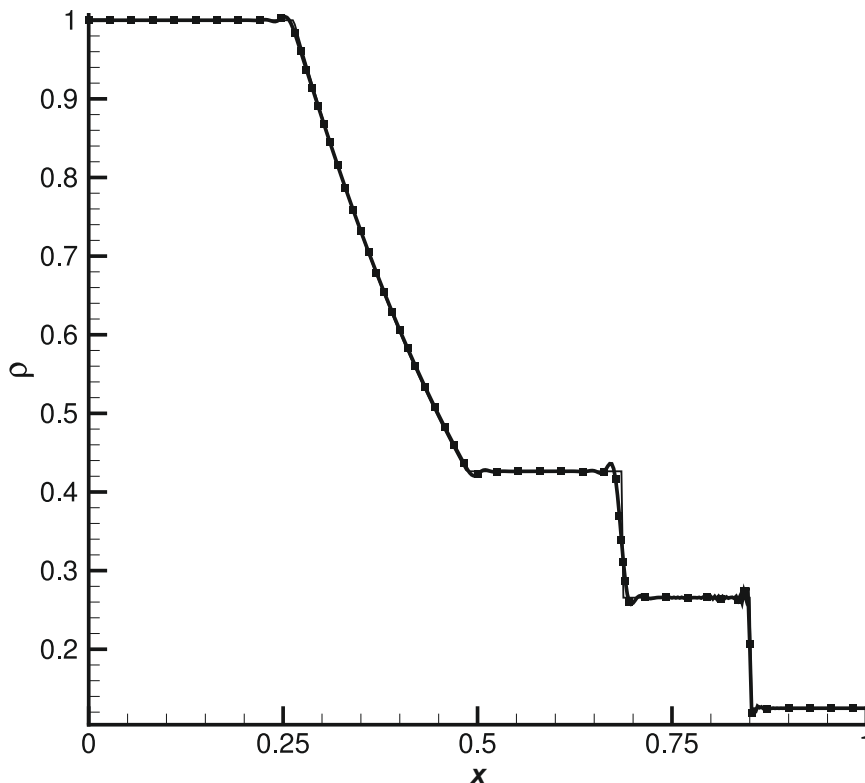
Проблема применения в широком диапазоне чисел Маха мультиоператорных аппроксимаций (в частности, рассмотренных выше), требует отдельного изучения. В данной статье ограничимся лишь несколькими примерами из этой области.

В качестве одного из них рассмотрим известную тестовую задачу о распаде разрыва (задачу Римана) [11], имеющую точное решение. Граничные условия для уравнения Эйлера в этом случае имеют вид

$$\rho = 1, u = 0, p = 1 \quad x < 0.5, \quad \rho = 0.125, u = 0, p = 0.1, \quad x > 0.5, \quad (11)$$

где ρ, u, p , соответственно, плотность, скорость и давление. Время, в которое предлагается получить решение, равно 0.2. Согласно точному решению, от центра области в направлении оси x распространяются ударная волна и с меньшей скоростью контактный разрыв.

Расчеты проводились для нескольких вариантов пар $\bar{c}_{\min}, \bar{c}_{\max}$, в том числе для диссипаций 1, 2, 3. Оказалось, что в случае диссипации 3 на решения, близкие к точным, накладываются коротковолновые колебания с длинами волн $\lambda > 2h$, т.е. более длинными, чем те, на которые эффективно действует мультиоператор. Их не удалось устранить увеличением постоянной C . Решения для диссипации 2 и другими диссипациями с похожими спектральными свойствами в графическом представлении оказались визуально мало различимыми. На фиг. 4 для диссипации 2 представлено распределение плотности в виде кривой с маркерами, а также точное решение в виде тонкой линии. Численное решение было получено на достаточно грубой сетке (число узлов $N = 200$). Видно, что небольшие отклонения от точного решения наблюдаются лишь в малых окрестностях разрывов. Заметим, что во всех расчетных вариантах возникающие в этих окрестностях паразитические колебания быстро затухают при удалении от разрывов. Это соответствует тому, что при выполнении вычислений (4) влияние значения искомой функции в каком-либо узле сетки x_i на вычисляемые значения в последующих узлах $x_j, j > i$, убывает как $O(a^{j-i})$, $a = c_k/(1 + c_k) < 1$, где $c_k > -1/2$ — любое из выбранных значений параметра c .



Фиг. 4. Задача о распаде разрыва: распределение плотности; кривая с маркерами и тонкая линия соответствуют численному и точному решениям.

Для оценки малых отличий численных решений от точных при удалении от разрывов рассмотрим решения в трех интервалах расчетной области. Они образованы следующим образом.левой границей первого и правой границей второго интервала являются границы расчетной области. Третий интервал расположен между разрывами производных, где гладкое решение определяется веером характеристик. Границы интервалов, близкие к точкам разрыва, расположены на удалении 0.1 от них.

В табл. 2 представлены среднеквадратичные отклонения численных решений от точных на этих интервалах для диссипации 1 и 2.

Из табл. 2 следует, что локальные скорости уменьшения малых погрешностей численных решений при удвоении числа узлов сетки заметно выше первого порядка вне области возмущений, создаваемой распадом разрыва. На интервалах 1 и 2 убывания погрешностей характеризуются многими порядками. Хотя на этих интервалах точное решение является тривиальным, тем не менее малость погрешностей и их быстрое убывание дает некоторую оценку точности численных решений. На интервале 3 погрешности убывают значительно медленнее (с локальными порядками между 3 и 4), но при этом остаются достаточно малыми. При сравнении двух диссипаций вторая кажется более выгодной, что указывает на возможность тонкой настройки мультиоператора диссипации. Однако такая настройка в данном случае влияет на точность решений в пределах очень малых величин и может рассматриваться как избыточная.

Ограниченные численные эксперименты показали, что применение линейного диссипативного мультиоператора позволяет получать численные решения с отсутствующими или несущественными проявлениями немонотонности в диапазоне умеренных интенсивностей разрывов. Например, в случае уравнений газовой динамики так происходило при числах Маха, не превышающих два [3]. Однако возможны ситуации, когда диссипативные свойства линейного оператора недостаточны для эффективной монотонизации решений. При этом возникновение значительных паразитических осцилляций в случае уравнений Эйлера или Навье–Стокса может приводить к потере устойчивости мультиоператорных схем. В частности, это может происходить при наличии сильных и экстремально сильных разрывов или очень больших градиентов точных решений.

2.1. Нелинейная монотонизация решений

В случае схем повышенного порядка аппроксимации существует множество подходов к устранению паразитических осцилляций численных решений, вызванных присущим таким схемам свойством немонотонности. Они основаны на той или иной коррекции получаемых в процессе счета сеточных функций или потоков через грани ячеек, образованных сетками. Поскольку реализации мультиоператорных и стандартных схем отличаются лишь алгоритмами вычислений сеточных функций и потоков, можно предположить, что при применении мультиоператоров также может быть использован наработанный опыт монотонизации. Возможно, что при этом не существует универсального подхода. На это указывает сравнение решений для нескольких популярных схем с разными монотонизаторами при их использовании в случае достаточно большого количества модельных задач с сильными разрывами [12]. Оказалось, что победителями этого соревнования для каждой задачи были разные схемы. Можно предположить, что зависимость эффективности монотонизаторов от конкретной задачи существует также в случае мультиоператорных схем. Этот вопрос требует специальных исследований и выходит за рамки данной статьи.

Единственным вариантом монотонизации численных решений в случае мультиоператорных схем было использование коррекции потоков, предложенное в ранних исследованиях [14], [13]. Причиной такого выбора, помимо простоты реализации, была абстрактная форма коррекции, при которой для ограничения потоков абстрактной схемы высокого порядка на каждом шаге по времени используются решения абстрактной монотон-

Таблица 2. Среднеквадратичные отклонения численного решения от точного на трех интервалах при удвоении числа узлов сетки N .

N	Интервал 1	Интервал 2	Интервал 3
200, диссип. 1	1.0e–5	2.6e–6	1.6e–4
400, диссип. 1	1.1e–6	1.8e–8	5.5e–5
800, диссип. 1	1.6e–8	1.0e–10	1.5e–5
200, диссип. 2	2.4e–6	4.5e–7	1.6e–4
400, диссип. 2	4.9e–9	2.5e–9	5.8e–5
800, диссип. 2	1.6e–14	6.3e–13	2.15e–5

ной схемы. Хотя такая форма допускает различные варианты “партнеров” в виде уже монотонизированных схем порядка выше первого, опыт их выбора в случае мультиоператорных аппроксимаций ограничен схемой первого порядка Лакса-Фридрихса (детали приведены в [2]). Методические расчеты в случае мультиоператорных схем меньших порядков показали, что использование ограничителей потоков в случае даже достаточно сильных разрывов позволяли получать решения задач Римана близких к точным без заметных искажений в виде пилообразных осцилляций [2]).

Рассмотрим возможность применения указанной выше коррекции потоков в случае рассматриваемой схемы 24-го порядка без какого-либо анализа ее эффективности. Приведем результаты вычислений для задачи об ударной волне, распространяющейся в покоящейся неоднородной среде. Эта задача, сформулированная в [15] (задача Шу-Ошера), рассматривалась ее авторами как имитация распространения ударных волн в фоновом турбулентном потоке. Ее версия [16] задавалась начальными данными для уравнений Эйлера вида

$$\begin{aligned} -1 \leq x \leq 1, \\ \rho = 3.857143, u = 2.629369, p = 10.33333, x < -0.8, \\ \rho = 1 + 0.2 \sin 5x, u = 0, p = 1, x \geq -0.8. \end{aligned}$$

Поскольку точное решение этой задачи отсутствует, в качестве реферсного решения в [16] рассматривалось численное решение, полученное на основе использованной в расчетах схемы повышенного порядка (схемы WENO) при большом числе узлов сетки.

Рассмотрим прежде всего численные решения этой задачи при применении схемы первого порядка Лакса-Фридрихса. Они не содержат артефактов и могут характеризовать структуру решений при уменьшении шагов сетки.

Вычисленные на основе этой схемы распределения плотности для последовательности сеток в момент времени $t = 0.47$, указанный в [16], представлены на фиг. 5.

Кривые на фиг. 5 описывают решения за волной при числах узлов $N = 1600, 3200, 12800$. Согласно этим решениям, за фронтом волны на некотором интервале присутствуют визуально гладкие колебания плотности с резко выраженными экстремумами. Основные различия между решениями наблюдаются в окрестности этих экстремумов, где максимальные значения плотности возрастают, а минимальные убывают при возрастании числа узлов сетки. Здесь сходимость оказывается достаточно медленной и даже увеличение числа узлов от 12800 до 25600, как показывают расчеты, не приводит к совпадению экстремальных значений на графиках.

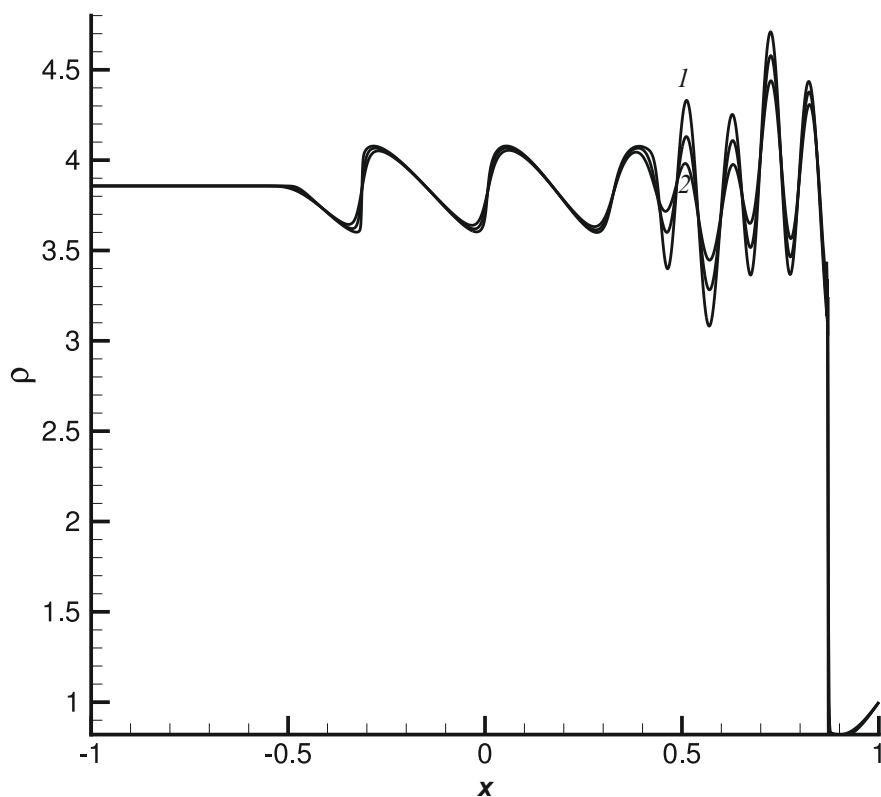
На фиг. 6 представлены распределения плотности, полученные на сетках с числами узлов $N = 400, 800, 1600$ в случае схемы 24-го порядка с коррекцией потоков [13]. Увеличение числа узлов до $N = 3200$ не изменило график решения при $N = 1600$.

На фиг. 7 приведены в увеличенном виде колебания плотности перед скачком в решении задачи Шу-Ошера [15], отличающейся от задачи Шу увеличенными размерами области и временем $t = 1.8$. Видно, что решения для чисел узлов сетки $N = 1600$ и $N = 800$ лишь незначительно отличаются в окрестностях острых экстремумов. Можно предположить, что основная коррекция решений с помощью схемы первого порядка происходит именно в этих окрестностях. Это дает основание для предположения о том, что более быстрой сходимости можно ожидать либо в случае более точной корректирующей схемы, либо при использовании более совершенных монотонизаторов. Поиск таких путей может быть предметом отдельных исследований.

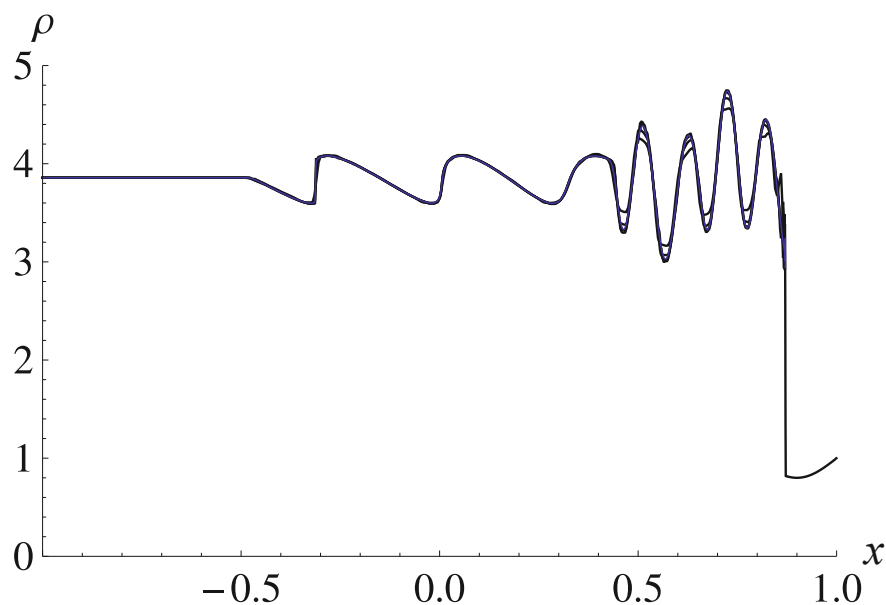
3. ЗАКЛЮЧЕНИЕ

В рамках исследования мультиоператорных аппроксимаций и схем, использующих экономично обрабатываемые двухточечные операторы, рассмотрены аппроксимации 24-го порядка первых производных в задачах с конвективными членами. Основное внимание уделено их спектральным свойствам. Было показано, что в результате оптимизации выбора параметров мультиоператоров можно достичь очень малых отличий аппроксимаций производных и точного дифференцирования в широком диапазоне волновых чисел гармоник, поддерживаемых сетками. Таким образом, помимо очень высокой точности в длиноволновом диапазоне, мультиоператорная схема обладает способностью правильно описывать мелкие детали точных решений исходных уравнений при разумных числах узлов сеток. Для иллюстрации этих свойств приведены примеры решения модельных задач.

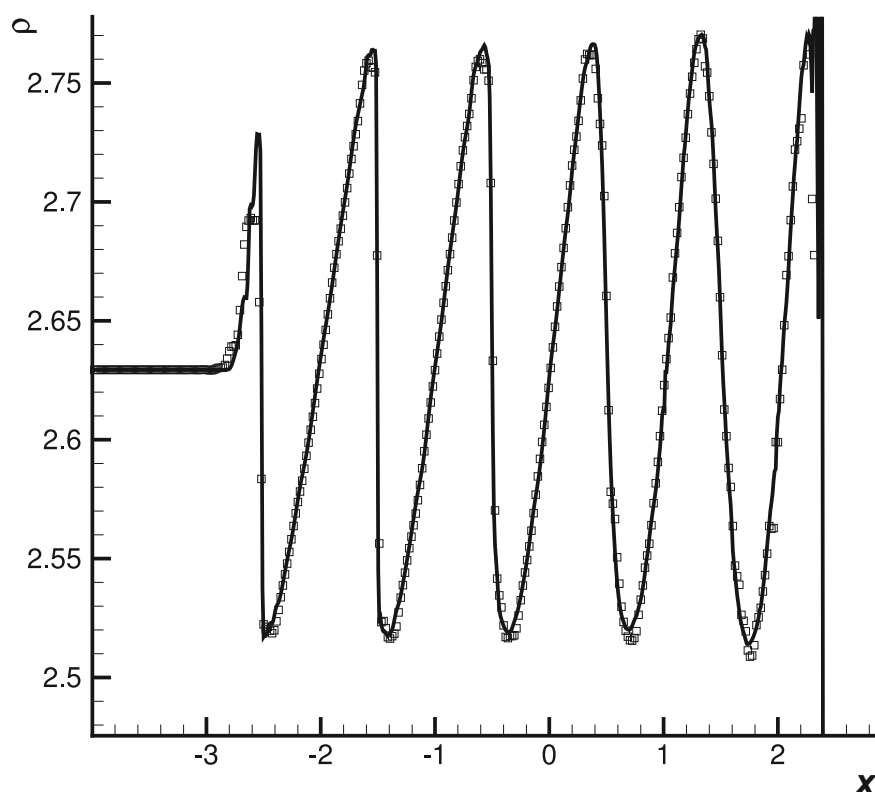
При рассмотрении возможности использования мультиоператорной схемы в случае разрывных решений, основанной на ее консервативности, были представлены результаты вычислений для задачи о распаде разрыва. Был также приведен пример применения нелинейной коррекции схемы, предназначенной для расчетов в случае сильных разрывов. Однако вопросы эффективности мультиоператорных схем и их нелинейных вариантов для задач газовой динамики при больших числах Маха в настоящее время не изучены и требуют специальных исследований.



Фиг. 5. Решения задачи Шу [16] с применением схемы первого порядка: распределения плотности при удвоении числа узлов сетки; кривые 1, 2 с максимальными и минимальными пиками значениями и промежуточная кривая соответствуют числам узлов 1280, 1600 и 3200.



Фиг. 6. Решение задачи Шу [16] с применением схемы 24-го порядка с коррекцией потоков через грани ячеек: распределения плотности при удвоении числа узлов сетки, $t = 0.47$; кривые с максимальными и минимальными пиками значениями и промежуточная кривая соответствуют числам узлов 1600, 400 и 800.



Фиг. 7. Решение задачи Шу-Ошера [15] с применением схемы 24-го порядка с коррекцией потоков через грани ячеек, $t = 1.8$: распределения плотности перед скачком при удвоении числа узлов сетки; сплошная линия и маркеры соответствуют числам узлов сетки 1600 и 800.

Важной чертой мультиоператорных схем рассмотренного класса является возможность достижения высокой параллельной эффективности в случае суперкомпьютерных вычислений. Она следует из того, что основную роль в вычислениях играют арифметические операции, связанные с обращением двухдиагональных матриц согласно процедуре бегущего счета.

СПИСОК ЛИТЕРАТУРЫ

1. *Tolstykh A.I.* Multioperator high-order compact upwind methods for CFD parallel calculations, in: D.R. Emerson et al. (Eds.), *Parallel Computational Fluid Dynamics*, Amsterdam: Elsevier. 1998. P. 383–390.
2. *Толстых А.И.* Компактные и мультиоператорные аппроксимации высокой точности для уравнений в частных производных. М.: Наука, 2015.
3. *Tolstykh A.I., Shirobokov D.A.* Fast calculations of screech using highly accurate multioperators-based schemes // *J. Appl. Acoustics*. 2013. V. 74. P. 102–109.
4. *Tolstykh A. I.* Development of arbitrary-order multioperators-based schemes for parallel calculations. 2. Families of compact approximations with two-diagonal inversions and related multioperators // *J. Comput. Phys.* 2008. V. 227. P. 2922–2940.
5. *Толстых А.И.* О семействах высокоточных мультиоператорных аппроксимаций производных, использующих двухточечные операторы // *Докл. АН*. 2017. Т. 473. № 12, С. 138–141.
6. *Tolstykh A.I., Shirobokov D.A.* Using 16-th Order Multioperators-Based Scheme for Supercomputer Simulation of the Initial Stage of Laminar-Turbulent Transitions // *Communications in Computer and Information Science*, Springer. 2021, 1510 CCIS. P. 270–282.
7. *Lipavskii I. Konshin.* Parallel Implementation of Multioperators-Based Scheme of the 16-th Order for Three-Dimensional Calculation of the Jet Flows. In: *Proc. of International Conference “Russian Supercomputing Days 2022”*, September 26–27, 2022, Moscow: MSU Publishing House. P. 1–13.

8. *Tolstykh A.I., Shirobokov D.A.* Observing production and growth of Tollmien-Schlichting waves in subsonic flat plate boundary layer via excitors-free high fidelity numerical simulation // *J. of Turbulence*. 2020. V. 21. № 11. P. 632–649.
9. *Tam C.K.W.* Problem 1-aliasing, In: *Fourth Computational Aeroacoustics (CAA) Workshop on benchmark problems*, 2004, NASA/CP-2004-2159.
10. *Adams N.A., Shariff K.* A high resolution Compact-ENO schemes for shock-turbulence interaction problems // *J. Comput. Phys.* 1996. V. 127. P. 27–51.
11. *Sod G.A.* A survey of several finite difference schemes for hyperbolic conservation laws // *J. Comput. Phys.* 1978. V. 27. P. 1–31.
12. *Liska R., Wendroff B.* Comparison of several difference schemes on 1D and 2D test problems for the Euler equations // *SIAM J. Sci. Comput.* 2003. V. 26. P. 995–1017.
13. *Zalesac S.T.* Fully multidimensional flux-corrected transport algorithms for fluids // *J. Comput. Phys.* 1979. V. 31. P. 335–362.
14. *Boris J.P., Book D.L.* Flux-Corrected Transport. I. SHASTA, a fluid transport algorithm that works // *J. Comput. Phys.* 1973, V. 11. P. 38–69.
15. *Shu C.W., Osher S.* Efficient implementation of essentially non-oscillatory shock capturing schemes II // *J. Comput. Phys.* 1989, V. 83. P. 32–78.
16. *Balsara D.S., Shu C.-W.* Monotonicity preserving weighted essentially non-oscillatory schemes with increasingly high order of accuracy // *J. Comput. Phys.* 2020. V. 160 P. 405–452.

ON 24TH-ORDER MULTI-OPERATOR APPROXIMATIONS IN SCHEMES FOR EQUATIONS WITH CONVECTIVE TERMS

A. I. Tolstykh*

119333 Moscow, Vavilov Str., 44, Federal Research Center Computer Science and Control, RAS, Russia

**e-mail: tol@ccas.ru*

Received 19.01.2024

Revised 04.04.2024

Accepted 31.05.2024

Abstract. As part of the study of multi-operator approximations and schemes using economically reversible two-point operators, approximations of the 24th order of the first derivatives in problems with convective terms are considered. The main attention is paid to the spectral properties characterizing their high accuracy and resolution. To illustrate these properties, examples of solving model problems are given. The possibilities of using such multi-operator schemes in the case of discontinuous solutions are considered.

Keywords: multioperators, 24th order approximation, Euler equations, problems with discontinuous solutions.

МЕТОДОЛОГИЯ БЫСТРОГО АВТОМАТИЧЕСКОГО ДИФФЕРЕНЦИРОВАНИЯ И УПРАВЛЕНИЕ ТЕПЛОВЫМИ ДИНАМИЧЕСКИМИ СИСТЕМАМИ

© 2024 г. В.И. Зубов^{1,*}

¹ 119333 Москва, ул. Вавилова, 44, ФИЦ ИУ РАН, Россия

*e-mail: vladimir.zubov@mail.ru

Поступила в редакцию 17.03.2024 г.

Переработанный вариант 27.03.2024 г.

Принята к публикации 31.05.2024 г.

В работе описывается методология быстрого автоматического дифференцирования и ее преимущества при численном решении задач оптимального управления. Представлены результаты, полученные при решении с помощью этой методологии задач оптимального управления тепловыми процессами с фазовыми переходами. Библ. 11. Фиг. 7.

Ключевые слова: оптимальное управление, градиент, быстрое автоматическое дифференцирование, уравнение теплопроводности.

DOI: 10.31857/S0044466924090031, **EDN:** WLDAIW

ВВЕДЕНИЕ

При численном решении задач управления сложными динамическими системами возникающая при этом задача оптимального управления обычно сводится к задаче нелинейного программирования. В ВЦ РАН более 50 лет проводятся работы по созданию и совершенствованию численных методов решения такого рода задач. Часто они решаются градиентными методами минимизации целевого функционала. Для этого необходимо уметь вычислять градиент целевой функции конечного числа переменных.

Эффективно и с машинной точностью это позволяет сделать методология быстрого автоматического дифференцирования (БАД-методология), связанная с именем Ю.Г. Евтушенко и впервые достаточно полно представленная в [1]. Дальнейшее ее развитие (см. [2]) привело к тому, что в настоящее время БАД-методология представляет собой современный аппарат, помогающий эффективно решать задачи оптимального управления сложными динамическими системами. В основе БАД-методологии лежит общий подход к дифференцированию сложных функций, возникающих в многоступенчатых процессах. По своей сути эта методология представляет собой метод множителей Лагранжа, который применяется к дискретному варианту задачи оптимального управления. БАД-методология позволяет получить такую аппроксимацию сопряженной задачи, которая согласована с аппроксимацией прямой задачи и с аппроксимацией целевого функционала.

Важность и эффективность использования БАД-методологии при решении задач оптимизации и оптимального управления продемонстрирована, например, в работе [3]. Там показано, с какими трудностями приходится сталкиваться при использовании метода приближенного численного дифференцирования даже для вычисления производных элементарных функций. Эти трудности проиллюстрированы в [3] на примере сложной и практически важной функции, представляющей энергию системы атомов, потенциал взаимодействия которых — потенциал Терсоффа.

В работах [1], [4] сформулировано и обосновано утверждение о том, что время, требуемое для определения градиента целевой функции с помощью БАД-методологии, не превышает времени, требуемого для вычисления трех значений самой функции.

На основе БАД-методологии был разработан и применен на практике эффективный подход к численному решению ряда задач управления динамическими системами. Среди них выделим задачи управления тепловыми процессами с фазовыми переходами. С такого рода задачами (задачи типа Стефана) сталкиваются во многих случаях, из которых важнейшими и наиболее распространенными являются случаи плавления и затвердевания. Существенной чертой таких задач является наличие движущейся поверхности раздела между двумя

фазами (жидкой и твердой). Закон движения этой поверхности заранее неизвестен, и его следует определять. Именно на этой поверхности происходит поглощение или выделение тепла, связанное с фазовым переходом. Термические свойства фаз по обеим сторонам движущейся поверхности могут оказаться различными. Задачи этого класса заметно сложнее тех, в которых отсутствует переход вещества из одной фазы в другую.

Что касается задач оптимального управления процессами с фазовыми переходами, то под ними понимается выбор тех или иных параметров процесса (управляющих параметров) таким образом, чтобы либо сам процесс протекал по сценарию, наиболее близкому к заданному, либо поведение границы раздела жидкой и твердой фаз или функция от значений температуры в некоторой области были наиболее близки к требуемым. К трудностям решения прямых задач здесь прибавляются еще специфические трудности решения задач оптимального управления.

БАД-методология показала свою эффективность и при решении обратных задач, связанных с определением термодинамических параметров исследуемых материалов (см., например, [5]).

Следует также отметить, что предложенный подход с применением БАД-методологии является универсальным и что он может быть успешно применен при исследовании других задач оптимального управления сложными системами.

1. О БАД-МЕТОДОЛОГИИ

В этом разделе приведем канонические соотношения, определяющие сущность БАД-методологии.

Пусть непрерывно-дифференцируемая вектор-функция $\Phi(z, u)$ задает отображение $\Phi : R^n \times R^r \rightarrow R^n$ и пусть имеется связь $\Phi(z, u) = 0_n$, где $z \in R^n$, $u \in R^r$, 0_n есть n -мерный нулевой вектор.

Пусть также задана непрерывно-дифференцируемая скалярная функция $W(z, u)$.

Если функцию $z = z(u)$ определять с помощью соотношения (связи) $\Phi(z, u) = 0_n$, то можно рассматривать сложную функцию $\Omega(u) = W(z(u), u)$. Определить градиент такой сложной функции не всегда просто. БАД-методология позволяет автоматизировать этот процесс.

Представим вычисление функции $z = z(u)$ формально как некоторый многошаговый процесс

$$z_i = F(i, Z_i, U_i), \quad 1 \leq i \leq n, \quad (1.1)$$

где Z_i — множество компонент z_j , встречающихся в правой части равенства (1.1), U_i — множество компонент u_j , встречающихся в правой части равенства (1.1).

Тогда градиент сложной функции $\Omega(u) = W(z(u), u)$ относительно переменной u определяется равенствами

$$\frac{\partial \Omega}{\partial u_i} = W_{u_i}(z, u) + \sum_{q \in K_i} F_{u_i}(q, Z_q, U_q) p_q,$$

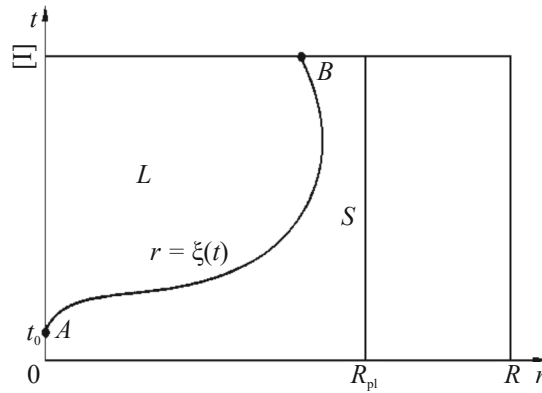
где множители $p_i \in R^n$ представляют собой решение сопряженной задачи

$$p_i = W_{z_i}(z, u) + \sum_{q \in Q_i} F_{z_i}(q, Z_q, U_q) p_q, \\ Q_i = \{j : 1 \leq j \leq n, \quad z_i \in Z_j\}, \quad K_i = \{j : 1 \leq j \leq n, \quad u_i \in U_j\}.$$

Важными преимуществами БАД-методологии являются следующие:

- она поставляет канонические формулы для вычисления градиента функции многих переменных, заданной на некотором многообразии;
- с помощью этих формул градиент вычисляется с машинной точностью;
- вычисление градиента функции с помощью формул БАД-методологии оказывается эффективным по сравнению с другими подходами.

Остановимся на рассмотрении некоторых задач оптимального управления, которые не удалось решить без применения БАД-методологии.



Фиг. 1.

2. УПРАВЛЕНИЕ ПРОЦЕССОМ СВАРКИ МЕТАЛЛОВ

Первая задача, иллюстрирующая применение БАД-методологии, формулируется следующим образом. Пусть имеется некий металлический образец цилиндрической формы и узкий источник энергии, расположенный вдоль оси цилиндра. Благодаря энергии, выделяемой источником, металлический образец нагревается, и часть его, ближайшая к источнику, может расплавиться. Управляя выделяемой энергией, мы можем управлять процессом плавления и кристаллизации металлического образца. Распределение по времени количества выделяемого источником тепла (мощность источника) выбирается в качестве управляющей функции. В задаче требуется найти такое распределение мощности источника по времени, при котором будет расплавлена масса вещества, не меньшая заданной массы, процесс кристаллизации будет протекать со скоростью, не превышающей заданную скорость, и суммарная энергия, выделенная источником за время всего процесса, будет минимальной.

Математическая модель процесса плавления и кристаллизации вещества представляет собой двухфазную начально-краевую задачу типа Стефана. Сформулированная задача исследовалась в рамках одномерной (с радиальной симметрией) нестационарной постановки.

В плоскости независимых переменных (r, t) рассмотрим прямоугольную область $Q = \{(r, t) : 0 < r < R, 0 < t \leq \Xi\}$ (фиг. 1). Область Q гладкой линией AB , уравнение которой есть $r = \xi(t)$, разбивается на две подобласти L (область жидкой фазы) и S (область твердой фазы). Линия AB — траектория движения фронта плавления и кристаллизации. Если $t_0 \geq 0$ — момент времени, при котором зарождается линия AB , то подобласти L и S определяются соотношениями:

$$\begin{aligned} L &= \{(r, t) : 0 < r < \xi(t), t_0 < t \leq \Xi\}, \\ S &= \{(r, t) : \xi(t) < r < R, 0 < t \leq \Xi\}. \end{aligned}$$

В области Q поставим следующую двухфазную задачу Стефана:

$$\begin{aligned} M_L &\equiv \rho_L C_L \frac{\partial T_L}{\partial t} - \frac{1}{r} \frac{\partial}{\partial r} \left(r k_L \frac{\partial T_L}{\partial r} \right) - F(r, t) = 0, & (r, t) \in L, \\ M_S &\equiv \rho_S C_S \frac{\partial T_S}{\partial t} - \frac{1}{r} \frac{\partial}{\partial r} \left(r k_S \frac{\partial T_S}{\partial r} \right) - F(r, t) = 0, & (r, t) \in S, \\ T_S(r, 0) &= T_{\text{in}}(r), & 0 < r < R, \\ T_L[\xi(t), t] &= T_S[\xi(t), t] = T_{\text{pl}}, & t_0 \leq t \leq \Xi, \\ \left[k_S \frac{\partial T_S}{\partial r} \right] \Big|_{(\xi(t)+0, t)} - \left[k_L \frac{\partial T_L}{\partial r} \right] \Big|_{(\xi(t)-0, t)} &= \rho_S \lambda \xi'(t), & t_0 \leq t \leq \Xi, \\ k_S \frac{\partial T_S}{\partial r} \Big|_R &= \alpha [T_{\text{ex}} - T_S(R, t)], & 0 < t \leq \Xi, \\ \frac{\partial T_L}{\partial r}(0, t) &= 0, & t_0 < t \leq \Xi, \quad \frac{\partial T_S}{\partial r}(0, t) = 0, & 0 < t < t_0. \end{aligned} \tag{2.1}$$

Здесь $T(r, t)$ — температура вещества в точке с координатами (r, t) ; ρ, C, k — плотность вещества, его теплоемкость и коэффициент теплопроводности соответственно; λ — теплота плавления вещества; нижние индексы L и S указывают на принадлежность величины жидкой или твердой фазы соответственно; T_{pl} — температура

плавления вещества; $T_{\text{in}}(r)$ — начальная температура вещества, $T_{\text{in}}(r) \leq T_{\text{pl}}$; α — коэффициент теплообмена с окружающей средой; T_{ex} — температура окружающей среды.

Источник подводимого тепла $F(r, t)$ представим в виде $F(r, t) = \phi(r) f(t)$, где $\phi(r)$ — некоторая заданная функция, описывающая распределение выделяемой энергии по пространству.

Пусть $\xi(t)$ — траектория движения фронта раздела фаз, соответствующая источнику $f(t)$, $t \in [0, \Xi]$, и ξ_f — максимальное значение величины $\xi(t)$ при $t_0 \leq t \leq \Xi$. Будем говорить, что функция $f(t)$ принадлежит классу $K(\Xi)$, если она:

- определена и кусочно-непрерывна на отрезке $[0, \Xi]$,
- имеет кусочно-непрерывную производную,
- удовлетворяет ограничениям $0 \leq f(t) \leq f_{\text{max}}$ для всех $t \in [0, \Xi]$,
- соответствующее ей $\xi_f \geq R_{\text{pl}}$, где $R_{\text{pl}} (R_{\text{pl}} < R)$ — заданная величина,
- $\xi'(t) \geq -d^2$ для всех $t \in [0, \Xi - \beta^2]$.

Ограничение сверху на функцию $f(t)$ может отсутствовать. Для заданной конечной величины f_{max} значение величины Ξ не должно быть меньше некоторого определенного значения, иначе класс $K(\Xi)$ окажется пустым.

Вариационную задачу можно сформулировать следующим образом: среди функций $f(t)$ из класса $K(\Xi)$ найти такую функцию $f_{\text{opt}}(t)$, что функционал

$$J = \int_0^{\Xi} f(t) dt \quad (2.2)$$

достигает на ней минимального значения.

При решении прямой задачи осуществляется переход от неизвестной функции температуры $T(r, t)$ к функции теплосодержания $E(r, t)$, которая определяется через температуру следующим соотношением:

$$E(T) = \begin{cases} \rho_S C_S T, & T < T_{\text{pl}}, \\ \rho_L C_L (T - T_{\text{pl}}) + \rho_S \lambda + \rho_S C_S T_{\text{pl}}, & T_{\text{pl}} \leq T. \end{cases}$$

Функция $E(T)$ в точке плавления вещества T_{pl} терпит разрыв. Если считать функцию теплосодержания $E(r, t)$ основной, а температуру определять через нее с помощью равенства

$$T(E) = \begin{cases} E \rho_S^{-1} C_S^{-1}, & E < E_- = \rho_S C_S T_{\text{pl}}, \\ T_{\text{pl}}, & E_- \leq E \leq E_+ = E_- + \rho_S \lambda, \\ \frac{E + (\rho_L C_L - \rho_S C_S) T_{\text{pl}} - \rho_S \lambda}{\rho_L C_L}, & E_+ < E, \end{cases}$$

то температура как функция теплосодержания будет уже функцией непрерывной (эта функция является обратной к функции $E(r, t)$).

Коэффициент теплопроводности в общем случае зависит от температуры и при переходе от твердой фазы к жидкой (при температуре плавления) испытывает скачок. В предлагаемом алгоритме коэффициент теплопроводности рассматривается как функция теплосодержания и определяется формулой:

$$\Omega(E) = k[T(E)] = \begin{cases} k_S, & E < E_-, \\ k_S + \frac{(E - E_-)(k_L - k_S)}{(E_+ - E_-)}, & E_- \leq E \leq E_+, \\ k_L, & E_+ < E. \end{cases}$$

Задача (2.1) в терминах функции теплосодержания $E(r, t)$ принимает вид

$$\begin{aligned} \frac{\partial E}{\partial t} &= \frac{1}{r} \frac{\partial}{\partial r} \left(r \Omega \frac{\partial T(E)}{\partial r} \right) + F(r, t), & (r, t) \in Q, \\ E(r, 0) &= E[T_{\text{in}}(r)], & 0 < r < R, \\ \frac{\partial E}{\partial r} \Big|_{r=0} &= 0, & 0 \leq t \leq \Xi, \\ \Omega(E) \frac{\partial T(E)}{\partial r} \Big|_{r=R} &= \alpha \{ T_{\text{ex}} - T[E(R, t)] \}, & 0 \leq t \leq \Xi. \end{aligned} \quad (2.3)$$

Для аппроксимации краевой задачи (2.3) в области Q вводится неравномерная сетка $\omega = \{r_i, t_j\}$. Использование интегроинтерполяционного метода и неявной аппроксимации по времени приводит к системе конечно-разностных уравнений, аппроксимирующих краевую задачу (2.3). Получившаяся система нелинейных алгебраических уравнений решалась с помощью нового итерационного алгоритма, предложенного в [6]. Заканчивается процесс решения задачи (2.3) выделением траектории движения фронта раздела фаз.

Решение оптимизационной задачи проводилось численно с использованием градиентных методов спуска. Ограничения, накладываемые на решение задачи, учитывались с помощью метода штрафных функций. В результате этого минимизировался функционал $I = J + g(\xi_f) + \Theta$, где J — целевой функционал (2.2), $g(r) = A_0 \cdot (r - R_{pl})^2$ — штрафной функционал, отвечающий за выполнение условия $\xi_f = R_{pl}$, и $\Theta = \int_0^{\infty} A(t) \cdot \left(\frac{d\xi}{dt} + d^2 \right) dt$ — штрафной функционал, обеспечивающий допустимую скорость остывания,

$$A(t) = \begin{cases} 0, & \left(\frac{d\xi}{dt} + d^2 \right) \geq 0, \\ A_0(t), & \left(\frac{d\xi}{dt} + d^2 \right) < 0. \end{cases}$$

Для аппроксимации обобщенного функционала I применялся метод прямоугольников, в результате чего получалась функция конечного числа переменных $\tilde{I}(f_1, \dots, f_M)$.

Компоненты градиента аппроксимированной целевой функции $\tilde{I}$ по компонентам вектора управления $\{f^j\}$ вычислялись с помощью методологии быстрого автоматического дифференцирования:

$$\frac{\partial \tilde{I}}{\partial f^j} = \tau^j + \sum_{i=0}^w \tau^j p_i^j \phi_i, \quad 1 \leq j \leq M,$$

где w — порядковый номер узла пространственной сетки, удовлетворяющий условию

$$\phi(r) = \begin{cases} \phi_w(r) \neq 0, & 0 \leq r \leq r_w, \\ 0, & r > r_w. \end{cases}$$

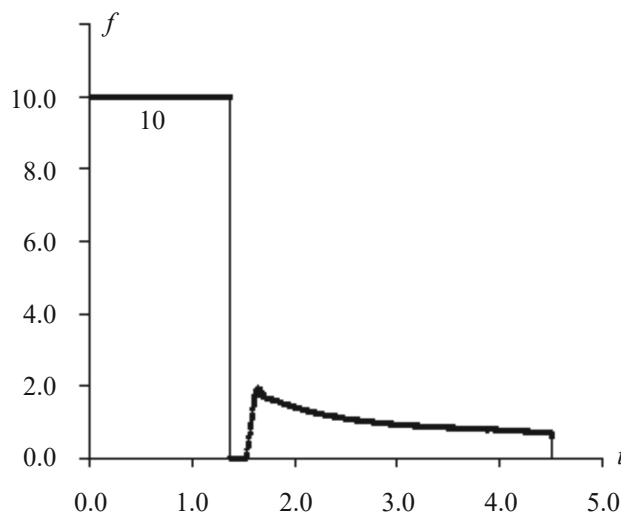
Переменные p_i^j — сопряженные переменные (импульсы), которые должны определяться в результате решения системы уравнений, аппроксимирующих непрерывную сопряженную задачу. Но не любая (даже высокой точности) аппроксимация сопряженной задачи позволяет получить точное значение градиента целевой функции. Аппроксимация сопряженной задачи должна быть согласованной с аппроксимацией прямой задачи (2.3) и с аппроксимацией целевого функционала (2.2). Такую согласованную аппроксимацию сопряженной задачи позволяет автономно получить БАД-методология. Вывод формул аппроксимации сопряженной задачи достаточно громоздок и приведен в работе [7]. Отметим, что использование БАД-методологии позволяет вычислять градиент целевого функционала с машинной точностью.

Анализ результатов проведенных численных расчетов оптимизационной задачи позволил сделать вывод о структуре оптимального управления (фиг. 2):

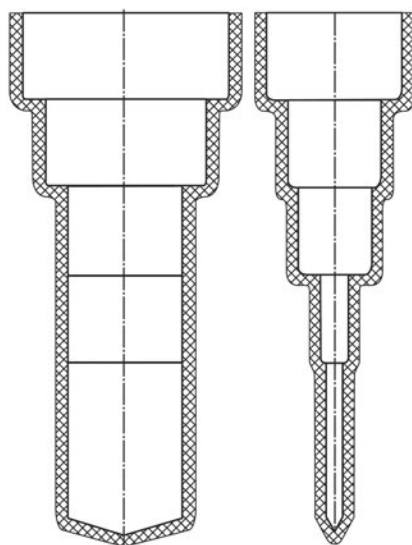
- оптимальное управление состоит из двух основных частей;
- первая часть (ответственная, в основном за плавление вещества) совпадает с верхним участком краевого экстремума $f(t) \equiv f_{\max}$;
- вторая часть оптимального управления (ответственная за кристаллизацию вещества) меньше первой части (если сравнивать их средние значения) и отделяется от нее небольшим участком с $f(t) \equiv 0$. Следует отметить, что включение источника на этапе кристаллизации происходит раньше того момента времени, когда скорость движения фронта кристаллизации достигает максимально допустимой величины (нередко это включение источника происходит на этапе плавления).

Основываясь на результатах проведенных исследований, был предложен алгоритм приближенного построения оптимального управления в рассмотренной выше задаче управления процессом плавления и кристаллизации вещества. А именно, в том диапазоне параметров, в котором проводились исследования, оптимальное управление может быть определено в результате последовательного решения двух вспомогательных задач: вначале задачи плавления, а затем, используя ее результаты как начальные данные, задачи кристаллизации. Использование такого разбиения при решении полной вариационной задачи существенно экономит затраты на получение оптимального решения.

Следует подчеркнуть, что применение БАД-методологии при решении этой задачи оказало существенное влияние на скорость получения результата и на его точность.



Фиг. 2.



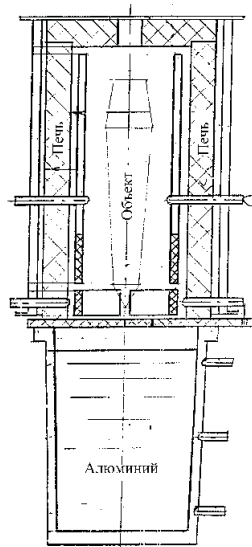
Фиг. 3.

3. УПРАВЛЕНИЕ ПРОЦЕССОМ КРИСТАЛЛИЗАЦИИ МЕТАЛЛА В ЛИТЕЙНОМ ДЕЛЕ

Другая задача, при решении которой использование БАД-методологии оказалось необходимым, относится к литейному делу.

Одним из важных этапов литейного дела является кристаллизация расплавленного металла. От того, как протекал процесс затвердевания жидкого металла, зависит качество полученного образца. Установлено, что для получения образца хорошего качества желательно, чтобы поверхность раздела фаз была бы как можно ближе к плоской, и чтобы скорость ее движения была бы небольшой.

Процесс кристаллизации металла начинается с того, что в рабочую полость литейной формы с определенной конфигурацией внешней границы и рабочей полости (продольные проекции реального объекта представлены на фиг. 3) заливают жидкий металл. Литейная форма подогрета до заданной температуры T_{form} , а залитый в нее металл до температуры T_{met} . После этого под воздействием изменяющихся внешних условий начинается постепенное охлаждение объекта (литейная форма с залитым в нее металлом). Для управления этим процессом используется специальная установка (фиг. 4). Она состоит из верхней и нижней частей. Верхняя часть представляет собой плавильную печь, внутри которой перемещается объект. Она моделируется двумя вертикальными, расположенными друг против друга стенками, которые соединены сверху горизонтальной стенкой ("крыша").



Фиг. 4.

Стенки печи и “крыша” разогреты до заданной сравнительно высокой температуры. Нижняя часть установки является охладителем и состоит из большой емкости, заполненной жидким алюминием.

В работе рассматривается вариант, когда две боковые стенки литейной формы (с тех сторон, где отсутствуют стенки печи) изготовлены из теплоизолированного материала. Такая модель литейной формы описывает ситуацию, когда в рассматриваемой установке обрабатываются несколько литейных форм, расположенных в ряд недалеко друг от друга. Литейная форма, заполненная жидким металлом, медленно погружается в охладитель. Жидкий алюминий имеет сравнительно низкую температуру, благодаря чему происходит кристаллизация металла. Однако объект получает тепло от стенок плавильной печи, что не позволяет процессу кристаллизации протекать слишком быстро.

Процесс остывания металла и литейной формы описывается нелинейным уравнением теплопроводности:

$$\frac{\partial G}{\partial t} = \frac{\partial}{\partial x} \left(K \frac{\partial T}{\partial x} \right) + \frac{\partial}{\partial y} \left(K \frac{\partial T}{\partial y} \right) + \frac{\partial}{\partial z} \left(K \frac{\partial T}{\partial z} \right), \quad (x, y, z) \in Q. \quad (3.1)$$

Здесь Q — область литейной формы с залитым в нее металлом, имеющая кусочно-гладкую границу Γ ; T — температура вещества в точке с координатами (x, y, z) в момент времени t . Функция теплосодержания $G(T(x, y, z, t))$ определяется соотношениями:

$$G(T(x, y, z, t)) = \begin{cases} G_1(T), & (x, y, z) \in \text{металлу}, \\ G_2(T), & (x, y, z) \in \text{форме}, \end{cases}$$

$$G_1(T) = \begin{cases} \rho_S c_S T, & T < T_1, \\ \left[\rho_S c_S + \rho_S \lambda (T_2 - T_1)^{-1} \right] T - \rho_S \lambda T_1 (T_2 - T_1)^{-1}, & T_1 \leq T < T_2, \\ \rho_L c_L (T - T_2) + \rho_S c_S T_2 + \rho_S \lambda, & T \geq T_2, \end{cases} \quad G_2(T) = \rho_F c_F T,$$

λ — удельная теплота плавления. Коэффициент теплопроводности имеет вид:

$$K(T) = \begin{cases} K_1(T), & (x, y, z) \in \text{металлу}, \\ k_F, & (x, y, z) \in \text{форме}, \end{cases} \quad K_1(T) = \begin{cases} k_S, & T < T_1, \\ \frac{k_L - k_S}{T_2 - T_1} T + \frac{k_S T_2 - k_L T_1}{T_2 - T_1}, & T_1 \leq T < T_2, \\ k_L, & T \geq T_2. \end{cases}$$

Константы c_S , c_L , c_F , ρ_S , ρ_L , ρ_F , k_S , k_L , k_F , T_1 , T_2 заданы. Термодинамические коэффициенты претерпевают разрыв на границе металл-форма. На этой границе требуется выполнение двух условий: непрерывность температуры и потока тепла. Интервал $[T_1, T_2]$ — узкий интервал температур, где термодинамические коэффициенты и функция теплосодержания изменяются очень быстро.

При пересечении поверхности раздела фаз температура остается непрерывной, а тепловой поток претерпевает скачок такой величины: $\left[K \frac{\partial T}{\partial n} \right] = \rho_S \lambda v_n$, (v_n — скорость поверхности раздела фаз).

Процесс остывания литейной формы и находящегося в ней металла происходит благодаря воздействию окружающей среды на объект. Одним из механизмов переноса тепла в данной задаче является тепловое излучение (см. [8]), а именно, потеря тепловой энергии тела за счет собственного излучения, приобретение энергии за счет излучения окружающего жидкого алюминия, приобретение энергии за счет излучения стенок печи и “крыши”, приобретение энергии за счет излучения поверхности жидкого алюминия. Кроме того, учитывается обмен тепловой энергией за счет теплопередачи между жидким алюминием и объектом.

На границе Γ области Q задаются условия теплообмена с внешней средой. Как отмечалось выше, эти условия зависят от рассматриваемой точки поверхности и времени. Однако все условия теплообмена можно записать в таком общем виде:

$$\tilde{\alpha}T + \tilde{\beta}\frac{\partial T}{\partial \mathbf{n}} = \tilde{\gamma}. \quad (3.2)$$

В соотношении (3.2) $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{\gamma}$ — известные функции координат (x, y, z) точки поверхности Γ и температуры $T(x, y, z, t)$, а под $\frac{\partial T}{\partial \mathbf{n}} = T_{\mathbf{n}}$ понимается производная от температуры T по направлению внешней нормали $\mathbf{n}$ к поверхности Γ .

Решение прямой задачи состоит в определении функции $T(x, y, z, t)$, удовлетворяющей уравнению (3.1) в рассматриваемой области Q , условиям (3.2) на внешней границе Γ области Q , двум условиям на поверхности раздела фаз (температура должна быть непрерывной, а тепловой поток должен испытывать указанный выше скачок) и двум условиям на границе металл — форма (непрерывность температуры и теплового потока).

Для решения прямой задачи исследуемый объект аппроксимировался телом, которое состоит из конечного числа параллелепипедов. Это тело образно помещается в некоторый вспомогательный параллелепипед, размеры которого совпадают с габаритами исследуемого объекта. Система координат связана с движущейся литейной формой. Ось Oz направлена вертикально вверх, ось Ox располагается в горизонтальной плоскости и направлена слева направо, а ось Oy выбрана так, чтобы система координат $Oxyz$ была бы правой. Начало O этой системы координат — это левая, ближайшая к наблюдателю вершина вспомогательного параллелепипеда, расположенная на его дне.

Вводятся временная сетка и две пространственные разностные сетки (в общем случае, неравномерные). Основная сетка построена таким образом, что все внешние поверхности аппроксимирующего тела, а также все границы металл — форма являются одной из координатных поверхностей этой сетки. Поверхности вспомогательной сетки параллельны поверхностям основной сетки и узлы этой сетки расположены на серединах отрезков, соединяющие узлы основной сетки. Плоскости $x = \tilde{x}_n$, $y = \tilde{y}_i$, $z = \tilde{z}_l$ делят объект на так называемые элементарные объемы, или элементарные ячейки. Элементарной ячейке приписываются индексы (n, i, l) , если вершина этой ячейки, ближайшая к началу координат, совпадает с узловой точкой $(\tilde{x}_n, \tilde{y}_i, \tilde{z}_l)$. Пусть V_{nil} — объем такой элементарной ячейки, а S_{nil} — ее внешняя поверхность.

Считается, что в рамках одной элементарной ячейки температура среды не зависит от пространственных координат, а зависит только от времени. Обозначается эта температура как $T_{nil}(t)$.

Любая элементарная ячейка либо полностью заполнена только одной средой (металл или литейная форма), либо одна ее часть заполнена одной средой, а другая ее часть другой средой. Обозначим через V_{nil}^1 часть объема V_{nil} элементарной ячейки, которая заполнена металлом, а через V_{nil}^2 — часть объема V_{nil} , которая заполнена материалом формы. Аналогично, S_{nil}^1 — часть внешней поверхности S_{nil} элементарной ячейки, которая прилегает к объему V_{nil}^1 , а S_{nil}^2 — часть внешней поверхности S_{nil} , которая прилегает к объему V_{nil}^2 .

В том случае, когда рассматриваемый объект параллелепипед, все элементарные ячейки также имеют форму параллелепипеда. Если же рассматривается объект сложной геометрической формы, то на стыке разных частей объекта возникают новые элементарные ячейки более сложной конфигурации. Ячейки такого вида всегда выходят на внешнюю границу литейной формы, и в этом случае усложняется конфигурация поверхности S_{nil}^2 .

Для любого объема V с внешней границей S , выделенного из объекта, справедлив закон теплового баланса, который состоит в том, что изменение теплосодержания в объеме V за фиксированный промежуток времени τ равняется количеству тепла, проходящему через поверхность S объема V за этот же промежуток времени. Поэтому, учитывая независимость температуры вещества от пространственных координат в пределах расчетной ячейки, имеем

$$\begin{aligned} & \left[V_{nil}^1 G_1(T_{nil}^{j+1}) + V_{nil}^2 G_2(T_{nil}^{j+1}) \right] - \left[V_{nil}^1 G_1(T_{nil}^j) + V_{nil}^2 G_2(T_{nil}^j) \right] = \\ & = \int_{t^j}^{t^{j+1}} \left[\iint_{S_{nil}^1} K_1(\tilde{T}_{nil}(t)) \cdot (\tilde{T}_n(t))_{nil} ds + \iint_{S_{nil}^2} K_2(\tilde{T}_{nil}(t)) \cdot (\tilde{T}_n(t))_{nil} ds \right] dt, \quad (3.3) \end{aligned}$$

где $T_{nil}^j = T_{nil}(t^j)$, а $K_1(\tilde{T}_{nil}(t)) \cdot (\tilde{T}_n(t))_{nil}$ и $K_2(\tilde{T}_{nil}(t)) \cdot (\tilde{T}_n(t))_{nil}$ — плотности теплового потока через внешнюю поверхность расчетной ячейки для металла и для материала формы соответственно.

В основе алгоритма численного решения прямой задачи лежит использование уравнения теплового баланса и переход от формулировки задачи в терминах температуры к формулировке в терминах теплосодержания. Дальнейшие преобразования соотношения (3.3) осуществляются подобно тому, как это было сделано в первом разделе. Однако теперь необходимо учитывать тот факт, что рассматриваемая расчетная область неоднородна (содержит и металл, и материал формы). Для того чтобы лучше учитывать геометрию расчетных ячеек, и то, как они заполнены, введено понятие так называемой “суммарной плотности теплосодержания” в ячейке.

Пусть $\Phi_{nil} = V_{nil}^2/V_{nil}$ — объемная доля материала формы в расчетной ячейке с индексами (n, i, l) , а $M_{nil} = V_{nil}^1/V_{nil}$ — объемная доля металла в этой ячейке. Тогда “суммарная плотность теплосодержания” в ячейке с индексами (n, i, l) в момент времени t^j определяется формулой: $E_{nil}^j = M_{nil}G_1(T_{nil}^j) + \Phi_{nil}G_2(T_{nil}^j)$. Принимая во внимание выражения для $G_1(T)$ и $G_2(T)$, получаем следующее представление для новой функции теплосодержания $E_{nil}^j(T_{nil}^j)$:

$$E_{nil}^j(T_{nil}^j) = \begin{cases} b_{nil}T_{nil}^j, & T_{nil}^j < T_1, \\ p_{nil}^1T_{nil}^j - p_{nil}^2, & T_1 \leq T_{nil}^j < T_2, \\ r_{nil}^1T_{nil}^j + r_{nil}^2, & T_{nil}^j \geq T_2, \end{cases}$$

где

$$b_{nil} = M_{nil}\rho_{SCS} + \Phi_{nil}\rho_{FCF}, \quad p_{nil}^1 = M_{nil}\left(\rho_{SCS} + \frac{\rho_S\lambda}{(T_2 - T_1)}\right) + \Phi_{nil}\rho_{FCF},$$

$$p_{nil}^2 = \frac{M_{nil}\rho_S\lambda T}{(T_2 - T_1)}, \quad r_{nil}^1 = M_{nil}\rho_{LC}L + \Phi_{nil}\rho_{FCF}, \quad r_{nil}^2 = M_{nil}(\rho_S\lambda + (\rho_{SCS} - \rho_{LC}L)T_2).$$

Определим температуру T_{nil}^j как функцию, зависящую от теплосодержания E_{nil}^j :

$$T_{nil}^j \equiv \beta(E_{nil}^j) = \begin{cases} E_{nil}^j/b_{nil}, & E_{nil}^j < b_{nil}T_1, \\ (E_{nil}^j + p_{nil}^2)/p_{nil}^1, & b_{nil}T_1 \leq E_{nil}^j < r_{nil}^1T_2 + r_{nil}^2, \\ (E_{nil}^j - r_{nil}^2)/r_{nil}^1, & E_{nil}^j \geq r_{nil}^1T_2 + r_{nil}^2. \end{cases}$$

Эта функция является обратной к функции $E_{nil}^j(T_{nil}^j)$. Коэффициенты теплопроводности $K_1(T_{nil}^j)$ и $K_2(T_{nil}^j)$ также представим как функции от E_{nil}^j :

$$K_1(T_{nil}^j) \equiv W_1(E_{nil}^j) = \begin{cases} k_S, & E_{nil}^j < E_1, \\ (E_2 - E_1)^{-1}[(k_L - k_S)E_{nil}^j + (k_SE_2 - k_LE_1)], & E_1 \leq E_{nil}^j < E_2, \\ k_L, & E_{nil}^j \geq E_2, \end{cases}$$

$$K_2(T_{nil}^j) \equiv W_2(E_{nil}^j) = \begin{cases} k_{\Phi_1}, & E_{nil}^j < E_3, \\ (E_4 - E_3)^{-1}[(k_{\Phi_2} - k_{\Phi_1})E_{nil}^j + (k_{\Phi_1}E_4 - k_{\Phi_2}E_3)], & E_3 \leq E_{nil}^j < E_4, \\ k_{\Phi_2}, & E_{nil}^j \geq E_4, \end{cases}$$

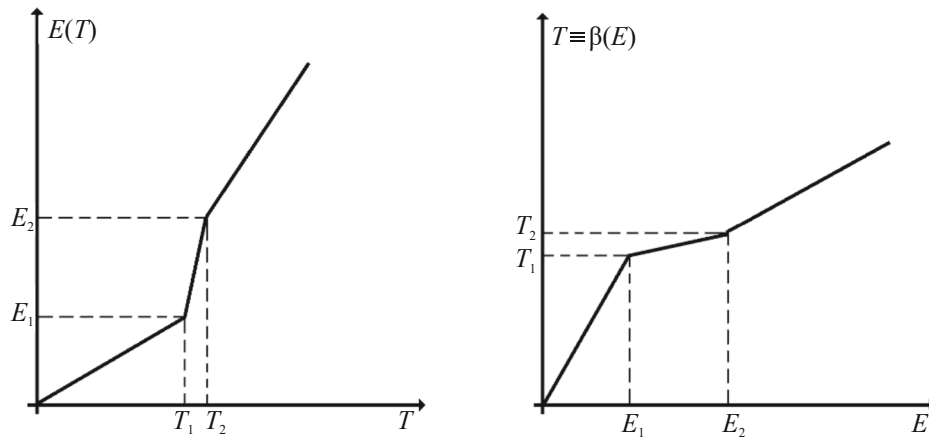
$$E_1 \equiv \rho_{SCS}T_1, \quad E_2 \equiv \rho_{SCS}T_2 + \rho_S\lambda, \quad E_3 \equiv \rho_{FCF}(T_3 - \delta), \quad E_4 \equiv \rho_{FCF}(T_3 + \delta).$$

Величина $\delta \ll T_3$ — произвольная, но фиксированная (в расчетах полагалось $\delta = 20$).

Функция $E_{nil}^j(T_{nil}^j)$ как функция, зависящая от температуры, в металле ведет себя так же, как и функция теплосодержания $G_1(T)$, т.е. в узком интервале температур $[T_1, T_2]$ меняется очень быстро (фиг. 5). По этой причине итерационные алгоритмы решения полученных систем конечно-разностных уравнений плохо сходятся.

Температура T_{nil}^j как функция от E_{nil}^j , меняется не так быстро (фиг. 5), и при выполнении определенных условий численные алгоритмы решения систем уравнений гарантированно сходятся (см. [6]). Из этих соображений в равенстве (3.3) перейдем от переменной $T_{nil}(t)$ к переменной $E_{nil}(t)$:

$$V_{nil} \cdot (E_{nil}^{j+1} - E_{nil}^j) = \int_{t^j}^{t^{j+1}} \left[\iint_{S_{nil}^1} H_1(E_{nil}(t)) ds + \iint_{S_{nil}^2} H_2(E_{nil}(t)) ds \right] dt, \quad (3.4)$$



Фиг. 5.

где

$$H_1(E_{nil}(t)) = W_1(\tilde{E}_{nil}(t)) \cdot \beta_n(\tilde{E}_{nil}(t)), \quad H_2(E_{nil}(t)) = W_2(\tilde{E}_{nil}(t)) \cdot \beta_n(\tilde{E}_{nil}(t)),$$

$$(n = \overline{n^*(m), N^*(m)}; \quad i = \overline{i^*(m), I^*(m)}; \quad l = \overline{l^*(m), L^*(m)}; \quad j = \overline{0, J-1}; \quad m = \overline{1, M}).$$

Соотношение (3.4) есть не что иное, как уравнение теплового баланса, записанное в терминах новой функции $E_{nil}(t)$ для каждой расчетной ячейки исследуемого объекта. Дискретизация по времени и аппроксимация по пространству соотношения (3.4) приводят к системе конечно-разностных уравнений относительно неизвестной E_{nil}^{j+1} .

Для дискретизации по времени уравнения (3.4) использовалась схема Писмена–Рекфорда.

Для схемы Писмена–Рекфорда характерен тот факт, что рассматриваемый интервал времени $(t^j, t^j + \tau)$ разбивается на три подинтервала: $(t^j, t^j + \tau/3)$, $(t^j + \tau/3, t^j + 2\tau/3)$ и $(t^j + 2\tau/3, t^j + \tau)$. В указанном выше соотношении $E_{nil}^{j+1/3} = E_{nil}(t^j + \tau/3)$, $E_{nil}^{j+2/3} = E_{nil}(t^j + 2\tau/3)$.

Для того чтобы более компактно записать эту разностную схему, введем следующие обозначения:

$$\tilde{\Lambda} E = \iint_{S_{nil}^{1x+}} H_1(E(t)) ds + \iint_{S_{nil}^{2x+}} H_2(E(t)) ds, \quad \tilde{\Lambda}_x E = \iint_{S_{nil}^{1x+} \cup S_{nil}^{1x-}} H_1(E) ds + \iint_{S_{nil}^{2x+} \cup S_{nil}^{2x-} \cup S_{nil}^{2xd}} H_2(E) ds,$$

$$\tilde{\Lambda}_y E = \iint_{S_{nil}^{1y+} \cup S_{nil}^{1y-}} H_1(E) ds + \iint_{S_{nil}^{2y+} \cup S_{nil}^{2y-} \cup S_{nil}^{2yd}} H_2(E) ds, \quad \tilde{\Lambda}_z E = \iint_{S_{nil}^{1z+} \cup S_{nil}^{1z-}} H_1(E) ds + \iint_{S_{nil}^{2z+} \cup S_{nil}^{2z-} \cup S_{nil}^{2zd}} H_2(E) ds.$$

Тогда временная аппроксимация уравнения теплового баланса по схеме Писмена–Рекфорда принимает вид:

$$V_{nil} (E_{nil}^{j+1} - E_{nil}^j) = \frac{2\tau}{3} \tilde{\Lambda}_x E_{nil}^{j+1/3} + \frac{\tau}{3} \tilde{\Lambda}_x E_{nil}^{j+2/3} +$$

$$+ \frac{\tau}{3} \tilde{\Lambda}_y E_{nil}^j + \frac{2\tau}{3} \tilde{\Lambda}_y E_{nil}^{j+2/3} + \frac{\tau}{3} \tilde{\Lambda}_z E_{nil}^j + \frac{\tau}{3} \tilde{\Lambda}_z E_{nil}^{j+1/3} + \frac{\tau}{3} \tilde{\Lambda}_z E_{nil}^{j+1}, \quad (3.5)$$

$$(n = \overline{n^*(m), N^*(m)}; \quad i = \overline{i^*(m), I^*(m)}; \quad l = \overline{l^*(m), L^*(m)}; \quad j = \overline{0, J-1}; \quad m = \overline{1, M}).$$

В левую часть уравнения (3.5) добавим и вычтем значения: $V_{nil} E_{nil}^{j+1/3}$ и $V_{nil} E_{nil}^{j+2/3}$. После этого разобьем это уравнение на три равенства, и тем самым образуем три одномерных по пространству подзадачи:

x направление

$$V_{nil} \cdot (E_{nil}^{j+1/3} - E_{nil}^j) = \frac{\tau}{3} \tilde{\Lambda}_x E_{nil}^{j+1/3} + \frac{\tau}{3} \tilde{\Lambda}_y E_{nil}^j + \frac{\tau}{3} \tilde{\Lambda}_z E_{nil}^j,$$

y направление

$$V_{nil} \cdot (E_{nil}^{j+2/3} - E_{nil}^{j+1/3}) = \frac{\tau}{3} \tilde{\Lambda}_y E_{nil}^{j+2/3} + \frac{\tau}{3} \tilde{\Lambda}_x E_{nil}^{j+1/3} + \frac{\tau}{3} \tilde{\Lambda}_z E_{nil}^{j+1/3},$$

z направление

$$V_{nil} \cdot (E_{nil}^{j+1} - E_{nil}^{j+2/3}) = \frac{\tau}{3} \tilde{\Lambda}_z E_{nil}^{j+1} + \frac{\tau}{3} \tilde{\Lambda}_x E_{nil}^{j+2/3} + \frac{\tau}{3} \tilde{\Lambda}_y E_{nil}^{j+2/3},$$

$$(n = \overline{n^*(m)}, N^*(m); \quad i = \overline{i^*(m)}, I^*(m); \quad l = \overline{l^*(m)}, L^*(m); \quad j = \overline{0, J-1}; \quad m = \overline{1, M}).$$

Пространственная аппроксимация указанных выше подзадач приведена в работе [9].

Полученная в результате пространственной аппроксимации первой подзадачи, система нелинейных уравнений состоит из множества других систем. Для каждой из них значения индексов i и l — фиксированы. Эти подсистемы уравнений решаются независимо друг от друга предложенным в [10] новым итерационным методом с привлечением метода прогонки. Значения коэффициентов во всех уравнениях определяются по полученным значениям функции теплосодержания в момент времени $(j + 1/3)$ на предыдущей итерации. Итерационный процесс прекращается при достижении требуемой точности. При численном решении задачи требовалось, чтобы максимальное относительное отклонение между двумя последующими итерациями не превосходило 10^{-12} .

Системы нелинейных уравнений, полученные в результате пространственной аппроксимации второй и третьей подзадач, решаются подобным образом. В качестве начальных данных для них используются значения функции теплосодержания, полученные в результате решения предыдущей подзадачи.

При проведении численных экспериментов, связанных с определением температурного поля объекта, основное внимание обращалось на динамику фронта кристаллизации. Многочисленные расчеты прямой задачи показали, что скорость перемещения литейной формы в рассматриваемой промышленной установке оказывает сильное влияние на эволюцию поверхности раздела фаз в металле. Этим параметром сравнительно легко управлять на практике. Поэтому в качестве функции, управляющей процессом, выбиралась скорость движения литейной формы относительно плавильной печи. Для того чтобы определить управляющую функцию, удовлетворяющую технологическим требованиям к процессу кристаллизации, была сформулирована задача оптимального управления. Эта задача состоит в определении такого режима остывания расплавленного металла, при котором поверхность раздела фаз имеет предписанную технологами форму и движется со скоростью, близкой к заданной.

Если движением литейной формы в рассматриваемой установке не управлять, то внутри детали могут формироваться и схлопываться пузыри, состоящие из жидкого металла. В результате получается металлический образец плохого качества.

В качестве целевого функционала выбран следующий:

$$I(u) = \frac{1}{(t_2 - t_1)} \int_{t_1}^{t_2} \iint_{S(z_*(t))} [Z_{pl}(x, y, t) - z_*(t)]^2 dx dy dt. \quad (3.6)$$

Здесь t_1 — момент времени, при котором зарождается фронт кристаллизации; t_2 — момент времени, при котором наступает полное затвердевание металла; $(x, y, Z_{pl}(x, y, t))$ — реальные координаты поверхности раздела фаз в момент времени t ; $(x, y, z_*(t))$ — желаемые координаты поверхности раздела фаз в момент времени t ; S — проекция поверхности раздела фаз на плоскость XOY .

Функционал (3.6) представляет собой среднее по времени среднеквадратичное отклонение реальной поверхности раздела фаз от требуемой. Он призван следить как за выпрямлением поверхности раздела фаз, так и за скоростью ее перемещения.

Пусть A_* и B_* — некоторые заданные константы (A_* есть z -координата, определяющая начальное расположение объекта относительно печи; B_* есть z -координата, определяющая расположение объекта при его максимально допустимом погружении в охладитель). Будем говорить, что функция $u(t)$ принадлежит классу K , если $u(t)$ непрерывная, кусочно-гладкая при $t \in [0, \infty)$ и удовлетворяет ограничениям $A_* \leq u(t) \leq B_*$, $u(0) = A_*$. Задача оптимального управления состоит в выборе такого управления $u(t) \in K$, при котором функционал (3.6) достигает минимального значения.

Функционал (3.6) аппроксимируется функцией $F(u)$ с помощью метода трапеций:

$$I(u) \cong F(u) = \frac{1}{2(t_2 - t_1)} \left(\tau^{j_1+1} f^{j_1} + \sum_{j=j_1+1}^{j_2-1} (\tau^j + \tau^{j+1}) f^j + \tau^{j_2} f^{j_2} \right).$$

Здесь j_1 — номер узла временной сетки, который соответствует моменту времени t_1 ; j_2 — номер узла временной сетки, который соответствует моменту времени t_2 ; $f^j = \sum_{n=0}^{n_1} \sum_{i=0}^{i_1} (Z_{ni}^j - z_*^j)^2 h_n^x h_i^y$; n_1 и i_1 — номера узлов сетки вдоль осей OX и OY соответственно, которые определяют границы площади $\tilde{S}$ ($\tilde{S}$ — четверть наибольшего горизонтального сечения части литейной формы, которая заполнена металлом); $Z_{ni}^j = Z_{pl}(x_n, y_i, t^j)$, $z_*^j = z_*(t^j)$.

В численных расчетах величины Z_{ni}^j ($j = \overline{0, J-1}$; $n = \overline{0, n_1}$, $i = \overline{0, i_1}$) на каждом j -м временном слое определяются по температурному полю, полученному в результате решения начально-краевой задачи Стефана. Если x_n, y_i, z_l — координаты узловых точек пространственной расчетной сетки, тогда для любой точки $(x_n, y_i) \in \tilde{S}$, $(n = \overline{0, n_1}, i = \overline{0, i_1})$ находим такой индекс l_* , который удовлетворяет следующему условию:

$$\left(\beta(E_{ni, l_*+1}^j) \leq T_{pl} \leq \beta(E_{ni, l_*}^j) \right) \cup \left(\beta(E_{ni, l_*}^j) \leq T_{pl} \leq \beta(E_{ni, l_*+1}^j) \right). \quad (3.7)$$

В этом случае полагаем

$$Z_{ni}^j = \frac{(z_{l_*+1} - z_{l_*})T_{pl} + (z_{l_*}\beta_{ni, l_*+1}^j - z_{l_*+1}\beta_{ni, l_*}^j)}{\beta_{ni, l_*+1}^j - \beta_{ni, l_*}^j}.$$

Если же условие (3.7) не выполняется, то полагаем $Z_{ni}^j = z_*^j$.

Указанная выше задача оптимального управления для различных параметров установки (температуры печи, температуры жидкого алюминия, глубины погружения объекта в жидкий алюминий) решалась численно с помощью градиентного метода минимизации функционала качества. Компоненты градиента целевой функции вычислялись с помощью БАД-методологии. Вывод формул для вычисления этого градиента и используемого при этом дискретного варианта сопряженной задачи для указанной здесь оптимизационной задачи достаточно громоздкий. Все необходимые соотношения представлены в работе [9].

Следует отметить важную особенность, связанную с применением БАД-методологии в этой задаче. Предложенный алгоритм решения прямой задачи предполагает указанную выше последовательность решения подзадач: x — направление, y — направление, z — направление. Что касается решения сопряженной задачи, то БАД-методология требует совершенно другой последовательности решения расщепленных задач, а именно: y — направление, x — направление, z — направление. Другой порядок решения подзадач приводит к неточному значению градиента функционала. Угадать нужную последовательность решения подзадач сопряженной задачи, по-видимому, нереально.

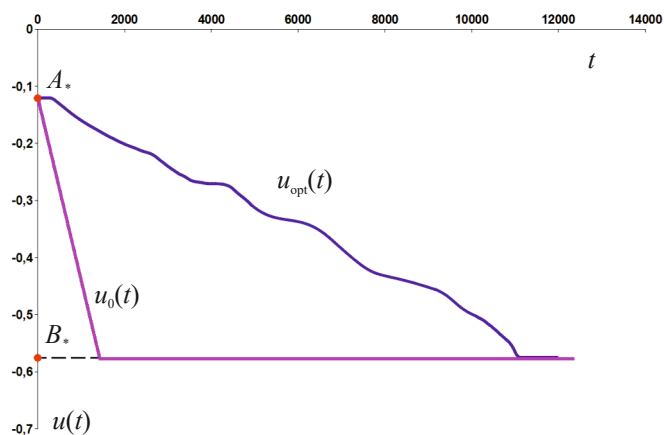
Отметим также, что решить рассматриваемую задачу, используя другие методы вычисления градиента целевой функции (например, метод конечных разностей), оказалось практически невозможным. Спуск на один-два шага в направлении антиградиента приводит к останову процесса минимизации (функция возрастает в направлении антиградиента). Поэтому вычисление точного значения градиента целевой функции с помощью БАД-методологии представляется необходимым, неизбежным элементом решения сложных задач оптимального управления, и это несмотря на громоздкость и трудности получения дискретного варианта сопряженной задачи и формул для вычисления градиента.

Для иллюстрации работоспособности и эффективности предложенного алгоритма ниже приводятся результаты одного из многочисленных проведенных численных экспериментов. Расчеты проводились для образца, сечения которого представлены на фиг. 3. Его размеры и все остальные параметры задачи даны в [11]. Температура стенок печи равнялась 1920°K . Координата $z_*(t)$ требуемой поверхности раздела фаз менялась во времени с постоянной скоростью 2 мм/мин. В качестве начального управления $u_0(t)$ выбрано перемещение литейной формы с постоянной скоростью 25 мм/мин (фиг. 6). Значение целевого функционала при этом равнялось $I(u_0) = 8.56$. На фиг. 7 проиллюстрировано отличие реальной поверхности раздела фаз Z_{pl} (кривая линия) от требуемой z_* (горизонтальная прямая линия) в некоторый момент времени t .

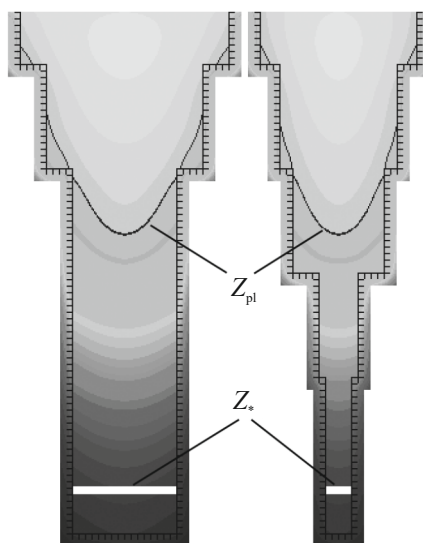
В результате оптимизации значение целевого функционала уменьшилось более чем в 3500 раз и составило $I(u_{\text{opt}}(t)) = 0.0024$. Оптимальное управление $u_{\text{opt}}(t)$ показано на фиг. 6. Его использование позволило существенно выпрямить поверхность раздела фаз и одновременно обеспечило ее движение с требуемой скоростью. При таком управлении реальная поверхность раздела фаз практически совпала с требуемой.

4. ЗАКЛЮЧЕНИЕ

В работе описывается предложенная Ю.Г. Евтушенко методология быстрого автоматического дифференцирования и демонстрируются ее применение к решению тепловых задач. Делается акцент на то, что эта методология позволяет эффективно и с машинной точностью вычислять градиент любой, сколь угодно сложной целевой функции конечного числа переменных. Большой опыт решения разнообразных задач оптимизации



Фиг. 6.



Фиг. 7.

позволяет утверждать, что в настоящее время предложенная Ю.Г. Евтушенко БАД-методология представляет собой современный аппарат, помогающий эффективно решать задачи оптимального управления разнообразными сложными динамическими системами.

СПИСОК ЛИТЕРАТУРЫ

1. *Evtushenko Y.G.* Computation of exact gradients in distributed dynamic systems // Optimization Methods and Software. 1998. V. 9. P. 45–75.
2. *Евтушенко Ю.Г., Зубов В.И.* Об обобщенной методологии быстрого автоматического дифференцирования // Ж. вычисл. матем. и матем. физ. 2016. Т. 56. № 11. С. 1847–1862.
3. *Албу А.Ф., Горчаков А.Ю., Зубов В.И.* БАД-методология и дифференцирование сложной функции // Ж. вычисл. матем. и матем. физ. 2023. Т. 63. № 1. С. 61–73.
4. *Албу А.Ф., Зубов В.И.* Об эффективности решения задач оптимального управления с помощью методологии быстрого автоматического дифференцирования // Труды Института математики и механики УрО РАН. 2015. Т. 21. № 4. С. 20–29.

5. *Зубов В.И., Албу А.Ф.* Идентификация коэффициента теплопроводности по результатам эксперимента // Монография. М.: МАКС Пресс, 2023. 188 с. ISBN 978-5-317-07020-5.
6. *White R.E.* An enthalpy formulation of the Stephan problem // SIAM J. Numer. Anal. 1982. Vol. 19. No 6. P. 1129–1157.
7. *Alla Albu, Vladimir Zubov.* Modeling and optimization of melting and solidification process // BULETINUL ACADEMIEI DE STIINTE A REPUBLICII MOLDOVA. MATEMATICA. 2004. Number 3(46). P. 91–109.
8. *Албу А.Ф.* Расчет теплового излучения при моделировании процесса кристаллизации вещества в литейном деле // Информационные технологии и вычисл. системы. 2015. Т. 65. № 1. С. 47–55.
9. *Албу А.В., Албу А.Ф., Зубов В.И.* Вычисление градиента функционала в одной задаче оптимального управления сложной динамической системой // Ж. вычисл. матем. и матем. физ. 2011. Т. 51. № 5. С. 814–833.
10. *Албу А.Ф., Зубов В.И.* О модификации одной схемы для расчета процесса плавления // Ж. вычисл. матем. и матем. физ. 2001. Т. 41. № 9. С. 1434–1443.
11. *Албу А.Ф., Зубов В.И.* О влиянии параметров установки на управление процессом кристаллизации вещества в литейном деле // Ж. вычисл. матем. и матем. физ. 2013. Т. 53. № 2. С. 238–248.

METHODOLOGY OF RAPID AUTOMATIC DIFFERENTIATION AND CONTROL OF THERMAL DYNAMIC SYSTEMS

V. I. Zubov*

119333 Moscow, Vavilov Str., 44, Federal Research Center Computer Science and Control, RAS, Russia

*e-mail: vladimir.zubov@mail.ru

Received: 17.03.2024

Revised: 27.03.2024

Accepted: 31.05.2024

Abstract. The paper describes the methodology of rapid automatic differentiation and its advantages in the numerical solution of optimal control problems. The results obtained in solving problems of optimal control of thermal processes with phase transitions using this methodology are presented.

Keywords: optimal control, gradient, fast automatic differentiation, thermal conductivity equation.

СИНТЕЗ РЕГУЛЯТОРА ДЛЯ ЛИНЕЙНО-КВАДРАТИЧНОЙ ЗАДАЧИ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

© 2024 г. А.С. Антипин^{1,*}, Е.В. Хорошилова^{2,**}

¹ 119333 Москва, ул. Вавилова, 44, ФИЦ ИУ РАН, Россия

² 119991 Москва, Ленинские горы, МГУ им. М.В. Ломоносова, Россия

*e-mail: asantip@yandex.ru

**e-mail: khorelena@gmail.com

Поступила в редакцию 19.02.2024 г.

Переработанный вариант 19.02.2024 г.

Принята к публикации 31.05.2024 г.

В гильбертовом пространстве рассматривается линейно-квадратичная задача оптимального управления с закрепленным левым концом и подвижным правым концом, на фиксированном отрезке времени. Целевой функционал представляет собой сумму интегральной и терминальной компонент квадратичного вида. Каждая из компонент ищет свой минимум на своем допустимом множестве независимо друг от друга. На правом конце отрезка времени мы имеем задачу линейного программирования. Решение этой задачи неявно определяет терминальное условие для управляемой динамики. Предлагается седловой подход для решения задачи, который сводится к вычислению седловой точки функции Лагранжа. В основе подхода лежат седловые неравенства по обеим группам переменных: прямых и двойственных. Эти неравенства представляют собой достаточные условия оптимальности. Формулируется метод вычисления седловой точки функции Лагранжа. Доказывается сходимости по прямым и двойственным переменным, а именно: слабая сходимость по управлениям, сильная сходимость по фазовым и сопряженным траекториям, а также по терминальным переменным краевой задачи. На базе седлового подхода строится синтез управления, т. е. обратная связь при наличии ограничений на управления в форме выпуклого замкнутого множества. Это новый результат, поскольку в классическом случае в теории линейного регулятора аналогичное утверждение доказывается при отсутствии ограничений на управления, что дает возможность использовать матричное уравнение Риккати. При наличии ограничений на управление эти рассуждения уже не проходят. Поэтому в основе полученного результата лежит понятие опорной плоскости ко множеству управлений. Библ. 20.

Ключевые слова: терминальное управление, краевая задача, функция Лагранжа, седловые методы, синтез управления, сходимость.

DOI: 10.31857/S0044466924090049, EDN: WKZSWO

ВВЕДЕНИЕ

Известно, что принцип максимума относится к числу наиболее мощных и сильных инструментов решения задач оптимального управления. Однако принцип максимума, с точки зрения разработки теории методов решения задач оптимального управления, представляет собой всего лишь необходимое условие оптимальности. Последнее означает, что решение, найденное с помощью принципа максимума, может как быть решением исходной задачи, но, в равной степени, может им и не быть. В этом случае вычислительный метод не может быть признан обоснованным с точки зрения доказательных вычислений. На практике для обоснования полученного решения часто опираются на опыт и интуицию экспертов. Однако такого рода экспертиза носит субъективный характер и не является доказательной.

Достаточные условия оптимальности представляются естественным выходом из положения. Основные идеи этого подхода обсуждались в [1], но там они не получили четкого обоснования, лишь предполагалось, что на базе седлового метода можно синтезировать линейный регулятор (обратную связь). В настоящей статье доказывается, что механизм обратной связи целиком «зашит» в вычислительную технологию седлового градиентного метода [2–12]. Фактически между этими процессами существует внутренняя связь.

Обоснованность подхода базируется на доказанных теоремах, в основе которых лежит теория двойственности. Все теоремы работают на классах методов и для классов задач. Сходимость вычислительных процессов гарантируется по всем компонентам решения, а именно: по управлениям (слабая), по фазовым и сопряженным траекториям (сильная), по терминальным переменным сходимость также сильная.

1. ПОСТАНОВКА ЗАДАЧИ

В работе рассматривается задача терминального управления с линейной динамикой на конечном отрезке времени

$$\frac{d}{dt}x(t) = D(t)x(t) + B(t)u(t), \quad x(t_0) = x_0, \quad t_0 \leq t \leq t_1, \quad (1)$$

где $D(t)$, $B(t)$ суть $n \times n$, $n \times r$ -функциональные матрицы. Левый конец $x(t_0)$ фазовой переменной закреплён и равен x_0 , правый конец $x(t_1)$ подвижен в пределах выпуклого многогранника $M = \{x_1 \in \mathbb{R}^n : G_1 x_1 \leq g_1\}$, G_1 – заданная матрица размерности $m \times n$, $m < n$, вектор g_1 – фиксирован. Управление $u(t) \in U \subset \mathbb{L}_2^n$, U – выпуклое замкнутое ограниченное множество. Для любых управлений из U и заданного $x(t_0)$ система дифференциальных уравнений (1) порождает траектории $x(t)$, правые концы которых описывают множество достижимости $X_1 = X(t_1) \subset \mathbb{R}^n$. Задача рассматривается в гильбертовом пространстве.

Под решением дифференциальной системы (1) будем понимать любую пару $(x(t), u(t))$, удовлетворяющую тождественно условию

$$x(t) = x(t_0) + \int_{t_0}^t (D(\tau)x(\tau) + B(\tau)u(\tau))d\tau, \quad t_0 \leq t \leq t_1.$$

В [13, Кн. 1, с. 443] показано, что любому управлению $u(t) \in U$ в линейной дифференциальной системе отвечает единственная траектория $x(t)$, удовлетворяющая этому тождеству. Траектория $x(t)$ в условиях тождества является абсолютно непрерывной функцией. Класс абсолютно непрерывных функций представляет собой линейное многообразие, всюду плотное в $\mathbb{L}_2^n[t_0, t_1]$. Этот класс будем обозначать через $\mathbb{AC}^n[t_0, t_1]$. Для любой пары функций $(x(t), u(t)) \in \mathbb{AC}^n[t_0, t_1] \times U$ выполняется формула Ньютона–Лейбница и формула интегрирования по частям. Скалярные произведения и нормы определяются, соответственно, как

$$\langle x(\cdot), y(\cdot) \rangle = \int_{t_0}^{t_1} \langle x(t), y(t) \rangle dt, \quad \|x(\cdot)\|^2 = \int_{t_0}^{t_1} |x(t)|^2 dt,$$

где

$$\langle x(t), y(t) \rangle = \sum_{i=1}^n x_i(t)y_i(t), \quad |x(t)|^2 = \sum_{i=1}^n x_i^2(t), \quad t_0 \leq t \leq t_1.$$

Будем понимать обозначение $x(\cdot)$ как элемент гильбертова пространства (если это надо подчеркнуть), а $x(t)$ как значение конечномерного вектора в точке t . Далее там, где не будет разночтений, наряду с обозначениями $x(\cdot)$, $y(\cdot)$ будем также использовать $x(t)$, $y(t)$.

Минимизируемый функционал представлен суммой интегральной и терминальной компонент:

$$J(u) = \frac{1}{2} \int_{t_0}^{t_1} (\langle Q_1(t)x(t), x(t) \rangle + \langle Q_2(t)u(t), u(t) \rangle) dt + \langle Sx_1, x_1 \rangle,$$

где S – фиксированная положительно-полуопределённая симметричная $n \times n$ матрица; $Q_1(t)$, $Q_2(t)$ – непрерывные положительно-полуопределённые симметричные матрицы размерностей $n \times n$, $n \times r$ соответственно.

С учетом сказанного, сформулируем задачу терминального управления в виде

$$\begin{aligned} (x^*(t), u^*(t), x_1^*) \in \operatorname{Argmin} \Big\{ & \frac{1}{2} \int_{t_0}^{t_1} (\langle Q_1(t)x(t), x(t) \rangle + \langle Q_2(t)u(t), u(t) \rangle) dt + \\ & + \langle Sx_1, x_1 \rangle \Big| G_1 x_1 \leq g_1, \quad x_1 \in X_1 \subset \mathbb{R}^n, \quad x(t) \in \mathbb{AC}^n[t_0, t_1], \quad u(t) \in U \Big\}; \end{aligned} \quad (2)$$

$$\frac{d}{dt}x(t) = D(t)x(t) + B(t)u(t), \quad t_0 \leq t \leq t_1, \quad x(t_0) = x_0, \quad x(t_1) = x_1^*.$$

В теории оптимального управления существуют несколько видов достаточных условий. Наиболее известные из них – это достаточные условия по В.Ф. Кротову, основанные на неравенствах Гамильтона–Якоби [14], по В.В. Величенко, опирающиеся на теорию поля экстремалей [15]. В нашем случае мы используем достаточные условия седлового типа, которые следуют из теории двойственности и были получены нами в случае линеаризации выпуклых и гладких задач терминального управления. Они были названы седловыми, так как порождаются седловыми неравенствами относительно прямых и двойственных переменных функции Лагранжа в задачах линейного и выпуклого программирования.

Следуя методологии теории двойственности, линеаризуем линейно-квадратичную задачу (2) в точке $(x_1^*, x^*(t), u^*(t))$, т. е. в решении задачи, тогда получим

$$\begin{aligned} (x_1^*, x^*(t), u^*(t)) \in \operatorname{Argmin} \Bigg\{ & \int_{t_0}^{t_1} (\langle Q_1(t)x^*(t), x(t) - x^*(t) \rangle + \langle Q_2(t)u^*(t), u(t) - u^*(t) \rangle) dt + \\ & + \langle Sx_1^*, x_1 - x_1^* \rangle \Big| G_1 x_1 \leq g_1, x_1 \in X_1, x(t) \in \mathbb{AC}^n[t_0, t_1], u(t) \in U \Bigg\}; \\ & \frac{d}{dt}x(t) = D(t)x(t) + B(t)u(t), \quad t_0 \leq t \leq t_1, x(t_0) = x_0, x(t_1) = x_1^*. \end{aligned}$$

Исключив константы в целевом функционале, приходим к более простой, но эквивалентной задаче:

$$\begin{aligned} (x_1^*, x^*(t), u^*(t)) \in \operatorname{Argmin} \Bigg\{ & \int_{t_0}^{t_1} (\langle Q_1(t)x^*(t), x(t) \rangle + \langle Q_2(t)u^*(t), u(t) \rangle) dt + \\ & + \langle Sx_1^*, x_1 \rangle \Big| G_1 x_1 \leq g_1, x_1 \in X_1, x(t) \in \mathbb{AC}^n[t_0, t_1], u(t) \in U \Bigg\}; \\ & \frac{d}{dt}x(t) = D(t)x(t) + B(t)u(t), \quad t_0 \leq t \leq t_1, x(t_0) = x_0, x(t_1) = x_1^*. \end{aligned} \quad (3)$$

В этой задаче требуется найти управление $u^*(t) \in U$ такое, что отвечающая ему траектория $x^*(t)$ соединяет начальную точку x_0 с терминальной точкой $x^*(t_1) = x_1^* \in X_1$, при этом целевой функционал принимает минимальное значение. В [13, Кн.2, с. 653] доказано, что решение $(x_1^*, x^*(t), u^*(t)) \in \mathbb{R}^n \times \mathbb{AC}^n[t_0, t_1] \times U$ данной задачи существует.

2. ПРЯМАЯ И ДВОЙСТВЕННАЯ ФОРМЫ ЛАГРАНЖИАНА

Задача (3), аналог задачи линейного программирования, сформулирована в гильбертовом пространстве. Проводя аналогии между постановками задач оптимизации в конечномерных и функциональных пространствах, введем для (3) функцию Лагранжа

$$\begin{aligned} \mathcal{L}(p_1, \psi(t); x_1, x(t), u(t)) = & \int_{t_0}^{t_1} (\langle Q_1(t)x^*(t), x(t) \rangle + \langle Q_2(t)u^*(t), u(t) \rangle) dt + \\ & + \langle Sx_1^*, x_1 \rangle + \langle p_1, G_1 x_1 - g_1 \rangle + \int_{t_0}^{t_1} \langle \psi(t), D(t)x(t) + B(t)u(t) - \frac{d}{dt}x(t) \rangle dt \end{aligned} \quad (4)$$

при всех $(p_1, \psi(t)) \in \mathbb{R}_+^m \times \Psi_2^n[t_0, t_1]$, $(x_1, x(t), u(t)) \in \mathbb{R}^n \times \mathbb{AC}^n[t_0, t_1] \times U$. Здесь $\Psi_2^n[t_0, t_1]$ — линейное многообразие абсолютно непрерывных функций из сопряженного пространства. Отметим, что вектор $(x_1^*, x^*(t), u^*(t))$ играет двойную роль: как точка линеаризации и как решение задачи.

По теореме Куна–Таккера, седловая точка $(p_1^*, \psi^*(t); x_1^*, x^*(t), u^*(t))$ функции Лагранжа образована прямыми $(x_1^*, x^*(t), u^*(t))$ и двойственными $(p_1^*, \psi^*(t))$ переменными. (Недавно в [16] было предложено новое доказательство теоремы Куна–Таккера).

По определению, седловая точка удовлетворяет системе неравенств

$$\mathcal{L}(p_1, \psi(t); x_1^*, x^*(t), u^*(t)) \leq \mathcal{L}(p_1^*, \psi^*(t); x_1^*, x^*(t), u^*(t)) \leq \mathcal{L}(p_1^*, \psi^*(t); x_1, x(t), u(t)),$$

или, в развернутом виде,

$$\begin{aligned} & \int_{t_0}^{t_1} (\langle Q_1(t)x^*(t), x^*(t) \rangle + \langle Q_2(t)u^*(t), u^*(t) \rangle) dt + \langle Sx_1^*, x_1^* \rangle + \langle p_1, G_1 x_1^* - g_1 \rangle + \\ & + \int_{t_0}^{t_1} \langle \psi(t), D(t)x^*(t) + B(t)u^*(t) - \frac{d}{dt}x^*(t) \rangle dt \leq \\ & \leq \int_{t_0}^{t_1} (\langle Q_1(t)x^*(t), x^*(t) \rangle + \langle Q_2(t)u^*(t), u^*(t) \rangle) dt + \langle Sx_1^*, x_1^* \rangle + \langle p_1^*, G_1 x_1^* - g_1 \rangle + \\ & + \int_{t_0}^{t_1} \langle \psi^*(t), D(t)x^*(t) + B(t)u^*(t) - \frac{d}{dt}x^*(t) \rangle dt \leq \end{aligned}$$

$$\begin{aligned} \leq & \int_{t_0}^{t_1} (\langle Q_1(t)x^*(t), x(t) \rangle + \langle Q_2(t)u^*(t), u(t) \rangle) dt + \langle Sx_1^*, x_1 \rangle + \langle p_1^*, G_1x_1 - g_1 \rangle + \\ & + \int_{t_0}^{t_1} \langle \Psi^*(t), D(t)x(t) + B(t)u(t) - \frac{d}{dt}x(t) \rangle dt \end{aligned} \quad (5)$$

для всех $(p_1, \Psi(t)) \in \mathbb{R}_+^m \times \Psi_2^n[t_0, t_1]$, $(x_1, x(t), u(t)) \in \mathbb{R}^n \times \mathbb{AC}^n[t_0, t_1] \times U$. Несложными преобразованиями из системы (5) всегда можно вернуться к задаче (3), т. е. система (5) порождает исходную задачу относительно прямых переменных.

Введем понятие двойственного лагранжиана и покажем, что он порождает двойственную задачу. Переходя к сопряженным линейным операторам

$$\langle \Psi, Dx \rangle = \langle D^T \Psi, x \rangle, \quad \langle \Psi, Bu \rangle = \langle B^T \Psi, u \rangle$$

и используя формулу интегрирования по частям на отрезке $[t_0, t_1]$

$$\langle \Psi(t_1), x(t_1) \rangle - \langle \Psi(t_0), x(t_0) \rangle = \int_{t_0}^{t_1} \langle \frac{d}{dt} \Psi(t), x(t) \rangle dt + \int_{t_0}^{t_1} \langle \Psi(t), \frac{d}{dt} x(t) \rangle dt,$$

выпишем сопряженную (двойственную) по отношению к (4) функцию Лагранжа:

$$\begin{aligned} \mathcal{L}^T(p_1, \Psi_1, \Psi(t); x_1, x(t), u(t)) = & \int_{t_0}^{t_1} \langle Q_1(t)x^*(t) + D^T(t)\Psi(t) + \frac{d}{dt}\Psi(t), x(t) \rangle dt + \\ & + \int_{t_0}^{t_1} \langle Q_2(t)u^*(t) + B^T(t)\Psi(t), u(t) \rangle dt + \langle Sx_1^* + G_1^T p_1 - \Psi_1, x_1 \rangle + \langle p_1, -g_1 \rangle + \langle \Psi_0, x_0 \rangle \end{aligned} \quad (6)$$

для всех $(p_1, \Psi_1, \Psi(t)) \in \mathbb{R}_+^m \times \mathbb{R}^n \times \Psi_2^n[t_0, t_1]$, $(x_1, x(t), u(t)) \in \mathbb{R}^n \times \mathbb{AC}^n[t_0, t_1] \times U$, где $\Psi_0 = \Psi(t_0)$, $\Psi_1 = \Psi(t_1)$. Оба лагранжиана (4) и (6) имеют одно и то же множество седловых точек $(p_1^*, \Psi_1^*, \Psi^*(t); x_1^*, x^*(t), u^*(t))$. Седловая система (5) в сопряженном варианте имеет вид (чтобы убедиться в этом, достаточно в (5) перейти к сопряженным операторам):

$$\begin{aligned} & \int_{t_0}^{t_1} \langle Q_1(t)x^*(t) + D^T(t)\Psi(t) + \frac{d}{dt}\Psi(t), x^*(t) \rangle dt + \\ & + \int_{t_0}^{t_1} \langle Q_2(t)u^*(t) + B^T(t)\Psi(t), u^*(t) \rangle dt + \langle Sx_1^* + G_1^T p_1 - \Psi_1, x_1^* \rangle + \langle p_1, -g_1 \rangle + \langle \Psi_0, x_0 \rangle \leq \\ & \leq \int_{t_0}^{t_1} \langle Q_1(t)x^*(t) + D^T(t)\Psi^*(t) + \frac{d}{dt}\Psi^*(t), x^*(t) \rangle dt + \\ & + \int_{t_0}^{t_1} \langle Q_2(t)u^*(t) + B^T(t)\Psi^*(t), u^*(t) \rangle dt + \langle Sx_1^* + G_1^T p_1^* - \Psi_1^*, x_1^* \rangle + \langle p_1^*, -g_1 \rangle + \langle \Psi_0, x_0 \rangle \leq \\ & \leq \int_{t_0}^{t_1} \langle Q_1(t)x^*(t) + D^T(t)\Psi^*(t) + \frac{d}{dt}\Psi^*(t), x(t) \rangle dt + \\ & + \int_{t_0}^{t_1} \langle Q_2(t)u^*(t) + B^T(t)\Psi^*(t), u(t) \rangle dt + \langle Sx_1^* + G_1^T p_1^* - \Psi_1^*, x_1 \rangle + \langle p_1^*, -g_1 \rangle + \langle \Psi_0, x_0 \rangle \end{aligned} \quad (7)$$

для всех $(p_1, \Psi_1, \Psi(t)) \in \mathbb{R}_+^m \times \mathbb{R}^n \times \Psi_2^n[t_0, t_1]$, $(x_1, x(t), u(t)) \in \mathbb{R}^n \times \mathbb{AC}^n[t_0, t_1] \times U$.

Из правого неравенства системы (7) с учетом того, что члены $\langle \Psi_0, x_0 \rangle$ взаимно уничтожаются, имеем

$$\begin{aligned} & \int_{t_0}^{t_1} \langle Q_1(t)x^*(t) + D^T(t)\Psi^*(t) + \frac{d}{dt}\Psi^*(t), x^*(t) \rangle dt + \\ & + \int_{t_0}^{t_1} \langle Q_2(t)u^*(t) + B^T(t)\Psi^*(t), u^*(t) \rangle dt + \langle Sx_1^* + G_1^T p_1^* - \Psi_1^*, x_1^* \rangle \leq \\ & \leq \int_{t_0}^{t_1} \langle Q_1(t)x^*(t) + D^T(t)\Psi^*(t) + \frac{d}{dt}\Psi^*(t), x(t) \rangle dt + \end{aligned}$$

$$+ \int_{t_0}^{t_1} \langle Q_2(t)u^*(t) + B^T(t)\psi^*(t), u(t) \rangle dt + \langle Sx_1^* + G_1^T p_1^* - \psi_1^*, x_1 \rangle$$

при всех $(x_1, x(t), u(t)) \in \mathbb{R}^n \times \mathbb{AC}^n[t_0, t_1] \times U$. В силу независимого изменения переменных x_1 , $x(t)$ и $u(t)$ в пределах своих допустимых подпространств (множеств), можно положить в последнем неравенстве, например, $x(t) = x^*(t)$, $u(t) = u^*(t)$ и получить $\langle Sx_1^* + G_1^T p_1^* - \psi_1^*, x_1^* \rangle \leq \langle Sx_1^* + G_1^T p_1^* - \psi_1^*, x_1 \rangle$. Таким образом, неравенство распадается на три независимых вариационных неравенства:

$$\int_{t_0}^{t_1} \langle Q_1(t)x^*(t) + D^T(t)\psi^*(t) + \frac{d}{dt}\psi^*(t), x^*(t) - x(t) \rangle dt \leq 0, \quad x(t) \in \mathbb{AC}^n[t_0, t_1], \quad (8)$$

$$\int_{t_0}^{t_1} \langle Q_2(t)u^*(t) + B^T(t)\psi^*(t), u^*(t) - u(t) \rangle dt \leq 0, \quad u(t) \in U, \quad (9)$$

$$\langle Sx_1^* + G_1^T p_1^* - \psi_1^*, x_1^* - x_1 \rangle \leq 0, \quad x_1 \in \mathbb{R}^n. \quad (10)$$

В (10) линейная функция достигает конечного экстремума на всем пространстве $\mathbb{R}^n$, что возможно лишь в случае, когда ее градиент обращается в нуль:

$$\psi_1^* = Sx_1^* + G_1^T p_1^* \quad (11)$$

(условие трансверсальности). Аналогично, в неравенстве (8) дифференцируемый функционал достигает конечного минимума при $x(t) = x^*(t)$ на всем линейном многообразии $\mathbb{AC}^n[t_0, t_1]$, что возможно лишь в случае, когда его градиент в этой точке также равен нулю:

$$Q_1(t)x^*(t) + D^T(t)\psi^*(t) + \frac{d}{dt}\psi^*(t) = 0 \quad (12)$$

(система дифференциальных уравнений для сопряженной траектории в двойственной задаче). Неравенство (9) остается без изменений.

Левое неравенство системы (7) с учетом (11), (12) принимает вид

$$\begin{aligned} & \int_{t_0}^{t_1} \langle Q_1(t)x^*(t) + D^T(t)\psi(t) + \frac{d}{dt}\psi(t), x^*(t) \rangle dt + \\ & + \int_{t_0}^{t_1} \langle Q_2(t)u^*(t) + B^T(t)\psi(t), u^*(t) \rangle dt + \langle Sx_1^* + G_1^T p_1 - \psi_1, x_1^* \rangle + \langle p_1, -g_1 \rangle \leq \\ & \leq \int_{t_0}^{t_1} \langle Q_2(t)u^*(t) + B^T(t)\psi^*(t), u^*(t) \rangle dt + \langle p_1^*, -g_1 \rangle. \end{aligned}$$

Рассматривая это неравенство при скалярных ограничениях

$$\langle Sx_1^* + G_1^T p_1 - \psi_1, x_1^* \rangle = 0, \quad \int_{t_0}^{t_1} \langle Q_1(t)x^*(t) + D^T(t)\psi(t) + \frac{d}{dt}\psi(t), x^*(t) \rangle dt = 0, \quad (13)$$

получим задачу максимизации

$$\langle p_1 - p_1^*, -g_1 \rangle + \int_{t_0}^{t_1} \langle B^T(t)(\psi(t) - \psi^*(t)), u^*(t) \rangle dt \leq 0$$

для всех $(p_1, \psi(t)) \in \mathbb{R}_+^m \times \Psi_2^n[t_0, t_1]$. Но из (11), (12) следует, что решение $(p_1^*, \psi^*(t))$ одновременно принадлежит более узкому множеству, чем (13). Поэтому указанная точка остается минимумом и на подмножестве решений системы (13).

Объединяя последнее неравенство с условиями (9), (13), получаем систему:

$$\begin{aligned} & (p_1^*, \psi_1^*, \psi^*(t)) \in \operatorname{Argmax} \left\{ \int_{t_0}^{t_1} \langle \psi(t), B(t)u^*(t) \rangle dt + \langle p_1, -g_1 \rangle \mid \right. \\ & \quad \left. \frac{d}{dt}\psi(t) + D^T(t)\psi(t) = -Q_1(t)x^*(t), \quad \psi_1 = Sx_1^* + G_1^T p_1, \right. \\ & \quad \left. p_1 \in \mathbb{R}_+^m, \quad \psi(t) \in \Psi_2^n[t_0, t_1] \right\}, \\ & \int_{t_0}^{t_1} \langle Q_2(t)u^*(t) + B^T(t)\psi^*(t), u^*(t) - u(t) \rangle dt \leq 0, \quad u(t) \in U. \end{aligned} \quad (14)$$

Эта задача двойственная по отношению к задаче (3). Она включает в себя задачу оптимизации относительно двойственных переменных и вариационное неравенство относительно управления. Исследование прямых и двойственных задач в совокупности порождает новые (седловые) подходы к их решению [5].

3. ДОСТАТОЧНЫЕ УСЛОВИЯ ОПТИМАЛЬНОСТИ СЕДЛОВОГО ТИПА И МЕТОД РЕШЕНИЯ

Рассматривая вместе левое неравенство седловой системы (5) и правое неравенство сопряженной седловой системы (7), мы приходим к следующей задаче:

$$\frac{d}{dt}x^*(t) = D(t)x^*(t) + B(t)u^*(t), \quad x^*(t_0) = x_0, \quad (15)$$

$$\langle p_1 - p_1^*, G_1 x_1^* - g_1 \rangle \leq 0, \quad p_1 \in \mathbb{R}_+^m, \quad (16)$$

$$\frac{d}{dt}\psi^*(t) + D^T(t)\psi^*(t) = -Q_1(t)x^*(t), \quad \psi_1^* = Sx_1^* + G_1^T p_1^*, \quad (17)$$

$$\int_{t_0}^{t_1} \langle Q_2(t)u^*(t) + B^T(t)\psi^*(t), u^*(t) - u(t) \rangle dt \leq 0, \quad u(t) \in U. \quad (18)$$

Эта дифференциальная система представляет собой седловое достаточное условие оптимальности. Сопоставляя систему (15)–(18) с принципом максимума, можно видеть, что основное и единственное различие заключается в неравенстве (18). В нашем выпуклом случае (18) представляет собой вариационное неравенство относительно множества U в гильбертовом пространстве. В гладком случае принципа максимума мы будем иметь параметрическое семейство конечномерных вариационных неравенств, зависящих от параметра времени $t \in [t_0, t_1]$.

Отметим, что вариационные неравенства (16), (18) можно записать в эквивалентной форме уравнений с оператором проектирования на соответствующие выпуклые замкнутые множества:

$$p_1^* = \pi_+(p_1^* + \alpha(G_1 x_1^* - g_1)), \quad (19)$$

$$u^*(t) = \pi_U(u^*(t) - \alpha(Q_2(t)u^*(t) + B^T(t)\psi^*(t))), \quad (20)$$

где $\pi_+(\cdot)$, $\pi_U(\cdot)$ — операторы проектирования текущего вектора на положительный ортант пространства $\mathbb{R}_+^m$ и на множество управлений U , $\alpha > 0$.

Используя (19), (20), выпишем для решения системы (15)–(18) метод простой итерации:

$$\frac{d}{dt}x^k(t) = D(t)x^k(t) + B(t)u^k(t), \quad x^k(t_0) = x_0, \quad (21)$$

$$p_1^{k+1} = \pi_+(p_1^k + \alpha(G_1 x_1^k - g_1)), \quad (22)$$

$$\frac{d}{dt}\psi^k(t) + D^T(t)\psi^k(t) = -Q_1(t)x^k(t), \quad \psi_1^k = Sx_1^k + G_1^T p_1^k, \quad (23)$$

$$u^{k+1}(t) = \pi_U(u^k(t) - \alpha(Q_2(t)u^k(t) + B^T(t)\psi^k(t))), \quad k = 0, 1, 2, \dots \quad (24)$$

В таком виде метод простой итерации не предназначен для вычисления седловых точек. Однако на его основе можно построить управляемый метод простой итерации, уже обладающий сходимостью к седловым точкам. В предлагаемом методе каждая итерация распадается на два полушага, за счет чего реализуется принцип обратной связи и обеспечивается сходимость. Общая концепция подхода изложена в [6]. Другие подходы градиентного типа рассматривались многими авторами, в основном, применительно к методам решения вариационных неравенств, в частности, отметим [17, 18].

Формулы управляемого итеративного экстраградиентного метода имеют следующий вид:

1) прогнозный полушаг

$$\frac{d}{dt}\bar{x}^k(t) = D(t)\bar{x}^k(t) + B(t)\bar{u}^k(t), \quad \bar{x}^k(t_0) = x_0, \quad (25)$$

$$\bar{p}_1^k = \pi_+(p_1^k + \alpha(G_1 \bar{x}_1^k - g_1)), \quad (26)$$

$$\frac{d}{dt}\bar{\psi}^k(t) + D^T(t)\bar{\psi}^k(t) = -Q_1(t)\bar{x}^k(t), \quad \bar{\psi}_1^k = S\bar{x}_1^k + G_1^T \bar{p}_1^k, \quad (27)$$

$$\bar{u}^k(t) = \pi_U(u^k(t) - \alpha(Q_2(t)u^k(t) + B^T(t)\bar{\psi}^k(t))); \quad (28)$$

2) основной полушаг

$$\frac{d}{dt}\bar{x}^k(t) = D(t)\bar{x}^k(t) + B(t)\bar{u}^k(t), \quad \bar{x}^k(t_0) = x_0, \quad (29)$$

$$p_1^{k+1} = \pi_+(p_1^k + \alpha(G_1 \bar{x}_1^k - g_1)), \quad (30)$$

$$\frac{d}{dt}\bar{\psi}^k(t) + D^T(t)\bar{\psi}^k(t) = -Q_1(t)\bar{x}^k(t), \quad \bar{\psi}_1^k = S\bar{x}_1^k + G_1^T\bar{p}_1^k, \quad (31)$$

$$u^{k+1}(t) = \pi_U(u^k(t) - \alpha(Q_2(t)\bar{u}^k(t) + B^T(t)\bar{\psi}^k(t))), \quad k = 0, 1, 2, \dots \quad (32)$$

Из формул этого процесса видно, что дифференциальные уравнения (25), (29) и (27), (31) используются только для вычисления функций $x^k(t)$ и $\bar{x}^k(t)$, $\psi^k(t)$ и $\bar{\psi}^k(t)$, поэтому процесс можно записать в более компактном виде

$$\bar{p}_1^k = \pi_+(p_1^k + \alpha(G_1x_1^k - g_1)), \quad (33)$$

$$p_1^{k+1} = \pi_+(p_1^k + \alpha(G_1\bar{x}_1^k - g_1)), \quad (34)$$

$$\bar{u}^k(t) = \pi_U(u^k(t) - \alpha(Q_2(t)u^k(t) + B^T(t)\psi^k(t))); \quad (35)$$

$$u^{k+1}(t) = \pi_U(u^k(t) - \alpha(Q_2(t)\bar{u}^k(t) + B^T(t)\bar{\psi}^k(t))), \quad t \in [t_0, t_1], \quad (36)$$

где $x^k(t)$, $\bar{x}^k(t)$, $\psi^k(t)$ и $\bar{\psi}^k(t)$ вычисляются в (25), (29), (27), (31).

Для получения вспомогательных оценок, необходимых для доказательства теоремы о сходимости метода, представим операторные уравнения (26), (30), (28), (32) в форме вариационных неравенств

$$\langle \bar{p}_1^k - p_1^k - \alpha(G_1x_1^k - g_1), p_1 - \bar{p}_1^k \rangle \geq 0, \quad (37)$$

$$\langle p_1^{k+1} - p_1^k - \alpha(G_1\bar{x}_1^k - g_1), p_1 - p_1^{k+1} \rangle \geq 0, \quad (38)$$

$$\int_{t_0}^{t_1} \langle \bar{u}^k(t) - u^k(t) + \alpha(Q_2(t)u^k(t) + B^T(t)\psi^k(t)), u(t) - \bar{u}^k(t) \rangle dt \geq 0, \quad (39)$$

$$\int_{t_0}^{t_1} \langle u^{k+1}(t) - u^k(t) + \alpha(Q_2(t)\bar{u}^k(t) + B^T(t)\bar{\psi}^k(t)), u(t) - u^{k+1}(t) \rangle dt \geq 0, \quad (40)$$

выполняемых при всех $p_1 \in \mathbb{R}_+^m$, $u(t) \in U$. Аналогично тому, как это делалось, например, в работах [3, 5], используя лемму Гронуолла [13], можно получить оценки:

$$|\bar{p}_1^k - p_1^{k+1}| \leq \alpha \|G_1\| |x_1^k - \bar{x}_1^k|, \quad (41)$$

$$\|\bar{u}^k(\cdot) - u^{k+1}(\cdot)\| \leq \alpha Q_{2\max} \|u^k(\cdot) - \bar{u}^k(\cdot)\| + \alpha B_{\max} \|\psi^k(\cdot) - \bar{\psi}^k(\cdot)\|, \quad (42)$$

$$|x^k(t) - \bar{x}^k(t)|^2 \leq e^{2D_{\max}(t_1-t_0)} B_{\max}^2(t_1-t_0) \|u^k(\cdot) - \bar{u}^k(\cdot)\|^2, \quad (43)$$

$$|x_1^k - \bar{x}_1^k|^2 \leq e^{2D_{\max}(t_1-t_0)} B_{\max}^2(t_1-t_0) \|u^k(\cdot) - \bar{u}^k(\cdot)\|^2, \quad (44)$$

$$|x^k(t) - x^*(t)|^2 \leq e^{2D_{\max}(t_1-t_0)} B_{\max}^2(t_1-t_0) \|u^k(\cdot) - u^*(\cdot)\|^2, \quad (45)$$

где $Q_{2\max} = \max\|Q_2(t)\|$, $B_{\max} = \max\|B(t)\|$, $D_{\max} = \max\|D(t)\|$, $t \in [t_0, t_1]$. Последняя из оценок означает, что ограниченное множество управлений линейный оператор переводит в ограниченное множество траекторий.

Так же получаются оценки отклонений сопряженных траекторий:

$$|\psi^k(t) - \bar{\psi}^k(t)|^2 \leq e^{2D_{\max}(t_1-t)} |\psi_1^k - \bar{\psi}_1^k|^2, \quad (46)$$

$$|\psi_0^k - \bar{\psi}_0^k|^2 \leq e^{2D_{\max}(t_1-t_0)} |\psi_1^k - \bar{\psi}_1^k|^2, \quad (47)$$

$$|\psi_1^k - \bar{\psi}_1^k|^2 \leq (\|S\| |x_1^k - \bar{x}_1^k| + \|G_1^T\| |p_1^k - \bar{p}_1^k|)^2, \quad (48)$$

$$|\psi_0^k - \bar{\psi}_0^k|^2 \leq 2e^{2D_{\max}(t_1-t_0)} (\|S\|^2 |x_1^k - \bar{x}_1^k|^2 + \|G_1^T\|^2 |p_1^k - \bar{p}_1^k|^2), \quad (49)$$

$$\begin{aligned} \|\psi^k(\cdot) - \bar{\psi}^k(\cdot)\|^2 &\leq (e^{2D_{\max}(t_1-t_0)} - 1) / (2D_{\max}) (\|S\| |x_1^k - \bar{x}_1^k| + \|G_1^T\| |p_1^k - \bar{p}_1^k|)^2 \leq \\ &\leq (e^{2D_{\max}(t_1-t_0)} - 1) / D_{\max} (\|S\|^2 |x_1^k - \bar{x}_1^k|^2 + \|G_1^T\|^2 |p_1^k - \bar{p}_1^k|^2), \end{aligned} \quad (50)$$

$$\|\psi^k(\cdot) - \psi^*(\cdot)\|^2 \leq (e^{2D_{\max}(t_1-t_0)} - 1) / (2D_{\max}) (\|S\| |x_1^k - x_1^*| + \|G_1^T\| |p_1^k - p_1^*|)^2. \quad (51)$$

Перейдем к формулировке и доказательству теоремы сходимости метода.

4. ДОКАЗАТЕЛЬСТВО СХОДИМОСТИ МЕТОДА

Покажем, что процесс (25)–(32) сходится к одному из решений исходной задачи.

Теорема 1. Если множество решений $(p_1^*, \psi_1^*, \Psi^*(\cdot); x_1^*, x^*(t), u^*(t))$ задачи (15)–(18) не пусто и принадлежит пространству $\mathbb{R}_+^m \times \mathbb{R}^n \times \Psi_2^n[t_0, t_1] \times \mathbb{R}^n \times \mathbb{AC}_2^n[t_0, t_1] \times U$, то последовательность $\{(p_1^k, \psi_1^k, \Psi^k(t); x_1^k, x^k(t), u^k(t))\}$, порожденная методом (25)–(32) с длиной шага α , выбранной из условия $0 < \alpha < \min \left\{ \frac{1}{\sqrt{\gamma_1 + \gamma_3 + \gamma_5}}, \frac{1}{\sqrt{\gamma_2 + \gamma_4}} \right\}$, где γ_i находятся в (64)–(68)¹, содержит подпоследовательность $\{(p_1^{k_i}, \psi_1^{k_i}, \Psi^{k_i}(t); x_1^{k_i}, x^{k_i}(t), u^{k_i}(t))\}$, которая сходится к решению задачи. При этом:

- 1) сходимость по управлениям — слабая, по фазовым и сопряженным траекториям, а также по переменным p_1 , ψ_1 — сильная;
- 2) последовательность $\left\{ |p_1^* - p_1^k|^2 + \|u^k(\cdot) - u^*(\cdot)\|_{L_2^r}^2 \right\}$ монотонно убывает на декартовом произведении $\mathbb{R}_+^m \times \mathbb{L}_2^n[t_0, t_1]$.

Доказательство. Основные усилия в теореме направлены на получение оценок $|u^k(t) - u^*(t)|^2$ и $|p_1^k - p_1^*|^2$.

1. Запишем оба уравнения (31) в виде вариационного неравенства

$$\int_{t_0}^{t_1} \langle Q_1(t) \bar{x}^k(t) + D^T(t) \bar{\psi}^k(t) + \frac{d}{dt} \bar{\psi}^k(t), x^*(t) - \bar{x}^k(t) \rangle dt + \langle S \bar{x}_1^k + G_1^T \bar{p}_1^k - \bar{\psi}_1^k, x_1^* - \bar{x}_1^k \rangle \geq 0.$$

Аналогично поступим с уравнением (17):

$$- \int_{t_0}^{t_1} \langle Q_1(t) x^*(t) + D^T(t) \psi^*(t) + \frac{d}{dt} \psi^*(t), x^*(t) - \bar{x}^k(t) \rangle dt - \langle S x_1^* + G_1^T p_1^* - \psi_1^*, x_1^* - \bar{x}_1^k \rangle \geq 0.$$

Сложим полученные неравенства

$$\begin{aligned} & \int_{t_0}^{t_1} \langle Q_1(t) (\bar{x}^k(t) - x^*(t)) + D^T(t) (\bar{\psi}^k(t) - \psi^*(t)) + \frac{d}{dt} (\bar{\psi}^k(t) - \psi^*(t)), x^*(t) - \bar{x}^k(t) \rangle dt + \\ & + \langle S (\bar{x}_1^k - x_1^*) + G_1^T (\bar{p}_1^k - p_1^*) - (\bar{\psi}_1^k - \psi_1^*), x_1^* - \bar{x}_1^k \rangle \geq 0. \end{aligned} \quad (52)$$

Используя формулу интегрирования по частям и учитывая, что $x_0^* = \bar{x}_0^k = x_0$, преобразуем дифференциальный член левой части (52) (это преобразование означает переход к сопряженному дифференциальному оператору). Сокращая затем подобные члены, получим неравенство

$$\begin{aligned} & \int_{t_0}^{t_1} \langle Q_1(t) (\bar{x}^k(t) - x^*(t)), x^*(t) - \bar{x}^k(t) \rangle dt + \\ & + \int_{t_0}^{t_1} \langle \bar{\psi}^k(t) - \psi^*(t), D(t) (x^*(t) - \bar{x}^k(t)) - \frac{d}{dt} (x^*(t) - \bar{x}^k(t)) \rangle dt + \\ & + \langle S (\bar{x}_1^k - x_1^*), x_1^* - \bar{x}_1^k \rangle + \langle G_1^T (\bar{p}_1^k - p_1^*), x_1^* - \bar{x}_1^k \rangle \geq 0. \end{aligned}$$

В силу симметричности и положительной полуопределенности матриц S и $Q_1(t)$ имеем

$$\langle S (\bar{x}_1^k - x_1^*), x_1^* - \bar{x}_1^k \rangle \leq 0, \quad \int_{t_0}^{t_1} \langle Q_1(t) (\bar{x}^k(t) - x^*(t)), x^*(t) - \bar{x}^k(t) \rangle dt \leq 0,$$

поэтому, отбрасывая эти члены и усиливая оценку, находим

$$\int_{t_0}^{t_1} \langle \bar{\psi}^k(t) - \psi^*(t), D(t) (x^*(t) - \bar{x}^k(t)) - \frac{d}{dt} (x^*(t) - \bar{x}^k(t)) \rangle dt + \langle G_1^T (\bar{p}_1^k - p_1^*), x_1^* - \bar{x}_1^k \rangle \geq 0. \quad (53)$$

2. Получим неравенство относительно переменной p_1 . Для этого положим $p_1 = p_1^{k+1}$ в (37):

$$\langle \bar{p}_1^k - p_1^k - \alpha (G_1 x_1^k - g_1), p_1^{k+1} - \bar{p}_1^k \rangle \geq 0.$$

¹ См. доказательство теоремы далее.

Добавим и вычтем $\alpha \langle G_1 \bar{x}_1^k - g_1, p_1^{k+1} - \bar{p}_1^k \rangle$:

$$\langle \bar{p}_1^k - p_1^k, p_1^{k+1} - \bar{p}_1^k \rangle + \alpha \langle (G_1 \bar{x}_1^k - g_1) - (G_1 x_1^k - g_1), p_1^{k+1} - \bar{p}_1^k \rangle - \alpha \langle G_1 \bar{x}_1^k - g_1, p_1^{k+1} - \bar{p}_1^k \rangle \geq 0.$$

Используя (41), оценим второе слагаемое

$$\langle \bar{p}_1^k - p_1^k, p_1^{k+1} - \bar{p}_1^k \rangle + \alpha^2 \|G_1\|^2 |\bar{x}_1^k - x_1^k|^2 - \alpha \langle G_1 \bar{x}_1^k - g_1, p_1^{k+1} - \bar{p}_1^k \rangle \geq 0.$$

Положим $p_1 = p_1^*$ в (38):

$$\langle p_1^{k+1} - p_1^k, p_1^* - p_1^{k+1} \rangle - \alpha \langle G_1 \bar{x}_1^k - g_1, p_1^* - p_1^{k+1} \rangle \geq 0.$$

Сложим полученные неравенства

$$\langle \bar{p}_1^k - p_1^k, p_1^{k+1} - \bar{p}_1^k \rangle + \langle p_1^{k+1} - p_1^k, p_1^* - p_1^{k+1} \rangle + \alpha^2 \|G_1\|^2 |\bar{x}_1^k - x_1^k|^2 - \alpha \langle G_1 \bar{x}_1^k - g_1, p_1^* - \bar{p}_1^k \rangle \geq 0.$$

Полагая $p_1 = \bar{p}_1^k$ в неравенстве (16), имеем

$$\alpha \langle p_1^* - \bar{p}_1^k, G_1 x_1^* - g_1 \rangle \geq 0.$$

Суммируем два последних неравенства

$$\langle \bar{p}_1^k - p_1^k, p_1^{k+1} - \bar{p}_1^k \rangle + \langle p_1^{k+1} - p_1^k, p_1^* - p_1^{k+1} \rangle + \alpha^2 \|G_1\|^2 |\bar{x}_1^k - x_1^k|^2 - \alpha \langle G_1 (\bar{x}_1^k - x_1^*), p_1^* - \bar{p}_1^k \rangle \geq 0. \quad (54)$$

3. Продолжим получение оценок. Рассмотрим неравенства относительно управлений. Положим $u(t) = u^{k+1}(t)$ в (39)

$$\int_{t_0}^{t_1} \langle \bar{u}^k(t) - u^k(t) + \alpha(Q_2(t)u^k(t) + B^T(t)\psi^k(t)), u^{k+1}(t) - \bar{u}^k(t) \rangle dt \geq 0.$$

Разобьем левую часть на три отдельных слагаемых:

$$\begin{aligned} & \int_{t_0}^{t_1} \langle \bar{u}^k(t) - u^k(t), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \alpha \int_{t_0}^{t_1} \langle Q_2(t)u^k(t), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \\ & + \alpha \int_{t_0}^{t_1} \langle B^T(t)\psi^k(t), u^{k+1}(t) - \bar{u}^k(t) \rangle dt \geq 0, \end{aligned}$$

добавим и вычтем $\bar{u}^k(t)$ под знаком второго скалярного произведения, и аналогично добавим и вычтем $\bar{\psi}^k(t)$ под знаком третьего скалярного произведения:

$$\begin{aligned} & \int_{t_0}^{t_1} \langle \bar{u}^k(t) - u^k(t), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \alpha \int_{t_0}^{t_1} \langle Q_2(t)(u^k(t) - \bar{u}^k(t)), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \\ & + \alpha \int_{t_0}^{t_1} \langle Q_2(t)\bar{u}^k(t), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \alpha \int_{t_0}^{t_1} \langle B^T(t)(\psi^k(t) - \bar{\psi}^k(t)), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \\ & + \alpha \int_{t_0}^{t_1} \langle B^T(t)\bar{\psi}^k(t), u^{k+1}(t) - \bar{u}^k(t) \rangle dt \geq 0. \end{aligned} \quad (55)$$

Положим $u = u^*(t)$ в (40)

$$\int_{t_0}^{t_1} \langle u^{k+1}(t) - u^k(t) + \alpha(Q_2(t)\bar{u}^k(t) + B^T(t)\bar{\psi}^k(t)), u^*(t) - u^{k+1}(t) \rangle dt \geq 0. \quad (56)$$

Сложим (55) и (56), тогда получим

$$\begin{aligned} & \int_{t_0}^{t_1} \langle \bar{u}^k(t) - u^k(t), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \int_{t_0}^{t_1} \langle u^{k+1}(t) - u^k(t), u^*(t) - u^{k+1}(t) \rangle dt + \\ & + \alpha \int_{t_0}^{t_1} \langle Q_2(t)(u^k(t) - \bar{u}^k(t)), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \alpha \int_{t_0}^{t_1} \langle Q_2(t)\bar{u}^k(t), u^*(t) - \bar{u}^k(t) \rangle dt + \\ & + \alpha \int_{t_0}^{t_1} \langle B^T(t)(\psi^k(t) - \bar{\psi}^k(t)), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \alpha \int_{t_0}^{t_1} \langle B^T(t)\bar{\psi}^k(t), u^*(t) - \bar{u}^k(t) \rangle dt \geq 0. \end{aligned} \quad (57)$$

Подставляя $u(t) = \bar{u}^k(t)$ в вариационное неравенство (18), имеем

$$-\alpha \int_{t_0}^{t_1} \langle Q_2(t)u^*(t) + B^T(t)\psi^*(t), u^*(t) - \bar{u}^k(t) \rangle dt \geq 0. \quad (58)$$

Суммируем (57) и (58)

$$\begin{aligned} & \int_{t_0}^{t_1} \langle \bar{u}^k(t) - u^k(t), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \int_{t_0}^{t_1} \langle u^{k+1}(t) - u^k(t), u^*(t) - u^{k+1}(t) \rangle dt + \\ & + \alpha \int_{t_0}^{t_1} \langle Q_2(t)(u^k(t) - \bar{u}^k(t)), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \alpha \int_{t_0}^{t_1} \langle Q_2(t)(\bar{u}^k(t) - u^*(t)), u^*(t) - \bar{u}^k(t) \rangle dt + \\ & + \alpha \int_{t_0}^{t_1} \langle B^T(t)(\psi^k(t) - \bar{\psi}^k(t)), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \alpha \int_{t_0}^{t_1} \langle \bar{\psi}^k(t) - \psi^*(t), B(t)(u^*(t) - \bar{u}^k(t)) \rangle dt \geq 0. \end{aligned} \quad (59)$$

Умножим (53) на α и сложим с (59)

$$\begin{aligned} & \alpha \langle G_1^T(\bar{p}_1^k - p_1^*), x_1^* - \bar{x}_1^k \rangle + \alpha \int_{t_0}^{t_1} \langle \bar{\psi}^k(t) - \psi^*(t), D(t)(x^*(t) - \bar{x}^k(t)) + B(t)(u^*(t) - \bar{u}^k(t)) - \frac{d}{dt}(x^*(t) - \bar{x}^k(t)) \rangle dt + \\ & + \int_{t_0}^{t_1} \langle \bar{u}^k(t) - u^k(t), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \int_{t_0}^{t_1} \langle u^{k+1}(t) - u^k(t), u^*(t) - u^{k+1}(t) \rangle dt + \\ & + \alpha \int_{t_0}^{t_1} \langle Q_2(t)(u^k(t) - \bar{u}^k(t)), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \alpha \int_{t_0}^{t_1} \langle Q_2(t)(\bar{u}^k(t) - u^*(t)), u^*(t) - \bar{u}^k(t) \rangle dt + \\ & + \alpha \int_{t_0}^{t_1} \langle B^T(t)(\psi^k(t) - \bar{\psi}^k(t)), u^{k+1}(t) - \bar{u}^k(t) \rangle dt \geq 0. \end{aligned} \quad (60)$$

В силу (15) и (29) первый из интегралов обнуляется и, отбрасывая отрицательный член $\alpha \int_{t_0}^{t_1} \langle Q_2(t)(\bar{u}^k(t) - u^*(t)), u^*(t) - \bar{u}^k(t) \rangle dt$, имеем

$$\begin{aligned} & \alpha \langle G_1^T(\bar{p}_1^k - p_1^*), x_1^* - \bar{x}_1^k \rangle + \int_{t_0}^{t_1} \langle \bar{u}^k(t) - u^k(t), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \\ & + \int_{t_0}^{t_1} \langle u^{k+1}(t) - u^k(t), u^*(t) - u^{k+1}(t) \rangle dt + \alpha \int_{t_0}^{t_1} \langle Q_2(t)(u^k(t) - \bar{u}^k(t)), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \\ & + \alpha \int_{t_0}^{t_1} \langle B^T(t)(\psi^k(t) - \bar{\psi}^k(t)), u^{k+1}(t) - \bar{u}^k(t) \rangle dt \geq 0. \end{aligned} \quad (61)$$

Сложим (60) и (61)

$$\begin{aligned} & \langle \bar{p}_1^k - p_1^k, p_1^{k+1} - \bar{p}_1^k \rangle + \langle p_1^{k+1} - p_1^k, p_1^* - p_1^{k+1} \rangle + \alpha^2 \|G_1\|^2 |\bar{x}_1^k - x_1^k|^2 + \\ & + \int_{t_0}^{t_1} \langle \bar{u}^k(t) - u^k(t), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \int_{t_0}^{t_1} \langle u^{k+1}(t) - u^k(t), u^*(t) - u^{k+1}(t) \rangle dt + \\ & + \alpha \int_{t_0}^{t_1} \langle Q_2(t)(u^k(t) - \bar{u}^k(t)), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \alpha \int_{t_0}^{t_1} \langle B^T(t)(\psi^k(t) - \bar{\psi}^k(t)), u^{k+1}(t) - \bar{u}^k(t) \rangle dt \geq 0. \end{aligned} \quad (62)$$

4. Умножая неравенство (62) на два и используя тождество $2\langle y_1 - y_3, y_3 - y_2 \rangle = |y_1 - y_2|^2 - |y_1 - y_3|^2 - |y_3 - y_2|^2$, разложим первые четыре скалярных произведения в сумму квадратов. Сократив затем подобные члены и умножив на минус единицу, перепишем неравенство в виде

$$\begin{aligned} & |p_1^{k+1} - \bar{p}_1^k|^2 + |\bar{p}_1^k - p_1^k|^2 + |p_1^{k+1} - p_1^*|^2 - 2\alpha^2 \|G_1\|^2 |\bar{x}_1^k - x_1^k|^2 + \\ & + \|u^{k+1}(\cdot) - u^*(\cdot)\|^2 + \|\bar{u}^k(\cdot) - u^{k+1}(\cdot)\|^2 + \|u^k(\cdot) - \bar{u}^k(\cdot)\|^2 + \\ & + 2\alpha \int_{t_0}^{t_1} \langle Q_2(t)(\bar{u}^k(t) - u^k(t)), u^{k+1}(t) - \bar{u}^k(t) \rangle dt + \end{aligned}$$

$$+2\alpha \int_{t_0}^{t_1} \langle B^T(t)(\bar{\psi}^k(t) - \psi^k(t)), u^{k+1}(t) - \bar{u}^k(t) \rangle dt \leq \\ \leq |p_1^k - p_1^*|^2 + \|u^k(\cdot) - u^*(\cdot)\|^2.$$

Используя неравенство Коши—Буняковского, оценим оставшиеся слагаемые в форме скалярных произведений:

$$|p_1^{k+1} - \bar{p}_1^k|^2 + |\bar{p}_1^k - p_1^k|^2 + |p_1^{k+1} - p_1^*|^2 - 2\alpha^2 \|G_1\|^2 |\bar{x}_1^k - x_1^k|^2 + \\ + \|u^{k+1}(\cdot) - u^*(\cdot)\|^2 + \|\bar{u}^k(\cdot) - u^{k+1}(\cdot)\|^2 + \|u^k(\cdot) - \bar{u}^k(\cdot)\|^2 - \\ - 2\alpha Q_{2\max} \|\bar{u}^k(\cdot) - u^k(\cdot)\| \|u^{k+1}(\cdot) - \bar{u}^k(\cdot)\| - \\ - 2\alpha B_{\max} \|\bar{\psi}^k(\cdot) - \psi^k(\cdot)\| \|u^{k+1}(\cdot) - \bar{u}^k(\cdot)\| \leq |p_1^k - p_1^*|^2 + \|u^k(\cdot) - u^*(\cdot)\|^2. \quad (63)$$

а) Используя (42) и неравенство $2ab \leq a^2 + b^2$, получим

$$2\alpha B_{\max} \|\bar{\psi}^k(\cdot) - \psi^k(\cdot)\| \|u^{k+1}(\cdot) - \bar{u}^k(\cdot)\| \leq \\ \leq \alpha^2 B_{\max} (Q_{2\max} + 2B_{\max}) \|\bar{\psi}^k(\cdot) - \psi^k(\cdot)\|^2 + \alpha^2 B_{\max} Q_{2\max} \|u^k(\cdot) - \bar{u}^k(\cdot)\|^2.$$

В силу (43) и (50)

$$\|\psi^k(\cdot) - \bar{\psi}^k(\cdot)\|^2 \leq \\ \leq \left((e^{2D_{\max}(t_1-t_0)} - 1) / D_{\max} \right) \|S\|^2 e^{2D_{\max}(t_1-t_0)} B_{\max}^2(t_1 - t_0) \|u^k(\cdot) - \bar{u}^k(\cdot)\|^2 + \\ + \left((e^{2D_{\max}(t_1-t_0)} - 1) / D_{\max} \right) \|G_1^T\|^2 |p_1^k - \bar{p}_1^k|^2 = \lambda_1 \|u^k(\cdot) - \bar{u}^k(\cdot)\|^2 + \lambda_2 |p_1^k - \bar{p}_1^k|^2,$$

где $\lambda_1 = ((e^{2D_{\max}(t_1-t_0)} - 1) / D_{\max}) \|S\|^2 e^{2D_{\max}(t_1-t_0)} B_{\max}^2(t_1 - t_0)$,
 $\lambda_2 = ((e^{2D_{\max}(t_1-t_0)} - 1) / D_{\max}) \|G_1^T\|^2$, ПОЭТОМУ ИМЕЕМ

$$2\alpha B_{\max} \|\bar{\psi}^k(\cdot) - \psi^k(\cdot)\| \|u^{k+1}(\cdot) - \bar{u}^k(\cdot)\| \leq \\ \leq \alpha^2 B_{\max} ((Q_{2\max} + 2B_{\max}) \lambda_1 + Q_{2\max}) \|u^k(\cdot) - \bar{u}^k(\cdot)\|^2 + \\ + \alpha^2 B_{\max} (Q_{2\max} + 2B_{\max}) \lambda_2 |p_1^k - \bar{p}_1^k|^2 = \alpha^2 \gamma_1 \|u^k(\cdot) - \bar{u}^k(\cdot)\|^2 + \alpha^2 \gamma_2 |p_1^k - \bar{p}_1^k|^2, \\ \text{где } \gamma_1 = B_{\max} ((Q_{2\max} + 2B_{\max}) \lambda_1 + Q_{2\max}), \gamma_2 = B_{\max} (Q_{2\max} + 2B_{\max}) \lambda_2. \quad (64)$$

б) Аналогично,

$$2\alpha Q_{2\max} \|\bar{u}^k(\cdot) - u^k(\cdot)\| \|u^{k+1}(\cdot) - \bar{u}^k(\cdot)\| \leq \alpha^2 \gamma_3 \|u^k(\cdot) - \bar{u}^k(\cdot)\|^2 + \alpha^2 \gamma_4 |p_1^k - \bar{p}_1^k|^2, \\ \text{где } \gamma_3 = Q_{2\max} (2Q_{2\max} + B_{\max} + B_{\max} \lambda_1), \gamma_4 = Q_{2\max} B_{\max} \lambda_2. \quad (65)$$

в) В силу (50)

$$2\alpha^2 \|G_1\|^2 |\bar{x}_1^k - x_1^k|^2 \leq 2\alpha^2 \|G_1\|^2 e^{2D_{\max}(t_1-t_0)} B_{\max}^2(t_1 - t_0) \|u^k(\cdot) - \bar{u}^k(\cdot)\|^2 = \\ = \alpha^2 \gamma_5 \|u^k(\cdot) - \bar{u}^k(\cdot)\|^2, \text{ где } \gamma_5 = 2\|G_1\|^2 e^{2D_{\max}(t_1-t_0)} B_{\max}^2(t_1 - t_0). \quad (66)$$

С учетом полученных оценок неравенство (63) примет вид

$$|p_1^{k+1} - p_1^*|^2 + \|u^{k+1}(\cdot) - u^*(\cdot)\|^2 + |p_1^{k+1} - \bar{p}_1^k|^2 + (1 - \alpha^2(\gamma_2 + \gamma_4)) |\bar{p}_1^k - p_1^k|^2 + \\ + \|\bar{u}^k(\cdot) - u^{k+1}(\cdot)\|^2 + (1 - \alpha^2(\gamma_1 + \gamma_3 + \gamma_5)) \|u^k(\cdot) - \bar{u}^k(\cdot)\|^2 \leq \\ \leq |p_1^k - p_1^*|^2 + \|u^k(\cdot) - u^*(\cdot)\|^2. \quad (67)$$

Выбирая значение параметра α из условия

$$0 < \alpha < \min \left\{ \frac{1}{\sqrt{\gamma_1 + \gamma_3 + \gamma_5}}, \frac{1}{\sqrt{\gamma_2 + \gamma_4}} \right\}, \quad (68)$$

обеспечим положительность всех слагаемых в левой части (67). Отбрасывая в (67) часть положительных членов, получим

$$|p_1^{k+1} - p_1^*|^2 + \|u^{k+1}(\cdot) - u^*(\cdot)\|^2 \leq |p_1^k - p_1^*|^2 + \|u^k(\cdot) - u^*(\cdot)\|^2, \quad (69)$$

что означает монотонное убывание $\{|p_1^k - p_1^*|^2 + \|u^k(\cdot) - u^*(\cdot)\|^2\}$.

5. Просуммируем неравенство (67) от $k = 0$ до $k = N$:

$$|p_1^{N+1} - p_1^*|^2 + \|u^{N+1}(\cdot) - u^*(\cdot)\|^2 + \sum_{k=0}^N |p_1^{k+1} - \bar{p}_1^k|^2 + (1 - \alpha^2(\gamma_2 + \gamma_4)) \sum_{k=0}^N |\bar{p}_1^k - p_1^k|^2 + \sum_{k=0}^N \|\bar{u}^k(\cdot) - u^{k+1}(\cdot)\|^2 + (1 - \alpha^2(\gamma_1 + \gamma_3 + \gamma_5)) \sum_{k=0}^N \|u^k(\cdot) - \bar{u}^k(\cdot)\|^2 \leq |p_1^0 - p_1^*|^2 + \|u^0(\cdot) - u^*(\cdot)\|^2. \quad (70)$$

Из неравенства (70) следует ограниченность при любом N

$$|p_1^{N+1} - p_1^*|^2 + \|u^{N+1}(\cdot) - u^*(\cdot)\|^2 \leq |p_1^0 - p_1^*|^2 + \|u^0(\cdot) - u^*(\cdot)\|^2, \quad (71)$$

а также сходимость рядов

$$\sum_{k=0}^{+\infty} |p_1^{k+1} - \bar{p}_1^k|^2 < \infty, \quad \sum_{k=0}^{+\infty} |\bar{p}_1^k - p_1^k|^2 < \infty, \\ \sum_{k=0}^{+\infty} \|\bar{u}^k(\cdot) - u^{k+1}(\cdot)\|^2 < \infty, \quad \sum_{k=0}^{+\infty} \|u^k(\cdot) - \bar{u}^k(\cdot)\|^2 < \infty$$

и, следовательно, стремление к нулю величин

$$|p_1^{k+1} - \bar{p}_1^k| \rightarrow 0, \quad |\bar{p}_1^k - p_1^k| \rightarrow 0, \quad \|\bar{u}^k(\cdot) - u^{k+1}(\cdot)\| \rightarrow 0, \quad \|u^k(\cdot) - \bar{u}^k(\cdot)\| \rightarrow 0, \quad (72)$$

откуда по неравенству треугольника получим

$$|p_1^{k+1} - p_1^k| \rightarrow 0, \quad \|u^{k+1}(\cdot) - u^k(\cdot)\| \rightarrow 0. \quad (73)$$

Из (47), (48), (56) тогда следует, что

$$|x^k(t) - \bar{x}^k(t)| \rightarrow 0, \quad |x_1^k - \bar{x}_1^k| \rightarrow 0, \quad \|\psi^k(\cdot) - \bar{\psi}^k(\cdot)\| \rightarrow 0. \quad (74)$$

Кроме того, из (73) следует ограниченность последовательностей

$$|p_1^k - p_1^*| \leq \text{const}, \quad \|u^k(\cdot) - u^*(\cdot)\| \leq \text{const}, \quad (75)$$

а из (49) и (57) — ограниченность последовательностей

$$\|x^k(\cdot) - x^*(\cdot)\| \leq \text{const}, \quad |x_1^k - x_1^*| \leq \text{const}, \quad \|\psi^k(\cdot) - \psi^*(\cdot)\| \leq \text{const}. \quad (76)$$

6. Поскольку последовательность $\{(p_1^k, \psi_1^k, \psi^k(t); x_1^k, x^k(t), u^k(t))\}$ ограничена на декартовом произведении пространств $\mathbb{R}_+^n \times \mathbb{R}^n \times \Psi_2^n[t_0, t_1] \times \mathbb{R}^n \times \mathbb{A}\mathbb{C}^n[t_0, t_1] \times \mathbb{U}$, то она слабо компактна. Это означает, что существует подпоследовательность $\{(p_1^{k_i}, \psi_1^{k_i}, \psi^{k_i}(t); x_1^{k_i}, x^{k_i}(t), u^{k_i}(t))\}$ и точка $(p_1', \psi_1', \psi'(t); x_1', x'(t), u'(t))$, которая является ее слабым пределом. Покажем, что слабый предел этой последовательности является решением задачи (15)–(18).

Заметим, что все линейные дифференциальные операторы системы (25)–(32) являются слабо непрерывными (за исключением уравнений (28), (32)) и поэтому допускают переход к слабому пределу. Перейдем, например, в основном полушаге (29)–(32) к слабому пределу при $k_i \rightarrow \infty$ с учетом (72)–(76) и $x_1 \in \mathbb{R}^n$, получим

$$\frac{d}{dt} x'(t) = D(t)x'(t) + B(t)u'(t), \quad x'(t_0) = x_0, \quad p_1' = \pi_+(p_1' + \alpha(G_1 x_1' - g_1)), \\ \frac{d}{dt} \psi'(t) + D^T(t)\psi'(t) = -Q_1(t)x'(t), \quad \psi_1' = Sx_1' + G_1^T p_1'. \quad (77)$$

Точка $(p_1', \psi_1', \psi'(t); x_1', x'(t), u'(t))$, являясь решением системы (77), одновременно является решением задачи (2). Задача (2), в свою очередь, порождает седловую функцию (4). Эта функция имеет седловую точку, удовлетворяющую системе (5). Компоненты $(x_1', x'(t), u'(t))$ седловой точки являются решением задачи (2). Компоненты $(p_1', \psi_1', \psi'(t))$ являются решением двойственной задачи. Эта задача порождается двойственной, или сопряженной, функцией Лагранжа. Чтобы получить двойственную функцию Лагранжа, достаточно в лагранжиане (4) перейти к сопряженным линейным операторам. Тогда получим сопряженную функцию Лагранжа (6), которая также имеет седловую точку. Эта седловая точка удовлетворяет сопряженной системе неравенств (7).

Понятно, что седловые точки обеих систем совпадают. Система (14) в явном виде представляет двойственную по отношению к (3) задачу. Поскольку решения обеих систем есть одно и то же множество, то выполняется равенство

$$u'(t) = \pi_U(u'(t) - \alpha(Q_2(t)u'(t) + B^T(t)\psi'(t))). \quad (78)$$

Теперь, если к системе (77) добавить уравнение (78) или его вариационный аналог из (18) и затем перейти к пределу при $k_i \rightarrow \infty$, то получим полную предельную систему, совпадающую с (15)–(18). Из последнего следует, что предельная система совпадает с (25)–(32) и поэтому можно считать, что

$$(p'_1, \psi'_1, \psi'(t); x'_1, x'(t), u'(t)) = (p_1^*, \psi_1^*, \psi^*(t); x_1^*, x^*(t), u^*(t)),$$

т. е. любая предельная точка процесса (25)–(32) является решением задачи.

Теорема доказана.

5. СИНТЕЗ УПРАВЛЕНИЯ ДЛЯ ЛИНЕЙНОЙ ДИФФЕРЕНЦИАЛЬНОЙ СИСТЕМЫ ПРИ НАЛИЧИИ ОГРАНИЧЕНИЙ НА УПРАВЛЕНИЕ

Покажем, как на основе предложенной теории можно сконструировать (синтезировать) управление линейной системой (1) при наличии ограничений на управление. С этой целью выпишем систему (15)–(18):

$$\frac{d}{dt}x^*(t) = D(t)x^*(t) + B(t)u^*(t), \quad t_0 \leq t \leq t_1, \quad x^*(t_0) = x_0, \quad (79)$$

$$\langle p_1 - p_1^*, G_1 x_1^* - g_1 \rangle \leq 0, \quad p_1 \in \mathbb{R}_+^m, \quad (80)$$

$$\frac{d}{dt}\psi^*(t) + D^T(t)\psi^*(t) = -Q_1(t)x^*(t), \quad \psi_1^* = Sx_1^* + G_1^T p_1^*, \quad (81)$$

$$\int_{t_0}^{t_1} \langle Q_2(t)u^*(t) + B^T(t)\psi^*(t), u^*(t) - u(t) \rangle dt \leq 0, \quad u(t) \in U. \quad (82)$$

Как уже отмечалось, решение этой системы представляет собой седловую точку. Это значит, что если известно x_0 и управление $u(t)$, то можно решить дифференциальное уравнение (79) и найти траекторию $x(t)$. Затем можно вычислить ее значение на правом конце $x_1 = x(t)|_{t=t_1}$ и найти p_1 из вариационного неравенства (80). Сформировав условие трансверсальности (второе конечномерное уравнение из (81)), мы можем решить сопряженную дифференциальную систему (первое дифференциальное уравнение из (81)) и найти сопряженную траекторию $\psi(t)$, с помощью которой можно сформировать нормаль опорной плоскости $\int_{t_0}^{t_1} \langle Q_2(t)u(t) + B^T(t)\psi(t), u(t) \rangle dt$ для функционального множества U . Далее остается только найти точку $u(t) \in U$, которая является опорной для нашего линейного функционала. Если начальные данные для уравнения (79) были оптимальными, то найденное управление совпадает с оптимальным $u^*(t)$. В этом случае решение системы (79)–(82) найдено.

Эта система имеет очень простую алгоритмическую структуру. Фактически из первого уравнения системы от двух переменных нужно выразить одну через другую. С помощью полученной переменной из условия трансверсальности определить терминальное значение для сопряженного дифференциального уравнения. Затем решить это уравнение. С помощью найденного решения формируем нормаль плоскости на выпуклом замкнутом множестве управлений U и находим опорную точку.

Эта содержательная схема носит алгоритмический характер, поэтому представим ее в виде некоторого алгоритма. С этой целью рассмотрим следующие отображения (операторы).

1. Введем оператор $K_1(u) = x(t_1; u(t))$, который каждому управлению $u(t) \in U$ из уравнения (79) ставит в соответствие правый конец $x_1 = x(t_1; u(t)) \in \mathbb{R}^n$ траектории $x(t; u(t))$. В [13] показано, что таким образом определенный оператор, действующий из гильбертова пространства в конечномерное, является однозначным, линейным и ограниченным (непрерывным).

2. Условие трансверсальности для каждого x_1 однозначно определяет терминальное условие $\psi_1 = Sx_1 + G_1^T p_1$ для сопряженного дифференциального уравнения (81). Это уравнение порождает сопряженную траекторию $\psi(t) = K_2(x) = \psi(x_1; t)$. Оба отображения зависят от параметров, которые согласовывают области определения операторов так, чтобы второй был определен на образе первого. В методе, который описан в этой работе, это делается автоматически. Поэтому представляется естественным рассмотреть суперпозицию операторов.

3. Теперь введем оператор $K_3(\psi) = VI(U, \psi)$, который каждому элементу ψ сопряженного пространства $\Psi_2^n[t_0, t_1]$ ставит в соответствие решение вариационного неравенства (82). Вектор $\psi(t)$ порождает нормаль линейного функционала $Q_2(t)u^*(t) + B^T(t)\psi^*(t)$, определенного на множестве управлений U .

Все три отображения зависят от параметров, которые согласовываются с помощью терминальных условий таким образом, чтобы все отображения порождали суперпозицию. Эта суперпозиция есть отображение, которое преобразует множество U в себя, и его неподвижная точка является решением исходной задачи. Нетрудно видеть, что эту систему отображений можно преобразовать в управление по замкнутому контуру

$$u(t) = K(t_1)x(t),$$

где $u(t) = K(t_1)x(t) = K_3(K_2(K_1(u)))$. Тогда система (15)–(18) с управлением в виде обратной связи принимает вид

$$\frac{d}{dt}x(t) = D(t)x(t) + B(t)u(t), \quad x(t_0) = x_0, \quad u(t) = K(t_1)x(t). \quad (83)$$

Соответственно, вычислительный процесс (25)–(32) может рассматриваться как итеративный процесс решения системы (83) с управлением в виде обратной связи. Обратную связь можно интерпретировать как градиентный спуск к седловой точке функции Лагранжа задачи терминального управления. Фактически синтез управления вложен в вычислительный процесс (25)–(32).

6. ТЕОРИЯ ЛИНЕЙНОГО РЕГУЛЯТОРА. СИНТЕЗ УПРАВЛЕНИЯ ДЛЯ ЛИНЕЙНОЙ ДИФФЕРЕНЦИАЛЬНОЙ СИСТЕМЫ: КЛАССИЧЕСКИЙ ПОДХОД

Отметим один важный частный случай этой задачи, который известен как теория линейного регулятора [19, 20]. В этом случае ограничения на управления отсутствуют, т. е. $U \equiv \mathbb{L}_2^r[t_0, t_1]$. Тогда задача (2) принимает вид

$$\begin{aligned} \frac{d}{dt}x(t) = D(t)x(t) + B(t)u(t), \quad t_0 \leq t \leq t_1, \quad x(t_0) = x_0, \\ (x_1^*, x^*(t), u^*(t)) \in \operatorname{Argmin} \left\{ \frac{1}{2} \int_{t_0}^{t_1} (\langle Q_1(t)x(t), x(t) \rangle + \langle Q_2(t)u(t), u(t) \rangle) dt + \right. \\ \left. + \frac{1}{2} \langle Sx_1, x_1 \rangle \mid u(\cdot) \in \mathbb{L}_2^r[t_0, t_1] \right\}. \end{aligned} \quad (84)$$

Дифференциальная система (15)–(18) для этой задачи, которая фактически является необходимым и достаточным условием оптимальности, принимает форму системы дифференциальных и конечномерных уравнений. Преобразуем эту систему

$$\frac{d}{dt}x(t) = D(t)x(t) + B(t)u(t), \quad x(t_0) = x_0, \quad (85)$$

$$\frac{d}{dt}\psi(t) + D^T(t)\psi(t) = -Q_1(t)x(t), \quad \psi_1 = Sx_1^* + Q_2(t)u(t) + B^T(t)\psi(t) = 0, \quad (86)$$

$$u(t) = -Q_2^{-1}(t)B^T(t)\psi(t). \quad (87)$$

Прежде всего, из этой системы необходимо исключить фазовую переменную $\psi(t)$. С этой целью в [19] было сделано предположение, что должна существовать связь типа

$$\psi(t) = P(t)x(t), \quad (88)$$

где $P(t)$ – матричная непрерывно дифференцируемая функция соответствующей размерности. Поскольку в системе (85)–(87) имеется производная этой функции, которая нам понадобится в дальнейших рассуждениях, то продифференцируем (88):

$$\frac{d}{dt}\psi(t) = \frac{dP(t)}{dt}x(t) + P(t)\frac{dx(t)}{dt}. \quad (89)$$

Сопоставляя уравнения (89) и (86), получаем

$$\frac{dP(t)}{dt}x(t) + P(t)\frac{dx(t)}{dt} = -Q_1(t)x(t) - D^T(t)P(t)x(t). \quad (90)$$

Итак, мы проследили воздействие (88) на (86). Теперь посмотрим на результат взаимодействия (88) с (85). Для этого подставим $\psi(t)$ из (88) в (87), а затем найденное $u(t)$ подставим в (85), тогда получим

$$\frac{d}{dt}x(t) = D(t)x(t) - B(t)Q_2^{-1}(t)B^T(t)P(t)x(t), \quad x(t_0) = x_0. \quad (91)$$

Исключая производную $\frac{d}{dt}x(t)$ из (90) и (91), находим

$$\frac{dP(t)}{dt}x(t) + P(t)(D(t)x(t) - B(t)Q_2^{-1}(t)B^T(t)P(t)x(t)) = -Q_1(t)x(t) - D^T(t)P(t)x(t).$$

Отсюда имеем

$$\left(\frac{dP(t)}{dt} + P(t)D(t) + D^T(t)P(t) - P(t)B(t)Q_2^{-1}(t)B^T(t)P(t) + Q_1(t)\right)x(t) = 0.$$

Это равенство должно выполняться для произвольных $x(t)$, следовательно, множитель, стоящий в больших скобках, должен быть равен нулю. Отсюда получаем матричное уравнение Риккати с граничным условием (86)

$$\frac{dP(t)}{dt} = -P(t)D(t) - D^T(t)P(t) + P(t)B(t)Q_2^{-1}(t)B^T(t)P(t) + Q_1(t) = 0, \quad P(t_1) = S.$$

Из уравнения видно, что матрица $P(t)$ размерности $n \times n$ должна быть симметрической и положительно-определенной. Из (87) и (88) имеем

$$u(t) = -Q_2^{-1}(t)B^T(t)P(t)x(t),$$

т. е. матрица $K(T)$ из (83) имеет структуру $K(T) = -Q_2^{-1}(t)B^T(t)P(t)$.

7. ЗАКЛЮЧЕНИЕ

В статье рассматривается задача терминального управления с конечномерной краевой задачей на правом конце временного интервала. Задача имеет выпуклую структуру, что позволяет в рамках теории двойственности разработать теорию седловых методов для решения терминальных проблем управления.

Доказывается монотонная сходимостъ вычислительного процесса на декартовом произведении прямых и двойственных переменных, при этом доказывается слабая сходимостъ по управлениям, сильная сходимостъ по фазовым и сопряженным траекториям, а также по терминальным переменным краевой задачи. На базе седлового подхода строится синтез управления, т. е. обратная связь при наличии ограничений на управления в форме выпуклого замкнутого множества. Это новый результат, поскольку в классическом случае в теории линейного регулятора аналогичное утверждение доказывается лишь при отсутствии ограничений на управления. В основе теории линейного регулятора лежат матричные уравнения Риккати, в то время как в основе полученного результата лежит понятие опорной функции (отображения).

СПИСОК ЛИТЕРАТУРЫ

1. Антипин А.С., Хорошилова Е.В. О синтезе обратной связи для задачи терминального управления // Ж. вычисл. матем. и матем. физ. 2018. Т. 58. № 12. С. 1973–1991.
2. Антипин А.С. Терминальное управление краевыми моделями // Ж. вычисл. матем. и матем. физ. 2014. Т. 54. № 2. С. 257–285.
3. Antipin Anatoly S, Khoroshilova Elena V. Linear Programming and Dynamics // Ural Mathematical Journal. 2015. V. 1. № 1. P. 3–18.
4. Antipin Anatoly, Khoroshilova Elena. Saddle point approach to solving problem of optimal control with fixed ends // J. of Global Optimization. 2016. P. 3–17.
5. Antipin Anatoly, Khoroshilova Elena. On methods of terminal control with boundary-value problems: Lagrange approach. In book "Optimization and Application in Control and Data Science". Series Title: Springer Optimization and Its Applications. 2016. P. 17–49.
6. Antipin A. Sufficient condition and evidence-based solution // The Proc. of the conference "Constructive Nonsmooth Analysis and Related Topics" (dedicated to the memory of V.F. Demyanov) (CNSA), 2017. P. 1–3.
7. Antipin A.S, Khoroshilova E.V. Lagrangian as a tool for solving linear optimal control problems with state constraints // Оптимальное управление и дифференциальные игры. Материалы Международной конференции, посвященной 110-летию со дня рождения Льва Семеновича Понтрягина. 2018. С. 23–26.

8. Antipin A.S., Khoroshilova E.V. Controlled dynamic model with boundary-value problem of minimizing a sensitivity function // *Optim. Lett.* 2019. V. 13. № 3. P. 451–473.
9. Антипин А.С., Хорошилова Е.В. Динамика, фазовые ограничения и линейное программирование // *Ж. вычисл. матем. и матем. физ.* 2020. Т. 60. № 2. С. 177–196.
10. Antipin A.S., Khoroshilova E.V. Optimal Control of Two Linear Programming Problems // XII International Conference Optimization and Applications (XII OPTIMA–2021). In: Olenev N.N., Evtushenko Y.G., Jacimovic M., Khachay M., Malkova V. (eds) *Optimization and Applications*. LNCS, 2021. Springer, V. 1378. P. 151–164.
11. Antipin A.S., Khoroshilova E.V. A proven method for optimal control problems with linear dynamics and phase constraints // In: *Dynamical systems: stability, control, optimization: Proc. of the International scientific conference in memory of Professor R.F. Gabasov, Minsk, October 5–10. 2021*. P. 56–58.
12. Antipin A.S., Khoroshilova E.V. Terminal Control of Multi-Agent System // *Lecture Notes in Computer Science*, 2022. V. 1378. P. 5–16.
13. Васильев Ф.П. Методы оптимизации: Кн. 1, 2. М.: МЦНМО, 2011.
14. Кротов В.Ф., Гурман В.И. Методы и задачи оптимального управления. М: Наука. Гл. ред. физ.-мат. лит-ры, 1973.
15. Ащепков Л.Т., Величенко В.В. Оптимальное управление. Курс лекций. Владивосток: Изд-во Дальневосточного университета, 1989.
16. Евтушенко Ю.Г., Третьяков А.А. Новое доказательство теорем Куна–Таккера и Фаркаша // *Ж. вычисл. матем. и матем. физ.* Т. 58. № 7 (2018). С. 1084–1088.
17. Antipin A.S., Jacimovic M. and Mijajlovic N. Extragradient method for solving quasivariational inequalities // *Optimization*. 2017.
18. Антипин А.С., Ячимович В., Ячимович М. Динамика и вариационные неравенства // *Ж. вычисл. матем. и матем. физ.* 2017. Т. 57. № 5. С. 783–800.
19. Сейдж Э.П., Уайт Ч.С. Оптимальное управление системами. Пер. с англ. М: Радио и связь, 1982. 392 с.
20. Калман Р., Фалб П., Арбиб М. Очерки по математической теории систем. М.: URSS, 2004.

SYNTHESIS OF A REGULATOR FOR A LINEAR-QUADRATIC OPTIMAL CONTROL PROBLEM

A. S. Antipin^{a,*}, E. V. Khoroshilova^{b,**}

^a 119333 Moscow, Vavilova str, 44, Dorodnitsyn Computing Center, Federal Research Center Computer Science and Control, RAS, Russia

^b 119991 Moscow, Leninskie Gory, Lomonosov Moscow State University,

*e-mail: asantip@yandex.ru

**e-mail: khorelena@gmail.com

Received: 19.02.2024

Revised: 19.02.2024

Accepted: 31.05.2024

Abstract. In Hilbert space, we consider a linear-quadratic optimal control problem with a fixed left end and a movable right end, for a fixed period of time. The target functional is the sum of the integral and terminal components of the quadratic form. Each of the components searches for its minimum on its permissible set independently of each other. At the right end of the time interval, we have a linear programming problem. The solution to this problem implicitly defines a terminal condition for controlled dynamics. A saddle approach is proposed to solve the problem, which boils down to calculating the saddle point of the Lagrange function. The approach is based on saddle inequalities in both groups of variables: direct and dual. These inequalities represent sufficient conditions for optimality. A method for calculating the saddle point of the Lagrange function is formulated. Convergence in direct and dual variables is proved, namely: weak convergence in controls, strong convergence in phase and conjugate trajectories, as well as in terminal variables of the boundary value problem. On the basis of the saddle approach, control synthesis is built, i.e. feedback in the presence of control constraints in the form of a convex closed set. This is a new result, since in the classical case, in the theory of a linear regulator, a similar statement is proved in the absence of control constraints, which makes it possible to use the Riccati matrix equation. If there are restrictions on management, these arguments no longer pass. Therefore, the basis of the obtained result is the concept of a reference plane to a set of controls.

Keywords: terminal control, boundary value problem, Lagrange function, saddle methods, control synthesis, convergence.

Посвящается 85-летию Юрия Гавриловича Евтушенко

СИНТЕЗ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ В МОДЕЛИ РАМСЕЕВСКОГО ТИПА¹⁾

© 2024 г. Н. В. Трусов^{1,3,*}, А. А. Шананин^{1,2,3,**}

¹119333 Москва, ул. Вавилова, 44, ФИЦ ИУ РАН, Россия

²141701 Долгопрудный, М.о., Институтский пер., 9, МФТИ, Россия

³119991 Москва, Ленинские горы, 1, Московский центр фундаментальной и прикладной математики, Россия

*e-mail: trunick.10.96@gmail.com

**e-mail: alexshan@yandex.ru

Поступила в редакцию 06.03.2024 г.

Переработанный вариант 06.03.2024 г.

Принята к публикации 31.05.2024 г.

В работе исследуется математическое описание экономического поведения репрезентативного домашнего хозяйства на несовершенном рынке потребительских кредитов и депозитов. Модель формализована в виде задачи оптимального управления на конечном временном горизонте. Получена классификация поведения социальных слоев в зависимости от параметров экономической конъюнктуры. Построен синтез оптимального управления на бесконечном временном горизонте. Библ. 24. Фиг. 23.

Ключевые слова: модель рамсеевского типа, несовершенный рынок потребительского кредита, синтез оптимального управления, принцип максимума, математическое моделирование.

DOI: 10.31857/S0044466924090052, EDN: WKXWXN

1. ВВЕДЕНИЕ

В работе исследуется математическое описание экономического поведения рационального домашнего хозяйства на рынке потребительских кредитов и депозитов. Моделирование экономического поведения домашних хозяйств опирается на концепцию рационального репрезентативного экономического агента и восходит к работе Ф. Рамсея [1]. Модель формализована в виде задачи оптимального управления на конечном временном горизонте. Домашнее хозяйство максимизирует дисконтированное потребление с постоянным отвращением к риску, управляя динамикой своих расходов в зависимости от текущих параметров экономической конъюнктуры и поведенческих характеристик самого домашнего хозяйства. Модели рамсеевского типа в форме задач оптимального управления исследовались, например, в [2], [3]. В настоящей работе рассматривается несовершенный рынок, когда процентная ставка по кредитам отличается от процентной ставки по депозитам. Различие процентных ставок по кредитам и депозитам приводит к негладкости правой части дифференциального уравнения на фазовую переменную. Применяя принцип максимума Понтрягина в форме Ф. Кларка [4] к задаче оптимального управления с негладкой правой частью дифференциального уравнения на фазовую переменную, мы получаем область в пространстве фазовой и сопряженной переменных, в которой при описании динамики возникает дополнительное управление, что вносит существенное изменение в анализ задачи. При рассмотрении задачи оптимального управления на бесконечном временном горизонте удастся построить синтез. Синтез позволяет определить оптимальное управление в зависимости от текущего значения фазовой переменной и параметров экономической конъюнктуры. Он зависит от текущих процентных ставок и от поведенческих характеристик репрезентативного домашнего хозяйства. Данная работа является продолжением статьи [5].

Исследование экономического поведения домашнего хозяйства на несовершенном рынке кредитов и депозитов мотивировано обстоятельствами российской экономики. Так, на протяжении более десяти лет отношение процентных ставок по потребительскому кредиту к депозитам варьировалось в диапазоне от 2.5 до 3.5. Проблема задолженности населения по потребительскому кредиту является актуальной. К концу 2023 г. задолженность населения по потребительскому кредиту превысила 15.3 трлн.руб., что составляет около 10% ВВП.

¹⁾ Работа выполнена при финансовой поддержке РНФ (грант 24-11-00329).

Исследования [6], [7] показывают, что около половины заемщиков имеют высокую долговую нагрузку, а у 20% возникают просрочки по обслуживанию имеющихся кредитов. В [8] отмечено, что просроченная задолженность по кредитам выше в слаборазвитых регионах с низкими доходами. Большинство заемщиков имеют низкие реальные доходы и вынуждены занимать новый потребительский кредит для обслуживания ранее выданных. Такая ситуация приводит к образованию финансовой пирамиды. В 2019 г. в правительстве РФ обсуждалась проблема потребительского кредита как актива коммерческих банков [9], [10]. В аналитической записке ЦБ РФ [11] проводится анализ статистических данных “Всероссийское обследование домохозяйств по потребительским финансам” [12] за 2022 г. Результаты анализа говорят о большом разрыве в доходах и расходах домашних хозяйств. Малообеспеченные домашние хозяйства, как правило, не располагают достаточными средствами для покрытия расходов, поэтому вынуждены осуществлять займы по потребительскому кредиту. Среди них большинство тратит значительную часть своего дохода на покрытие выплат по кредитам. Авторы [11] подчеркивают, что неустойчивое финансовое положение домохозяйств из низкодходных групп и семей с детьми в большей степени должно рассматриваться как социальная проблема.

Работа устроена следующим образом. В разд. 2 представлено математическое описание экономического поведения рационального домашнего хозяйства. В разд. 3 изложено построение синтеза оптимального управления на бесконечном временном горизонте. В заключении (разд. 4) сформулированы основные результаты работы.

2. МОДЕЛЬ ЭКОНОМИЧЕСКОГО ПОВЕДЕНИЯ ДОМАШНЕГО ХОЗЯЙСТВА

Мы полагаем, что доходы S_0 репрезентативного домашнего хозяйства растут с темпом $\gamma \in \mathbb{R}$, т.е. $S_0(t) = Se^{\gamma t}$, $S > 0$. Динамика ликвидных средств $M_0(t)$ домашнего хозяйства описывается дифференциальным уравнением

$$\frac{dM_0(t)}{dt} = S_0(t) - p(t)C(t) + H_L(t) - H_D(t), \quad (1)$$

где $C(t)$ — осуществляемые расходы домашним хозяйством, $p(t)$ — индекс потребительских цен, растущий с темпом j . Функции $H_L(t)$ и $H_D(t)$ характеризуют взаимодействие домашнего хозяйства с рынками потребительского кредитования и сбережений в форме депозитов соответственно. Если функция $H_L(t) > 0$, то домашнее хозяйство осуществляет займ по потребительскому кредиту у коммерческого банка. Если $H_L(t) < 0$, то домашнее хозяйство осуществляет платеж по кредитной задолженности. Аналогично, если $H_D(t) > 0$, то домашнее хозяйство увеличивает депозитарный счет. Если $H_D(t) < 0$, то происходит снятие ликвидных средств с депозитарного счета. Пусть $L(t) \geq 0$ — счет кредитной задолженности, $D(t) \geq 0$ — сбережения в форме депозитов. Изменение кредитной задолженности описывается уравнением

$$\frac{dL(t)}{dt} = H_L(t) + r_L L(t), \quad (2)$$

где r_L — процентная ставка по потребительскому кредиту. Изменение сбережений в форме депозитов описывается уравнением

$$\frac{dD(t)}{dt} = H_D(t) + r_D D(t), \quad (3)$$

где r_D — процентная ставка по депозитам.

Отсутствие арбитража на рынке потребительского кредита предполагает, что $r_L > r_D > 0$. Запас ликвидных средств $M_0(t)$, необходимый домашнему хозяйству для осуществления потребительских расходов $p(t)C(t)$, моделируется законом Фишера $M_0(t) = \theta p(t)C(t)$, где $\frac{1}{\theta}$ — скорость обращения денег, $\theta > 0$. Такой подход использовался в [2], [13]. Для дальнейшего изложения удобно ввести величину, характеризующую финансовое состояние домашнего хозяйства $X(t) = M_0(t) + D(t) - L(t)$. Из рационального поведения домашнего хозяйства следует, что оно не осуществляет займы по потребительскому кредиту и не сберегает в форме депозитов одновременно, поэтому $L(t) = (M_0(t) - X(t))_+$, $D(t) = (X(t) - M_0(t))_+$, где $(a)_+ = \max\{a, 0\} \forall a \in \mathbb{R}$. В силу уравнений (1)–(3), финансовое состояние домашнего хозяйства описывается дифференциальным уравнением

$$\frac{dX}{dt} = S_0 - \frac{1}{\theta} M_0 - r_L (M_0 - X)_+ + r_D (X - M_0)_+$$

с начальным условием

$$X(0) = X_0.$$

Домашнее хозяйство стремится максимизировать дисконтированное потребление с коэффициентом дисконтирования $\delta_0 > 0$ и постоянным коэффициентом отвращения к риску $\rho > 0$ на временном интервале $[0, T]$,

управляя своими расходами, т.е.

$$\int_0^T \frac{(C(t))^{1-\rho}}{1-\rho} e^{-\delta_0 t} dt \rightarrow \max_{C \geq 0}.$$

В данной работе мы полагаем, что показатель отвращения к риску ρ положителен и не ограничен сверху. Мотивировкой данного предположения являются материалы исследований, опубликованные в работе [14], в которой были идентифицированы социальные слои населения России с постоянным отвращением к риску из диапазона (4, 6).

Будем говорить, что финансовое состояние $X(t)$ является ликвидным, если существует управление, обеспечивающее выполнение условия $X(T) \geq 0$. Иными словами, к конечному моменту времени домашнее хозяйство должно расплатиться со своими кредитами.

Пусть $x(t) = X(t)e^{-\gamma t}$, $M(t) = M_0(t)e^{-\gamma t}$, $\delta = \delta_0 + (1 - \rho)\gamma$. Также, предположим, что $r_L > r_D > (\gamma)_+$, $\delta > ((1 - \rho)r_L)_+$. Учитывая связь потребительских расходов с ликвидными средствами, задача оптимального управления на конечном временном горизонте формулируется в следующем виде:

$$\int_0^T \frac{M^{1-\rho}}{1-\rho} e^{-(\delta-(1-\rho)\gamma)t} dt \rightarrow \max_M, \quad (4)$$

$$\frac{dx}{dt} = S - \gamma x - \frac{1}{\theta} M - r_L (M - x)_+ + r_D (x - M)_+, \quad (5)$$

$$x(0) = x_0, x(T) \geq 0, \quad (6)$$

$$M(t) \geq 0. \quad (7)$$

Замечание 1. Чтобы обеспечить условие $x(T) \geq 0$, необходимо выполнение неравенства $x > -\frac{S}{r_L - \gamma}$. При $T \rightarrow +\infty$ данное условие является не только необходимым, но и достаточным. В случае, если финансовое состояние $x \leq -\frac{S}{r_L - \gamma}$, домашнее хозяйство не имеет возможности расплатиться с кредитами при сохранении текущей экономической конъюнктуры и образуется финансовая пирамида.

Замечание 2. Предельный случай $\rho = 1$ основан на том, что предел $\lim_{\rho \rightarrow 1} \frac{M^{1-\rho}-1}{1-\rho} = \ln M$. Этому пределу соответствует функционал $\int_0^T \ln M e^{-\delta t} dt \rightarrow \max_M$.

Теорема 1. Пусть $S + (r_L - \gamma)x_0 > 0$, $T > \left(\frac{1}{r_L} \ln \left(\frac{S}{S + (r_L - \gamma)x_0}\right)\right)_+$. Если задача оптимального управления имеет вогнутый функционал, то задача оптимального управления (4)–(7) имеет решение.

Доказательство теоремы 1 для случая $\rho \in (0, 1)$ можно найти в [5]. Более общий случай $\rho > 0$ не нарушает структуру доказательства. В силу выполнения условия $x(T) \geq 0$, множество допустимых управлений ограничено в пространстве L_1 . По теореме Комлоша (см. [15]) из ограниченной в норме пространства L_1 последовательности функций $M_i(t)$, $i = 1, 2, \dots$, можно выделить подпоследовательность $M_{i_n}(t)$, $n = 1, 2, \dots$, средние по Чезаро которой сходятся почти всюду на отрезке $[0, T]$ к функции $M_{\text{opt}}(t)$. Так как исходная функция $\frac{M^{1-\rho}}{1-\rho}$ является вогнутой при $\rho > 0$ (при $\rho = 1$ подынтегральная дробно-степенная функция переходит в логарифмическую, которая также является вогнутой) на множестве $M \geq 0$, то предельная функция $M_{\text{opt}}(t)$ будет являться максимумом функционала (4).

3. ПОСТРОЕНИЕ СИНТЕЗА ЗАДАЧИ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

Негладкий анализ в теории оптимального управления исследовался в работах [4], [16], [17]. В силу негладкости правой части дифференциального уравнения (5) применяется принцип максимума Понтрягина в форме Ф. Кларка [4]. С его помощью выделяются три режима экономического поведения домашнего хозяйства: заимствования, не взаимодействия с банковской системой и сбережения в форме депозитов в зависимости от финансового состояния домашнего хозяйства и параметров экономической конъюнктуры. Результаты сформулированы в теореме 2.

Теорема 2. Если $\{x(t), M(t) | t \in [0, T]\}$ — решение задачи оптимального управления (4)–(7), то существует абсолютно непрерывная функция $\{\varphi(t) | t \in [0, T]\}$.

1. Если

$$x^\rho(t)\varphi(t) < \frac{\theta}{1 + \theta r_L},$$

то оптимальное управление

$$M(t) = \left[\frac{\theta}{(1 + \theta r_L) \varphi(t)} \right]^{\frac{1}{\rho}}, \quad (8)$$

и

$$\frac{dx}{dt} = S + (r_L - \gamma)x - \frac{1 + \theta r_L}{\theta} \left(\frac{\theta}{(1 + \theta r_L) \varphi} \right)^{\frac{1}{\rho}}, \quad (9)$$

$$\frac{d\varphi}{dt} = (\delta + \rho\gamma - r_L) \varphi. \quad (10)$$

2. Если

$$\frac{\theta}{1 + \theta r_L} < x^\rho(t) \varphi(t) < \frac{\theta}{1 + \theta r_D},$$

то оптимальное управление

$$M(t) = x(t), \quad (11)$$

и

$$\frac{dx}{dt} = S - \frac{1 + \theta\gamma}{\theta} x, \quad (12)$$

$$\frac{d\varphi}{dt} = (\delta + \rho\gamma - u(t)) \varphi, \quad r_D \leq u(t) \leq r_L. \quad (13)$$

3. Если

$$x^\rho(t) \varphi(t) > \frac{\theta}{1 + \theta r_D},$$

то оптимальное управление

$$M(t) = \left[\frac{\theta}{(1 + \theta r_D) \varphi(t)} \right]^{\frac{1}{\rho}}, \quad (14)$$

и

$$\frac{dx}{dt} = S + (r_D - \gamma)x - \frac{1 + \theta r_D}{\theta} \left(\frac{\theta}{(1 + \theta r_D) \varphi} \right)^{\frac{1}{\rho}}, \quad (15)$$

$$\frac{d\varphi}{dt} = (\delta + \rho\gamma - r_D) \varphi. \quad (16)$$

Кроме того, выполняется условие трансверсальности

$$x(T) = 0.$$

Доказательство теоремы 2 можно найти в [5].

Структуру оптимального управления удобно исследовать на плоскости $Ox\varphi$. Режимы заимствования, автономный и сбережения разделены на плоскости $Ox\varphi$ двумя гиперболами: $\varphi_{r_L}(x) = \frac{\theta}{1 + \theta r_L} x^{-\rho}$ и $\varphi_{r_D}(x) = \frac{\theta}{1 + \theta r_D} x^{-\rho}$. Легко видеть, что $\varphi_{r_D}(x) > \varphi_{r_L}(x)$, $\lim_{x \rightarrow 0+0} \varphi_{r_L}(x) = +\infty$, $\lim_{x \rightarrow 0+0} \varphi_{r_D}(x) = +\infty$, $\lim_{x \rightarrow +\infty} \varphi_{r_L}(x) = 0$, $\lim_{x \rightarrow +\infty} \varphi_{r_D}(x) = 0$. Область заимствования определяется множеством

$$\Omega_B : \left\{ (x, \varphi) : -\frac{S}{r_L - \gamma} < x \leq 0 \cup x > 0, \varphi < \varphi_{r_L}(x) \right\}.$$

Область

$$\Omega_A : \{(x, \varphi) : x > 0, \varphi_{r_L}(x) < \varphi < \varphi_{r_D}(x)\}$$

соответствует автономному поведению домашнего хозяйства. Область сбережений определяется множеством

$$\Omega_S : \{(x, \varphi) : x > 0, \varphi > \varphi_{r_D}(x)\}.$$

Утверждение 1. На плоскости $Ox\varphi$ имеется траектория $\varphi_z(x)$, для которой выполнено условие $\frac{dx}{dt} = 0$. Она имеет вид

$$\varphi_z(x) = \begin{cases} \left(\frac{\theta}{1+\theta r_L}\right)^{1-\rho} (S + (r_L - \gamma)x)^{-\rho}, & \text{если } x \leq \frac{S\theta}{1+\gamma\theta}, 1+\gamma\theta > 0, \\ \left(\frac{\theta}{1+\theta r_D}\right)^{1-\rho} (S + (r_D - \gamma)x)^{-\rho}, & \text{если } x > \frac{S\theta}{1+\gamma\theta}, 1+\gamma\theta > 0, \\ \left(\frac{\theta}{1+\theta r_L}\right)^{1-\rho} (S + (r_L - \gamma)x)^{-\rho}, & \text{если } 1+\gamma\theta < 0. \end{cases} \quad (17)$$

Доказательство. Рассмотрим случай $1 + \gamma\theta > 0$. Приравняем правую часть дифференциального уравнения (9) к нулю и выразим сопряженную переменную φ . Получим, что в области заимствования, траектория, удовлетворяющая состоянию $\frac{dx}{dt} = 0$ определяется функцией

$$\left(\frac{\theta}{1+\theta r_L}\right)^{1-\rho} (S + (r_L - \gamma)x)^{-\rho}. \quad (18)$$

В автономном режиме условие $\frac{dx}{dt} = 0$ реализуется в точке $x = \frac{S\theta}{1+\gamma\theta}$ (см. (12)). В области сбережения условие $\frac{dx}{dt} = 0$ определяется функцией

$$\left(\frac{\theta}{1+\theta r_D}\right)^{1-\rho} (S + (r_D - \gamma)x)^{-\rho}. \quad (19)$$

Легко убедиться, что точка $x = \frac{S\theta}{1+\gamma\theta}$ является точкой пересечения функции (18) с гиперболой $\varphi_{r_L}(x)$, а также функции (19) с гиперболой $\varphi_{r_D}(x)$. Таким образом, мы получаем выражение (17).

Рассмотрим случай $1 + \gamma\theta < 0$. Покажем, что функция (18) лежит ниже гиперболы $\varphi_{r_L}(x)$, т.е. $\varphi(x) < \varphi_{r_L}(x)$ для любого $x > 0$:

$$\left(\frac{\theta}{1+\theta r_L}\right)^{1-\rho} (S + (r_L - \gamma)x)^{-\rho} < \frac{\theta}{1+\theta r_L} x^{-\rho}.$$

Разделим обе части на положительную величину $\frac{\theta}{1+\theta r_L}$ и возведем в отрицательную степень $-\frac{1}{\rho}$. Получим

$$\frac{\theta}{1+\theta r_L} (S + (r_L - \gamma)x) > x.$$

Перенесем x в левую часть. После преобразований получим

$$S\theta - (1 + \gamma\theta)x > 0.$$

Так как $1 + \gamma\theta < 0$, $x > 0$, то левая часть больше правой. Значит, $\varphi_z(x) < \varphi_{r_L}(x)$ при любом $x > 0$. Утверждение доказано.

Отметим, что траектория $\varphi_z(x)$ также имеет вид гиперболы. В случае $1 + \gamma\theta > 0$ она терпит разрыв I рода и пересекает гиперболы $\varphi_{r_L}(x)$ и $\varphi_{r_D}(x)$ в точке $x = \frac{S\theta}{1+\gamma\theta}$. Гипербола $\varphi_z(x)$ обладает свойствами

$$\lim_{x \rightarrow -\frac{S}{r_L - \gamma} + 0} \varphi_z(x) = +\infty, \quad \lim_{x \rightarrow +\infty} \varphi_z(x) = 0.$$

В классической модели Рамсея рассматривается совершенный рынок потребительского кредита, когда процентные ставки по кредитам и депозитам совпадают ($r_L = r_D$). В этом случае не возникает дополнительного управления $u(t) \in [r_D, r_L]$ в правой части дифференциального уравнения (13), задача оптимального управления становится гладкой, для решения которой могут быть применены стандартные методы решения [18], [19]. Мы изучаем случай несовершенного рынка, когда ставка по потребительскому кредиту превышает ставку по депозитам. Согласно принципу максимума Понтрягина в форме Ф. Кларка [4], из-за негладкости дифференциального уравнения (5), возникает дополнительное управление в дифференциальном уравнении на сопряженную переменную (13). Выбор дополнительного управления $u(t)$ в правой части дифференциального уравнения (13), который максимизирует функционал исходной задачи (4) формулирует.

Утверждение 2 (принцип Беллмана). Максимальное значение функционала (4) достигается при $u(t) \equiv r_D$ в дифференциальном уравнении (13) в области автономного режима.

Доказательство. На плоскости (x, φ) движение траектории $\varphi(x)$ описывается дифференциальными уравнениями (9), (10), (12), (13), (15), (16). Зафиксируем начальное значение точки $x(0) = x_0$. Будем двигаться в обратном времени до пересечения оптимальной траектории $\varphi(x)$ с одной из гипербол $\varphi_{r_L}(x)$, $\varphi_{r_D}(x)$.

Без ограничения общности, будем считать, что оптимальная траектория $\varphi(x)$ пересечет гиперболу $\varphi_{r_L}(x)$ из режима заимствования (случай, когда оптимальная траектория $\varphi(x)$ пересекает гиперболу $\varphi_{r_D}(x)$ из области сбережения рассматривается аналогично). Обозначим точку пересечения оптимальной траектории $\varphi(x)$ с нижней гиперболой $\varphi_{r_L}(x)$ точкой A . Дальнейшее движение оптимальной траектории $\varphi(x)$ из точки A происходит в автономной области и зависит от выбора функции $u(t)$ в дифференциальном уравнении (13). Отметим, что движение траектории вдоль оси Ox не зависит от сопряженной функции $\varphi(t)$, таким образом, не зависит от выбора функции $u(t)$ в автономном режиме (12). Рассмотрим значения функций $\varphi(t)$, удовлетворяющие дифференциальному уравнению (13), при двух управлениях $u(t)$. Обозначим их через $u_1(t)$ и $u_2(t)$. Положим $u_1(t) \equiv r_D$, $u_2(t) = r_D + \hat{u}(t)$, где $\hat{u}(t) \in [0, r_L - r_D]$. Пусть $\varphi_1(t)$ соответствует значению $u_1(t)$, а $\varphi_2(t)$ соответствует значению $u_2(t)$. Выпишем решение дифференциального уравнения (13) для функции $\varphi_2(t)$ в обратном времени:

$$\varphi_2(t) = \varphi_0 \exp \left[\int_{t_0}^t (\delta + \rho\gamma - r_D - \hat{u}(\tau)) d\tau \right],$$

где t_0 соответствует моменту времени, когда оптимальная траектория $\varphi(x)$ находится в точке A . Заметим, что

$$\varphi_2(t) = \varphi_1(t) \cdot \exp \left[- \int_{t_0}^t \hat{u}(\tau) d\tau \right].$$

Множитель $\exp \left[- \int_{t_0}^t \hat{u}(\tau) d\tau \right] \geq 1$, так как $t < t_0$, $\hat{u}(\tau) \geq 0$. Таким образом, $\varphi_2(t) \geq \varphi_1(t)$. Причем равенство достигается тогда и только тогда, когда $u_1(t) \equiv u_2(t)$, т.е. $\hat{u}(\tau) \equiv 0$. Рассмотрим $\hat{u}(\tau) \neq 0$. Тогда существует момент времени $\hat{t}$ такой, что для любого $t < \hat{t}$ функция $\varphi_2(t) > \varphi_1(t)$. Обозначим траекторию $\varphi_1(x)$, которая удовлетворяет дифференциальным уравнениям (12), (13) при $u(t) \equiv u_1(t)$, траекторию $\varphi_2(x)$, которая удовлетворяет дифференциальным уравнениям (12), (13) при $u(t) \equiv u_2(t)$. Согласно принципу максимума Понтрягина в форме Ф. Кларка, обе траектории $\varphi_1(x)$ и $\varphi_2(x)$ являются оптимальными. Пусть траектория $\varphi_1(x)$ пересекает гиперболу $\varphi_{r_D}(x)$ в точке C , а траектория $\varphi_2(x)$ пересекает гиперболу $\varphi_{r_D}(x)$ в точке B . Так как $\varphi_2(t) > \varphi_1(t)$, то точка C лежит правее точки B на оси Ox . Пусть точка B имеет координаты $(x_B, \varphi_{r_D}^B)$, а точка C имеет координаты $(x_C, \varphi_{r_D}^C)$. Зафиксируем $\hat{x} \in (x_B, x_C)$. Траектория $\varphi_1(x)$ при $x = \hat{x}$ находится в автономной области, в то время как траектория $\varphi_2(x)$ при $x = \hat{x}$ находится в области сбережения. Так как в автономной области $M = x$, а в области сбережения $M < x$, то значение функционала (4) больше при выборе оптимальной траектории $\varphi_1(x)$. В силу произвольного выбора функции $\varphi_2(t)$, удовлетворяющего допустимому множеству дополнительного управления $u(t)$ в автономной области, максимальное значение функционала (4) будет достигаться при $u(t) \equiv r_D$. На фиг.1 отображена область $ABCD$, содержащая всевозможные допустимые управления $u(t)$ в правой части дифференциального уравнения (13). Утверждение доказано.

Утверждение 3. Пусть $r_L > \delta - \frac{\rho}{\theta}$, оптимальная траектория $\varphi(x)$ пересекает гиперболу $\varphi_{r_L}(x)$ в точке $x = \hat{x}$, $\hat{x} \geq \frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}}$. Если $x_{c1} > \hat{x}$, то на отрезке $x \in [\hat{x}, x_{c1}]$ оптимальная траектория $\varphi(x)$ движется по гиперболе $\varphi_{r_L}(x)$.

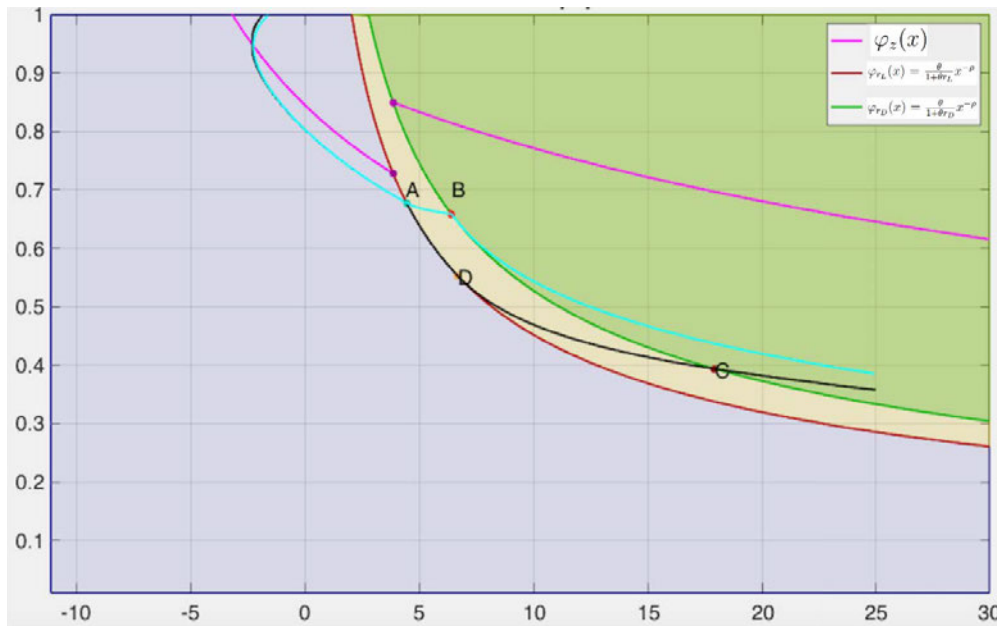
Причем если $\delta - \frac{\rho}{\theta} < r_D$, то $x_{c1} = \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}$. В случае $r_D < \delta - \frac{\rho}{\theta}$, $x_{c1} = +\infty$.

Более того, при движении оптимальной траектории $\varphi(x)$ по гиперболе $\varphi_{r_L}(x)$, дополнительное управление в правой части дифференциального уравнения (13) определяется выражением

$$u(x) = \rho \left(\frac{S}{x} - \frac{1 + \gamma\theta}{\theta} \right) + \delta + \rho\gamma. \quad (20)$$

Доказательство. Без ограничения общности, рассмотрим случай, когда траектория $\varphi(x)$ пересекает гиперболу $\varphi_{r_L}(x)$ при движении в обратном времени из области заимствования, $\delta - \frac{\rho}{\theta} < r_L < \delta + \rho\gamma$, $\varphi(x) < \varphi_z(x)$. Остальные случаи рассматриваются аналогично. Из условий следует, что в областях заимствования и автономного поведения $\frac{dx}{dt} < 0$, $\frac{d\varphi}{dt} > 0$. Движение оптимальной траектории $\varphi(x)$ в области заимствования удовлетворяет дифференциальным уравнениям (9), (10), а в области автономного поведения — дифференциальным уравнениям (12), (13). Изучим движение оптимальной траектории $\varphi(x)$ в обратном времени. Оптимальная траектория $\varphi(x)$ попадает на гиперболу $\varphi_{r_L}(x)$ в точке $A = \left(\hat{x}, \frac{\theta}{1 + \theta r_L} \hat{x}^{-\rho} \right)$. Подставляя координаты точки A в правые части дифференциальных уравнений (9), (10), найдем вектор направления движения оптимальной траектории $\varphi(x)$ из области заимствования на границе $\varphi_{r_L}(x)$:

$$\bar{V}_B = \left\{ S - \frac{1 + \gamma\theta}{\theta} \hat{x}, (\delta + \rho\gamma - r_L) \frac{\theta}{1 + \theta r_L} \hat{x}^{-\rho} \right\}.$$



Фиг. 1. Движение траекторий $\varphi(x)$ в зависимости от выбора функции $u(t)$. Черная траектория соответствует выбору функции $u(t) \equiv r_D$, бирюзовая траектория соответствует выбору функции $u(t) \equiv r_L$. Красным, зеленым и розовым отображены гиперболы $\varphi_{r_L}(x)$, $\varphi_{r_D}(x)$, $\varphi_z(x)$ соответственно.

Вектор направления движения оптимальной траектории $\varphi(x)$ из автономной области на границе $\varphi_{r_L}(x)$:

$$\bar{V}_A = \left\{ S - \frac{1+\gamma\theta}{\theta} \hat{x}, (\delta + \rho\gamma - r_D) \frac{\theta}{1+\theta r_L} \hat{x}^{-\rho} \right\}.$$

Очевидно, что вектор $\bar{V}_A$ лежит выше вектора $\bar{V}_B$. Покажем, что уравнение касательной к гиперболе $\varphi_{r_L}(x)$ в точке $x = \hat{x}$ лежит между векторами $\bar{V}_A$ и $\bar{V}_B$. Уравнение касательной к гиперболе $\varphi_{r_L}(x)$ в точке $x = \hat{x}$ определяется выражением

$$\varphi_{\tan}(x) = \frac{\theta}{1+\theta r_L} \hat{x}^{-\rho} \cdot \left[-\rho \cdot \left(\frac{x}{\hat{x}} - 1 \right) + 1 \right].$$

Для начала покажем, что касательная $\varphi_{\tan}(x)$ в точке $x = \hat{x}$ лежит выше вектора $\bar{V}_B$. Для этого достаточно показать, что разность углов наклона вектора $\bar{V}_B$ и касательной $\varphi_{\tan}(x)$ в точке $x = \hat{x}$ меньше нуля. Обозначим эту разность через $f(\hat{x})$:

$$f(\hat{x}) = \frac{\theta}{1+\theta r_L} \hat{x}^{-\rho-1} \left[\frac{\delta + \rho\gamma - r_L}{\frac{S}{\hat{x}} - \frac{1+\gamma\theta}{\theta}} + \rho \right].$$

Функция $f(\hat{x})$ обращается в ноль при $\hat{x}^* = \frac{S\rho}{r_L - \delta + \frac{\rho}{\theta}}$. При $\hat{x} > \frac{S\rho}{r_L - \delta + \frac{\rho}{\theta}}$ функция $f(\hat{x}) < 0$. Таким образом, мы показали, что касательная $\varphi_{\tan}(x)$ в точке $x = \hat{x}$ лежит выше вектора $\bar{V}_B$.

В случае $\delta - \frac{\rho}{\theta} < r_D$, аналогично можно показать, что при $\frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}} \leq \hat{x} < \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}$ касательная $\varphi_{\tan}(x)$ в точке $x = \hat{x}$ лежит ниже вектора $\bar{V}_A$. С одной стороны, оптимальная траектория $\varphi(x)$ пересекает гиперболу $\varphi_{r_L}(x)$ из области заимствования и попадает в автономную область. С другой стороны, попадая в автономную область, оптимальная траектория $\varphi(x)$ пересекает гиперболу $\varphi_{r_L}(x)$ и попадает обратно в область заимствования. Таким образом, движение оптимальной траектории $\varphi(x)$ будет проходить по гиперболе $\varphi_{r_L}(x)$ на отрезке $x \in \left[\hat{x}, \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}} \right]$. Вычислим дополнительное управление в правой части дифференциального уравнения (13) на гиперболе $\varphi_{r_L}(x)$, удовлетворяющее движению оптимальной траектории. Гипербола $\varphi_{r_L}(x)$ задается уравнением

$$x^\rho \varphi = C, \quad (21)$$

где $C = \frac{\theta}{1+\theta r_L}$ — константа. Продифференцируем уравнение (21) в силу системы (12), (13). Получим

$$\rho x^{\rho-1} \left[S - \frac{1+\gamma\theta}{\theta} x \right] \varphi + x^\rho [\delta + \rho\gamma - u] \varphi = 0.$$

Учитывая, что $x > 0$, $\varphi > 0$, получаем управление на гиперболе $\varphi_{r_L}(x)$, которое задается уравнением (20). Покажем, что для любого $x \in \left[\hat{x}, \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}\right]$ управление (20) не выходит за границы $[r_D, r_L]$. Легко видеть, что функция $u(x)$ монотонно убывает по переменной x . Подставляя левое граничное значение $\hat{x}_L = \frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}}$, получаем, что $u(\hat{x}_L) = r_L$. Подставляя правое граничное значение $\hat{x}_R = \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}$, получаем, что $u(\hat{x}_D) = r_D$.

Рассмотрим случай $r_D < \delta - \frac{\rho}{\theta}$. Покажем, что вектор $\bar{V}_A$ лежит выше касательной $\varphi_{\tan}(x)$ в точке $x = \hat{x}$ для любого $\hat{x} \geq \frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}}$. Пусть

$$g(\hat{x}) = \frac{\theta}{1 + \theta r_L} \hat{x}^{-\rho-1} \left[\frac{\delta + \rho\gamma - r_D}{\frac{S}{\hat{x}} - \frac{1+\gamma\theta}{\theta}} + \rho \right]$$

есть разность углов наклона вектора $\bar{V}_A$ и касательной $\varphi_{\tan}(x)$ в точке $x = \hat{x}$. Функцию $g(\hat{x})$ можно представить в виде произведения двух функций: $g(\hat{x}) = g_1(\hat{x}) \cdot g_2(\hat{x})$:

$$g_1(\hat{x}) = \frac{\theta}{1 + \theta r_L} \hat{x}^{-\rho-1},$$

$$g_2(\hat{x}) = \frac{\delta + \rho\gamma - r_D}{\frac{S}{\hat{x}} - \frac{1+\gamma\theta}{\theta}} + \rho.$$

Отметим, что $g_1(\hat{x}) > 0$ для любого $\hat{x} > 0$, а производная $\frac{dg_2(\hat{x})}{d\hat{x}} > 0$:

$$\frac{dg_2(\hat{x})}{d\hat{x}} = \frac{S\theta^2(\delta + \rho\gamma - r_D)}{(S\theta - (1 + \gamma\theta)^2)\hat{x}^2} > 0.$$

Функция $g_2(\hat{x})$ монотонно возрастает, имеет единственный отрицательный корень $\hat{x}^* = \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}$, значит, при $\hat{x} > 0$ функция $g(\hat{x}) > 0$. В рассматриваемом случае оптимальная траектория не сойдет с гиперболы $\varphi_{r_L}(x)$ при $x > \hat{x}$. Управление (20) остается допустимым: при $x \rightarrow +\infty$, $u(x) \rightarrow \delta - \frac{\rho}{\theta} > r_D$. Утверждение доказано.

Из доказанного утверждения можно сформулировать два следствия, доказательства которых очевидны.

Следствие 1. Если оптимальная траектория $\varphi(x)$ движется по гиперболе $\varphi_{r_L}(x)$ на отрезке $x \in [x_L, x_R]$, то в точках $x = x_L$ и $x = x_R$ оптимальная траектория касается гиперболы $\varphi_{r_L}(x)$.

Следствие 2. Движение оптимальной траектории $\varphi(x)$ по гиперболе $\varphi_{r_L}(x)$ возможно только на отрезке $x \in \left[\frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}}, \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}\right]$. Причем, если $r_D < \delta - \frac{\rho}{\theta}$, то движение оптимальной траектории $\varphi(x)$ по гиперболе $\varphi_{r_L}(x)$ возможно только на множестве $x \in \left[\frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}}, +\infty\right)$.

Утверждение 4. На гиперболе $\varphi_{r_D}(x)$ нет особого режима у оптимальной траектории $\varphi(x)$.

Доказательство. Пусть оптимальная траектория $\varphi(x)$ попадает на гиперболу $\varphi_{r_D}(x)$ в точке $A(\hat{x}, \hat{\varphi})$, причем $\hat{\varphi} = \frac{\theta}{1 + \theta r_D} \hat{x}^{-\rho}$. Подставляя эту точку в правую часть дифференциального уравнения (15), получаем, что в точке A скорость по фазовой координате x равна

$$\left. \frac{dx}{dt} \right|_{x=\hat{x}, \varphi=\hat{\varphi}} = S - \frac{1 + \gamma\theta}{\theta} \hat{x},$$

что соответствует скорости по фазовой координате x в автономной области (12). Согласно утверждению 2, в автономной области $u(t) \equiv r_D$ в правой части дифференциального уравнения (13). Таким образом, скорости по координате φ в автономной области и области сбережения совпадают. Мы получили, что при переходе из автономной области в область сбережения (и обратно) через гиперболу $\varphi_{r_D}(x)$ скорости координат x и φ траектории равны. Таким образом, траектория $\varphi(x)$ пересекает гиперболу $\varphi_{r_D}(x)$ в точке. Утверждение доказано.

Пусть $y(\bar{x}, \beta)$ — решение уравнения

$$\left[\frac{1 + \theta r_L}{1 + \theta r_D} \right]^{\frac{\beta}{\rho}} \cdot \frac{y(\bar{x}, \beta)}{\bar{x}} = \left[\left(y(\bar{x}, \beta) - \frac{S\theta}{1 + \gamma\theta} \right) \left(\bar{x} - \frac{S\theta}{1 + \gamma\theta} \right)^{-1} \right]^{\frac{\theta[\delta + \rho\gamma - r_D]}{(1 + \gamma\theta)\rho}}, \quad (22)$$

и пусть $M_r(\bar{x}, r)$ — решение уравнения

$$x + \frac{S}{r - \gamma} = \frac{(1 + \theta r)\rho}{\theta(\delta - (1 - \rho)r)} M_r(\bar{x}, r) + \left[\frac{\bar{x}}{M_r(\bar{x}, r)} \right]^{\frac{\rho(r - \gamma)}{\delta + \rho\gamma - r}} \cdot \left[\frac{S}{r - \gamma} - \frac{r - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r} \bar{x} \right]. \quad (23)$$

Оптимальное управление M зависит от сопряженной переменной φ в режиме заимствования (8) и в режиме сбережения (10). На конечном временном промежутке $[0, T]$, решая дифференциальные уравнения (8), (9), (11), (12), (14), (15) в соответствии с начальным условием $x(0) = x_0$ и условием трансверсальности $x(T) = 0$, мы получаем оптимальное управление $M(\varphi; T)$, где $\varphi = \varphi(t, x(t); T)$. Устремив $T \rightarrow +\infty$, можно получить предельную траекторию $\varphi(x)$. Подставляя ее в оптимальное управление $M(\varphi)$, мы получим синтез оптимального управления $M(x)$. Синтез позволяет определить оптимальное управление в зависимости от фазовой координаты x и от параметров экономической конъюнктуры.

Переключение режимов (заимствования, автономного и сбережения) происходят в зависимости от соотношения величин $r_L, r_D, \delta - \frac{\rho}{\theta}, \delta + \rho\gamma$. Они определяют тип поведения домашнего хозяйства. Справедлива следующая теорема, описывающая синтез оптимального управления при $T \rightarrow +\infty$.

Теорема 3 (синтез задачи оптимального управления).

1. Если $r_L < \delta - \frac{\rho}{\theta}$, то домашнее хозяйство находится в режиме заимствования.

Тип 1.1. Синтез оптимального управления определяется выражением

$$M(x; S, r_L, \theta, \rho, \gamma, \delta) = \frac{\theta(\delta - (1 - \rho)r_L)}{\rho(1 + \theta r_L)} \left(x + \frac{S}{r_L - \gamma} \right).$$

2. Если $r_D < \delta - \frac{\rho}{\theta} < r_L$, то домашнее хозяйство находится либо в режиме заимствования, либо в автономном режиме.

Тип 2.1. Если $r_D < \delta - \frac{\rho}{\theta} < r_L < \delta + \rho\gamma$, то синтез оптимального управления определяется выражением

$$M(x; S, r_L, \theta, \rho, \gamma, \delta) = \begin{cases} \frac{\theta(\delta - (1 - \rho)r_L)}{\rho(1 + \theta r_L)} \left(x + \frac{S}{r_L - \gamma} \right), & \text{если } x < \frac{S(\delta - (1 - \rho)r_L)}{(r_L - \gamma)(r_L - \delta + \frac{\rho}{\theta})}, \\ x, & \text{если } x \geq \frac{S(\delta - (1 - \rho)r_L)}{(r_L - \gamma)(r_L - \delta + \frac{\rho}{\theta})}. \end{cases}$$

Тип 2.2. Если $r_D < \delta - \frac{\rho}{\theta} < \delta + \rho\gamma < r_L$ или $\max\{r_D, \delta + \rho\gamma\} < \delta - \frac{\rho}{\theta} < r_L$, то синтез оптимального управления определяется выражением

$$M(x; S, r_L, \theta, \rho, \gamma, \delta) = \begin{cases} M_r(x_{p_1}, r_L), & \text{если } x < x_{p_1}, \\ x, & \text{если } x \geq x_{p_1}, \end{cases}$$

где $x_{p_1} = \frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}}$. Более того, уравнение (23) имеет единственное решение $M_r(x_{p_1}, r_L)$ на множестве $((x)_+, x_{p_1})$.

3. Если $\delta - \frac{\rho}{\theta} < r_D$, то домашнее хозяйство находится в одном из трех режимов: заимствования, автономном или сбережения.

Тип 3.1. Если $\delta - \frac{\rho}{\theta} < r_D < r_L < \delta + \rho\gamma$, то синтез оптимального управления определяется выражением

$$M(x; S, r_L, r_D, \theta, \rho, \gamma, \delta) = \begin{cases} \frac{\theta(\delta - (1 - \rho)r_L)}{\rho(1 + \theta r_L)} \left(x + \frac{S}{r_L - \gamma} \right), & \text{если } x < x_{p_1}, \\ x, & \text{если } x_{p_1} \leq x \leq x_{p_2}, \\ M_r(x_{p_2}, r_D), & \text{если } x > x_{p_2}, \end{cases}$$

где $x_{p_1} = \frac{S(\delta - (1 - \rho)r_L)}{(r_L - \gamma)(r_L - \delta + \frac{\rho}{\theta})}$, $x_{c_1} = \max\left\{x_{p_1}, \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}\right\}$, $x_{p_2} = y(x_{c_1}, -1)$. Более того, уравнение (22) имеет единственное решение $y(x_{c_1}, -1)$ на множестве $(x_{c_1}, +\infty)$, а уравнение (23) имеет единственное решение $M_r(x_{p_2}, r_D)$ на множестве (x_{p_2}, x) .

Тип 3.2. Если $\delta - \frac{\rho}{\theta} < r_D < \delta + \rho\gamma < r_L$, то синтез оптимального управления определяется выражением

$$M(x; S, r_L, r_D, \theta, \rho, \gamma, \delta) = \begin{cases} M_r(x_{p_1}, r_L), & \text{если } x < x_{p_1}, \\ x, & \text{если } x_{p_1} \leq x \leq x_{p_2}, \\ M_r(x_{p_2}, r_D), & \text{если } x > x_{p_2}, \end{cases}$$

где $x_{p_1} = \frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}}$, $x_{c_1} = \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}$, $x_{p_2} = y(x_{c_1}, -1)$. Более того, уравнение (22) имеет единственное решение $y(x_{c_1}, -1)$ на множестве $(x_{c_1}, +\infty)$, уравнение (23) имеет единственное решение $M_r(x_{p_1}, r_L)$ на множестве $((x)_+, x_{p_1})$ и единственное решение $M_r(x_{p_2}, r_D)$ на множестве (x_{p_2}, x) .

Тип 3.3. Если $\max\{\delta - \frac{\rho}{\theta}, \delta + \rho\gamma\} < r_D$, то синтез оптимального управления определяется выражением

$$M(x; S, r_L, r_D, \theta, \rho, \gamma, \delta) = \begin{cases} M_r(x_{p_1}, r_L), & \text{если } x < x_{p_1}, \\ x, & \text{если } x_{p_1} \leq x \leq x_{p_2}, \\ \frac{\theta(\delta - (1-\rho)r_D)}{\rho(1+\theta r_D)} \cdot \left(x + \frac{S}{r_D - \gamma}\right), & \text{если } x > x_{p_2}, \end{cases}$$

где $x_{p_2} = \frac{S}{r_D - \gamma} \cdot \frac{\delta - (1-\rho)r_D}{r_D - \delta + \frac{\rho}{\theta}}$, $x_{p_1} = \min\left\{\frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}}, y(x_{p_2}, 1)\right\}$. Более того, уравнение (23) имеет единственное решение $M_r(x_{p_1}, r_L)$ на множестве $((x)_+, x_{p_1})$, а уравнение (22) имеет единственное решение $y(x_{p_2}, 1)$ на множестве $(0, x_{p_2})$.

Прежде чем приступить к доказательству теоремы 3, сформулируем и докажем ряд вспомогательных лемм.

Лемма 1. Если функция $f(x)$ непрерывна на отрезке $[a, b]$, $a, b \in \mathbb{R}$, $f(a) \cdot f(b) < 0$, существует и единственен экстремум функции $f(x)$ на множестве (a, b) , то существует единственная точка $c \in (a, b)$: $f(c) = 0$.

Доказательство. Без ограничения общности, будем считать, что $f(a) < 0$, $f(b) > 0$, точка экстремума является точкой максимума (случай, когда точка экстремума является точкой минимума, а также случай $f(a) > 0$, $f(b) < 0$ доказываются аналогично). Пусть $d = \arg \max_{(a,b)} \{f(x)\}$. Тогда, на множестве $[a, d]$ функция $f(x)$ мо-

нотонно возрастает, а на множестве $(d, b]$ функция $f(x)$ монотонно убывает. Из монотонности следует, что $f(d) > f(b)$. Так как $f(b) > 0$, то на отрезке $[d, b]$ функция $f(x)$ не имеет корней. Функция $f(x)$ монотонна, на концах отрезка $[a, d]$ принимает значения разных знаков, поэтому существует единственная точка $c \in (a, b)$, которая является корнем функции $f(x)$. Лемма доказана.

Лемма 2. Если функция $f(x)$ непрерывна на отрезке $[a, b]$, дифференцируема на интервале (a, b) , $a, b \in \mathbb{R}$, принимает значение $0 < f(a) \leq b - a$, а ее производная $\frac{df(x)}{dx} < -1$ на интервале (a, b) , то существует единственная точка $c \in [a, b]$: $f(c) = 0$.

Доказательство. Пусть $g(x) = -x + b$. Легко видеть, что функция $f(x) \leq g(x)$, причем равенство может достигаться только в точке $x = a$. В точке $x = b$ функция $g(b) = 0$, а так как $f(x) < g(x)$, $x \in (a, b)$, то $f(b) < 0$. Так как производная функции имеет определенный знак на интервале (a, b) , то функция $f(x)$ является монотонной. На концах отрезка $[a, b]$ функция $f(x)$ принимает значения разных знаков, поэтому существует единственная точка $c \in (a, b)$, которая является корнем функции $f(x)$. Лемма доказана.

Лемма 3. Если функция $f(x)$ непрерывна на отрезке $[a, b]$, дифференцируема на интервале (a, b) , $a, b \in \mathbb{R}$, принимает значение $0 > f(b) \geq -b + a$, а ее производная $\frac{df(x)}{dx} < -1$ на интервале (a, b) , то существует единственная точка $c \in [a, b]$: $f(c) = 0$.

Доказательство. Пусть $g(x) = -x + a$. Легко видеть, что функция $f(x) \geq g(x)$, причем равенство может достигаться только в точке $x = b$. В точке $x = a$ функция $g(a) = 0$, а так как $f(x) > g(x)$, $x \in [a, b)$, то $f(a) > 0$. Так как производная функции имеет определенный знак на интервале (a, b) , то функция $f(x)$ является монотонной. На концах отрезка $[a, b]$ функция $f(x)$ принимает значения разных знаков, поэтому существует единственная точка $c \in (a, b)$, которая является корнем функции $f(x)$. Лемма доказана.

Лемма 4. 1. Если $\max\{\delta - \frac{\rho}{\theta}, \delta + \rho\gamma\} < r_D$, то уравнение $y(x_{p_2}, 1)$ имеет единственное решение на множестве $(0, x_{p_2})$, где $x_{p_2} = \frac{S}{r_D - \gamma} \cdot \frac{\delta - (1-\rho)r_D}{r_D - \delta + \frac{\rho}{\theta}}$.

2. Если $\delta - \frac{\rho}{\theta} < r_D < r_L < \delta + \rho\gamma$, то уравнение $y(x_{c_1}, -1)$ имеет единственное решение на множестве $(x_{c_1}, +\infty)$, где $x_{c_1} = \max\left\{x_{p_1}, \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}\right\}$, $x_{p_1} = \frac{S(\delta - (1-\rho)r_L)}{(r_L - \gamma)(r_L - \delta + \frac{\rho}{\theta})}$.

3. Если $\delta - \frac{\rho}{\theta} < r_D < \delta + \rho\gamma < r_L$, то уравнение $y(x_{c_1}, -1)$ имеет единственное решение на множестве $(x_{c_1}, +\infty)$, где $x_{c_1} = \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}$.

Доказательство. Докажем каждый из пунктов леммы.

1. Пусть $\hat{y} = y(x_{p_2}, 1)$. Подставим $\beta = 1$, $\bar{x} = x_{p_2}$ в уравнение (22). Вычтем из левой части уравнения (22) правую и рассмотрим функцию

$$f(\hat{y}) = \left[\frac{1 + \theta r_L}{1 + \theta r_D}\right]^{\frac{1}{\rho}} \cdot \frac{\hat{y}}{x_{p_2}} - \left[\left(\hat{y} - \frac{S\theta}{1 + \gamma\theta}\right) \left(x_{p_2} - \frac{S\theta}{1 + \gamma\theta}\right)^{-1}\right]^{\frac{\theta[\delta + \rho\gamma - r_D]}{(1 + \gamma\theta)\rho}}.$$

По условию, $x_{p_2} = \frac{S}{r_D - \gamma} \cdot \frac{\delta - (1 - \rho)r_D}{r_D - \delta + \frac{\rho}{\theta}}$. Покажем, что $x_{p_2} < \frac{S\theta}{1 + \gamma\theta}$:

$$\frac{S}{r_D - \gamma} \cdot \frac{\delta - (1 - \rho)r_D}{r_D - \delta + \frac{\rho}{\theta}} < \frac{S\theta}{1 + \gamma\theta}.$$

Так как в рассматриваемом случае $r_D > \delta - \frac{\rho}{\theta}$, то обе части неравенства положительны. Умножим обе части неравенства на величину $\frac{1}{S} \cdot (r_D - \gamma) (r_D - \delta + \frac{\rho}{\theta}) (1 + \gamma\theta)$. После преобразований, получим

$$(r_D\theta + 1) [\rho\gamma - r_D + \delta] < 0. \quad (24)$$

Так как $r_D > \max \left\{ \delta - \frac{\rho}{\theta}, \delta + \rho\gamma \right\}$, то выражение в квадратных скобках (24) отрицательно. Значит, неравенство (24) верно, и $x_{p_2} < \frac{S\theta}{1 + \gamma\theta}$. Определим знаки функции $f(\hat{y})$ в точках $\hat{y} = 0$ и $\hat{y} = x_{p_2}$:

$$f(0) = - \left[\frac{S\theta}{1 + \gamma\theta} \left(\frac{S\theta}{1 + \gamma\theta} - x_{p_2} \right)^{-1} \right]^{\frac{\theta[\delta + \rho\gamma - r_D]}{(1 + \gamma\theta)\rho}} < 0,$$

$$f(x_{p_2}) = \left[\frac{1 + \theta r_L}{1 + \theta r_D} \right]^{\frac{1}{\rho}} - 1 > 0.$$

Последнее верно, так как $1 + \theta r_L > 1 + \theta r_D$.

Отметим, что функция $f(\hat{y})$ имеет единственный экстремум. Если этот экстремум находится вне множества $(0, x_{p_2})$, то функция $f(\hat{y})$ монотонна на отрезке $[0, x_{p_2}]$, $f(0) \cdot f(x_{p_2}) < 0$, значит, существует единственное значение $\hat{y}^* \in (0, x_{p_2}) : f(\hat{y}^*) = 0$. Это значение и будет является корнем уравнения $y(x_{p_2}, 1)$. Если экстремум функции $f(\hat{y})$ находится внутри множества $(0, x_{p_2})$, то, согласно лемме 1, функция $f(\hat{y})$ также имеет единственный корень на множестве $(0, x_{p_2})$.

2. Перейдем к доказательству второго пункта леммы. Пусть $\check{y} = y(x_{c_1}, -1)$. Подставим $\beta = -1$, $\bar{x} = x_{c_1}$ в уравнение (22). Вычтем из левой части уравнения (22) правую и рассмотрим функцию

$$g(\check{y}) = \left[\frac{1 + \theta r_D}{1 + \theta r_L} \right]^{\frac{1}{\rho}} \cdot \frac{\check{y}}{x_{c_1}} - \left[\left(\check{y} - \frac{S\theta}{1 + \gamma\theta} \right) \left(x_{c_1} - \frac{S\theta}{1 + \gamma\theta} \right)^{-1} \right]^{\frac{\theta[\delta + \rho\gamma - r_D]}{(1 + \gamma\theta)\rho}}.$$

По условию, $x_{c_1} = \max \left\{ \frac{S(\delta - (1 - \rho)r_L)}{(r_L - \gamma)(r_L - \delta + \frac{\rho}{\theta})}, \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}} \right\}$. Сравним значение x_{c_1} со значением $\frac{S\theta}{1 + \gamma\theta}$. Пусть $x_{c_1} = \frac{S(\delta - (1 - \rho)r_L)}{(r_L - \gamma)(r_L - \delta + \frac{\rho}{\theta})}$. Покажем, что в данном случае $x_{c_1} > \frac{S\theta}{1 + \gamma\theta}$:

$$\frac{S(\delta - (1 - \rho)r_L)}{(r_L - \gamma)(r_L - \delta + \frac{\rho}{\theta})} > \frac{S\theta}{1 + \gamma\theta}.$$

Так как в рассматриваемом случае $r_L > \delta - \frac{\rho}{\theta}$, то обе части неравенства положительны. Умножим обе части неравенства на величину $\frac{1}{S} \cdot (r_L - \gamma) (r_L - \delta + \frac{\rho}{\theta}) (1 + \gamma\theta)$. После преобразований, получим

$$(r_L\theta + 1) [\rho\gamma - r_L + \delta] > 0. \quad (25)$$

Так как $r_L > \delta - \frac{\rho}{\theta}$, то выражение в квадратных скобках (25) положительно. Значит, $x_{c_1} > \frac{S\theta}{1 + \gamma\theta}$. Если $x_{c_1} = \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}$, то $x_{c_1} \geq \frac{S(\delta - (1 - \rho)r_L)}{(r_L - \gamma)(r_L - \delta + \frac{\rho}{\theta})} > \frac{S\theta}{1 + \gamma\theta}$. Таким образом, у функции $g(\check{y})$ коэффициент при множителе $\check{y}^{\frac{\theta[\delta + \rho\gamma - r_D]}{(1 + \gamma\theta)\rho}}$ отрицателен, а при множителе $\check{y}$ положителен. Сравним положительные степени переменной $\check{y}$ у функции $g(\check{y})$. Покажем, что

$$1 > \frac{\theta[\delta + \rho\gamma - r_D]}{(1 + \gamma\theta)\rho}.$$

Перенесем левую часть в правую. После преобразований, получим

$$0 > \delta - \frac{\rho}{\theta} - r_D.$$

По условию $r_D > \delta - \frac{\rho}{\theta}$, значит, неравенство $1 > \frac{\theta[\delta + \rho\gamma - r_D]}{(1 + \gamma\theta)\rho}$ верно. Отсюда следует, что $g(\tilde{y}) \rightarrow +\infty$ при $\tilde{y} \rightarrow +\infty$. В точке $\tilde{y} = x_{c_1}$ функция

$$g(x_{c_1}) = \left[\frac{1 + \theta r_D}{1 + \theta r_L} \right]^{\frac{1}{\rho}} - 1 < 0.$$

Отметим, что функция $g(\tilde{y})$ имеет единственный экстремум. Если этот экстремум находится вне множества $(x_{c_1}, +\infty)$, то функция $g(\tilde{y})$ монотонна на $[x_{c_1}, +\infty)$, $g(x_{c_1}) \cdot f(\tilde{y}) < 0$ при $\tilde{y} \rightarrow +\infty$, значит, существует единственное значение $\tilde{y}^* \in (x_{c_1}, +\infty) : g(\tilde{y}^*) = 0$. Это значение и будет является корнем уравнения $y(x_{c_1}, -1)$. Если экстремум функции $g(\tilde{y})$ находится внутри множества $(x_{c_1}, +\infty)$, то, согласно лемме 1, функция $g(\tilde{y})$ также имеет единственный корень на множестве $(x_{c_1}, +\infty)$.

3. Перейдем к доказательству последнего пункта леммы. Покажем, что $x_{c_1} > \frac{S\theta}{1 + \gamma\theta}$, где $x_{c_1} = \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}$:

$$\frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}} > \frac{S\theta}{1 + \gamma\theta}.$$

Умножим обе части неравенства на положительную величину $\frac{1}{S} \cdot (r_D - \delta + \frac{\rho}{\theta})(1 + \gamma\theta)$. После преобразований, получим

$$\delta + \rho\gamma > r_D.$$

По условию, $r_D < \delta + \rho\gamma$, значит, $x_{c_1} > \frac{S\theta}{1 + \gamma\theta}$. Дальнейшее доказательство повторяет доказательство из п. 2.

Лемма доказана.

Лемма 5. Если $r_L > \max\{\delta - \frac{\rho}{\theta}, \delta + \rho\gamma\}$, то уравнение $M_r(x_{p_1}, r_L)$, где $0 < x_{p_1} \leq \frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}}$, имеет единственный корень на множестве $((x)_+, x_{p_1})$.

Доказательство. Пусть $M_{r_L} = M_r(x_{p_1}, r_L)$. Подставим $\bar{x} = x_{p_1}$, $r = r_L$ в уравнение (23). Вычтем из левой части уравнения (23) правую и рассмотрим функцию

$$f(M_{r_L}) = x + \frac{S}{r_L - \gamma} - \frac{(1 + \theta r_L)\rho}{\theta(\delta - (1 - \rho)r_L)} M_{r_L} + \left[\frac{M_{r_L}}{x_{p_1}} \right]^{\frac{\rho(r_L - \gamma)}{r_L - \delta - \rho\gamma}} \cdot \left[\frac{r_L - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_L} x_{p_1} - \frac{S}{r_L - \gamma} \right].$$

Рассмотрим случай, когда $x \leq 0$. Тогда корень функции $f(M_{r_L})$ будем искать на множестве $[0, x_{p_1})$. В точке $M_{r_L} = 0$ функция

$$f(0) = x + \frac{S}{r_L - \gamma} > 0,$$

а в точке $M_{r_L} = x_{p_1}$ функция

$$\begin{aligned} f(x_{p_1}) &= x + \frac{S}{r_L - \gamma} - \frac{(1 + \theta r_L)\rho}{\theta(\delta - (1 - \rho)r_L)} x_{p_1} + \frac{r_L\theta - \delta\theta + \rho}{\theta(\delta - (1 - \rho)r_L)} x_{p_1} - \frac{S}{r_L - \gamma} = \\ &= x + \frac{(1 - \rho)r_L - \delta}{\delta - (1 - \rho)r_L} x_{p_1} = x - x_{p_1} < 0. \end{aligned} \quad (26)$$

Вычислим производную функции $f(M_{r_L})$:

$$\frac{df(M_{r_L})}{dM_{r_L}} = -\frac{(1 + \theta r_L)\rho}{\theta(\delta - (1 - \rho)r_L)} + \frac{\rho(r_L - \gamma)}{r_L - \delta - \rho\gamma} \cdot \frac{1}{x_{p_1}} \cdot \left[\frac{M_{r_L}}{x_{p_1}} \right]^{\frac{\delta - (1 - \rho)r_L}{r_L - \delta - \rho\gamma}} \cdot \left[\frac{r_L - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_L} x_{p_1} - \frac{S}{r_L - \gamma} \right].$$

Если $\frac{r_L - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_L} x_{p_1} - \frac{S}{r_L - \gamma} < 0$, то функция $f(M_{r_L})$ монотонна, $f(0) \cdot f(x_{p_1}) < 0$, значит, существует единственный корень на интервале $(0, x_{p_1})$. Если $\frac{r_L - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_L} x_{p_1} - \frac{S}{r_L - \gamma} > 0$, то функция $f(M_{r_L})$ имеет единственный экстремум. Если этот экстремум находится вне множества $(0, x_{p_1})$, то функция $f(M_{r_L})$ монотонна на $[0, x_{p_1}]$, $f(0) \cdot f(x_{p_1}) < 0$, значит, существует единственный корень на интервале $(0, x_{p_1})$. Если экстремум функции $f(M_{r_L})$ находится внутри множества $(0, x_{p_1})$, то, согласно лемме 1, функция $f(M_{r_L})$ также имеет единственный корень на множестве $(0, x_{p_1})$.

Перейдем к рассмотрению случая, когда $x > 0$. В этом случае корень функции $f(M_{r_L})$ будем искать на множестве (x, x_{p_1}) .

Пусть $\frac{r_L - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_L} x_{p_1} - \frac{S}{r_L - \gamma} < 0$. Получим оценку снизу на функцию $f(M_{r_L})$ при $M_{r_L} = x$:

$$\begin{aligned} f(x) &= x + \frac{S}{r_L - \gamma} - \frac{(1 + \theta r_L) \rho}{\theta (\delta - (1 - \rho)r_L)} x - \\ &\quad - \left[\frac{x}{x_{p_1}} \right]^{\frac{\rho(r_L - \gamma)}{r_L - \delta - \rho\gamma}} \cdot \left[\frac{S}{r_L - \gamma} - \frac{r_L - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_L} x_{p_1} \right] > \left\{ \frac{x}{x_{p_1}} < 1, \text{ т.к. } x < x_{p_1} \right\} > \\ &\geq \frac{\theta \delta - \rho - \theta r_L}{\theta (\delta - (1 - \rho)r_L)} x + \frac{r_L \theta - \delta \theta + \rho}{\theta (\delta - (1 - \rho)r_L)} x_{p_1} = \frac{r_L - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_L} (x_{p_1} - x) > 0. \end{aligned}$$

В точке $M_{r_L} = x_{p_1}$ функция $f(M_{r_L})$ положительна (26). Аналогично, функция $f(M_{r_L})$ монотонна, $f(x) \cdot f(x_{p_1}) < 0$, значит, существует единственный корень на интервале (x, x_{p_1}) .

Пусть $\frac{r_L - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_L} x_{p_1} - \frac{S}{r_L - \gamma} > 0$. Производная функции $f(M_{r_L})$ удовлетворяет выражению (26). Вычислим значение производной функции $f(M_{r_L})$ в точке $M_{r_L} = x_{p_1}$. После упрощений подобных слагаемых, получим

$$\left. \frac{df(M_{r_L})}{dM_{r_L}} \right|_{M_{r_L}=x_{p_1}} = \frac{\rho(1 + \gamma\theta)x_{p_1} - \rho S \theta}{\theta(r_L - \delta - \rho\gamma)x_{p_1}}.$$

Покажем, что $\left. \frac{df(M_{r_L})}{dM_{r_L}} \right|_{M_{r_L}=x_{p_1}} < -1$:

$$\frac{\rho(1 + \gamma\theta)x_{p_1} - \rho S \theta}{\theta(r_L - \delta - \rho\gamma)x_{p_1}} < -1.$$

Разделим обе части неравенства на θ , приведем подобные слагаемые и выразим x_{p_1} . Получим

$$x_{p_1} < \frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}}.$$

По условию, $x_{p_1} \leq \frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}}$, значит, $\left. \frac{df(M_{r_L})}{dM_{r_L}} \right|_{M_{r_L}=x_{p_1}} \leq -1$. Пусть $x < z < x_{p_1}$. Тогда справедлива следующая оценка:

$$\begin{aligned} \left. \frac{df(M_{r_L})}{dM_{r_L}} \right|_{M_{r_L}=z} &= -\frac{(1 + \theta r_L) \rho}{\theta (\delta - (1 - \rho)r_L)} + \frac{\rho(r_L - \gamma)}{r_L - \delta - \rho\gamma} \cdot \frac{1}{x_{p_1}} \cdot \left[\frac{z}{x_{p_1}} \right]^{\frac{\delta - (1 - \rho)r_L}{r_L - \delta - \rho\gamma}} \cdot \\ &\cdot \left[\frac{r_L - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_L} x_{p_1} - \frac{S}{r_L - \gamma} \right] < -\frac{(1 + \theta r_L) \rho}{\theta (\delta - (1 - \rho)r_L)} + \frac{\rho(r_L - \gamma)}{r_L - \delta - \rho\gamma} \cdot \frac{1}{x_{p_1}} \cdot \\ &\cdot \left[\frac{r_L - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_L} x_{p_1} - \frac{S}{r_L - \gamma} \right] = \left. \frac{df(M_{r_L})}{dM_{r_L}} \right|_{M_{r_L}=x_{p_1}} \leq -1. \end{aligned}$$

Таким образом, мы получили, что на отрезке $[x, x_{p_1}]$, функция $f(M_{r_L})$ в точке $M_{r_L} = x_{p_1}$: $f(x_{p_1}) = x - x_{p_1} < 0$, производная $\left. \frac{df(M_{r_L})}{dM_{r_L}} \right|_{M_{r_L}=x_{p_1}} < -1$ на интервале (x, x_{p_1}) , значит, выполняются условия леммы 3, и существует единственный корень функции $f(M_{r_L})$ на интервале (x, x_{p_1}) . Лемма доказана.

Лемма 6. Если $\delta - \frac{\rho}{\theta} < r_D < \delta + \rho\gamma$, $x > x_{p_2} \geq \frac{S\rho}{r_D - \delta + \frac{\rho}{\theta}}$, то уравнение $M_r(x_{p_2}, r_D)$ имеет единственное решение на множестве (x_{p_2}, x) .

Доказательство. Пусть $M_{r_D} = M_r(x_{p_2}, r_D)$. Подставим $\bar{x} = x_{p_2}$, $r = r_D$ в уравнение (23). Вычтем из левой части уравнения (23) правую и рассмотрим функцию

$$f(M_{r_D}) = x + \frac{S}{r_D - \gamma} - \frac{(1 + \theta r_D) \rho}{\theta (\delta - (1 - \rho)r_D)} M_{r_D} + \left[\frac{x_{p_2}}{M_{r_D}} \right]^{\frac{\rho(r_D - \gamma)}{\delta + \rho\gamma - r_D}} \cdot \left[\frac{r_D - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_D} x_{p_2} - \frac{S}{r_D - \gamma} \right].$$

Определим знак функции $f(M_{r_D})$ в точке $M_{r_D} = x_{p_2}$. Получим

$$f(x_{p_2}) = x - x_{p_2} > 0.$$

Рассмотрим случай, когда $\frac{r_D - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_D} x_{p_2} - \frac{S}{r_D - \gamma} > 0$. Определим знак функции $f(M_{r_D})$ в точке $M_{r_D} = x$. Справедлива следующая оценка:

$$\begin{aligned} f(x) &= x + \frac{S}{r_D - \gamma} - \frac{(1 + \theta r_D) \rho}{\theta (\delta - (1 - \rho)r_D)} x + \left[\frac{x_{p_2}}{x} \right]^{\frac{\rho(r_D - \gamma)}{\delta + \rho\gamma - r_D}} \cdot \left[\frac{r_D - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_D} x_{p_2} - \frac{S}{r_D - \gamma} \right] \leq \\ &\leq \left\{ \frac{x_{p_2}}{x} < 1, \text{ т.к. } x_{p_2} < x \right\} \leq x - \frac{(1 + \theta r_D) \rho}{\theta (\delta - (1 - \rho)r_D)} x + \frac{r_D - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_D} x_{p_2} = \\ &= \frac{r_D - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_D} (x_{p_2} - x) < 0. \end{aligned}$$

Покажем, что экстремума у функции $f(M_{r_D})$ на множестве $(0, +\infty)$ нет. Выпишем производную:

$$\frac{df(M_{r_D})}{dM_{r_D}} = -\frac{(1 + \theta r_D) \rho}{\theta (\delta - (1 - \rho)r_D)} - \frac{\rho(r_D - \gamma)}{\delta + \rho\gamma - r_D} \cdot \frac{1}{M_{r_D}} \cdot \left[\frac{x_{p_2}}{M_{r_D}} \right]^{\frac{\rho(r_D - \gamma)}{\delta + \rho\gamma - r_D}} \cdot \left[\frac{r_D - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_D} x_{p_2} - \frac{S}{r_D - \gamma} \right]. \quad (27)$$

Очевидно, что $\frac{df(M_{r_D})}{dM_{r_D}} < 0$ на всей области определения, т.е. функция $f(M_{r_D})$ убывающая, не имеющая экстремумов. Таким образом, функция $f(M_{r_D})$ монотонна на множестве (x_{p_2}, x) , имеет разные знаки на концах этого множества, значит, существует единственный корень функции $f(M_{r_D})$ внутри множества (x_{p_2}, x) .

Перейдем к рассмотрению случая, когда $\frac{r_D - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_D} x_{p_2} - \frac{S}{r_D - \gamma} < 0$. Производная функции $f(M_{r_D})$ удовлетворяет выражению (27). Вычислим значение производной функции $f(M_{r_D})$ в точке $M_{r_D} = x_{p_2}$. После упрощений подобных слагаемых, получим

$$\left. \frac{df(M_{r_D})}{dM_{r_D}} \right|_{M_{r_D}=x_{p_2}} = \frac{\rho S \theta - \rho(1 + \gamma \theta) x_{p_2}}{\theta (\delta + \rho\gamma - r_D) x_{p_2}}.$$

Покажем, что $\left. \frac{df(M_{r_D})}{dM_{r_D}} \right|_{M_{r_D}=x_{p_2}} < -1$:

$$\frac{\rho S \theta - \rho(1 + \gamma \theta) x_{p_2}}{\theta (\delta + \rho\gamma - r_D) x_{p_2}} < -1.$$

Разделим обе части неравенства на θ , приведем подобные слагаемые и выразим x_{p_2} . Получим

$$\frac{S \rho}{r_D - \delta + \frac{\rho}{\theta}} < x_{p_2}.$$

По условию, $\frac{S \rho}{r_D - \delta + \frac{\rho}{\theta}} \leq x_{p_2}$, значит, $\left. \frac{df(M_{r_D})}{dM_{r_D}} \right|_{M_{r_D}=x_{p_2}} \leq -1$.

Пусть $z > x_{p_2}$. Тогда справедлива следующая оценка:

$$\begin{aligned} \left. \frac{df(M_{r_D})}{dM_{r_D}} \right|_{M_{r_D}=z} &= -\frac{(1 + \theta r_D) \rho}{\theta (\delta - (1 - \rho)r_D)} + \frac{\rho(r_D - \gamma)}{\delta + \rho\gamma - r_D} \cdot \frac{1}{z} \cdot \left[\frac{x_{p_2}}{z} \right]^{\frac{\rho(r_D - \gamma)}{\delta + \rho\gamma - r_D}} \times \\ &\times \left[\frac{S}{r_D - \gamma} - \frac{r_D - \delta + \frac{\rho}{\theta}}{\delta - (1 - \rho)r_D} x_{p_2} \right] < \left\{ \frac{1}{z} < \frac{1}{x_{p_2}} \right\} < \left. \frac{df(M_{r_D})}{dM_{r_D}} \right|_{M_{r_D}=x_{p_2}} \leq -1. \end{aligned}$$

Таким образом, мы получили, что на отрезке $[x_{p_2}, x]$ функция $f(M_{r_D})$ в точке $M_{r_D} = x_{p_2}$: $f(x_{p_2}) = x - x_{p_2} > 0$, производная $\frac{df(M_{r_D})}{dM_{r_D}} < -1$ на интервале (x_{p_2}, x) , значит, выполняются условия леммы 2, и существует единственный корень функции $f(M_{r_D})$ на интервале (x_{p_2}, x) . Лемма доказана.

Для качественного анализа экономического поведения репрезентативного домашнего хозяйства на несовершенном рынке кредитов и депозитов было разработано специализированное Приложение [20]. В зависимости от задаваемых параметров модели: экономической конъюнктуры r_L, r_D, γ и поведенческих характеристик репрезентативного домашнего хозяйства θ, ρ, δ , Приложение строит поле скоростей на плоскости $Ox\varphi$. Гиперболы $\varphi_{r_L}(x)$ и $\varphi_{r_D}(x)$ разделяют плоскость $Ox\varphi$ на три области: заимствования, не взаимодействия с банками и

сбережения. Приложение позволяет интерактивно строить траектории $\varphi(x; x_0, T)$ по заданной начальной точке $x(0) = x_0$ для конечного временного горизонта T , подбирая начальное значение $\varphi(0) = \varphi_0$ так, чтобы удовлетворить условию трансверсальности $x(T) = 0$. Визуализация поля скоростей координат x, φ на плоскости $Ox\varphi$ и построения траекторий $\varphi(x; x_0, T)$ позволяет на качественном уровне изучать структуру оптимальной траектории $\varphi(x)$, которая является предельной для $\varphi(x; x_0, T)$ при $T \rightarrow +\infty$.

Опираясь на утверждения и леммы выше, приступим к доказательству теоремы 3.

Доказательство. Рассмотрим каждый из типов поведения теоремы.

Тип 1.1. Зафиксируем начальное значение $x(0) = x_0$. Нам необходимо найти такое значение сопряженной переменной в начальный момент времени $\varphi(0) = \varphi_0$, чтобы удовлетворить условию трансверсальности в конечный момент времени $x(T) = 0$. В режиме заимствования сопряженная функция удовлетворяет дифференциальному уравнению (10). Решением этого дифференциального уравнения является функция

$$\varphi(t) = \varphi_0 e^{[\delta + \rho\gamma - r_L]t}. \quad (28)$$

Подставим решение (28) в дифференциальное уравнение (9). Решив его, получим траекторию изменения фазовой координаты x в зависимости от начальных значений φ_0, x_0 :

$$x(t; \varphi_0, x_0) = -\frac{S}{r_L - \gamma} + \left(\frac{\theta}{1 + \theta r_L}\right)^{\frac{1-\rho}{\rho}} \varphi_0^{-\frac{1}{\rho}} \cdot \frac{\rho}{\delta - (1-\rho)r_L} e^{\frac{r_L - \delta - \rho\gamma}{\rho}t} + e^{(r_L - \gamma)t} \left[x_0 + \frac{S}{r_L - \gamma} - \left(\frac{\theta}{1 + \theta r_L}\right)^{\frac{1-\rho}{\rho}} \varphi_0^{-\frac{1}{\rho}} \cdot \frac{\rho}{\delta - (1-\rho)r_L} \right]. \quad (29)$$

Выразим φ_0 из уравнения (29), разделив уравнение (29) на величину $e^{(r_L - \gamma)T}$ и устремив $T \rightarrow +\infty$. Получим

$$\varphi_0 = \left(\frac{\theta}{1 + \theta r_L}\right)^{1-\rho} \left[\frac{(r_L - \gamma)\rho}{(x_0(r_L - \gamma) + S)(\delta - (1-\rho)r_L)} \right]^\rho.$$

Так как начальное значение x_0 произвольно из допустимого множества $\left(-\frac{S}{r_L - \gamma}, +\infty\right)$, мы получаем оптимальную траекторию $\varphi(x)$:

$$\varphi(x) = \left(\frac{\theta}{1 + \theta r_L}\right)^{1-\rho} \left[\frac{(r_L - \gamma)\rho}{(x(r_L - \gamma) + S)(\delta - (1-\rho)r_L)} \right]^\rho. \quad (30)$$

Подставляя оптимальную траекторию $\varphi(x)$ в уравнение (8), мы получаем синтез оптимального управления в области заимствования

$$M(x; S, r_L, \theta, \rho, \gamma, \delta) = \frac{\theta(\delta - (1-\rho)r_L)}{\rho(1 + \theta r_L)} \left(x + \frac{S}{r_L - \gamma}\right).$$

Покажем, что оптимальная траектория $\varphi(x)$ лежит ниже гиперболы $\varphi_{r_L}(x)$, т.е. $\varphi(x) < \varphi_{r_L}(x)$ для любого $x > 0$:

$$\left(\frac{\theta}{1 + \theta r_L}\right)^{1-\rho} \left[\frac{(r_L - \gamma)\rho}{(x(r_L - \gamma) + S)(\delta - (1-\rho)r_L)} \right]^\rho < \frac{\theta}{1 + \theta r_L} x^{-\rho}. \quad (31)$$

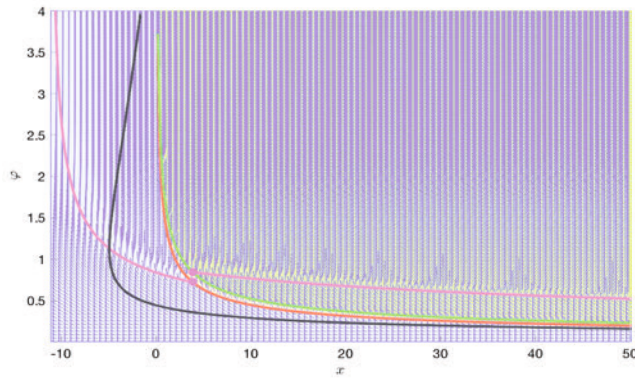
Разделим обе части неравенства на $\frac{\theta}{1 + \theta r_L}$ и возведем их в степень $\frac{1}{\rho}$. После преобразований, получим

$$x(r_L - \gamma)[\rho + \theta(r_L - \delta)] < S\theta(\delta - (1-\rho)r_L).$$

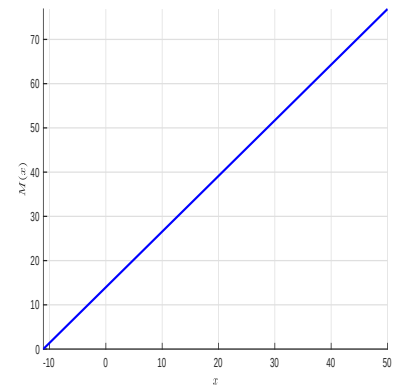
Разделив обе части неравенства на величину θ , получим

$$x(r_L - \gamma) \left[r_L - \delta + \frac{\rho}{\theta} \right] < S(\delta - (1-\rho)r_L).$$

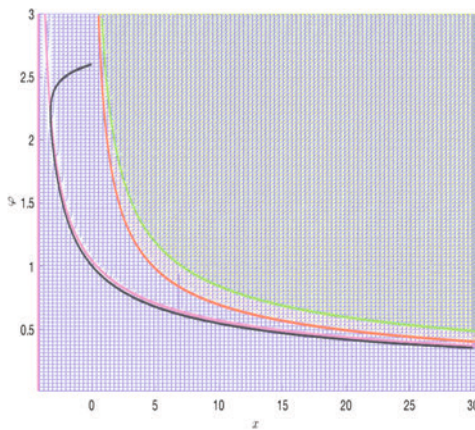
По условию, $r_L < \delta - \frac{\rho}{\theta}$, значит, неравенство верно, оптимальная траектория $\varphi(x)$ не пересекает гиперболу $\varphi_{r_L}(x)$, т.е. реализуется только режим заимствования. Для наглядности отобразим вид оптимальной траектории на плоскости $Ox\varphi$ и синтез оптимального управления на фиг. 2, 3. Пусть $S = 1, r_L = 0.1, r_D = 0.05, \gamma = 0.01, \theta = 4, \rho = 0.5, \delta = 0.2701, x_0 = 50, T = 20$. Оптимальная траектория $\varphi(x)$ устроена следующим образом. Она лежит в области заимствования ниже гиперболы $\varphi_z(x)$, с течением времени фазовая координата x уменьшается. За любое конечное время T траектория $\varphi(x; T)$ разворачивается вокруг гиперболы $\varphi_z(x)$, чтобы удовлетворить условию трансверсальности $x(T) = 0$.



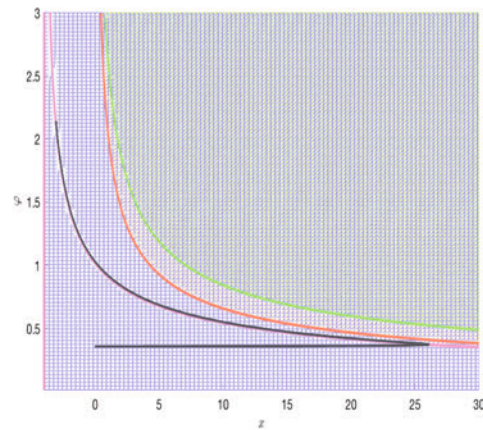
Фиг. 2. Красным, зеленым и розовым отображены гиперболы $\varphi_{r_L}(x)$, $\varphi_{r_D}(x)$, $\varphi_z(x)$ соответственно. Черным отображена траектория $\varphi(x; x_0, T)$. Синим отображено поле скоростей на плоскости $Ox\varphi$.



Фиг. 3. Синтез оптимального управления.



Фиг. 4. Красным, зеленым и розовым отображены гиперболы $\varphi_{r_L}(x)$, $\varphi_{r_D}(x)$, $\varphi_z(x)$ соответственно. Черным отображена траектория $\varphi(x; x_0, T)$. Синим отображено поле скоростей на плоскости $Ox\varphi$.



Фиг. 5. Красным, зеленым и розовым отображены гиперболы $\varphi_{r_L}(x)$, $\varphi_{r_D}(x)$, $\varphi_z(x)$ соответственно. Черным отображена траектория $\varphi(x; x_0, T)$. Синим отображено поле скоростей на плоскости $Ox\varphi$.

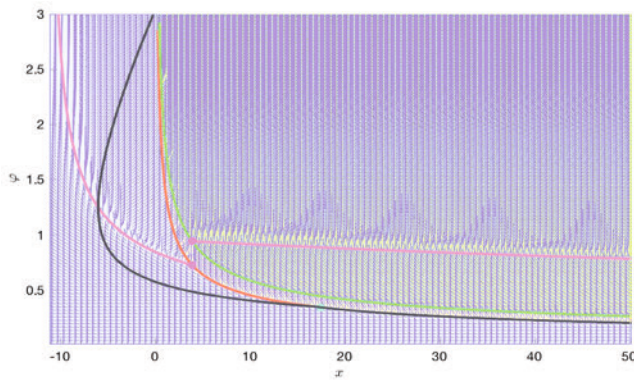
Для полноты картины продемонстрируем как устроена оптимальная траектория $\varphi(x)$ в случаях, когда $1 + \gamma\theta < 0$. Неравенство $1 + \gamma\theta < 0$ эквивалентно неравенству $\delta + \rho\gamma < \delta - \frac{\rho}{\theta}$. Согласно утверждению 1, в данном случае $\varphi_z(x) < \varphi_{r_L}(x)$.

Фиг. 4 соответствует условию $r_L < \delta + \rho\gamma$, где $S = 1$, $r_L = 0.14$, $r_D = 0.1$, $\gamma = -0.1$, $\theta = 11.4$, $\rho = 0.5$, $\delta = 0.2$, $x_0 = 30$, $T = 200$. Принцип движения оптимальной траектории $\varphi(x)$ на плоскости $Ox\varphi$ аналогичен фиг. 2. Фиг. 5 соответствует условию $\delta + \rho\gamma < r_L < \delta - \frac{\rho}{\theta}$, где $S = 1$, $r_L = 0.153$, $r_D = 0.1$, $\gamma = -0.1$, $\theta = 11.4$, $\rho = 0.5$, $\delta = 0.2$, $x_0 = -3$, $T = 600$. Оптимальная траектория $\varphi(x)$ лежит в области заимствования между гиперболами $\varphi_{r_L}(x)$ и $\varphi_z(x)$. С течением времени фазовая координата x увеличивается. За любое конечное время T траектория $\varphi(x; T)$ разворачивается вокруг гиперболы $\varphi_z(x)$, чтобы удовлетворить условию трансверсальности $x(T) = 0$.

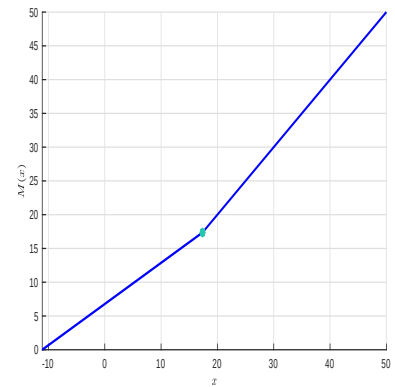
Тип 2.1. По условию, $r_L > \delta - \frac{\rho}{\theta}$, значит, оптимальная траектория $\varphi(x)$ пересечет гиперболу $\varphi_{r_L}(x)$ (см. (31)). Приравняв функции $\varphi(x)$ и $\varphi_{r_L}(x)$ друг к другу, получаем точку входа x_{p_1} оптимальной траектории $\varphi(x)$ из режима заимствования в автономный режим, где

$$x_{p_1} = \frac{S(\delta - (1 - \rho)r_L)}{(r_L - \gamma)(r_L - \delta + \frac{\rho}{\theta})}. \quad (32)$$

Так как $r_D < \delta - \frac{\rho}{\theta}$, то, по условию утверждения 3, оптимальная траектория $\varphi(x)$ после попадания на гиперболу $\varphi_{r_L}(x)$ в точке $x = x_{p_1}$ остается на ней для любого $x \geq x_{p_1}$. Мы получили синтез оптимального управления, соответствующий типу 2.1. На фиг. 6, 7 изображены оптимальная траектория и синтез оптимального управления



Фиг. 6. Красным, зеленым и розовым отображены гиперболы $\varphi_{r_L}(x)$, $\varphi_{r_D}(x)$, $\varphi_z(x)$ соответственно. Черным отображена траектория $\varphi(x; x_0, T)$. Синим отображено поле скоростей на плоскости $Ox\varphi$.



Фиг. 7. Синтез оптимального управления.

при $S = 1$, $r_L = 0.1$, $r_D = 0.0201$, $\gamma = 0.01$, $\theta = 4$, $\rho = 0.5$, $\delta = 0.1567$, $x_0 = 50$, $T = 40$. Точка касания оптимальной траектории $\varphi(x)$ с гиперболой $x_{p_1} = 17.4$.

Продолжим доказательство теоремы, приступив к рассмотрению типа 3.1.

Тип 3.1. Покажем, что при условиях типа 3.1. достигается режим сбережения. В рассматриваемом случае $r_D > \delta - \frac{\rho}{\theta}$, таким образом оптимальная траектория $\varphi(x)$ сходит с гиперболы $\varphi_{r_L}(x)$ в автономный режим в обратном времени в точке (x_{c_1}, φ_{c_1}) , где, согласно утверждению 3, $x_{c_1} = \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}$, $x_{c_1} \geq x_{p_1}$, x_{p_1} — определяется выражением (32), $\varphi_{c_1} = \frac{\theta}{1 + \theta r_L} x_{c_1}^{-\rho}$. Найдем решения дифференциальных уравнений (12), (13) в обратном времени при начальных условиях $x(0) = x_{c_1}$, $\varphi(0) = \varphi_{c_1}$. В соответствии с утверждением 2, положим $u(t) \equiv r_D$ в правой части дифференциального уравнения (13). Получим

$$x(t) = \frac{S\theta}{1 + \gamma\theta} + \left(x_{c_1} - \frac{S\theta}{1 + \gamma\theta}\right) e^{-\frac{1+\gamma\theta}{\theta}t}, \quad t \leq 0, \quad (33)$$

$$\varphi(t) = \varphi_{c_1} e^{[\delta + \rho\gamma - r_D]t}, \quad t \leq 0. \quad (34)$$

Условие пересечения оптимальной траектории $\varphi(x)$ с гиперболой $\varphi_{r_D}(x)$ означает выполнение равенства $\varphi(t) = \varphi_{r_D}(x(t))$. Подставим в него уравнения (33), (34). Учитывая, что $\varphi_{c_1} = \frac{\theta}{1 + \theta r_L} x_{c_1}^{-\rho}$, получим

$$\frac{\theta}{1 + \theta r_L} x_{c_1}^{-\rho} \cdot e^{[\delta + \rho\gamma - r_D]t} = \frac{\theta}{1 + \theta r_D} \cdot \left[\frac{S\theta}{1 + \gamma\theta} + \left(x_{c_1} - \frac{S\theta}{1 + \gamma\theta}\right) e^{-\frac{1+\gamma\theta}{\theta}t} \right]^{-\rho} \quad (35)$$

Выразим e^t из уравнения (33):

$$e^t = \left[\left(x - \frac{S\theta}{1 + \gamma\theta}\right) \left(x_{c_1} - \frac{S\theta}{1 + \gamma\theta}\right)^{-1} \right]^{-\frac{\theta}{1+\gamma\theta}}. \quad (36)$$

Подставим выражение (36) в (35) и возведем обе части в степень $-\frac{1}{\rho}$. После преобразований, получим

$$\left[\frac{1 + \theta r_L}{1 + \theta r_D} \right]^{-\frac{1}{\rho}} \cdot \frac{x}{x_{c_1}} = \left[\left(x - \frac{S\theta}{1 + \gamma\theta}\right) \left(x_{c_1} - \frac{S\theta}{1 + \gamma\theta}\right)^{-1} \right]^{\frac{\theta[\delta + \rho\gamma - r_D]}{(1 + \gamma\theta)\rho}}. \quad (37)$$

Уравнение (37) соответствует $y(\bar{x}, \beta)$ при $\beta = -1$, $\bar{x} = x_{c_1}$. Согласно п.2 леммы 4, уравнение (37) имеет единственное решение на множестве $(x_{c_1}, +\infty)$. Обозначим решение уравнения (37) через x_{p_2} .

Найдем синтез оптимального управления в области сбережения при $x > x_{p_2}$. Согласно утверждению 4, точка входа оптимальной траектории $\varphi(x)$ в область сбережения имеет координаты (x_{p_2}, φ_{p_2}) , где $\varphi_{p_2} = \frac{\theta}{1 + \theta r_D} x_{p_2}^{-\rho}$. В области сбережения оптимальная траектория удовлетворяет дифференциальным уравнениям (15), (16). Представим решения дифференциальных уравнений (15), (16) в обратном времени из точки (x_{p_2}, φ_{p_2}) :

$$\varphi(t) = \varphi_{p_2} e^{[\delta + \rho\gamma - r_D]t}, \quad (38)$$

$$x(t) = -\frac{S}{r_D - \gamma} + \left(\frac{\theta}{1 + \theta r_D} \right)^{\frac{1-\rho}{\rho}} \Phi_{p_2}^{-\frac{1}{\rho}} \cdot \frac{\rho}{\delta - (1-\rho)r_D} e^{\frac{r_D - \delta - \rho\gamma}{\rho} t} + e^{(r_D - \gamma)t} \left[x_{p_2} + \frac{S}{r_D - \gamma} - \left(\frac{\theta}{1 + \theta r_D} \right)^{\frac{1-\rho}{\rho}} \Phi_{p_2}^{-\frac{1}{\rho}} \cdot \frac{\rho}{\delta - (1-\rho)r_D} \right]. \quad (39)$$

Выразим из уравнения (38) e^t и подставим его в уравнение (39). После преобразований, получим связь фазовой координаты x с сопряженной функцией φ :

$$x = -\frac{S}{r_D - \gamma} + \frac{(1 + \theta r_D) \rho}{\theta (\delta - (1-\rho)r_D)} \cdot \left(\frac{\theta}{1 + \theta r_D} \right)^{\frac{1}{\rho}} \cdot \varphi^{-\frac{1}{\rho}} + \varphi^{-\frac{r_D - \gamma}{r_D - \delta - \rho\gamma}} \times \left(\frac{\theta}{1 + \theta r_D} \right)^{\frac{r_D - \gamma}{r_D - \delta - \rho\gamma}} \cdot x_{p_2}^{-\frac{\rho(r_D - \gamma)}{r_D - \delta - \rho\gamma}} \cdot \left[x_{p_2} + \frac{S}{r_D - \gamma} - \frac{(1 + \theta r_D) \rho}{\theta (\delta - (1-\rho)r_D)} \cdot x_{p_2} \right]. \quad (40)$$

Оптимальное управление в области сбережения удовлетворяет (14). Выразим из него φ :

$$\varphi = \frac{\theta}{1 + \theta r_D} \cdot M^{-\rho}.$$

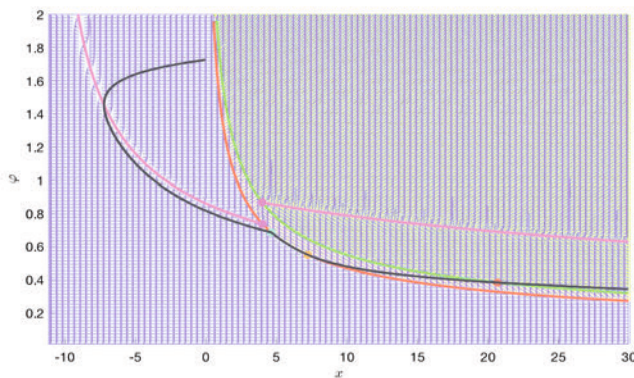
Последнее подставим в (40). Получим неявное уравнение, определяющее синтез оптимального управления в области сбережения:

$$x + \frac{S}{r_D - \gamma} = \frac{(1 + \theta r_D) \rho}{\theta (\delta - (1-\rho)r_D)} M + \left[\frac{x_{p_2}}{M} \right]^{\frac{\rho(r_D - \gamma)}{\delta - \rho\gamma - r_D}} \cdot \left[\frac{S}{r_D - \gamma} + \frac{\theta (\delta - r_D) - \rho}{\theta (\delta - (1-\rho)r_D)} x_{p_2} \right]. \quad (41)$$

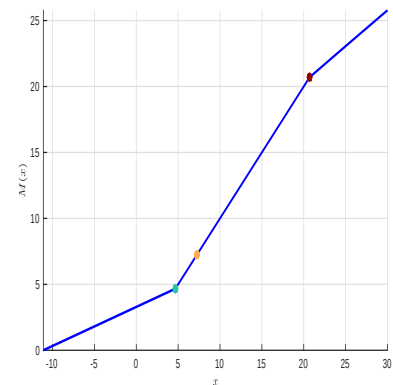
Согласно лемме 6, уравнение (41) имеет единственное решение на множестве (x_{p_2}, x) . Если уравнение (41) имеет решение на множестве $[0, x_{p_2}]$, то синтез оптимального управления $M(x)$ не монотонен. Это неверно, так как из монотонности функции цены для исходной задачи (4) при $T \rightarrow +\infty$ следует монотонность сопряженной переменной $\varphi(x)$, и, значит, синтез оптимального управления $M(x)$ также монотонен. Мы получили синтез оптимального управления, соответствующий типу 3.1.

На фиг. 8, 9 изображены оптимальная траектория и синтез оптимального управления при $S = 1$, $r_L = 0.1$, $r_D = 0.05$, $\gamma = 0.01$, $\theta = 4.2$, $\rho = 0.5$, $\delta = 0.1$, $x_0 = 30$, $T = 200$. Точки касания оптимальной траектории $\varphi(x)$ с гиперболой $\varphi_{r_L}(x)$ равны соответственно $x_{p_1} = 4.7$, $x_{c_1} = 7.2$. Оптимальная траектория $\varphi(x)$ пересекает гиперболу $\varphi_{r_D}(x)$ в точке $x_{p_2} = 20.7$.

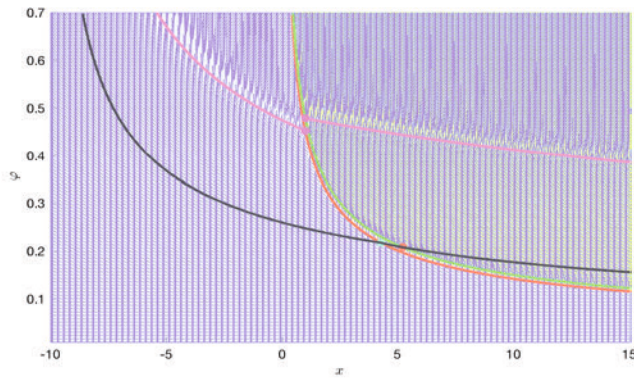
На фиг. 10, 11 изображены оптимальная траектория и синтез оптимального управления при $S = 1$, $r_L = 0.1098$, $r_D = 0.05$, $\gamma = 0.01$, $\theta = 1$, $\rho = 0.5$, $\delta = 0.2206$, $x_0 = 15$, $T = 80$. На гиперболе $\varphi_{r_L}(x)$ нет особого режима, оптимальная траектория $\varphi(x)$ пересекает ее в точке $x_{p_1} = 4.3$. Оптимальная траектория $\varphi(x)$ пересекает гиперболу $\varphi_{r_D}(x)$ в точке $x_{p_2} = 5.2$. Отметим, что в типе 3.1. движение оптимальной траектории $\varphi(x)$ по



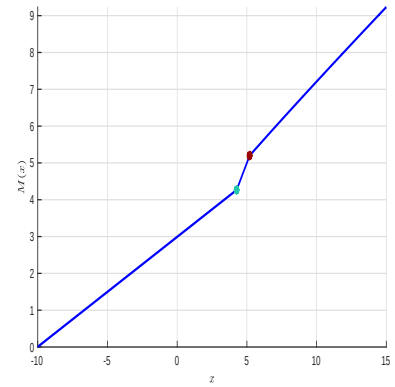
Фиг. 8. Красным, зеленым и розовым отображены гиперболы $\varphi_{r_L}(x)$, $\varphi_{r_D}(x)$, $\varphi_z(x)$ соответственно. Черным отображена траектория $\varphi(x; x_0, T)$. Синим отображено поле скоростей на плоскости $Ox\varphi$.



Фиг. 9. Синтез оптимального управления.



Фиг. 10. Красным, зеленым и розовым отображены гиперболы $\varphi_{r_L}(x)$, $\varphi_{r_D}(x)$, $\varphi_z(x)$ соответственно. Черным отображена траектория $\varphi(x; x_0, T)$. Синим отображено поле скоростей на плоскости Oxf .



Фиг. 11. Синтез оптимального управления.

гиперболе $\varphi_{r_L}(x)$ имеет место быть в случае, когда выполнено неравенство

$$(r_L - r_D)(\delta - (1 - \rho)r_L) > \left(r_L - \delta + \frac{\rho}{\theta}\right)(r_L - \delta - \rho\gamma).$$

Приступим к доказательству типа 2.2.

Тип 2.2. Рассмотрим случай $r_D < \delta - \frac{\rho}{\theta} < \delta + \rho\gamma < r_L$. В этом случае гипербола $\varphi_z(x)$ пересекает гиперболу $\varphi_{r_L}(x)$ в точке $\hat{x} = \frac{S\theta}{1+\gamma\theta}$. Согласно утверждению 3 и следствиям 1,2, оптимальная траектория $\varphi(x)$ касается гиперболы $\varphi_{r_L}(x)$ в точке $x_{p_1} = \frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}}$. Покажем, что $x_{p_1} < \hat{x}$. Если это так, то

$$\frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}} < \frac{S\theta}{1 + \gamma\theta}.$$

Обе части последнего неравенства умножим на положительную дробь $\frac{(1+\gamma\theta)(r_L - \delta + \frac{\rho}{\theta})}{S}$. Перенесем все в левую часть. После преобразований, получим

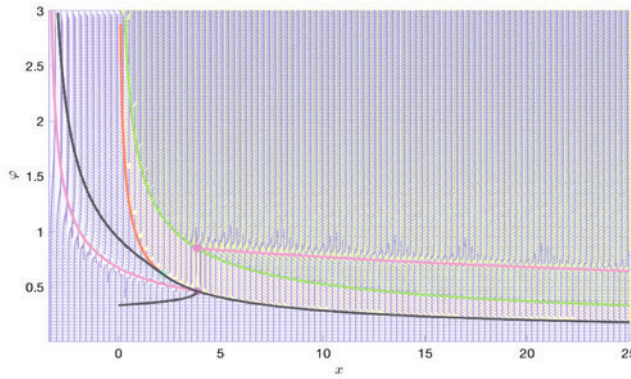
$$\theta(\delta + \rho\gamma - r_L) < 0,$$

что верно в рассматриваемом случае. Таким образом, $x_{p_1} < \hat{x}$. Так как $r_D < \delta - \frac{\rho}{\theta}$, то, согласно утверждению 3, для любого $x > x_{p_1}$, оптимальная траектория $\varphi(x)$ будет двигаться по гиперболе $\varphi_{r_L}(x)$ и останется в автономном режиме. Найдём синтез оптимального управления в области заимствования. Рассуждения аналогичны доказательству типа 3.1.: Из точки $(x_{p_1}, \varphi_{r_L}(x_{p_1}))$ выпустим траекторию в обратном времени, удовлетворяющую дифференциальным уравнениям (9), (10). Аналогично рассуждениям (38)-(40), получим неявное уравнение, определяющее синтез оптимального управления в области заимствования:

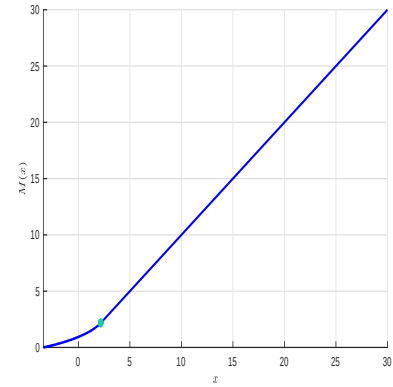
$$x + \frac{S}{r_L - \gamma} = \frac{(1 + \theta r_L)\rho}{\theta(\delta - (1 - \rho)r_L)}M + \left[\frac{x_{p_1}}{M}\right]^{\frac{\rho(r_L - \gamma)}{\delta + \rho\gamma - r_L}} \cdot \left[\frac{S}{r_L - \gamma} + \frac{\theta(\delta - r_L) - \rho}{\theta(\delta - (1 - \rho)r_L)}x_{p_1}\right]. \quad (42)$$

Согласно лемме 5, уравнение (42) имеет единственное решение на множестве $((x)_+, x_{p_1})$. Аналогично рассуждениям доказательства типа 3.1., любой другой корень уравнения (42), лежащий на луче $[x_{p_1}, +\infty)$ не является синтезом оптимального управления в области заимствования.

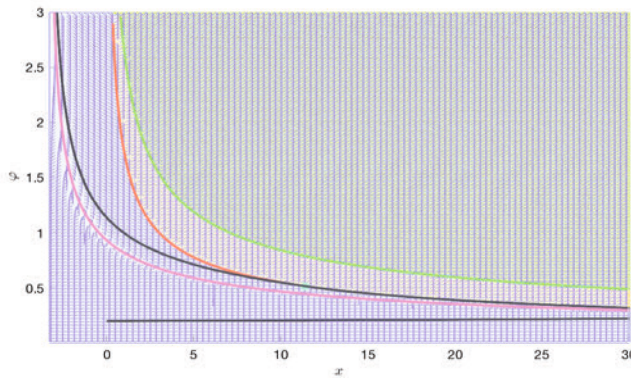
Оптимальная траектория состоит из двух частей. С одной стороны, она движется из области заимствования, касается гиперболы $\varphi_{r_L}(x)$ в точке $x_{p_1} < \hat{x}$, после чего начинает движение по гиперболе $\varphi_{r_L}(x)$. В точке $\hat{x}$ на гиперболе $\varphi_{r_L}(x)$ меняется направление движения вдоль фазовой координаты x . За любое конечное время T траектория $\varphi(x; T)$ развернется вокруг гиперболы $\varphi_z(x)$ в точке $\bar{x} < \hat{x}$ и удовлетворит условию трансверсальности $x(T) = 0$. При $T \rightarrow +\infty$, оптимальная траектория $\varphi(x)$ спустится по гиперболе $\varphi_{r_L}(x)$ в точку $A(\hat{x}, \varphi_{r_L}(\hat{x}))$. Легко убедиться, что точка A является точкой покоя системы в рассматриваемом случае. Действительно, относительно точки A движение траектории вдоль фазовой координаты x меняет направление, так как в нее входит гипербола $\varphi_z(x)$. Движение траектории вдоль координаты φ также меняет направление: при рассматриваемых соотношениях процентных ставок относительно величины $\delta + \rho\gamma$, в автономной области имеем $\frac{d\varphi}{dt} > 0$, а в области заимствования $\frac{d\varphi}{dt} < 0$. С другой стороны, при $x > \hat{x}$ оптимальная траектория $\varphi(x)$ движется по гиперболе



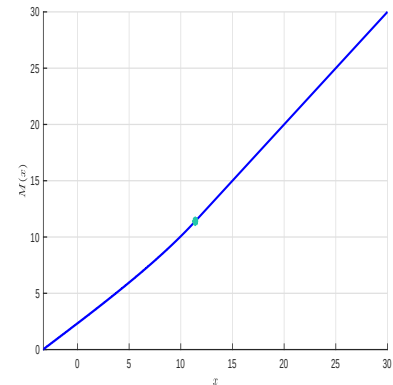
Фиг. 12. Красным, зеленым и розовым отображены гиперболы $\varphi_{r_L}(x)$, $\varphi_{r_D}(x)$, $\varphi_z(x)$ соответственно. Черным отображена траектория $\varphi(x; x_0, T)$. Синим отображено поле скоростей на плоскости $Ox\varphi$.



Фиг. 13. Синтез оптимального управления.



Фиг. 14. Красным, зеленым и розовым отображены гиперболы $\varphi_{r_L}(x)$, $\varphi_{r_D}(x)$, $\varphi_z(x)$ соответственно. Черным отображена траектория $\varphi(x; x_0, T)$. Синим отображено поле скоростей на плоскости $Ox\varphi$.



Фиг. 15. Синтез оптимального управления.

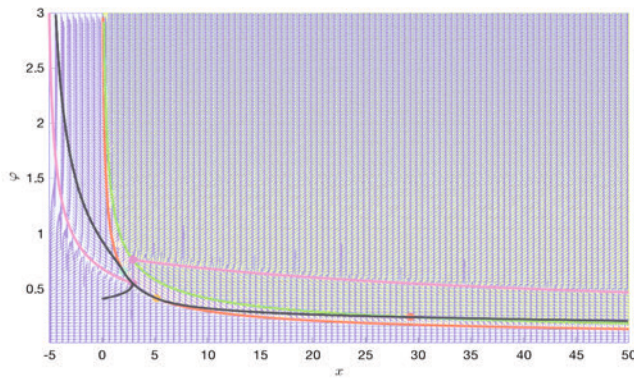
$\varphi_{r_L}(x)$ в направлении точки $\hat{x}$. За любое конечное время T траектория $\varphi(x; T)$ сойдет в область заимствования, чтобы удовлетворить условию трансверсальности $x(T) = 0$. При $T \rightarrow +\infty$, оптимальная траектория $\varphi(x)$ поднимется по гиперболе $\varphi_{r_L}(x)$ в неподвижную точку A .

Рассмотрим случай $\max\{r_D, \delta + \rho\gamma\} < \delta - \frac{\rho}{\theta} < r_L$. Из неравенств следует, что $\delta + \rho\gamma < \delta - \frac{\rho}{\theta}$, т.е. $1 + \gamma\theta < 0$. Согласно утверждению 1, в данном случае $\varphi_z(x) < \varphi_{r_L}(x)$. Аналогично, оптимальная траектория $\varphi(x)$ касается гиперболы $\varphi_{r_L}(x)$ в точке $x_{p1} = \frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}}$. Так как $r_D < \delta - \frac{\rho}{\theta}$, то, согласно утверждению 3, для любого $x > x_{p1}$ оптимальная траектория $\varphi(x)$ будет двигаться по гиперболе $\varphi_{r_L}(x)$ и останется в автономном режиме. Движение по гиперболе определено однозначно ($\varphi_z(x) < \varphi_{r_L}(x)$). За любое конечное время T траектория $\varphi(x; T)$ развернется вокруг гиперболы $\varphi_z(x)$ в точке $\tilde{x} < \hat{x}$ и удовлетворит условию трансверсальности $x(T) = 0$.

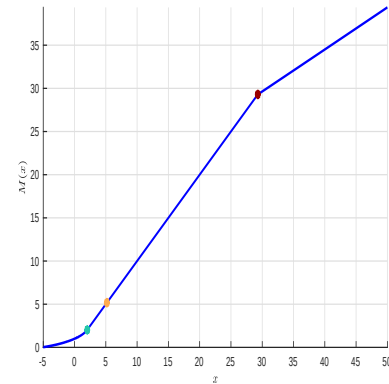
На фиг. 12, 13 изображены оптимальная траектория и синтез оптимального управления для случая $r_D < \delta - \frac{\rho}{\theta} < \delta + \rho\gamma < r_L$. Пусть $S = 1$, $r_L = 0.3$, $r_D = 0.05$, $\gamma = 0.01$, $\theta = 4$, $\rho = 0.5$, $\delta = 0.1959$, $x_0 = -3$ и $x_0 = 25$, $T = 70$. Точка касания оптимальной траектории $\varphi(x)$ с гиперболой $\varphi_{r_L}(x)$ $x_{p1} = 2.2$.

На фиг. 14, 15 изображены оптимальная траектория и синтез оптимального управления для случая $\max\{r_D, \delta + \rho\gamma\} < \delta - \frac{\rho}{\theta} < r_L$. Пусть $S = 1$, $r_L = 0.2$, $r_D = 0.1$, $\gamma = -0.1$, $\theta = 11.4$, $\rho = 0.5$, $\delta = 0.2$, $x_0 = -3$, $T = 70$. Точка касания оптимальной траектории $\varphi(x)$ с гиперболой $\varphi_{r_L}(x)$ $x_{p1} = 11.4$.

Тип 3.2. Построение синтеза в типе 3.2. отличается от построения синтеза в типе 2.2. тем, что в рассматриваемом случае $r_D > \delta - \frac{\rho}{\theta}$. Оптимальная траектория $\varphi(x)$ сходит с гиперболы $\varphi_{r_L}(x)$ в автономный режим в обратном времени в точке (x_{c1}, φ_{c1}) , где, согласно утверждению 3, $x_{c1} = \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}$. Покажем, что $x_{c1} > \hat{x}$, где



Фиг. 16. Красным, зеленым и розовым отображены гиперболы $\varphi_{r_L}(x)$, $\varphi_{r_D}(x)$, $\varphi_z(x)$ соответственно. Черным отображена траектория $\varphi(x; x_0, T)$. Синим отображено поле скоростей на плоскости $Ox\varphi$.



Фиг. 17. Синтез оптимального управления.

$$\hat{x} = \frac{S\theta}{1+\gamma\theta}.$$

$$\frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}} > \frac{S\theta}{1 + \gamma\theta}.$$

Обе части последнего неравенства умножим на положительную дробь $\frac{(1+\gamma\theta)(r_L - \delta + \frac{\rho}{\theta})}{S}$. Перенесем все в левую часть. После преобразований, получим

$$\theta(\delta + \rho\gamma - r_D) > 0,$$

что верно в рассматриваемом случае. Таким образом, $x_{c_1} > \hat{x}$. Покажем, что при условиях типа 3.2. достигается режим сбережения. Аналогично доказательству типа 3.1., получаем, что точка попадания оптимальной траектории на гиперболу $\varphi_{r_D}(x)$ определяется уравнением. (37). Оно соответствует $y(\bar{x}, \beta)$ при $\beta = -1$, $\bar{x} = x_{c_1}$. Согласно п. 3 леммы 4, уравнение (37) имеет единственное решение на множестве $(x_{c_1}, +\infty)$. Обозначим решение уравнения (37) через x_{p_2} . Нахождение синтеза оптимального управления в области сбережения повторяет рассуждения доказательства типа 3.1. В области сбережения синтез определяется неявным уравнением (41). Согласно лемме 6, уравнение (41) имеет единственное решение на множестве (x_{p_2}, x) . Если уравнение (41) имеет решение на множестве $[0, x_{p_2}]$, то синтез оптимального управления $M(x)$ не монотонен. Такого быть не может (см. доказательство типа 3.1.), значит, значение синтеза оптимального управления в области сбережения определяется однозначно на множестве (x_{p_2}, x) . Мы получили синтез оптимального управления, соответствующий типу 3.2.

На фиг. 16, 17 изображены оптимальная траектория и синтез оптимального управления. Пусть $S = 1$, $r_L = 0.2$, $r_D = 0.05$, $\gamma = 0.01$, $\theta = 3$, $\rho = 0.5$, $\delta = 0.12$, $x_0 = -5$ и $x_0 = 50$, $T = 50$. Точки касания оптимальной траектории $\varphi(x)$ с гиперболой $\varphi_{r_L}(x)$ равны соответственно $x_{p_1} = 2$, $x_{c_1} = 5.2$. Оптимальная траектория $\varphi(x)$ пересекает гиперболу $\varphi_{r_D}(x)$ в точке $x_{p_2} = 29.3$.

Тип 3.3. Отличие типа 3.3. от типа 3.1. состоит в том, что в типе 3.1. $\frac{d\varphi}{dt} > 0$ во всех трех областях, а в типе 3.3. $\frac{d\varphi}{dt} < 0$ во всех трех областях. В некотором смысле, можно сказать, что тип 3.3. “зеркален” типу 3.1. За любое конечное время T траектория $\varphi(x; T)$ развернется вокруг гиперболы $\varphi_z(x)$ и удовлетворит условию трансверсальности $x(T) = 0$. При $T \rightarrow +\infty$ оптимальная траектория $\varphi(x)$ будет лежать выше траектории разворота $\varphi_z(x)$. Рассуждая аналогично доказательству типа 1.1., повторяя шаги (28)–(30) с точностью до процентной ставки, получаем, что в области сбережения оптимальная траектория $\varphi(x)$ задается формулой

$$\varphi(x) = \left(\frac{\theta}{1 + \theta r_D} \right)^{1-\rho} \left[\frac{(r_D - \gamma)\rho}{(x(r_D - \gamma) + S)(\delta - (1 - \rho)r_D)} \right]^\rho. \quad (43)$$

Подставляя оптимальную траекторию $\varphi(x)$ в уравнение (14), мы получаем синтез оптимального управления в области сбережения

$$M(x; S, r_D, \theta, \rho, \gamma, \delta) = \frac{\theta(\delta - (1 - \rho)r_D)}{\rho(1 + \theta r_D)} \left(x + \frac{S}{r_D - \gamma} \right).$$

Согласно утверждению 4, на гиперболе $\varphi_{r_D}(x)$ нет особого режима оптимальной траектории, траектория $\varphi(x)$ пересекает гиперболу $\varphi_{r_D}(x)$ в точке x_{p_2} . Приравняв функцию (43) к $\varphi_{r_D}(x)$, получаем точку пересечения

оптимальной траектории гиперболы $\varphi_{r_D}(x)$. Получаем

$$x_{p_2} = \frac{S(\delta - (1 - \rho)r_D)}{(r_D - \gamma)(r_D - \delta + \frac{\rho}{\theta})}. \quad (44)$$

В автономной области движение оптимальной траектории подчиняется дифференциальным уравнениям (12)–(13). Найдем точку касания оптимальной траектории $\varphi(x)$ с гиперболой $\varphi_{r_L}(x)$, двигаясь в обратном времени из точки $A(x_{p_2}, \frac{\theta}{1+\theta r_D} x_{p_2}^{-\rho})$ на гиперболе $\varphi_{r_D}(x)$. Условие пересечения оптимальной траектории $\varphi(x)$ с гиперболой $\varphi_{r_L}(x)$ означает выполнение равенства $\varphi(t) = \varphi_{r_L}(x(t))$, где

$$x(t) = \frac{S\theta}{1 + \gamma\theta} + \left(x_{p_2} - \frac{S\theta}{1 + \gamma\theta}\right) e^{-\frac{1+\gamma\theta}{\theta}t}, \quad t \leq 0, \quad (45)$$

$$\varphi(t) = \varphi_{p_2} e^{[\delta + \rho\gamma - r_D]t}, \quad t \leq 0.$$

Учитывая, что $\varphi_{p_2} = \frac{\theta}{1+\theta r_D} x_{p_2}^{-\rho}$, получим

$$\frac{\theta}{1 + \theta r_D} x_{p_2}^{-\rho} \cdot e^{[\delta + \rho\gamma - r_D]t} = \frac{\theta}{1 + \theta r_L} \cdot \left[\frac{S\theta}{1 + \gamma\theta} + \left(x_{p_2} - \frac{S\theta}{1 + \gamma\theta}\right) e^{-\frac{1+\gamma\theta}{\theta}t} \right]^{-\rho}. \quad (46)$$

Выразим e^t из уравнения (45):

$$e^t = \left[\left(x - \frac{S\theta}{1 + \gamma\theta}\right) \left(x_{p_2} - \frac{S\theta}{1 + \gamma\theta}\right)^{-1} \right]^{-\frac{\theta}{1+\gamma\theta}}. \quad (47)$$

Подставим выражение (47) в (46) и возведем обе части в степень $-\frac{1}{\rho}$. После преобразований, получим

$$\left[\frac{1 + \theta r_L}{1 + \theta r_D} \right]^{\frac{1}{\rho}} \cdot \frac{x}{x_{p_2}} = \left[\left(x - \frac{S\theta}{1 + \gamma\theta}\right) \left(x_{p_2} - \frac{S\theta}{1 + \gamma\theta}\right)^{-1} \right]^{\frac{\theta[\delta + \rho\gamma - r_D]}{(1+\gamma\theta)\rho}}. \quad (48)$$

Уравнение (48) соответствует $y(\bar{x}, \beta)$ при $\beta = 1$, $\bar{x} = x_{p_2}$. Согласно п.1 леммы 4, уравнение (48) имеет единственное решение на множестве $(0, x_{p_2})$. Обозначим решение уравнения (48) через x_{c_1} .

Согласно следствию 2, движение оптимальной траектории $\varphi(x)$ по гиперболе $\varphi_{r_L}(x)$ возможно на отрезке $x \in \left[\frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}}, \left[\frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}} \right] \right]$. Если $x_{c_1} < \frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}}$, то оптимальная траектория $\varphi(x)$ пересекает гиперболу $\varphi_{r_L}(x)$ без движения по ней. Покажем, что $x_{c_1} < \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}$. Очевидно, что $x_{c_1} < x_{p_2}$. Покажем, что $x_{p_2} < \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}$:

$$\frac{S(\delta - (1 - \rho)r_D)}{(r_D - \gamma)(r_D - \delta + \frac{\rho}{\theta})} < \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}.$$

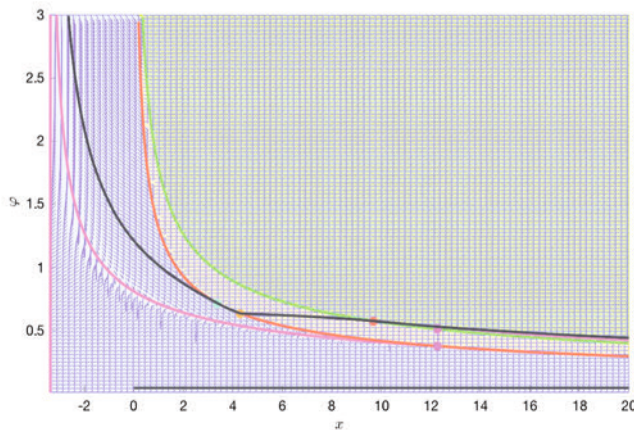
Обе части неравенства умножим на положительную дробь $\frac{(r_D - \gamma)(r_D - \delta + \frac{\rho}{\theta})}{S}$. Перенесем все слагаемые в левую часть. После преобразований получим

$$\delta + \rho\gamma < r_D.$$

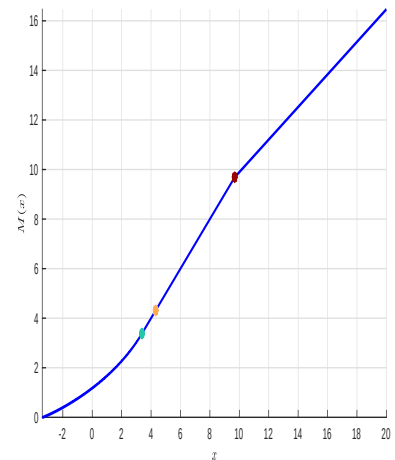
Последнее неравенство верно в рассматриваемом случае. Так как $x_{c_1} < x_{p_2}$, $x_{p_2} < \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}$, то $x_{c_1} < \frac{\rho S}{r_D - \delta + \frac{\rho}{\theta}}$. Обозначим точку начала движения оптимальной траектории $\varphi(x)$ в обратном времени в области заимствования через $x_{p_1} = \min \left\{ \frac{\rho S}{r_L - \delta + \frac{\rho}{\theta}}, x_{c_1} \right\}$. По условию, $r_D > \max \left\{ \delta - \frac{\rho}{\theta}, \delta + \rho\gamma \right\}$. Так как $r_L > r_D$, то в рассматриваемом случае справедливо неравенство $r_L > \max \left\{ \delta - \frac{\rho}{\theta}, \delta + \rho\gamma \right\}$. Последнее выполняется в типе 3.2., поэтому построение синтеза оптимального управления в области заимствования аналогично типу 3.2.

Рассмотрим случай, когда справедлива следующая цепочка неравенств: $\delta - \frac{\rho}{\theta} < \delta + \rho\gamma < r_D$. На фиг. 18, 19 изображены оптимальная траектория и синтез оптимального управления. Пусть $S = 1$, $r_L = 0.3$, $r_D = 0.2$, $\gamma = 0.0021$, $\theta = 12.6$, $\rho = 0.5$, $\delta = 0.1918$, $x_0 = -3$, $T = 100$. Точки касания оптимальной траектории $\varphi(x)$ с гиперболой $\varphi_{r_L}(x)$ равны соответственно $x_{p_1} = 3.4$, $x_{c_1} = 4.3$. Оптимальная траектория $\varphi(x)$ пересекает гиперболу $\varphi_{r_D}(x)$ в точке $x_{p_2} = 9.7$.

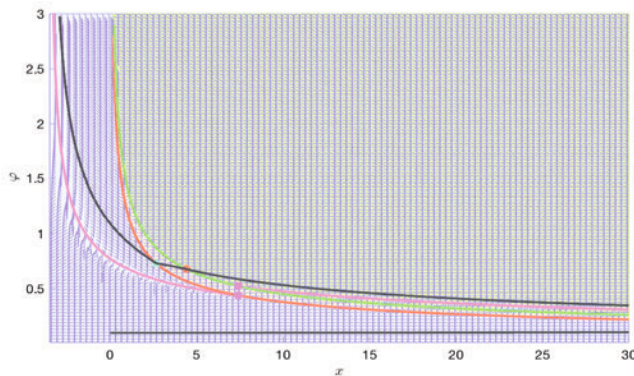
На фиг. 20, 21 изображены оптимальная траектория и синтез оптимального управления при $S = 1$, $r_L = 0.2954$, $r_D = 0.2273$, $\gamma = 0.01$, $\theta = 8$, $\rho = 0.5$, $\delta = 0.2$, $x_0 = -3$, $T = 100$. На гиперболе $\varphi_{r_L}(x)$ нет особого



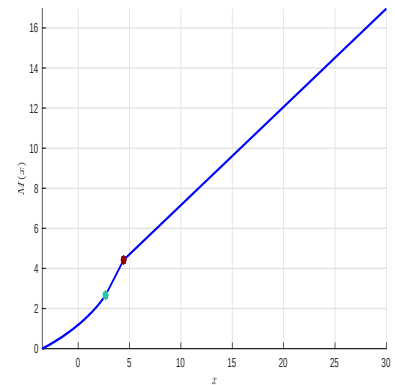
Фиг. 18. Красным, зеленым и розовым отображены гиперболы $\varphi_{r_L}(x)$, $\varphi_{r_D}(x)$, $\varphi_z(x)$ соответственно. Черным отображена траектория $\varphi(x; x_0, T)$. Синим отображено поле скоростей на плоскости $Ox\varphi$.



Фиг. 19. Синтез оптимального управления.



Фиг. 20. Красным, зеленым и розовым отображены гиперболы $\varphi_{r_L}(x)$, $\varphi_{r_D}(x)$, $\varphi_z(x)$ соответственно. Черным отображена траектория $\varphi(x; x_0, T)$. Синим отображено поле скоростей на плоскости $Ox\varphi$.



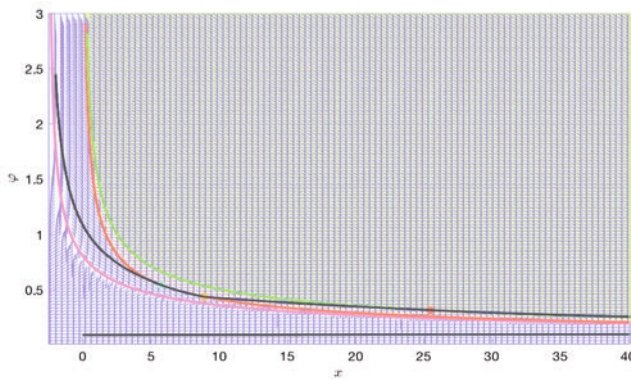
Фиг. 21. Синтез оптимального управления.

режима, оптимальная траектория $\varphi(x)$ пересекает ее в точке $x_{p_1} = 2.7$. Оптимальная траектория $\varphi(x)$ пересекает гиперболу $\varphi_{r_D}(x)$ в точке $x_{p_2} = 4.4$.

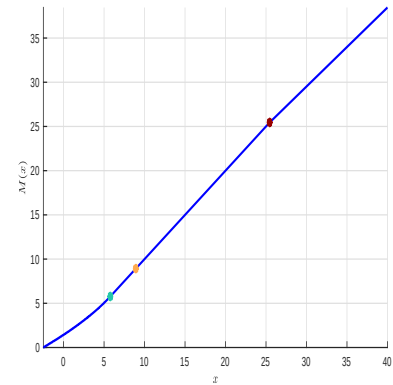
Случай, когда справедлива цепочка неравенств $\delta + \rho\gamma < \delta - \frac{\rho}{\theta} < r_D$ отличается от предыдущего тем, что гипербола $\varphi_z(x)$ лежит ниже гиперболы $\varphi_{r_L}(x)$. Покажем, что оптимальная траектория также будет находиться в трех областях. Для этого достаточно показать, что оптимальная траектория $\varphi(x)$ при $x > x_{p_2}$ будет находиться в области сбережения. Предположим, что это не так. Пусть оптимальная траектория находится только в области заимствования. Тогда она определяется функцией (30) и справедливо неравенство $\varphi(x) < \varphi_{r_L}(x)$. Как и в доказательстве типа 1.1., это приводит к неравенству

$$x(r_L - \gamma) \left[r_L - \delta + \frac{\rho}{\theta} \right] < S(\delta - (1 - \rho)r_L). \quad (49)$$

По условию, $r_L > \delta - \frac{\rho}{\theta}$, обе части неравенства (49) положительны. Для любого $x > \frac{S(\delta - (1 - \rho)r_L)}{(r_L - \gamma)[r_L - \delta + \frac{\rho}{\theta}]}$ неравенство (49) будет нарушаться, значит, предположение о том, что оптимальная траектория $\varphi(x)$ находится только в области заимствования неверное. Предположим, что оптимальная траектория $\varphi(x)$ находится в двух областях: заимствования и автономной. Движение оптимальной траектории $\varphi(x)$ в автономной области определяется дифференциальными уравнениями (9), (10), их решения определяются формулами (33), (34). Зафиксируем начальную точку $x(0) = x_0$, $\varphi(0) = \varphi_0$ в автономной области. Выразим зависимость $t(x)$ и подставим в траекторию



Фиг. 22. Красным, зеленым и розовым отображены гиперболы $\varphi_{r_L}(x)$, $\varphi_{r_D}(x)$, $\varphi_z(x)$ соответственно. Черным отображена траектория $\varphi(x; x_0, T)$. Синим отображено поле скоростей на плоскости Oxf .



Фиг. 23. Синтез оптимального управления.

$\varphi(t)$. Получим, что оптимальная траектория $\varphi(x)$ в автономной области описывается функцией

$$\varphi(x) = \varphi_0 \left[\left(x - \frac{S\theta}{1+\gamma\theta} \right) \left(x_0 - \frac{S\theta}{1+\gamma\theta} \right)^{-1} \right]^{\frac{\theta[r_D - \delta - \rho\gamma]}{1+\gamma\theta}}. \quad (50)$$

При наших предположениях, функция (50) лежит ниже $\varphi_{r_D}(x)$, т.е.

$$\varphi_0 \left[\left(x - \frac{S\theta}{1+\gamma\theta} \right) \left(x_0 - \frac{S\theta}{1+\gamma\theta} \right)^{-1} \right]^{\frac{\theta[r_D - \delta - \rho\gamma]}{1+\gamma\theta}} < \frac{\theta}{1+\theta r_D} x^{-\rho}. \quad (51)$$

Отметим, что степень множителя x в левой части неравенства (51) положительна: $\frac{\theta[r_D - \delta - \rho\gamma]}{1+\gamma\theta} > 0$, а в правой части неравенств отрицательна: $-\rho < 0$. Значит, существует значение $\hat{x}$ такое, что при любом $x > \hat{x}$ неравенство (51) будет нарушаться. Значит, предположение о том, что оптимальная траектория $\varphi(x)$ не находится в области сбережения неверное. В области сбережения оптимальная траектория $\varphi(x)$ определяется функцией (43). Покажем, что в области сбережения $\varphi(x) > \varphi_{r_D}(x)$: Аналогично (49), это приводит к неравенству

$$x(r_D - \gamma) \left[r_D - \delta + \frac{\rho}{\theta} \right] > S(\delta - (1 - \rho)r_D). \quad (52)$$

Неравенство (52) верно для любых $x > x_{p2}$, где x_{p2} определяется формулой (44).

На фиг. 22, 23 изображены оптимальная траектория и синтез оптимального управления для случая $\delta + \rho\gamma < \delta - \frac{\rho}{\theta} < r_D$. Пусть $S = 1$, $r_L = 0.3$, $r_D = 0.23$, $\gamma = -0.1$, $\theta = 12$, $\rho = 0.5$, $\delta = 0.255$, $x_0 = -2$, $T = 70$. Точки касания оптимальной траектории $\varphi(x)$ с гиперболой $\varphi_{r_L}(x)$ равны соответственно $x_{p1} = 5.8$, $x_{c1} = 8.9$. Оптимальная траектория $\varphi(x)$ пересекает гиперболу $\varphi_{r_D}(x)$ в точке $x_{p2} = 25.5$.

Мы рассмотрели все типы теоремы, для каждого из которых нашли оптимальную траекторию на плоскости Oxf , построили синтез оптимального управления $M(x; S, r_L, r_D, \theta, \rho, \gamma, \delta)$. Теорема доказана.

4. ЗАКЛЮЧЕНИЕ

В работе исследовалось математическое описание экономического поведения рационального домашнего хозяйства на несовершенном рынке потребительских кредитов и депозитов. Модель формализована в виде задачи оптимального управления на конечном временном горизонте. Негладкость правой части дифференциального уравнения на фазовую переменную мотивирует использовать принцип максимума Понтрягина в форме Ф. Кларка. При его применении возникает область, в которой домашние хозяйства не взаимодействуют с банковской системой. С областью не взаимодействия с банковской системой связана возможность особых режимов, исследование которых проделано в работе. Структура оптимального управления исследована на плоскости координат фазовой и сопряженной переменных. В зависимости от соотношений процентных ставок и поведенческих характеристик домашнего хозяйства, построена оптимальная траектория. При стремлении временного горизонта задачи оптимального управления к бесконечности построен синтез.

Математическая модель экономического поведения домашнего хозяйства может быть использована для идентификации поведенческих характеристик, как репрезентативного, так и группового поведения домашних хозяйств из различных социальных слоев. Модели группового поведения позволяют следить за экономическим поведением не репрезентативного домашнего хозяйства, а за поведением группового поведения большого количества домашних хозяйств. С математической точки зрения такие модели базируются на концепции игр среднего поля и представляют систему из связанных между собой уравнения Гамильтона–Якоби–Беллмана, которое описывает выбор стратегии поведения домашних хозяйств, и уравнения Колмогорова–Фоккера–Планка, которое описывает динамику состояния системы. В основе идентификации группового поведения домашних хозяйств лежит идентификация репрезентативного домашнего хозяйства, характерное для рассматриваемого социального слоя. Для идентификации используются статистические данные. Статистические данные Российского Мониторинга Экономического Положения и Здоровья Населения (РМЭЗ) [21], собираемые НИУ ВШЭ, и статистические данные Обследования Бюджетов Домашних Хозяйств (ОБДХ) [22], собираемые Росстатом [22], использовались авторами работы для калибровки модели экономического поведения репрезентативного домашнего хозяйства в [5], [23] и группового поведения домашних хозяйств в [24]. С их помощью на основе математической модели рамсеевского типа анализировалось состояние рынка потребительского кредита в России.

СПИСОК ЛИТЕРАТУРЫ

1. Ramsey F.P. A mathematical theory of savings // The Economic Journal. 1928. Vol. 152. No 38. P. 543–559.
2. Рудева А.В., Шананин А.А. Синтез управления в модифицированной модели Рамсея с учетом ограничения ликвидности // Дифференц. уравнения. 2009. Т. 45. № 12. С. 1799–1803.
3. Гималтдинов И.Ф. Исследование спроса на потребительские кредиты и наличные деньги // Матем. моделирование. 2012. Т. 24. № 2. С. 84–98.
4. Кларк Ф. Оптимизация и негладкий анализ. М.: Наука, 1988, 280 с.
5. Тарасенко М.В., Трусов Н.В., Шананин А.А. Математическое моделирование экономического положения домашних хозяйств в России // Ж. вычисл. матем. и матем. физ. 2021. Т. 61. № 6. С. 1034–1056.
6. Крупенский Н.А., Кузина О.Е. Перекредитованность россиян: миф или реальность? // Вопросы экономики. 2018. Т. 11. С. 85–104.
7. World Bank. Household Over-Indebtedness in Russia, 2020.
8. Зубаревич Н.В., Сафронов С. Г. Люди и деньги: доходы, потребление и финансовое поведение населения российских регионов в 2000–2017 гг. // Изв. РАН. Сер. географическая. 2019. Т. 5. С. 3–17.
9. Базанова Е. Минэкономразвития готово поддержать закредитованное население. Ведомости, 2019; <https://www.vedomosti.ru/economics/articles/2019/07/2>.
10. Синяков А., Ушакова Ю. Опыт двух кредитных бумов в России и текущий рост ипотеки: почему важно вовремя остановиться // Аналитическая записка Департамента исследований и прогнозирования Банка России, 2019, 17 с.
11. Бессонова Е., Цветкова А. Финансы российских домохозяйств в 2022 году // Аналитическая записка Департамента исследований и прогнозирования Банка России, 2023. 28 с.
12. Всероссийское обследование домохозяйств по потребительским финансам [Электронный ресурс]. 2013–2022. http://www.cbr.ru/ec_research/vserossiyskoe-obsledovanie-domokhozyaystv-po-potrebitel-skim-finansam/ (дата обращения: 05.03.2024)
13. Гурьев С. М. Модель формирования сбережений и спроса на деньги: I // Матема. моделирование. 1994. Т. 6. № 7. С. 25–40.
14. Коваль П.К., Полбин А.В. Оценка потребительского поведения домохозяйств в РФ // Вопросы экономики. 2022. Т. 3. С. 98–117.
15. Komlos J. A generalization of a problem of Steinhaus // Acta Mathematica Academiae Scientiarum Hungaricae. 1967. Vol. 18. P. 217–229.

16. *Clarke F.H., Ledyaev Yu.S., Stern R.J., Wolenski P.R.* Nonsmooth Analysis and Control Theory. N.Y.: Springer-Verlag, 1998. 276 p.
17. *Плотников В.И., Сумин М.И.* Необходимые условия в негладкой задаче оптимального управления // Математические заметки. 1982. Т. 32. № 2. С. 187–197.
18. *Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф.* Математическая теория оптимальных процессов. М.: Физматлит, 1961. 391 с.
19. *Ли Э., Маркус Л.* Основы теории оптимального управления. М.: Наука, 1972, 574 с.
20. Свидетельство о государственной регистрации программы для ЭВМ No 2023685827. “Моделирование экономического поведения домашних хозяйств в условиях несовершенного рынка кредитов”. Автор: Трусов Николай Всеволодович. Правообладатель: Федеральное государственное учреждение “Федеральный исследовательский центр “Информатика и управление” Российской академии наук” (ФИЦ ИУ РАН). Заявка No 2023685469. Дата государственной регистрации в Реестре программ для ЭВМ 30 ноября 2023г.
21. "Russia Longitudinal Monitoring survey, RLMS-HSE "Российский мониторинг экономического положения и здоровья населения НИУ ВШЭ (RLMS HSE) проводимый Национальным исследовательским университетом "Высшая школа экономики" и ООО "Демоскоп" при участии Центра народонаселения Университета Северной Каролины в Чапел Хилле и Института социологии Федерального научно-исследовательского социологического центра РАН.
22. Обследование бюджетов домашних хозяйств [Электронный ресурс] / Федеральная служба государственной статистики. 2015–2022. <https://obdx.gks.ru/> (дата обращения: 05.03.2024)
23. *Трусов Н.В., Шананин А.А.* Математическое моделирование рынка потребительского кредита в России в условиях санкций // Докл. АН. Математика, информатика, процессы управления. 2022. Т. 507. № 6. С. 71–80.
24. *Trusov N.V.* Identification of the household behavior modeling based on modified Ramsey model // Lobachevskii Journal of Mathematics. 2023. Vol. 44. No 1. P. 455–469.

OPTIMAL CONTROL SYNTHESIS IN A RAMSAY-TYPE MODEL

N. V. Trusov^{a,b,*}, A. A. Shaninin^{a,b,c,**}

^a 119333 Moscow, Vavilov Str., 44, Federal Research Center Computer Science and Control, RAS, Russia

^b 119991 Moscow, Moscow Center for Fundamental and Applied Mathematics, Lomonosov Moscow State University, Russia

^c 141701 Dolgoprudny, M.O., Institutsky lane, 9, MIPT, Russia

*e-mail: asantip@yandex.ru

**e-mail: khorelena@lmail.com

Received: 06.03.2024

Revised: 06.03.2024

Accepted: 31.05.2024

Abstract. The paper examines the mathematical description of the economic behavior of a representative household in an imperfect consumer credit and deposit market. The model is formalized as an optimal control problem on a finite time horizon. A classification of the behavior of social strata, dependence on the parameters of the economic environment is obtained. The synthesis of optimal control over an infinite time horizon is constructed.

Keywords: Ramsey type model, imperfect consumer credit market, synthesis of optimal control, maximum principle, mathematical modeling.

О МУЛЬТИПЛИКАТИВНОМ СВОЙСТВЕ ОПРЕДЕЛЯЮЩИХ ПОЛИНОМОВ

© 2024 г. С. А. Абрамов^{1,*}¹ 119333 Москва, ул. Вавилова, 44, ВЦ ФИЦ ИУ РАН, Россия

*e-mail: sergeyabramov@mail.ru

Поступила в редакцию 03.03.2024 г.

Переработанный вариант 05.04.2024 г.

Принята к публикации 31.05.2024 г.

Корни определяющего полинома, построенного для данного линейного обыкновенного дифференциального оператора, дают информацию об особенностях решений соответствующего однородного дифференциального уравнения. Обсуждаются операторы и уравнения, коэффициенты которых являются формальными лорановыми рядами. Такого же вида рассматриваются и решения. В этих предположениях описывается структура определяющего полинома произведения дифференциальных операторов. Это структурное (мультипликативное) свойство сохраняется и в случае сходящихся рядов. Библ. 8.

Ключевые слова: линейные обыкновенные однородные дифференциальные уравнения, определяющий полином, формальные лорановы ряды, индуцированные рекуррентные операторы.

DOI: 10.31857/S0044466924090067, EDN: WKRNAUK

1. ВВЕДЕНИЕ

Решения линейного обыкновенного дифференциального уравнения, имеющие вид формальных лорановых рядов, могут строиться с помощью индуцированного рекуррентного оператора, предоставляющего соотношение для коэффициентов такого рода решений. Уточним, что *формальным лорановым рядом* над полем K называется выражение $s(x)$ (формальная сумма — вопросы сходимости для таких рядов не рассматриваются), имеющее вид

$$\sum_{n=-\infty}^{\infty} c(n)x^n, \quad (1)$$

где все $c(n)$, т.е. коэффициенты ряда принадлежат K , и, при этом, существует $m \in \mathbb{Z}$ такое, что $c(n) = 0$ при $n < m$; для такого m ряд (1) можно переписать в виде $\sum_{n=m}^{\infty} c(n)x^n$, или в виде $c(m)x^m + c(m+1)x^{m+1} + \dots$. Подчеркнем, что и при наличии такого m ряд (1) мы считаем двусторонним, равно как и последовательность его коэффициентов, — в такой двусторонней последовательности все элементы, имеющие, в частности, отрицательные достаточно большие по абсолютной величине значения индекса равны нулю.

В некоторых публикациях по компьютерной алгебре употребляется термин “формальный мероморфный ряд”, мы в этой статье используем термин “формальный лоранов ряд”, как в большинстве публикаций.

Для формальных лорановых рядов определены основные арифметические операции и дифференцирование — подробности даются в п. 2.1.

В этом контексте рассмотрение рекуррентных операторов возникает в связи с тем, что если имеется дифференциальный оператор

$$L = a_r(x)D^r + \dots + a_1(x)D + a_0(x) \quad (2)$$

(здесь и далее D обозначает оператор дифференцирования ряда по x) и в роли коэффициентов оператора L выступают формальные лорановы ряды над некоторым полем K , то тогда этому L может быть сопоставлен *индуцированный* рекуррентный оператор

$$L^{(1)} = b_t\sigma^t + b_{t-1}\sigma^{t-1} + \dots \in K[n][[(\sigma^{-1})]], \quad (3)$$

в котором

- σ — оператор сдвига, т.е. $\sigma c(n) = c(n+1)$ для произвольной последовательности $\{c(n)\}_{n \in \mathbb{Z}}$,
- коэффициенты b_i суть полиномы от n над K , имеющие степень не выше, чем r — максимальной степени вхождения D в L ,
- оператор σ входит в $L^{(1)}$ с конечным числом целых неотрицательных показателей степени и, вообще говоря, с бесконечным числом целых отрицательных показателей. Подробности даются в п. 2.3.

Таким образом, $L^{(1)}$ представляется в виде формального лоранова ряда по σ^{-1} . При этом, если применение L к ряду (1) дает двусторонний ряд $\sum_{n \in \mathbb{Z}} \tilde{c}(n)x^n$, то двусторонняя последовательность коэффициентов $\{\tilde{c}(n)\}_{n \in \mathbb{Z}}$ получается применением $L^{(1)}$ к двусторонней последовательности $\{c(n)\}_{n \in \mathbb{Z}}$. В частности, любому решению вида (1) уравнения $L(y) = 0$ соответствует решение $\{\tilde{c}(n)\}_{n \in \mathbb{Z}}$ уравнения $L^{(1)}(c) = 0$ и наоборот. Существенно, что решением уравнения $L(y) = 0$ служит формальный ряд, в то время как решением уравнения $L^{(1)}(c) = 0$ — последовательность коэффициентов.

Оператор $L^{(1)}$ может использоваться в вычислении коэффициентов решения-ряда. Разумеется, сам индуцированный оператор $L^{(1)}$ требует построения. Но, если он уже построен (с учетом того, что для предполагаемых вычислений может требоваться лишь некоторый “фрагмент” оператора $L^{(1)}$, а не весь этот оператор, — в вычислениях оказывается нужным лишь конечное число членов оператора $L^{(1)}$, имеющих отрицательные показатели степени при σ), то, как будет установлено ниже, определяющий полином (см. п. 2.2) и некоторые другие вспомогательные объекты оказываются тоже построенными, коль скоро упомянутый “фрагмент” был достаточно представительным — они, эти вспомогательные объекты, без особых затрат извлекаются из $L^{(1)}$. Еще до прямого вычислительного использования рекуррентного оператора сам его вид позволяет находить некоторые характеристики решений-рядов, что дополнительно открывает возможность доказательства некоторых свойств коэффициентов решений, а также различных величин и объектов, сопоставляемых как самому дифференциальному уравнению, так и его решениям. Так, в п. 3.2 доказываются некоторые мультипликативные свойства определяющих полиномов (теоремы 1, 2).

Нелишним будет подчеркнуть, что в статье не предлагается новых методов решения дифференциальных уравнений. Цель статьи иная — доказать наличие полезных нетривиальных свойств для такого важного инструмента исследования и решения дифференциальных уравнений, как определяющий полином и, соответственно, определяющее уравнение.

В статье используются стандартные обозначения $K[x]$ для кольца полиномов и $K((x))$ для поля или кольца формальных лорановых рядов от x над заданным полем (или кольцом) K , т.е. полиномов и рядов с коэффициентами, принадлежащими полю или кольцу K .

2. ПРЕДВАРИТЕЛЬНЫЕ СВЕДЕНИЯ

2.1. Формальные лорановы ряды

Для краткости там, где это не вызывает недоразумений, мы будем вместо “формальный лоранов ряд” писать “лоранов ряд” или просто “ряд”. Ряд, все коэффициенты которого равны нулю, обозначается символом 0. Выделяют также ряд, все коэффициенты $c(n)$ которого, кроме $c(0)$, равны нулю, при этом $c(0) = 1$; он обозначается символом 1. Ряд $-s(x) = \sum_{n=-\infty}^{\infty} (-c(n))x^n$ является противоположным ряду $s(x) = \sum_{n=-\infty}^{\infty} c(n)x^n$. Ряд $\tilde{s}(x)$ является обратным ряду $s(x)$, если $s(x)\tilde{s}(x) = 1$; для любого ненулевого ряда $s(x)$ существует обратный $s^{-1}(x)$.

Если $s(x) \neq 0$, то, как уже было сказано в разд. 1, существует наименьшее $m \in \mathbb{Z}$ такое, что коэффициент $c(m)$ ряда $s(x)$ отличен от нуля. Это m будем обозначать через $\text{val } s(x)$ и называть *валуацией* ряда $s(x)$ (в [2, гл. 15], используется термин “младшая степень”). Коэффициент $c(m)$ называется *младшим коэффициентом* ряда $s(x)$ и обозначается через $\text{tc } s(x)$. По определению считаем, что $\text{val } (0) = \infty$, $\text{tc}(0) = 0$. Легко проверяется, что $\text{val } s(x)t(x) = \text{val } s(x) + \text{val } t(x)$, $\text{tc } s(x)t(x) = \text{tc } s(x) \cdot \text{tc } t(x)$ и $\text{val } (s(x) + t(x)) \geq \min\{\text{val } s(x), \text{val } t(x)\}$ для любых $s(x), t(x) \in K((x))$. Добавим, что $\text{val } s^{-1}(x) = -\text{val } s(x)$, $\text{tc } s^{-1}(x) = (\text{tc } s(x))^{-1}$ для любого $s(x) \in K((x)) \setminus \{0\}$.

При определенных таким образом операциях формальные лорановы ряды над полем K сами образуют поле $K((x))$. Это и другие свойства таких рядов см. в [2, гл. 15], [3, гл. 1, §1], [4, гл. 1].

Производная ряда $s(x) = \sum_{n=-\infty}^{\infty} c(n)x^n$ определяется как $Ds(x) = s'(x) = \sum_{n=-\infty}^{\infty} d(n)x^n$, где $d(n) = (n+1)c(n+1)$ для всех n , из чего следует, что коэффициент $d(-1)$ производной всегда равен нулю.

Каждый из рассматриваемых рядов имеет конечную валуацию, это позволяет применять к ряду дифференциальный оператор. Такое применение дает нулевой ряд, если и только если применение индуцированного оператора к последовательности коэффициентов данного ряда дает нулевую последовательность.

Далее мы используем обозначение $O(x^m)$, где $m \in \mathbb{Z}$, для не указываемого явно формального ряда, валюация которого больше или равна m .

2.2. Валюационный блок, приращение, определяющий полином

Определяющий полином, сопоставляемый уравнению $L(y) = 0$, где L имеет вид (2), содержит среди своих корней все валюации лорановых решений исходного уравнения. Такой полином может иметь и “лишние” корни. Известны алгоритмы проверки существования решения с заданной валюацией v и построения таких решений в случае существования. Лишние корни могут быть отбракованы этими алгоритмами.

Определение 1. Валюационным блоком оператора L будем называть любое множество целых чисел, содержащее валюации всех ненулевых формальных лорановых решений оператора L . (Не исключается, что валюационный блок содержит и какие-то целые числа, не являющиеся валюациями рассматриваемых решений.)

Определяющий полином для данного L может строиться по-разному. Обычно в таком построении участвует целое число, которое ниже будет именоваться приращением:

Определение 2. С оператором (2) свяжем *приращение*

$$\omega_L = \min_{0 \leq j \leq r} (\text{val } a_j - j). \quad (4)$$

Приравнивая нулю младший коэффициент в выражении для ряда $L(s(x))$, получаем алгебраическое определяющее уравнение $I_L(n) = 0$ (см. [1, §8]), левая часть которого дает определяющий полином:

$$I_L(n) = \sum_{\substack{0 \leq j \leq r \\ \text{val } a_j - j = \omega_L}} \text{tc}(a_j) n^j, \quad (5)$$

$$n^j = n(n-1) \dots (n-j+1).$$

Имеем $\text{val } L(s) \geq \omega_L + \text{val } s(x)$. Непосредственно проверяется, что если $\text{val } s(x) = n$, то

$$L(s(x)) = (\text{tc } s(x)) I_L(n) x^{n+\omega_L} + O(x^{n+\omega_L+1}),$$

т.е. ряд $L(s(x))$ имеет при $x^{n+\omega_L}$ коэффициент $\text{tc } s(x) I_L(n)$. Таким образом, для ненулевого $s(x)$ равенство $I_L(n) = 0$ является необходимым условием равенства $L(s(x)) = 0$.

Сказанное иллюстрируется проостым примером.

Пример 1. Пусть $L = D$. В соответствии с (4) имеем

$$\omega_L = -1, \quad I_L(n) = n. \quad (6)$$

Таким образом, $L(s(x)) = 0$ влечет $I_L(\text{val } s(x)) = 0$, т.е. $\text{val } s(x) = 0$. Тот вывод, который здесь может быть сделан непосредственно из ранее сказанного таков: если уравнение $D(y) = 0$ имеет ненулевое решение $s(x) \in K((x))$, то $s(x) = C + O(x)$, где C является ненулевой константой; решение $s(x) \in K((x))$ не может содержать x в отрицательной степени. Ни на что большее, в частности, на то, что $O(x)$ для $s(x)$ есть просто 0, равенства (6) не указывают.

2.3. Индуцированные рекуррентные операторы

Пусть K — поле характеристики 0, например, какое-то числовое поле, и σ — оператор сдвига по n : $\sigma(n) = n + 1$.

Предложение 1 (см. [5], [6]). (i) *Соответствие*

$$x \rightarrow \sigma^{-1}, \quad x^{-1} \rightarrow \sigma, \quad D \rightarrow (n+1)\sigma \quad (7)$$

задает изоморфизм $\mathcal{F}_D^\sigma$ кольца $K((x))[D]$ на кольцо $K[n][(\sigma^{-1})]$. Обратный изоморфизм задается так:

$$n \rightarrow xD, \quad \sigma \rightarrow x^{-1}, \quad \sigma^{-1} \rightarrow x. \quad (8)$$

(ii) Пусть $L \in K((x))[D]$, $R = \mathcal{F}_D^\sigma(L)$. Пусть $s(x) = \sum_n c(n)x^n$, $\bar{s}(x) = \sum_n \bar{c}(n)x^n$, $s(x), \bar{s}(x) \in K((x))$ так, что $L(s(x)) = \bar{s}(x)$. Тогда $R(c) = \bar{c}$, где $c, \bar{c}$ являются последовательностями коэффициентов рядов $s(x), \bar{s}(x)$.

Пример 2. Пусть $L = xD - (x - 1)$, тогда

$$L^{\textcircled{1}} = \sigma^{-1}(n+1)\sigma - (\sigma^{-1} - 1) = n+1 - \sigma^{-1}.$$

Пусть $y(x) = s(x)$, где $s(x)$ — ряд (1). Тогда $L(y) = 0$, если и только если $L^{\textcircled{1}}(c) = 0$, т.е. выполнение равенства $(n+1)c(n) - c(n-1) = 0$, $n \in \mathbb{Z}$, необходимо и достаточно для того, чтобы ряд (1) удовлетворял уравнению $L(y) = 0$. Нетрудно видеть, что $c(n) = 0$ при $n \leq -2$, а $c(-1)$ может быть выбрано произвольно, и $c(n) = \frac{c(-1)}{(n+1)!}$ при $n \geq -1$.

3. РЕКУРРЕНТНОЕ СООТНОШЕНИЕ ДЛЯ КОЭФФИЦИЕНТОВ РЕШЕНИЙ И ОПРЕДЕЛЯЮЩИЙ ПОЛИНОМ

3.1. Определяющий полином и приращение в индуцированном операторе

Предложение 2. Пусть $L^{\textcircled{1}}$ имеет вид (3). Тогда

- (i) $\omega_L = -t$,
- (ii) $I_L(n) = b_t(n-t) = b_t(n + \omega_L)$.

Доказательство. Обозначим через j_0 наибольшее целое такое, что $0 \leq j_0 \leq r$ и соответствующее ему значение $j_0 - \text{val } a_{j_0}$ максимально. Максимальная степень, в которой σ появляется в $L^{\textcircled{1}}$, есть $t = j_0 - \text{val } a_{j_0}$. Старшим по степени σ слагаемым в $L^{\textcircled{1}}$ служит

$$\begin{aligned} b_t(n)\sigma^t &= \sum_{\substack{0 \leq j \leq r \\ \text{val } a_j - j = -t}} \text{tc}(a_j) \sigma^{-\text{val } a_j} \underbrace{(n+1)\sigma \cdots (n+1)\sigma}_j = \\ &= \sum_{\substack{0 \leq j \leq r \\ \text{val } a_j - j = -t}} \text{tc}(a_j) \sigma^{-\text{val } a_j} (n+1) \cdots (n+j) \sigma^j = \\ &= \sum_{\substack{0 \leq j \leq r \\ \text{val } a_j - j = -t}} \text{tc}(a_j) \sigma^{j - \text{val } a_j} (n-j+1) \cdots n = \\ &= \sigma^t \left(\sum_{\substack{0 \leq j \leq r \\ \text{val } a_j - j = -t}} \text{tc}(a_j) n^{\underline{j}} \right). \end{aligned}$$

Коэффициент при σ^t будет полиномом

$$\sum_{\substack{0 \leq j \leq r \\ \text{val } a_j - j = \text{val } a_{j_0} - j_0}} \text{tc}(a_j) (n+t)^{\underline{j}} \quad (9)$$

от n . Этот полином имеет степень j_0 и, как следствие, является ненулевым. Очевидно, что максимальность $j - \text{val } a_j$ для $0 \leq j \leq r$ влечет минимальность $\text{val } a_j - j$, что соответствует определению (4) значения ω_L . Имеем (i).

Сравнение выписанного коэффициента (9) с (4) показывает, что справедливо (ii).

Пример 3. Вновь рассмотрим случай $L = D$. Имеем (6). Здесь $t = 1$, $b_t = n+1$. Равенства (i), (ii), очевидно, выполняются.

Из предложения 2 (ii) следует, что если определяющий полином имеет целые корни, то исходное уравнение $L(y) = 0$ имеет решение в виде ряда $s(x)$, валюация которого равна наибольшему из целых корней полинома $I(n)$. Таким образом, справедливо следующее

Предложение 3. Если определяющий полином (5) обладает целыми корнями, то множество решений исходного дифференциального уравнения, имеющих вид ненулевых формальных рядов, не пусто.

3.2. Мультипликативность определяющего полинома

Сформулируем и докажем основной результат статьи.

Теорема 1. Пусть L_1, L_2 — дифференциальные операторы вида (2). Тогда

- (i) $\omega_{L_1 L_2} = \omega_{L_1} + \omega_{L_2}$;
- (ii) $I_{L_1 L_2} = I_{L_1}(n + \omega_{L_2})I_{L_2}(n)$.

Доказательство. Согласно предложению 1(i) мы имеем

$$(L_1 L_2)^{\textcircled{1}} = L_1^{\textcircled{1}} L_2^{\textcircled{1}}.$$

Раскрываем произведение из правой части последнего равенства:

$$\begin{aligned} & (\sigma^{-\omega_{L_1}} I_{L_1}(n) + b_{t_1-1} \sigma^{-\omega_{L_1}-1} + \dots)(\sigma^{-\omega_{L_2}} I_{L_2}(n) + b_{t_2-1} \sigma^{\omega_{L_2}-1} + \dots) = \\ & = \sigma^{-\omega_{L_1} - \omega_{L_2}} I_{L_1}(n + \omega_{L_2}) I_{L_2}(n) + \dots \end{aligned}$$

Отсюда получаются (i), (ii).

Из приведенного доказательства теоремы 1 следует, что определяющий полином обладает мультипликативным свойством, указанным в п. (ii) этой теоремы.

Пример 4. Вернемся к примеру 1 и рассмотрим оператор $L^k = \underbrace{D \cdots D}_k$. По теореме 1 имеем

$$\omega_{L^k} = -k, \quad I_{L^k}(n) = n^k. \quad (10)$$

Теорема 1 (ii) допускает

Следствие. Пусть $L_1, L_2 \in k((x))[D]$ и $L = L_1 L_2$. Пусть $\{\alpha_1, \dots, \alpha_{N_1}\}, \{\beta_1, \dots, \beta_{N_2}\}$ суть множества корней определяющих полиномов I_{L_1}, I_{L_2} в некотором расширении $\tilde{K}$ поля K . Тогда множество корней полинома I_L в $\tilde{K}$ есть

$$\{\alpha_1 - \omega_{L_2}, \dots, \alpha_{N_1} - \omega_{L_2}, \beta_1, \dots, \beta_{N_2}\}.$$

3.3. Валюационные блоки

Теорема 2. Пусть L_1, L_2 — операторы, V_1, V_2 — валюационные блоки (см. определение 1) этих операторов. Пусть $L = L_1 L_2$ и пусть $\tilde{V}_1$ получено из V_1 вычитанием ω_{L_2} из каждого его элемента. Тогда $V = \tilde{V}_1 \cup V_2$ является валюационным блоком оператора L .

Доказательство. Валюации формальных лорановых решений операторов L_1, L_2 являются корнями полиномов I_{L_1}, I_{L_2} . Валюации всех решений оператора L являются корнями I_L . По теореме 1 (ii) валюации всех таких решений содержатся в V .

3.4. Заключительные пример и замечание

Пример 5. Для

$$L = (x^2 + O(x^3))D^3 - 1, \quad (11)$$

в соответствии с (4), (5) имеем:

$$\omega_L = -1, \quad I_L(n) = n^3 = n(n-1)(n-2). \quad (12)$$

Пользуясь теоремой 1, можно определить $\omega_{L^2}, I_{L^2}(n)$, не прибегая к явному виду оператора L^2 :

$$\omega_{L^2} = (-1) + (-1) = -2, \quad (13)$$

$$I_{L^2}(n) = ((n-1)(n-2)(n-3))n(n-1)(n-2) = n(n-1)^2(n-2)^2(n-3). \quad (14)$$

С помощью системы Maple (см. [7]) А.А. Рябенко нашла выражение для L^2 :

$$(x^4 + O(x^5))D^6 + (6x^3 + O(x^4))D^5 + (6x^2 + O(x^3))D^4 + O(x^2)D^3 + 1, \quad (15)$$

из которого можно также получить (13), (14) — слагаемые вида $O(\dots)$ в (15) не оказывают влияния на приращение и определяющий полином, сопоставляемые оператору L^2 .

Можно проверить правильность (12), рассматривая некоторый конкретный вариант $O(x^3)$, т.е. вариант “хвоста” ряда, входящего в (11).

Например, заменим $O(x^3)$ нулевым рядом. Таким образом, пусть $L = x^2 D^3 - 1$. Для оператора

$$L^2 = x^4 D^6 + 6x^3 D^5 + 6x^2 D^4 - 2x^2 D^3 + 1$$

получаем, исходя из (4) и (5):

$$\omega_{L^2} = -2,$$

и

$$I_{L^2}(n) = n^6 + 6n^5 + 6n^4 = n(n-1)^2(n-2)^2(n-3).$$

То же самое получается и по теореме 1 (см. (13), (14)).

Замечание 1. Доказанное мультипликативное свойство определяющих полиномов (теорема 1) сохраняется при переходе от формальных рядов к сходящимся в силу того, что правила сложения, умножения и дифференцирования одинаковы для обоих видов рядов.

Автор выражает признательность А.А.Рябенко за помощь в работе над статьей.

СПИСОК ЛИТЕРАТУРЫ

1. Коддингтон Э.А., Левинсон Н. Теория обыкновенных дифференциальных уравнений. М.: Изд-во иностр. лит., 1958.
2. Туганбаев А.А. Теория колец. Арифметические модули и кольца. М.: МЦНМО, 2009.
3. Картан А. Элементарная теория аналитических функций одного и нескольких комплексных переменных. М.: Изд-во иностр. лит., 1963.
4. Henrici P. Applied and computational complex analysis. Vol. 1. John Wiley & Sons, 1974.
5. Abramov S. EG—eliminations // J. of Difference Equations and Applications. 1999. V. 5. P. 393—433.
6. Abramov S., Petkovsek M., Ryabenko A. Special formal series solutions of linear operator equations // Discrete Math. 2000. V. 210. P. 3—25.
7. Maple online help: <http://www.maplesoft.com/support/help/>

ON THE MULTIPLICATIVE PROPERTY OF DEFINING POLYNOMIALS

S. A. Abramov*

119333 Moscow. Vavilova str, 44, Dorodnitsyn Computing Center,
Federal Research Center Computer Science and Control, RAS, Russia

*e-mail: sergeyabramov@mail.ru

Received: 03.03.2024

Revised: 05.04.2024

Accepted: 31.05.2024

Abstract. The roots of the indicial polynomial constructed for a given linear ordinary differential operator provide information about the features of solutions of the corresponding homogeneous differential equation. Operators and equations whose coefficients are formal Laurent series are discussed. Solutions of the same kind are considered. These assumptions describe the structure of the indicial polynomial of the product of differential operators. This structural (multiplicative) property is preserved in the case of converging series.

Keywords: linear ordinary homogeneous differential equations, defining polynomial, formal Laurent series, induced recurrent operators.

О ГРАНИЧНЫХ СВОЙСТВАХ КОНФОРМНЫХ
ОТОБРАЖЕНИЙ¹⁾© 2024 г. А. П. Солдатов^{1,2,*}¹ 119333 Москва, Вавилова, 44, ВЦ ФИЦ ИУ РАН, Россия² 360002 Нальчик, Долинск, ул. Балкарова, 2, Институт прикладной математики и автоматизации КБНЦ РАН, Россия

*e-mail: soldatov48@gmail.com

Поступила в редакцию 24.04.2024 г.

Переработанный вариант 24.04.2024 г.

Принята к публикации 31.05.2024 г.

Описаны классы C^1 — гладких областей, граничный контур которых является ляпуновским вне любой окрестности некоторой его точки, таких, что производная конформного отображения на единичный круг непрерывна в этой точке. Описание дано в терминах некоторых пространств для единичного касательного вектора на граничном контуре. Как следствие получены соответствующие результаты для кусочно-гладких областей. Библ. 6.

Ключевые слова: конформное отображение, гладкий контур, граничные свойства, обобщенные гильбертовы пространства.

DOI: 10.31857/S0044466924090073, EDN: WKJHUX

1. ФОРМУЛИРОВКА ОСНОВНОГО РЕЗУЛЬТАТА

Пусть односвязная область D ограничена гладким контуром Γ , и ω — конформное отображение этой области на единичный круг $\mathbb{D}$. Обозначим через $e_\Gamma(t)$, $t \in \Gamma$, единичный касательный вектор в точке t , ориентированный в положительном направлении (оставляющем область слева). Этот вектор рассматриваем как непрерывную комплекснозначную функцию на Γ . По определению контур Γ принадлежит классу $C^{1,\mu}$, если функция $e_\Gamma \in C^\mu(\Gamma)$. Если точное значение μ несущественно, этот контур называют ляпуновским.

Классическая теорема Келлога (см. [1]) утверждает, что если $\Gamma \in C^{1,\mu}$, то функция $\ln \omega'(z) \in C^\mu(\overline{D})$. В случае просто C^1 -гладкого контура функция $\ln \omega'(z)$ не является, вообще говоря, непрерывной вплоть до границы и может при приближении к границе обращаться в бесконечность. Тем не менее, по известной теореме Линделёфа (см. [1]) мнимая часть $v(z) = \arg \omega'(z)$ этой функции непрерывна в замкнутой области $\overline{D}$ и на границе удовлетворяет соотношению

$$\arg \omega'(t) = \arg e_\Gamma[\omega(t)] - \arg e_\Gamma(t), \quad t \in \Gamma. \quad (1.1)$$

Аналогичным свойством обладает и обратное отображение $\chi = \omega^{-1} : \mathbb{D} \rightarrow D$.

Поведение самой функции $\ln \chi'$ на границе единичного круга можно уточнить. По теореме Смирнова (см. [1]) она принадлежит классу Харди $H^1(\mathbb{D})$. В действительности

$$u + iv = \ln \chi' \in \cap_{p>1} H^p(\mathbb{D}). \quad (1.2)$$

В самом деле, функция $v \in C(\overline{\mathbb{D}})$, и, следовательно, по теореме Рисса для любых $p > 1$ и $0 < r < 1$ имеет место оценка

$$\int_0^{2\pi} |u(re^{i\theta})|^p d\theta \leq C \int_0^{2\pi} |v(re^{i\theta})|^p d\theta,$$

равномерная по r , что и доказывает (1.2).

Известно из [2] (см. гл. 3, с. 36), что функции $\phi \in H^p(\mathbb{D})$ допускают оценку

$$|\phi(\zeta)| \leq C |\phi|_{H^p} (1 - |\zeta|)^{-1/p}$$

¹⁾ Работа выполнена при финансовой поддержке Минобрнауки РФ в рамках реализации программы Московского центра фундаментальной и прикладной математики по соглашению № 075-15-2022-284 и поддержана грантом АР 19678182 Министерство науки и высшего образования Республики Казахстан.

с положительной постоянной C , зависящей только от p . Применительно к $\phi = \ln \chi'$ отсюда следует, что для любого $\varepsilon > 0$ выполнена оценка

$$|\ln \chi'(\zeta)| \leq C_\varepsilon (1 - |\zeta|)^{-\varepsilon}.$$

Другими словами, при $|\zeta| \rightarrow 1$ функция $\ln \chi'(\zeta)$ может обращаться в бесконечность логарифмического типа. Равенство

$$\ln \omega' \circ \chi = -\ln \chi' \quad (1.3)$$

показывает, что в некотором смысле подобным поведением обладает и функция $\ln \omega'(z)$ при приближении точки z к граничному контуру Γ .

Рассмотрим случай, когда контур Γ ляпуновский вне любой окрестности некоторой его фиксированной точки τ . Именно, пусть гладкий контур принадлежит классу $C_{\text{loc}}^{1,\mu}$ по отношению к этой точке в том смысле, что любая дуга $G \subseteq \Gamma$, не содержащая τ , принадлежит классу $C^{1,\mu}$. Этот факт указываем также записью $(\Gamma, \tau) \in C_{\text{loc}}^{1,\mu}$. Точно также под $C_{\text{loc}}^{\mu}(\Gamma, \tau)$ понимается класс функций, которые принадлежат $C^{\mu}(\Gamma)$ на любой дуге $G \subseteq \Gamma$, не содержащей τ . Аналогичный смысл имеет и класс $C_{\text{loc}}^{\mu}(\overline{D}, \tau)$ функций, принадлежащий $C^{\mu}(\overline{G})$ в любой подобласти $G \subseteq D$, граница которой не содержит τ . В этих обозначениях имеет место следующий локальный вариант теоремы Келлога.

Теорема 1.1. Если гладкий контур принадлежит классу $C_{\text{loc}}^{1,\mu}$ по отношению к некоторой своей точке τ , то $\ln \omega' \in C_{\text{loc}}^{\mu}(\overline{D}, \tau)$.

Доказательство в основном осуществляется по той же схеме, что и теорема Келлога в [1]. Заметим прежде всего, что условие $(\Gamma, \tau) \in C_{\text{loc}}^{1,\mu}$ в терминах единичного касательного вектора можем выразить в форме $e \in C_{\text{loc}}^{\mu}(\Gamma, \tau)$. Утверждается, что

$$\arg \omega' \in C_{\text{loc}}^{\mu}(\Gamma, \tau). \quad (1.4)$$

В самом деле, согласно теореме Келлога в соотношении (1.1) теоремы Линделефа в качестве $\mathbb{D}$ можно брать любую область с границей класса $C^{1,\mu}$. Пусть дуга $G \subseteq \Gamma$ не содержит τ , и область $\mathbb{D}$ этого типа выбрана так, что образ $\omega(G)$ является прямолинейным отрезком. Тогда соотношение (1.1) показывает, что $\arg \omega' \in C^{\mu}(G_0)$ на строго внутренней дуге $G_0 \subseteq G$. Тем самым соотношение (1.3) установлено.

Утверждение теоремы достаточно доказать для функции $\ln \chi'$, считая опять область $\mathbb{D}$ единичным кругом. В силу (1.2) на границе единичного круга угловые предельные значения функции χ' принадлежат L^p с любым $p > 1$. Поэтому в силу неравенства Гёльдера на дуге $G \subseteq \mathbb{T}$ с концами t_1, t_2 можем записать

$$|\chi(t_1) - \chi(t_2)| = \left| \int_G \chi'(t) d_1 t \right| \leq C |t_1 - t_2|^{1/p'},$$

где $d_1 t$ означает элемент длины дуги и $1/p' = 1 - 1/p$. Следовательно, функция $\chi \in C^v(\mathbb{T})$ для любого $0 < v < 1$. На основании (3), (4) отсюда следует, что граничное значение v^+ гармонической функции $v = \arg \chi'$ принадлежит классу $C^{\mu v}[\mathbb{T}, \omega(\tau)]$. Сама функция v восстанавливается интегралом Пуассона

$$v(\zeta) = \frac{1}{\pi} \int_{\mathbb{T}} \frac{1 - |\zeta|^2}{|t - \zeta|^2} v^+(t) d_1 t.$$

Равенство

$$\frac{1 - |\zeta|^2 + 2i \operatorname{Im}(\zeta \bar{t})}{|t - \zeta|^2} = \frac{t + \zeta}{t - \zeta} = \frac{1}{t - \zeta} - \frac{1}{2t}$$

показывает, что

$$v(z) = \operatorname{Re} \left[\frac{1}{\pi} \int_{\mathbb{T}} \frac{v^+(t) dt}{t - \zeta} \right].$$

Следовательно, с точностью до аддитивной мнимой постоянной функция $-i \ln \chi' = v - iu$ представляется интегралом типа Коши с плотностью $v^+ \in C(\mathbb{T}) \cap C_{\text{loc}}^{\mu v}[\mathbb{T}, \omega(\tau)]$. На основании классических граничных свойств интеграла типа Коши отсюда $\ln \chi' \in C_{\text{loc}}^{\mu v}[\overline{\mathbb{D}}, \omega(\tau)]$. В частности, на каждой дуге $G \subseteq \mathbb{T}$, не содержащей точки $\omega(\tau)$, функция χ удовлетворяет двустороннему условию Липшица

$$C^{-1} |t_1 - t_2| \leq |\chi(t_1) - \chi(t_2)| \leq C |t_1 - t_2|.$$

Поэтому в действительности из (1.4) следует, что $v^+ \in C_{\text{loc}}^{\mu}[\mathbb{T}, \omega(\tau)]$ и, значит, $v \in C_{\text{loc}}^{\mu}[\overline{\mathbb{D}}, \omega(\tau)]$, что завершает доказательство теоремы.

Из теоремы 1.1 с учетом предыдущих рассмотрений следует, что для гладкого контура $(\Gamma, \tau) \in C_{\text{loc}}^{1,\mu}$ функция $\ln \omega'$ допускает в точке τ особенность логарифмического характера. Возникает вопрос, при каких условиях эта

функция будет непрерывной в точке τ ? Конечно, если дополнительно $\Gamma \in C^{1,\varepsilon}$ с некоторым $\varepsilon > 0$, то на основании теорем 1.1 и Келлога функция $\ln \omega'$ также принадлежит классу $C^\varepsilon(\overline{D}) \cap C_{\text{loc}}^\mu(\mathbb{D}, \tau)$. Исследованию этого вопроса для контуров, не являющихся ляпуновскими, и посвящена настоящая статья.

Напомним, что $C^\mu(G)$, $0 < \mu < 1$, и $C^{0,1}(G)$ означают банаховы пространства функций, удовлетворяющие соответственно условиям Гёльдера с показателем μ и Липшица, которые банаховы относительно соответствующей нормы. В основе дальнейших построений лежит пространство $C_0^\mu(G, \tau)$ всех ограниченных непрерывных функций $\varphi(z)$, $z \in G$, $z \neq \tau$, которые после умножения на весовую функцию $|z - \tau|^\mu$ удовлетворяют условию Гёльдера с показателем μ . Относительно соответствующей нормы это пространство банахово. Нетрудно убедиться, что оно является банаховой алгеброй по умножению. Аналогичный смысл имеет и банахова алгебра $C_0^{0,1}(G, \tau)$. С возрастанием μ эти пространства монотонно убывают по вложению:

$$C_0^{0,1} \subseteq C_0^\nu \subseteq C_0^\mu, \quad \mu < \nu < 1. \quad (1.5)$$

Основные свойства пространства C_0^μ подробно описаны в [3]. В частности, отметим следующую лемму, представляющую практический пример элементов пространства C_0^μ .

Лемма 1.1. Пусть функция φ ограничена и непрерывно дифференцируема на дуге G с концом τ (вне этого конца), причем функция $|t - \tau|\varphi'(t)$ также ограничена. Тогда $\varphi \in C_0^{0,1}(G, \tau)$.

Аналогичное утверждение справедливо и в случае $G = \overline{D}$, где область D односвязна и ограничена кусочно-гладким контуром, содержащим точку τ , а под производной φ' функции φ понимается пара ее частных производных.

Эта лемма показывает, что если G допускает вне точки τ непрерывную ветвь аргумента $\arg(z - \tau)$, то эта ветвь принадлежит $C_0^\mu(G, \tau)$. В частности, в определении пространства C_0^μ можно вместо $|z - \tau|^\mu$ брать весовую функцию $(z - \tau)^\mu$, умножение на которую не выводит из класса аналитических функций. Из леммы 1.1 следует, что при приближении к точке τ функции $\varphi \in C_0^\mu(G, \tau)$ могут осциллировать, оставаясь ограниченными. Например, на отрезке $[0, 1]$ действительной оси функция $\varphi(t) = \sin(\ln t)$, $t > 0$, принадлежит классу $C_0^\mu([0, 1], 0)$.

В этой связи обозначим через $C_{(0)}^\mu(G, \tau)$ замкнутое подпространство функций $\varphi \in C_0^\mu$, непрерывных в точке τ . Введем в $C_{(0)}^\mu(G, \tau)$ подпространство $C_{(\lambda)}^\mu(G, \tau)$, $0 < \lambda < 1$, всех функций φ , для которых

$$\varphi_0(t) = [\varphi(t) - \varphi(\tau)]|t - \tau|^{-\lambda} \in C_0^\mu(G, \tau). \quad (1.6)$$

Относительно соответствующей нормы $|\varphi| = |\varphi(\tau)| + |\varphi_0|_{C_0^\mu}$ это пространство банахово. Заметим, что по определению пространства C_0^μ имеем вложение

$$C_{(\lambda)}^\mu(G, \tau) \subseteq C^\varepsilon(G), \quad \varepsilon = \min(\lambda, \mu), \quad (1.7)$$

которое при $\lambda = \mu$ переходит в равенство.

Введем далее в $C_{(0)}^\mu$ подпространство $C_{[\delta]}^\mu(G, \tau)$, $\delta > 1$, с помощью аналогичного (1.6) условия

$$\varphi_0(t) = [\varphi(t) - \varphi(\tau)](1 + |\ln |t - \tau||)^\delta \in C_0^\mu(G, \tau), \quad (1.8)$$

которое банахово относительно соответствующей нормы.

С помощью леммы 1.1 легко убедиться, что с возрастанием λ и δ введенные пространства монотонно убывают, причем имеют место вложения $C_{(\lambda)}^\mu \subseteq C_{[\delta]}^\mu$ для любых $\lambda < 1 < \delta$. В частности, можно ввести классы

$$C_{(+0)}^\mu = \bigcup_\lambda C_{(\lambda)}^\mu, \quad C_{[\infty]}^\mu = \bigcap_\delta C_{[\delta]}^\mu.$$

Отметим, что первый из этих классов был введен в [4], где он обозначался символом $C_{(0)}^\mu$, который здесь используется для других целей. Из определений следует, что

$$C_{(+0)}^\mu \subseteq C_{[\infty]}^\mu \subseteq C_{(0)}^\mu. \quad (1.9)$$

Согласно (1.7) гладкий контур Γ с касательным вектором $e \in C_{(+0)}^\mu(\Gamma, \tau)$ является ляпуновским. Ясно также, что функции $\varphi \in C_{[\infty]}^\mu(\Gamma, \tau)$ на гладком контуре удовлетворяют условию

$$\int_G \frac{|\varphi(t) - \varphi(\tau)|}{|t - \tau|} d_1 t < \infty, \quad (1.10)$$

где $d_1 t$ означает элемент длины дуги. В точке τ эти функции “почти” удовлетворяют условию Гёльдера, т.е. $\varphi(t) - \varphi(\tau) = O(1)(\ln |t - \tau|)^{-n}$ для любого натурального n . В этом смысле контур Γ с касательным вектором $e \in C_{[\infty]}^\mu(\Gamma, \tau)$ является ляпуновским.

Основная цель настоящей статьи — установить следующий результат.

Теорема 1.2. Пусть односвязная область D ограничена гладким контуром Γ с касательным вектором $e \in C_{[\infty]}^{\mu}(\Gamma, \tau)$ и ω осуществляет конформное отображение D на единичный круг $\mathbb{D}$.

Тогда функции $\ln \omega'(z)$ и $a(z) = [\omega(z) - \omega(\tau)]/(z - \tau)$ принадлежат классу $C_{[\infty]}^{\mu}(\overline{D}, \tau)$. Аналогичные утверждения справедливы и для обратного отображения $\chi = \omega^{-1}$.

Для ляпуновского контура Γ с касательным вектором $e \in C_{(+0)}^{\mu}(\Gamma, \tau)$ аналогичный результат по отношению к классу $C_{(+0)}^{\mu}(\overline{D}, \tau)$ был установлен в [4]. Доказательство сформулированной теоремы требует привлечения значительно более тонких средств и будет изложено в отдельном разделе в конце статьи. Это доказательство в равной степени будет проходить и для отмеченного выше случая $e \in C_{(+0)}^{\mu}(\Gamma, \tau)$.

2. ПРОСТРАНСТВА $C_{(0)}^{\mu}$, $C_{(\lambda)}^{\mu}$ И $C_{[\delta]}^{\mu}$

Пусть для простоты $\tau = 0$. Как показано в [3], банахово пространство $C_0^{\mu}(G, 0)$ можно определить эквивалентной нормой

$$|\varphi| = \sup_{t \neq 0} |\varphi(t)| + \sup_{1/2 \leq |t_1|/|t_2| \leq 2} |t_1|^{\mu} \frac{|\varphi(t_1) - \varphi(t_2)|}{|t_1 - t_2|^{\mu}}. \quad (2.1)$$

В частности, C^{μ} -норма

$$\alpha(r) = |\varphi|_{C^{\mu}(G \cap \{|t-\tau| \geq r\})}$$

ведет себя как $O(r^{-\mu})$ при $r \rightarrow 0$.

Выражение под знаком $\sup$ во втором слагаемом (2.1) однородно относительно растяжения $t \rightarrow qt$ с $q > 0$, поэтому эта норма также эквивалентна норме

$$|\varphi| = \sup_{k=0, \pm 1, \dots} |\varphi_k|_{C^{\mu}(G_k)}, \quad (2.2)$$

где положено $\varphi_k(t) = \varphi(2^{-k}t)$, $t \in G_k$, $G_k = \{t, 2^{-k}t \in G\}$.

С помощью нормы (2.1) легко убедиться, что класс C_0^{μ} инвариантен относительно соответствующих операций суперпозиции.

Лемма 2.1. Пусть $\varphi \in C_0^{\mu}(G, \tau)$, функция f удовлетворяет условию Липшица на множестве $\varphi(G)$ и отображение α , осуществляющее гомеоморфизм множества G_1 на G , удовлетворяет двустороннему условию Липшица

$$C^{-1}|t_1 - t_2| \leq |\alpha(t_1) - \alpha(t_2)| \leq C|t_1 - t_2|.$$

Тогда $f \circ \varphi \in C_0^{\mu}(G, \tau)$ и $\varphi \circ \alpha \in C_0^{\mu}(G_1, \tau_1)$, где $\alpha(\tau_1) = \tau$.

Конечно, эта лемма сохраняет свою силу и по отношению к замкнутому подпространству $C_0^{\mu}(G, \tau) \subseteq C_0^{\mu}(G, \tau)$, элементы которого непрерывны в точке τ .

Нетрудно показать, что гладкая дуга G с концом τ в окрестности этой точки является радиальной, т.е. любая окружность с центром в τ пересекает ее ровно в одной точке (и при том некасательно). Эквивалентное определение — дуга радиальна по отношению к своему концу τ , если она допускает параметризацию вида

$$\gamma(s) = \tau + se^{i\theta(s)}, \quad 0 < s \leq \varrho, \quad (2.3)$$

где вещественная функция $\theta \in C[0, \varrho]$ непрерывно дифференцируема на $(0, \varrho]$ и $s\theta'(s) \rightarrow 0$ при $s \rightarrow 0$.

С помощью леммы 2.1 гладкие дуги с касательным вектором $e \in C_{(0)}^{\mu}$ легко описать с помощью параметризаций.

Лемма 2.2. Касательный вектор гладкой дуги G с концом τ принадлежит классу $C_{(0)}^{\mu}(G, \tau)$ тогда и только тогда, когда она допускает гладкую параметризацию $\gamma: [0, \varrho] \rightarrow G$, $\gamma(0) = \tau$, с производной $\gamma' \in C_{(0)}^{\mu}([0, \varrho]; 0)$.

В случае радиальной дуги применительно к радиальной параметризации (2.3) это условие равносильно тому, что функции

$$\theta(s), \quad s\theta'(s) \in C_{(0)}^{\mu}([0, \varrho]; 0). \quad (2.4)$$

Доказательство. Не ограничивая общности можно считать $\tau = 0$. Пусть $e \in C_{(0)}^{\mu}(G, 0)$, и γ_0 есть естественная параметризация дуги G , т.е. $|\gamma_0'(s)| \equiv 1$ и $\gamma_0(0) = 0$. Тогда $e[\gamma_0(s)] = \gamma_0'(s)$, и на основании леммы 2.1 производная $\gamma_0' \in C_{(0)}^{\mu}([0, \varrho]; 0)$.

Обратно, пусть аналогичное условие выполнено для некоторой параметризации γ . Рассмотрим функцию $\beta(t)$, обратную к монотонно возрастающей функции

$$\alpha(t) = \int_0^t |\gamma'(s)| ds.$$

Очевидно, $\gamma_0 = \gamma \circ \beta$ является естественной параметризацией дуги G . На основании леммы 2.1 функции $|\gamma'|$, $1/|\gamma'|$ принадлежат классу $C_{(0)}^\mu$, а вместе с ними этому классу принадлежат также производная функции β и γ_0 . Но тогда в силу леммы 2.1 из равенства $e \circ \gamma_0 = \gamma'_0$ следует, что и $e \in C_{(0)}^\mu(G, 0)$.

Пусть далее дуга G радиальна и для ее радиальной параметризации (2.3) условие (2.4) выполнено. Тогда в силу леммы 2.1 функция $e^{i\theta(s)}$, а вместе с ней и производная

$$\gamma'(s) = e^{i\theta(s)}[1 + is\theta'(s)]$$

принадлежат $C_{(0)}^\mu([0, \varrho], 0)$.

Обратно, пусть дуга допускает параметризацию γ_1 с производной $\gamma'_1(t) \in C_{(0)}^\mu$. Утверждается, что производная функции $\alpha(t) = |\gamma_1(t)|$ также принадлежит классу $C_{(0)}^\mu$. В самом деле, в силу леммы 1.1 функции $f_1(t) = \gamma_1(t)/t$, а вместе с ней и функция

$$\frac{f_1(t)}{|f_1(t)|} = \frac{\gamma_1(t)}{|\gamma_1(t)|}$$

принадлежат $C_{(0)}^\mu$. Но тогда на основании леммы 2.1 и

$$\alpha'(t) = \operatorname{Re} \left[\gamma'_1(t) \frac{\gamma_1(t)}{|\gamma_1(t)|} \right] \in C_{(0)}^\mu.$$

Как и в случае первой части леммы, рассмотрим функцию $\beta(s)$, обратную к $\alpha(t)$, производная которой также принадлежит $C_{(0)}^\mu$. Тогда $\gamma = \gamma_1 \circ \beta$ совпадает с радиальной параметризацией (2.3), так что и ее производная $\gamma' \in C_{(0)}^\mu$. Как и выше, отсюда $f(t) = \gamma(s)/s \in C_{(0)}^\mu$ и на основании леммы 2.1 этим свойством обладает и функция $\theta(s) = \operatorname{Im} [\ln f(s)]$. Поскольку $sf'(s) = \gamma'(s) - f(s)$, условие (2.4) для γ выполнено, что завершает доказательство леммы.

Для пространства $C_0^\mu(G, \tau)$ связность компакта G несущественна. Как показано в [4], в предположении связности пространство $C_{(\lambda)}^\mu(G, \tau)$ можно аналогично (2.1) охарактеризовать с помощью нормы

$$|\varphi| = \sup_{t \neq 0} |\varphi(t)| + \sup_{1/2 \leq |t_1|/|t_2| \leq 2} |t_1|^{\mu-\lambda} \frac{|\varphi(t_1) - \varphi(t_2)|}{|t_1 - t_2|^\mu}. \quad (2.5)$$

Заметим, что при $\lambda = \mu$ она эквивалентна C^μ -норме, что согласуется с (1.7). Следующая лемма показывает, что аналогичным свойством обладает и пространство $C_{[\delta]}^\mu$, $\delta > 1$.

Лемма 2.3. Пусть компакт G связан, и $\tau = 0 \in G$. Тогда равенство

$$|\varphi| = |\varphi(0)| + \sup_{1/2 \leq |t_1|/|t_2| \leq 2} |t_1|^\mu (1 + |\ln |t_1||)^\delta \frac{|\varphi(t_1) - \varphi(t_2)|}{|t_1 - t_2|^\mu} \quad (2.6)$$

определяет эквивалентную норму в $C_{[\delta]}^\mu$, $\delta > 1$.

Доказательство. Положим для краткости $\sigma(t) = (1 + |\ln |t||)^\delta$ и заметим, что

$$|\sigma(t_1) - \sigma(t_2)| \leq C |\ln(|t_1|/|t_2|)|, \quad (2.7)$$

где постоянная $C > 0$ зависит только от G и δ .

Обозначим через $|\varphi|_0$ и $\{\varphi\}_\mu$ соответствующие слагаемые нормы (2.1), а $[\varphi]_\mu$ отвечает второму слагаемому нормы (2.6). Пусть $\varphi \in C_{[\delta]}^\mu$, так что по определению $\varphi_0(t) = \sigma(t)[\varphi(t) - \varphi(0)] \in C_0^\mu$. Тогда в этих обозначениях

$$\sigma(t_1)|t_1|^\mu \frac{|\varphi(t_1) - \varphi(t_2)|}{|t_1 - t_2|^\mu} \leq \{\varphi_0\}_\mu + |\varphi_0|_0 |t_1|^\mu \frac{|\sigma(t_1) - \sigma(t_2)|}{|t_1 - t_2|^\mu}.$$

С учетом (2.7) можем записать

$$|t_1|^\mu \frac{|\sigma(t_1) - \sigma(t_2)|}{|t_1 - t_2|^\mu} \leq C \frac{|\ln |u||}{|1 - u|^\mu}$$

для $u = t_2/t_1$, так что правая часть ограничена при $1/2 \leq |u| \leq 2$. Следовательно, полуорма $[\varphi]_\mu$ оценивается через норму пространства $C_{[\delta]}^\mu$.

Обратно, пусть эта полуорма конечна. Утверждается, что тогда функция φ имеет предел $\varphi(0)$ при $t \rightarrow 0$ и допускает оценку

$$|\varphi(t) - \varphi(0)| \leq C_0 \sigma^{-1}(t) [\varphi]_\mu, \quad (2.8)$$

где постоянная $C_0 > 0$ зависит только от G и δ .

В самом деле, пусть $G \subseteq \{|t| \leq \varrho\}$ и $G \cap \{|t| = \varrho\} \neq \emptyset$. В силу связности G можно выбрать последовательность точек $t_k \in G \cap \{|t| = 2^{-k}\varrho\}$, $k = 0, 1, \dots$. Тогда

$$\sigma(t_k) |\varphi(t_{k+1}) - \varphi(t_k)| \leq \langle \varphi \rangle_\mu |1 - u_k|^\mu,$$

где точка $u_k = t_{k+1}/t_k$ лежит на окружности $|u| = 1/2$. Поскольку последовательность $\sigma(t_k) = (1 + | -k \ln 2 + \ln \varrho |)^\delta$, $k \geq 1$, эквивалентна k^δ , на основании предыдущей оценки ряд $\sum |\varphi(t_{k+1}) - \varphi(t_k)|$ сходится, и существует предел $c = \lim \varphi(t_k)$ при $k \rightarrow \infty$. Кроме того, имеет место оценка

$$|\varphi(t_k) - c| \leq \sum_{s \geq k} |\varphi(t_{s+1}) - \varphi(t_s)| \leq C_1 \sigma^{-1}(t_k) [\varphi]_\mu.$$

Рассматривая точки t в слое $G \cap \{\varrho 2^{-k-1} \leq |t| \leq \varrho 2^{-k}\}$, в свою очередь, отсюда убеждаемся, что функция φ имеет предел $\varphi(0) = c$ при $t \rightarrow 0$ и допускает оценку (2.8).

Полагая $\varphi_0(t) = \sigma(t)[\varphi(t) - \varphi(0)]$, имеем

$$|t_1|^\mu \frac{|\varphi_0(t_1) - \varphi_0(t_2)|}{|t_1 - t_2|^\mu} \leq \langle \varphi \rangle_\mu + |t_1|^\mu |\varphi_0(t_2) - \varphi_0(0)| \frac{|\sigma(t_1) - \sigma(t_2)|}{|t_1 - t_2|^\mu}.$$

Как и выше с учетом (2.7), (2.8) отсюда убеждаемся, что норма пространства $C_{[\delta]}^\mu$ оценивается через норму (2.6).

С помощью норм (2.5) и (2.6) утверждения лемм 2.1 и 2.2 сохраняют свою силу (с теми же рассуждениями в доказательстве) на пространствах $C_{(\lambda)}^\mu$ и $C_{[\delta]}^\mu$.

Лемма 2.4. В утверждениях лемм 2.1 и 2.2 символ $C_{(0)}^\mu$ можно заменить на $C_{(\lambda)}^\mu$ и $C_{[\delta]}^\mu$.

Отметим еще следующее предложение.

Лемма 2.5. Пусть касательный вектор e дуги G с концом τ принадлежит $C_{[\infty]}^\mu(G, \tau)$.

Тогда, если функция φ дифференцируема на этой дуге и ее производная φ' принадлежит классу $C_{[\infty]}^\mu(G, \tau)$, то этому классу принадлежит и функция $a(t) = [\varphi(t) - \varphi(\tau)]/(t - \tau) \in C_{[\infty]}^\mu(G, \tau)$.

Доказательство. В силу леммы 2.4 утверждение достаточно установить для случая, когда $G = [0, 1]$ и $\tau = 0$. По условию

$$\varphi'(s) = \varphi'(0) + (1 - \ln s)^{-\delta} \varphi_0(s)$$

с некоторыми $\delta > 1$ и $\varphi_0 \in C_0^\mu([0, 1], 0)$. Запишем

$$a(s) = \varphi'(0) + \int_0^1 [1 - \ln(us)]^{-\delta} \varphi_0'(su) du = \varphi(0) + (1 - \ln s)^{-\delta} \psi_0(s),$$

где положено

$$\psi_0(s) = \int_0^1 f(s, u) \varphi_0(su) du, \quad f(s, u) = \left[\frac{1 - \ln(us)}{1 - \ln s} \right]^\delta.$$

На основании леммы 1.1 функция $f(s, u)$ по переменной s принадлежит $C_0^\mu([0, 1], 0)$ равномерно по u . С помощью нормы (2.1) легко убедиться, что аналогичным свойством обладает и $\varphi_0(su)$. Поэтому $\psi_0 \in C_0^\mu([0, 1], 0)$, что завершает доказательство.

3. ИНТЕГРАЛ ТИПА КОШИ В КЛАССЕ $C_{[\infty]}^\mu$

Пусть область D ограничена гладким контуром Γ , ориентированным положительно по отношению к D (т.е. по направлению, оставляющему область слева). Рассмотрим интеграл типа Коши

$$(I\varphi)(z) = \frac{1}{2\pi i} \int_\Gamma \frac{\varphi(t) dt}{t - z}, \quad z \in D, \quad (3.1)$$

и отвечающий ему сингулярный интеграл Коши

$$(S\varphi)(t_0) = \frac{1}{\pi i} \int_{\Gamma} \frac{\varphi(t)dt}{t - t_0}, \quad t_0 \in \Gamma. \quad (3.2)$$

Известно (см. [5]), что оператор I ограничен $C^u(\Gamma) \rightarrow C^u(\overline{D})$, и для его граничных значений $(I\varphi)^+(t_0)$ справедлива формула Сохоцкого–Племеля

$$2(I\varphi)^+(t_0) = \varphi(t_0) + (S\varphi)(t_0), \quad t_0 \in \Gamma. \quad (3.3)$$

В частности, оператор S ограничен в пространстве $C^u(\Gamma)$.

Как было установлено в [3], [4], этот результат справедлив и по отношению к пространству $C_{(\lambda)}^u$. Покажем, что он допускает некоторый аналог для пространства $C_{[\delta]}^u$. Заметим, что в силу (1.10) для плотности φ из этого пространства с $\delta > 1$ сингулярный интеграл существует и в точке $t_0 = \tau$.

Теорема 3.1. Пусть область D ограничена кусочно-гладким контуром Γ с единственной угловой точкой $\tau \in \Gamma$, причем ее угол в этой точке положителен. Тогда при $\delta > 2$ оператор I ограничен, $C_{[\delta]}^u(\Gamma, \tau) \rightarrow C_{[\delta-1]}^u(\overline{D}, \tau)$, и для его граничных значений справедлива формула Сохоцкого–Племеля (2.9).

Доказательство. Пусть для простоты $\tau = 0$ и $\varphi = I\varphi$, $\varphi \in C_{[\delta]}^u(\Gamma, 0)$. В силу формулы Коши, не ограничивая общности, можно считать $\varphi(0) = 0$, так что, заменяя φ на $\varphi(z) - \varphi(0)$, от (3.1) можем перейти к интегралу

$$\varphi(z) = \frac{1}{2\pi i} \int_{\Gamma} \frac{z}{t} \frac{\varphi(t)dt}{t - z}, \quad z \in D.$$

Нужно доказать, что оператор, определяемый этим равенством, ограничен из пространства $\{\varphi \in C_{[\delta]}^u(\Gamma, \tau), \varphi(0) = 0\}$ в пространство $\{\varphi \in C_{[\delta-1]}^u(\overline{D}, \tau), \varphi(0) = 0\}$. В соответствии с определением этих пространств дело сводится к ограниченности оператора

$$\varphi_0(z) = \frac{1}{2\pi i} \int_{\Gamma} \frac{z(1 + |\ln|z||)^{\delta-1}}{t(1 + |\ln|t||)^{\delta}} \frac{\varphi_0(t)dt}{t - z}, \quad z \in D, \quad (3.4)$$

из $C_0^u(\Gamma, 0)$ в $C_0^u(\overline{D}, 0)$.

Достаточно ограничиться случаем, когда Γ является радиальной дугой по отношению к своему концу $\tau = 0$, заданной параметризацией (2.3), а функция φ обращается в нуль в окрестности второго конца дуги Γ , отличного от $\tau = 0$. При этом роль D играет криволинейный сектор круга $|z| < \rho$ раствора меньше 2π , одной из боковых сторон которой служит Γ .

Для целого $k = 0, 1, \dots$ рассмотрим в области $D_k = (2^k D) \cap \{1/2 < |z| < 2\}$ функцию $\varphi_0(2^{-k}z)$, где здесь и ниже $2^k D = \{z, 2^{-k}z \in D\}$. Эту функцию можем представить в виде произведения двух сомножителей:

$$\tilde{\psi}_k(z) = z \left(\frac{1 + |\ln|2^{-k}z||}{1 + k} \right)^{\delta-1}, \quad z \in D_k,$$

и

$$\psi_k(z) = \frac{(1+k)^{\delta-1}}{2\pi i} \int_{2^k \Gamma} \frac{\varphi_0(2^{-k}t)dt}{t(1 + |\ln|2^{-k}t||)^{\delta}(t-z)}, \quad z \in D_k. \quad (3.5)$$

Очевидно, C^u -нормы первого сомножителя ограничены равномерно по k . Поэтому в соответствии с видом (2.2) нормы пространства C_0^u достаточно получить оценку

$$|\psi_k|_{C^u(\overline{D}_k)} \leq C |\varphi_0|_{C_0^u(\Gamma, 0)} \quad (3.6)$$

для второго множителя с некоторой постоянной $C > 0$, не зависящей от k .

С этой целью функцию (3.5) представим в виде суммы $\psi_k^0 + \psi_k^1$, где первое слагаемое отвечает интегрированию по дуге $\Gamma_k^0 = (2^k \Gamma) \cap \{1/4 < |t| < 4\}$.

Очевидно, кривая $2^k \Gamma$ радиальна и аналогично (2.3) описывается параметризацией $\gamma_k(s) = se^{i\theta_k(s)}$, $0 \leq s \leq 2^k \varrho$, с функцией $\theta_k(s) = \theta(2^{-k}s)$. Вспомня, что $\theta \in C[0, \varrho] \cap C^1(0, \varrho]$ и $s\theta'(s) \rightarrow 0$ при $s \rightarrow 0$, отсюда следует, что производная

$$|\gamma_k'| \leq \max_{0 \leq s \leq \varrho} |1 + s\theta'(s)|. \quad (3.7)$$

Из этих же соображений функция $\gamma_k(s)$, $1/4 \leq s \leq 4$, является параметризацией дуги Γ_k^0 и стремится к функции $se^{i\theta(0)}$ по норме пространства $C^1[1/4, 4]$. Поэтому аналогично теореме 3.10.1(a) из [3] устанавливается, что для функции ψ_k^0 справедлива оценка

$$|\psi_k^0|_{C^\mu(\overline{D}_k)} \leq C|\varphi_0|_{C^\mu(\Gamma_k^0)},$$

где постоянная $C > 0$ не зависит от k . Поскольку верхняя грань норм в правой части этой оценки, аналогично (1.7), определяет эквивалентную норму пространства $C_0^\mu(\Gamma, 0)$, отсюда

$$|\psi_k^0|_{C^\mu(\overline{D}_k)} \leq C|\varphi_0|_{C_0^\mu(\Gamma, 0)} \quad (3.8)$$

с некоторой новой постоянной, не зависящей от k .

Что касается функции ψ_k^1 , то она вместе со своей первой производной оценивается непосредственно. Пусть φ_0 означает $\sup$ -норму функции φ_0 , тогда с учетом (3.7) имеем оценки

$$|\psi_k^1(z)| + |(\psi_k^1)'(z)| \leq C|\varphi_0|_0(1+k)^{\delta-1}(I_k^0 + I_k^1), \quad (3.9)$$

где положено

$$I_k^0 = \int_0^{1/4} \frac{ds}{s[1 - \ln(2^{-k}s)]^\delta}, \quad I_k^1 = \int_4^{2^k \varrho} \frac{ds}{s^2[1 - \ln(2^{-k}s)]^\delta}.$$

Конечно, второй интеграл здесь рассматривается при $2^k \varrho \geq 4$, т.е. для достаточно больших k .

Очевидно, первый интеграл вычисляется явно:

$$(\delta - 1)I_k^0 = [1 - \ln(2^{-k-2})]^{1-\delta}. \quad (3.10^0)$$

Утверждается, что при условии $1 - \ln \varrho \geq 2\delta$ (которое всегда можно считать выполненным) второй интеграл допускает оценку

$$2I_k^1 \leq [1 - \ln(2^{-k+2})]^{-\delta}. \quad (3.10^1)$$

В самом деле, интегрирование по частям дает равенство

$$I_k^1 = - \left(\frac{1}{s[1 - \ln(2^{-k}s)]^\delta} \right) \Big|_4^{2^k \varrho} + \delta \int_4^{2^k \varrho} \frac{ds}{s^2[1 - \ln(2^{-k}s)]^{\delta+1}},$$

так что

$$\int_4^{2^k \varrho} \left(1 - \frac{\delta}{1 - \ln(2^{-k}s)} \right) \frac{ds}{s^2[1 - \ln(2^{-k}s)]^\delta} \leq \frac{1}{4[1 - \ln(2^{-k+2})]^\delta}.$$

Остается заметить, что выражение в больших круглых скобках не меньше $1/2$.

На основании (3.9), (3.10) заключаем теперь, что для функции ψ_k^1 справедлива аналогичная (2.13) оценка, что завершает доказательство (3.6). Следовательно, интеграл (3.4) определяет функцию $\phi_0 \in C_0^\mu(\overline{D}, 0)$ с соответствующей оценкой ее нормы, что завершает доказательство теоремы.

Формула Сохоцкого—Племеля (3.3) показывает, что вместе с I сингулярный интегральный оператор (3.2) ограничен: $C_{[\delta]}^\mu(\Gamma, \tau) \rightarrow C_{(0, \delta-1)}^\mu(\Gamma, \tau)$. Наряду с ним рассмотрим оператор $\overline{S}\varphi = (\overline{S\varphi})$, где черта справа означает комплексное сопряжение. Согласно (3.2) он действует по формуле

$$(\overline{S}\varphi)(t_0) = -\frac{1}{\pi i} \int_\Gamma \frac{\varphi(t) \overline{dt}}{\overline{t} - \overline{t_0}}, \quad t_0 \in \Gamma.$$

Если контур гладкий и принадлежит классу $C^{1,\nu}$, то, как хорошо известно, оператор $S + \overline{S}$ компактен в пространстве $C^\mu(\Gamma)$ при $0 < \mu < \nu$.

Лемма 3.1. Пусть в условиях теоремы 3.1 пара $(\Gamma, \tau) \in C_{(1)}^{1,\nu}$, $\mu < \nu < 1$. Тогда оператор $2K = S + \overline{S}$ компактен, $C_{[\delta]}^\mu(\Gamma, \tau) \rightarrow C_{[\delta-1]}^\mu(\Gamma, \tau)$, $\delta > 2$.

Доказательство. В явном виде оператор K действует по формуле

$$2(K\varphi)(t_0) = \frac{1}{\pi i} \int_\Gamma \left[\frac{dt}{t - t_0} - \frac{d\overline{t}}{\overline{t} - \overline{t_0}} \right] \varphi(t), \quad t_0 \in \Gamma.$$

Не ограничивая общности можно считать $\tau = 0$. Тогда как и при доказательстве теоремы 2.1 достаточно установить, что оператор

$$2(K_0\varphi)(t_0) = \frac{1}{\pi i} \int_\Gamma \frac{|t_0|}{|t|} \frac{(1 + |\ln |t_0||)^{\delta-1}}{(1 + |\ln |t||)^\delta} \left[\frac{dt}{t - t_0} - \frac{d\overline{t}}{\overline{t} - \overline{t_0}} \right] \varphi(t), \quad t_0 \in \Gamma,$$

компактен в пространстве $C_0^\mu(\Gamma, 0)$.

Очевидно, если функции f_1, f_2 непрерывно дифференцируемы и обращаются в нуль в окрестности точки $\tau = 0$, то оператор $f_1 K f_1$ компактен в $C_{[\delta]}^\mu(\Gamma, \tau)$. При достаточно малом $\varrho > 0$ гладкая кривая $\Gamma \cap \{|t| \leq \varrho\}$ состоит из двух радиальных дуг Γ_1, Γ_2 с общим концом $\tau = 0$. Выберем вещественную непрерывно дифференцируемую функцию $\chi(t)$ на Γ , которая тождественно равна 1 в окрестности $\tau = 0$ и обращается в нуль вне $\Gamma_1 \cup \Gamma_2$. Тогда утверждение леммы достаточно установить для оператора $\chi K_0 \chi$.

В самом деле, пусть функция χ_0 определяется аналогично χ с той разницей, что $\chi_0 \chi = \chi_0$. Тогда интегральные операторы $(1 - \chi) K_0 \chi_0$ и $\chi_0 K_0 (1 - \chi)$ равны нулю, поскольку их ядра тождественно равны нулю. Но тогда в силу сказанного выше операторы $K_0 - \chi K_0 = (1 - \chi) K_0 (1 - \chi_0)$ и $K_0 - K_0 \chi = (1 - \chi_0) K_0 (1 - \chi)$, а вместе с ними и $\chi K_0 - \chi K_0 \chi = \chi (K_0 - K_0 \chi)$ компактны в пространстве $C_{[\delta]}^\mu(\Gamma, \tau)$.

Рассмотрим радиальную параметризацию $\gamma_j = s e^{\theta_j(s)}$, $0 \leq s \leq \varrho$, дуги Γ_j , $j = 1, 2$, в которой согласно лемме 1.3 функции θ_j удовлетворяют условию (1.8). Удобно от оператора $\chi K_0 \chi$ перейти к операторной (2×2) -матрице $K^0 = (K_{ij}^0)_1^2$ по правилу

$$2[(\chi K_0 \chi)\varphi] \circ \gamma_i = \sum_{j=1,2} K_{ij}^0(\varphi \circ \gamma_j), \quad i = 1, 2.$$

В явном виде

$$(K_{ij}^0 \psi)(s_0) = \frac{1}{\pi} \int_0^\varrho \frac{s_0 (1 + |\ln s_0|)^{\delta-1}}{s (1 + |\ln s|)^\delta} \chi_{ij}^0(s_0, s) \frac{a_{ij}(s_0, s)}{s \pm s_0} \psi(s) ds, \quad 0 < s_0 \leq \varrho, \quad (3.11)$$

где выбирается верхний (нижний) знак для $i \neq j$ ($i = j$), положено

$$\chi_{ij}^0(s_0, s) = \chi[\gamma_i(s_0)] \chi[\gamma_j(s)], \quad a_{jj}(s_0, s) = \operatorname{Im} \left[\frac{\gamma_j'(s)(s - s_0)}{\gamma_j(s) - \gamma_j(s_0)} \right],$$

$$a_{ij}(s_0, s) = \operatorname{Im} \left[\frac{\gamma_j'(s)(s + s_0)}{\gamma_j(s) - \gamma_i(s_0)} \right], \quad i \neq j,$$

и учтено, что $|\gamma_i(s)| = s$.

Необходимо доказать, что операторы K_{ij}^0 компактны в пространстве $C_0^\mu([0, \varrho], 0)$. Обозначим $C_{(0)}^\nu([0, \varrho] \otimes [0, \varrho], 0)$ банахово пространство непрерывных в квадрате $0 \leq s_0, s \leq \varrho$ функций, которые по одной переменной принадлежат $C_{(0)}^\nu([0, \varrho], 0)$ равномерно по другой переменной, снабженное соответствующей нормой. Аналогичный смысл имеет банахово пространство $C_0^\nu([0, \varrho] \otimes [0, \varrho], 0)$ функций, непрерывных и ограниченных в полуоткрытом квадрате $0 < s_0, s \leq \varrho$.

Предварительно установим следующее вспомогательное утверждение.

Лемма 3.2. В обозначениях (3.11) функции

$$a_{jj} \in C_{(0)}^\nu([0, \varrho] \otimes [0, \varrho], 0), \quad a_{ij} \in C_0^\nu([0, \varrho] \otimes [0, \varrho], 0), \quad i \neq j,$$

причем

$$a_{jj}(t, t) \equiv 0, \quad \lim_{t+t_0 \rightarrow 0} a_{ij}(t_0, t) = 0. \quad (3.12)$$

Доказательство. Запишем

$$q_j(s_0, s) = \frac{\gamma_j(s) - \gamma_j(s_0)}{s - s_0} = \int_0^\varrho \gamma_j'[su + s_0(1 - u)] du.$$

Из этого равенства с помощью нормы (2.1) легко видеть, что функция q_j принадлежит классу $C_{(0)}^\nu([0, \varrho] \otimes [0, \varrho], 0)$ и всюду отлична от нуля, так что и обратная функция $1/q_j$ обладает этим свойством. Поскольку $q_j(s, s) = \gamma_j'(s)$ и $a_{jj} = \operatorname{Im}(\gamma_j'/q_j)$, приходим к справедливости леммы для a_{jj} .

Рассмотрим далее функцию

$$p_j(s) = \gamma_j(s)/s \in C_{(0)}^\nu([0, \varrho], 0), \quad |p_j(0)| = 1. \quad (3.13)$$

Тогда

$$p_{ij}(s_0, s) = -p_i(s_0)/p_j(s) \in C_{(0)}^\nu([0, \varrho] \otimes [0, \varrho], 0), \quad p_{ij}(0, 0) = 1.$$

Соответственно, функцию

$$q_{ij}(s_0, s) = \frac{\gamma_j(s) - \gamma_i(s_0)}{s + s_0} = p_j(s) \frac{s + p_{ij}(s_0, s)s_0}{s + s_0}$$

можно представить в виде

$$q_{ij}(s_0, s) = p_j(s) \left[1 + [p_{ij}(s_0, s) - 1] \frac{s_0}{s + s_0} \right] \in C_0^v([0, \varrho] \otimes [0, \varrho], 0).$$

Следовательно, функция q_{ij} принадлежит $C_0^v([0, \varrho] \otimes [0, \varrho], 0)$ и стремится к 1 при $s_0 + s \rightarrow 0$. Кроме того, функция q_{ij} ограничена от нуля по модулю и $a_{ij} = \text{Im}(\gamma'_j/p_{ij})$, что завершает доказательство леммы.

Продолжение доказательства леммы 3.1. Рассмотрим банахово пространство $C_0^v([0, \varrho] \otimes [0, \varrho], 0)$ и его замкнутое подпространство $C_{(0)}^v([0, \varrho] \otimes [0, \varrho], 0)$, которым согласно лемме 2.4 принадлежат функции a_{ij} . Как и в случае функций одной переменной, это пространство может быть определено с помощью эквивалентной нормы

$$|a| = \sup_{k,r=0,\pm 1,\dots} |a_{kr}|_{C^u(G_{kr})},$$

где положено

$$a_{kr}(s_0, s) = a(2^{-k}s_0, 2^{-r}s), \quad s_0, s \in G_{kr}, \quad G_{kr} = \{(s_0, s), 1/2 \leq 2^{-k}s_0, 2^{-r}s \leq 2\}.$$

С помощью данной нормы совершенно аналогично теореме 2.1 устанавливается, что операторы K_{ij}^0 , фигурирующие в (3.11), ограничены в пространстве $C_0^u([0, \varrho], 0)$, и их норма допускает оценку

$$|K_{ij}^0|_{\mathcal{L}(C_0^u)} \leq C|a_{ij}|_{C_0^v}. \quad (3.14)$$

Применим к пространству $C_0^v([0, \varrho] \otimes [0, \varrho], 0)$ теорему 2.7.2 из [3], согласно которой преобразование $a \rightarrow \tilde{a}$ по формуле $a(s_0, s) = \tilde{a}(-\ln s_0, -\ln s)$ осуществляет изоморфизм банахового пространства $C_0^v([0, \varrho] \otimes [0, \varrho], 0)$ на $C^v(J \times J)$, где $J = [-\ln \varrho, \infty)$. При этом преобразовании подпространство $C_{(0)}^v([0, \varrho] \otimes [0, \varrho], 0)$ переходит в подпространство C^v функций $\tilde{a}(t_0, t)$, допускающих пределы на бесконечности по каждой переменной t_0, t в отдельности. На основании теоремы 2.2.1 из [3], примененной к этим пространствам, для заданного $\mu < \nu_0 < \nu$ существуют такие последовательности функций

$$a_{ij,n} \in C^v([0, \varrho] \times [0, \varrho], \quad n = 1, 2,$$

что

$$\lim_{n \rightarrow \infty} |a_{ij,n} - a_{ij}|_{C^{\nu_0}} = 0, \quad (3.15)$$

и при каждом n выполнены условия

$$a_{ij,n}(s_0, s) \equiv 0 \quad \text{в окрестности точки } (0, 0),$$

$$a_{jj,n}(s_0, s) \equiv 0 \quad \text{в окрестности диагонали } s_0 = s.$$

Очевидно, операторы $K_{ij,n}^0$, определяемые по $a_{ij,n}$ аналогично (3.11), компактны в пространстве $C_0^u([0, \varrho], 0)$. С другой стороны, на основании (3.15) и оценки (3.14), где a_{ij} надо заменить на $a_{ij,n} - a_{ij}$ и ν на ν_0 , операторы $K_{ij,n}^0$ сходятся к K_{ij}^0 при $n \rightarrow \infty$ по операторной норме. Поэтому операторы $K_{ij,n}^0$ также компактны в пространстве $C_0^u([0, \varrho], 0)$, что завершает доказательство леммы 3.2.

4. ОСНОВНЫЕ РЕЗУЛЬТАТЫ

Доказательство теоремы 1.2 тесно связано с задачей Шварца, которая заключается в отыскании в области D аналитической функции ϕ по заданной вещественной части ее граничного значения:

$$\text{Re } \phi^+ = f. \quad (4.1)$$

Теорема 4.1. Пусть область D ограничена простым гладким контуром Γ с касательным вектором $\in C_{(1)}^{1,\nu}(\Gamma, \tau)$.

Тогда задача (4.1) с точностью до мнимой постоянной однозначно разрешима в классе $C_{[\infty]}^{\mu}(\overline{D}, \tau)$, $0 < \mu < \nu$.

Доказательство. Решение задачи ищем в виде интеграла типа Коши $I\phi$ с вещественной плотностью $\phi \in C_{[\infty]}^{\mu}(\Gamma, \tau)$. Тогда по формуле Сохоцкого–Племеля $2(\phi^+ + \overline{\phi^+}) = 2\phi + S\phi + \overline{S\phi}$, так что для ϕ приходим к уравнению $\phi + K\phi = 2f$, которое равносильно задаче (4.1). Пространство $C_{[\infty]}^{\mu}(\Gamma, \tau)$ можно рассматривать как счетно нормированное пространство. С топологической точки зрения оно представляет собой отделимое полное локально выпуклое пространство. В силу леммы 2.3 оператор K компактен в этом пространстве, так что к нему можно применить теорему Рисса из [6]. Поэтому достаточно убедиться, что однородное уравнение

$\varphi + K\varphi = 0$ имеет только нулевое решение. Но в этом случае $\operatorname{Re}(I\varphi)^+ = 0$, и, значит, $(I\varphi)(z) = i\xi$, $z \in D$, с некоторой вещественной постоянной ξ . Рассмотрим функцию $\phi_1 = I\varphi$ в бесконечной области $D_1 = \mathbb{C} \setminus \overline{D}$, исчезающую на бесконечности. По формуле Сохоцкого–Племеля $i\xi - \phi_1^- = \varphi$, так что

$$\operatorname{Im}\phi_1^- = \xi. \quad (4.2)$$

Поскольку $\phi_1(\infty) = 0$, равенство (4.2) возможно только в случае, когда $\xi = 0$ и $\varphi = 0$. Таким образом, оператор $1 + K$ обратим в пространстве $C_{[\infty]}^{\mu}(\Gamma, \tau)$, что завершает доказательство теоремы.

Доказательство теоремы 1.2. Рассмотрение проведем сначала для конформного $\omega_1 : D \rightarrow D_1$, где область D_1 ограничена контуром $\Gamma_1 \in C^{1,\mu}$, который содержит точку τ и прямолинеен в окрестности этой точки. По теореме Линделефа гармоническая в D функция $\arg \omega_1 \in C(\overline{D})$ и удовлетворяет аналогичному (1.1) соотношению. По условию единичный касательный вектор e принадлежит классу $C_{[\infty]}^{\mu}(\Gamma, \tau)$, так что с учетом теоремы 1.1 этому классу принадлежит и граничное значение $f = (\arg \omega_1')^+$. По теореме 3.1, где роль μ и ν играют соответственно $\mu/2$ и μ , существует аналитическая в D функция $\phi \in C_{[\infty]}^{\mu/2}(\overline{D}, \tau)$, для которой $\operatorname{Im} \phi^+ = f$. Поэтому непрерывная в $\overline{D}$ гармоническая функция $\operatorname{Im}(\ln \omega_1' - \phi)$ на границе области обращается в нуль, так что она равна нулю всюду в этой области. Следовательно, $\ln \omega_1' - \phi$ совпадает с мнимой постоянной, и, значит, $\ln \omega_1' \in C_{[\infty]}^{\mu/2}(\overline{D}, \tau)$.

Но тогда функция ω_1 и обратное отображение $\chi_1 = \omega_1^{-1}$ удовлетворяют условию Липшица. Лемма 2.1 справедлива и по отношению к пространству $C_{[\delta]}^{\mu}$. Нужно только принять во внимание, что $\alpha(t) - \alpha(\tau) \in C^{0,1}(G, \tau)$, и, значит, функция $a(t) = |\alpha(t) - \alpha(\tau)|/|t - \tau|$, а вместе с ней и $|\ln a(t)|^{\delta}$ принадлежат $C_0^0(G, \tau)$. Таким образом, функция $\arg \chi_1' = -\arg \omega_1' \circ \chi_1$ на границе области D_1 принадлежит классу $C_{[\infty]}^{\mu}(\Gamma_1, \tau)$, так что на основании теоремы 3.1 $\ln \chi_1' \in C_{[\infty]}^{\mu}(\overline{D}_1, \tau)$, и, значит, из тех же соображений $\ln \omega_1' \in C_{[\infty]}^{\mu}(\overline{D}, \tau)$.

Покажем, что тогда $\ln \omega' \in C_{[\infty]}^{\mu}(\overline{D}, \tau)$ и в общем случае. С этой целью рассмотрим конформное отображение α области D на D_1 , оставляющее неподвижной точку τ . Тогда $\omega = \omega_1 \circ \alpha$ и $\ln \omega' = \ln \omega_1' \circ \alpha + \ln \alpha'$. С учетом теоремы Келлога, примененной к α , и леммы 2.3 функция $\ln \omega' \in C_{[\infty]}^{\mu}(\overline{D}, \tau)$.

Что касается утверждения теоремы относительно функции a , то заметим, что в силу леммы 2.5 эта функция принадлежит классу $C_{[\infty]}^{\mu}(\Gamma, \tau)$ на границе области. Но тогда на основании теоремы 4.1 функция $a \in C_{[\infty]}^{\mu}(\overline{D}, \tau)$.

Обратимся к последнему утверждению теоремы относительно χ . В силу уже доказанного функция ω удовлетворяет двустороннему условию Липшица. Поэтому утверждение для $\ln \chi'$ вытекает из очевидного соотношения $\ln \chi' \circ \omega = -\ln \omega'$ и леммы 2.4. Аналогично устанавливается, что и функция $[\chi(\xi) - \chi(\tau')]/(\xi - \tau')$, где $\tau' = \omega(\tau)$, принадлежит классу $C_{[\infty]}^{\mu}(\mathbb{D}, \tau')$.

Исходя из теоремы 1.2, можно получить соответствующий результат для конформных отображений кусочно-гладких областей. Пусть область D ограничена кусочно-гладким контуром Γ , и конечное множество $F \subseteq \Gamma$ содержит все угловые точки контура. Ясно, что понимать под записью $(\Gamma, F) \in C_{(+0)}^{1,\mu}$ и $(\Gamma, F) \in C_{[\infty]}^{1,\mu}$. Выберем число $\varrho > 0$ столь малым, что для каждой точки $\tau \in F$ пересечение D с кругом $|z - \tau| < \varrho$ представляет собой криволинейный сектор S_{τ} с боковыми сторонами $\Gamma_{\tau}^{\pm}$ с общим концом в вершине τ , раствор этого сектора обозначим через θ_{τ} . В дальнейшем предполагается, что все эти растворы положительны, так что $0 < \theta_{\tau} \leq 2\pi$, $\delta_{\tau} = \pi/\theta_{\tau}$.

Аналог теоремы 1.1 справедлив и для кусочно-гладких областей.

Теорема 4.2. Пусть кусочно-гладкий контур Γ принадлежит классу $C_{\text{loc}}^{1,\mu}$ по отношению к множеству F .

Тогда $\ln \omega' \in C_{\text{loc}}^{\mu}(\overline{D}, F)$.

Доказательство. Зафиксируем точку $\tau \in F$ и рассмотрим область $\tilde{D}_1$, которая содержит D и ограничена контуром $\tilde{\Gamma}_1$ класса $C^{1,\mu}$ по отношению к τ , причем раствор $\tilde{\theta}$ соответствующего сектора $\tilde{D} \cap \{|z - \tau| < \tilde{\varrho}\}$ был заключен между θ_{τ} и π , а боковые стороны этого сектора были прямолинейны. Пусть произвольная область $\tilde{D}_2$ ограничена контуром $\tilde{\Gamma}_2$ класса $C^{1,\mu}$ по отношению к τ , причем раствор соответствующего сектора $\tilde{D}_2 \cap \{|z - \tau| < \tilde{\varrho}\}$ был равен $\tilde{\theta}$, а боковые стороны этого сектора были прямолинейны. Требуется, кроме того, чтобы степенная функция $\alpha(z) = \tau + (z - \tau)^{\delta_{\tau}}$ была однолистной в области $\tilde{D}_2$. Рассмотрим конформное отображение $\tilde{\omega} : \tilde{D}_1 \rightarrow \tilde{D}_2$, которое принадлежит классу $C^{1,\mu}$ в замыкании области $\tilde{D}_1$ и переводит D на некоторую область $\tilde{D}$. Эта область ограничена кусочно-гладким контуром класса $C^{1,\mu}$ по отношению к множеству $\tilde{F} = \tilde{\omega}(F)$, и в которой функция α однолистка. Очевидно, функция α_{τ} выпрямляет границу области $\tilde{D}$ в окрестности τ и также однолистка в этой области. Таким образом, дело сводится к доказательству теоремы для области $\alpha(\tilde{D})$, граница которой принадлежит классу $C^{1,\mu}$ по отношению к множеству $\alpha(\tilde{F} \setminus \{\tau\})$. Повторяя эту процедуру, после конечного числа шагов приходим к области, к которой остается применить теорему 1.1.

Для кусочно-гладкого контура Γ единичный касательный вектор e кусочно-непрерывен с возможными разрывами в точках $\tau \in F$. По определению $e \in C_{(+0)}^{\mu}(\Gamma, F)$, если для любой дуги $G \subseteq \Gamma$, пересекающейся с F по одному из своих концов τ , функция $e \in C_{(+0)}^{\mu}(G, \tau)$. Аналогичный смысл имеет запись $e \in C_{[\infty]}^{\mu}(\Gamma, F)$.

Если $e \in C_{(+0)}^{\mu}(\Gamma, F)$, то как установлено в [4], конформное отображение ω области D на единичный круг в каждом криволинейном секторе S_{τ} представимо в виде

$$\omega(z) - \omega(\tau) = a_{\tau}(z)(z - \tau)^{\delta_{\tau}}, \quad \omega'(z) = b_{\tau}(z)(z - \tau)^{\delta_{\tau}-1}, \quad (4.3)$$

где функции a_{τ}, b_{τ} принадлежат $C_{(+0)}^{\mu}(\overline{S}_{\tau}, \tau)$ и всюду отличны от нуля.

Следующая теорема распространяет этот факт и на класс $C_{[\infty]}^{\mu}(\Gamma, F)$.

Теорема 4.3. Пусть единичный касательный вектор e кусочно-гладкого контура Γ принадлежит классу $C_{[\infty]}^{\mu}(\Gamma, F)$.

Тогда (4.3) принадлежат $C_{[\infty]}^{\mu}(\overline{S}_{\tau}, \tau)$ и всюду отличны от нуля. Аналогичное утверждение справедливо и для обратного отображения $\chi = \omega^{-1}$ с заменой δ_{τ} на $1/\delta_{\tau}$.

Доказательство. Очевидно, утверждения теоремы достаточно установить для случая, когда роль единичного круга играет произвольная область D_1 с границей класса $C^{1,\mu}$. Выберем в D подобласть D_0 , которая по отношению к фиксированной точке $\tau \in F$ содержит сектор S_{τ} и принадлежит классу $C_{[1]}^{1,\mu}$. Тогда на основании теоремы 1.1 образ $D_1 = \omega(D_0)$ является областью с границей класса $C^{1,\mu}$. Не ограничивая общности можно считать, что $\tau = 0$ и степенная функция $\alpha(z) = z^{\delta}$ с $\delta = \delta_0$ однолистка в области D_0 . Тогда образ $D_2 = \alpha(D_0)$ является областью с гладкой границей, которая в силу приведенной ниже леммы 4.1 принадлежит классу $C_{[\infty]}^{1,\mu}$ (по отношению к точке $\tau = 0$). Запишем конформное отображение в виде $\omega(z) = \omega_2(z^{\delta})$ с соответствующим конформным отображением $\omega_2 : D_2 \rightarrow D_1$. На основании теоремы 1.2 функции $\omega_2'(z)$ и $a_2(z) = [\omega_2(z) - \omega_2(0)]/z$ принадлежат $C_{[\infty]}^{\mu}(\overline{D}_2, 0)$. При этом

$$\omega(z) - \omega(0) = \omega_2(z^{\delta}) - \omega_2(0), \quad \omega''(z) = \delta \omega_2'(z^{\delta}) z^{\delta-1}.$$

Следовательно, для коэффициентов (4.3) имеем выражения

$$a = a_2 \circ \alpha, \quad b = \delta \omega_2' \circ \alpha,$$

где $a_2(z) = [\omega_2(z) - \omega_2(0)]/z$. На основании теоремы 1.2 функции a_2 и ω_2' принадлежат классу $C_{[\infty]}^{\mu}(\overline{D}_2, 0)$. Поэтому остается заметить, что отображение $\varphi \rightarrow \varphi \circ \alpha$ ограничено $C_{[\infty]}^{\mu}(\overline{D}_2, 0) \rightarrow C_{[\infty]}^{\mu}(\overline{D}, 0)$. В самом деле, оно сводится к аналогичному утверждению применительно к пространству, которое станет очевидным, если воспользоваться эквивалентной нормой (2.2) этого пространства.

Лемма 4.1. Отображение $\alpha(z) = z^{\delta}$, переводящее радиальную дугу с концом $\tau = 0$ в дугу того же типа, инвариантно в классе $C_{[\infty]}^{\mu}$ их единичных касательных векторов.

Доказательство. Рассмотрим радиальную дугу G с параметризацией (2.3), где $\tau = 0$. На основании леммы 2.4 функции

$$\theta(s), \quad s\theta'(s) \in C_{[\infty]}^{\mu}([0, q]; 0). \quad (4.4)$$

Очевидно, дуга $\tilde{G} = \alpha(G)$ также радиальна и описывается аналогично (2.3) с помощью функции $\tilde{\theta}(s) = \delta\theta(s^{1/\delta})$, $0 \leq s \leq q^{\delta}$. Поэтому с учетом (4.4) эта функция также принадлежит классу $\tilde{\theta} \in C_{[\infty]}^{\mu}([0, q^{\delta}], 0)$. Поскольку $s\tilde{\theta}'(s) = s^{1/\delta}\theta'(s^{1/\delta})$, из тех же соображений этому классу принадлежит и функция $s\tilde{\theta}'(s)$, так что остается опять воспользоваться леммой 2.4.

СПИСОК ЛИТЕРАТУРЫ

1. Голузин Г.М. Геометрическая теория функций комплексного переменного. М.: Наука, 1972.
2. Duren P.L. Theory of H^p -spaces. New York-London: Acad. Press, 1970.
3. Солдатов А.П. Сингулярные интегральные операторы и эллиптические краевые задачи // Современ. матем. Фундамент. направления. 2017. Т. 63. С. 1–189.
4. Солдатов А.П. Аналог теоремы Келлога для кусочно-ляпуновских областей // Ж. вычисл. матем. и матем. физ. 2023. Т. 63. № 8. С. 102–111.
5. Мухелишвили Н.И. Сингулярные интегральные уравнения. М.: Наука, 1968.
6. Робертсон А., Робертсон В. Топологические векторные пространства. М.: Мир, 1967.

ON THE BOUNDARY PROPERTIES OF CONFORMAL MAPS

A. P. Soldatov^{a,b,*}^a 119333 Moscow, Vavilov Str., 44, Federal Research Center Computer Science and Control, RAS, Russia^b 360002 Nalchik, Dolinsk, Balkarova str., 2, Institute of Applied Mathematics and Automation, KBSC RAS, Russia

*e-mail: soldatov48@gmail.com

Received: 24.04.2024

Revised: 24.04.2024

Accepted: 31.05.2024

Abstract. Classes of C^1 -smooth regions are described, the boundary contour of which is Lyapunovian outside any neighborhood of a certain point, such that the derivative of the conformal map on the unit circle is continuous at this point. The description is given in terms of some spaces for a unit tangent vector on a boundary contour. As a consequence, the corresponding results are obtained for piecewise smooth regions.

Keywords: conformal mapping, smooth contour, boundary properties, generalized Helder spaces.

О ЧИСЛЕННОМ ГАШЕНИИ КОЛЕБАНИЙ СТРУНЫ С ИСПОЛЬЗОВАНИЕМ НЕСКОЛЬКИХ СТАЦИОНАРНЫХ АКТЬЮАТОРОВ

© 2024 г. И. Е. Михайлов^{1,2,*}¹ 119333 Москва, ул. Вавилова, 44, ФИЦ ИУ РАН, Россия² 105005 Москва, ул. 2-я Бауманская, 5, стр. 1, МГТУ им. Н.Э. Баумана, Россия

*e-mail: mikh_igor@mail.ru

Поступила в редакцию 17.03.2024 г.

Переработанный вариант 17.03.2024 г.

Принята к публикации 31.05.2024 г.

Ставится задача перевести струну из начального возмущенного состояния в состояние покоя за минимальное время. Гашение колебаний струны осуществляется с помощью нескольких стационарных актьюаторов. Минимизируемым функционалом является некоторый интеграл. Управление гашением колебаний осуществляется с помощью функции, входящей в правую часть гиперболического уравнения, описывающего поперечные колебания струны, и моделирующей действия актьюаторов. Разработаны вычислительные алгоритмы решения задачи на основе сеточного метода и градиентного метода отыскания минимума функций многих переменных, причем градиент вычисляется с помощью метода быстрого автоматического дифференцирования, предложенного Ю.Г. Евтушенко. Приводятся примеры расчетов гашения колебаний струны с помощью различного числа актьюаторов. Библ. 6. Фиг. 6.

Ключевые слова: гашение колебаний струны, быстрое автоматическое дифференцирование.

DOI: 10.31857/S0044466924090084, **EDN:** WJXWRJ

1. ПОСТАНОВКА ЗАДАЧИ

Поперечные колебания струны описываются гиперболическим уравнением

$$u_{tt} = a^2 u_{xx} + g(t, x), \quad t > 0, \quad 0 < x < l, \quad a > 0. \quad (1)$$

Начальные отклонение и скорость поперечного перемещения струны

$$u|_{t=0} = \varphi(x), \quad u_t|_{t=0} = \psi(x), \quad 0 \leq x \leq l, \quad (2)$$

будем рассматривать как заданные начальные возмущения. В качестве граничных условий возьмем условия закрепления

$$u|_{x=0} = 0, \quad u|_{x=l} = 0. \quad (3)$$

Задачу гашения колебаний струны сформулируем следующим образом: требуется найти функцию $g(t, x)$ из класса $L_2(0 \leq t \leq T, 0 \leq x \leq l)$, позволяющую перевести струну из состояния (2) в состояние покоя:

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

за минимальное время $T > 0$. Заметим, что условие (4) равносильно обращению следующего интеграла в нуль:

$$I(T) = \int_0^l (u^2 + u_t^2) dx = 0. \quad (5)$$

Методы гашения колебаний элементов сложных механических систем начали интенсивно развиваться в 70-х годах XX в. Наиболее значимыми были работы А. Бутковского [1], Д. Рассела [2], Д. Лагнесса [3], С.Ж. Асланова, И.Е. Михайлова, Л.А. Муравья [4], в которых рассматривалась задача гашения колебаний струны, и получены условия существования решения задачи.

В работе Бутковского [1] рассматривался точечный стационарный актюатор, помещенный в точку x_0 . Им было показано, что существует всюду плотное на $\{0, l\}$ множество точек $\{x_0\}$ таких, что помещенный в них актюатор не позволяет решить задачу гашения колебаний, например, на множестве решений уравнения

$$\sin \frac{k\pi x_0}{l} = 0, \quad k = 1, 2, \dots,$$

являющихся узлами стоячих волн решений

$$z_k(t, x) = (A_k \cos \omega_k t + B_k \sin \omega_k t) \sin \frac{k\pi x_0}{l}, \quad k = 1, 2, \dots,$$

однородного уравнения

$$u_{tt} = a^2 u_{xx}, \quad t > 0, \quad 0 < x < l, \quad a > 0,$$

где $\omega_k = \frac{k\pi a}{l}$. Поэтому при способе управления, рассмотренном в работе [1], задача может быть неразрешимой.

В работе Рассела [2] функция $g(t, x)$ представлялась в виде $g(t, x) = u(t)f(x)$. Были получены условия на функцию $f(x)$, которые позволяют решить задачу гашения колебаний с помощью одной управляющей функции $u(t)$. Однако функция $f(x)$ распределена на всем интервале ($\text{supp } f(x) = (0, l)$), поэтому для решения задачи управление должно осуществляться вдоль всей длины струны, что не позволяет использовать результаты работы [2] в практических приложениях при достаточно длинной струне.

В работе Лагнесса [3] было показано, что колебания можно погасить с помощью бесконечного числа управляющих функций $u_1(t), u_2(t), \dots, u_k(t), \dots$, если $g(t, x)$ представима в виде бесконечного ряда $\sum_{k=1}^{\infty} u_k(t)f_k(x)$, определенных на небольшом участке струны $[\alpha, \beta]$, причем, если $\beta - \alpha$ стремится к нулю, то амплитуда управляющей функции стремится к бесконечности, что указывает на невозможность использования управляющей функции Бутковского.

Все эти ученые использовали метод разделения переменных Фурье и сводили задачу к тригонометрической проблеме моментов — бесконечной системе интегральных уравнений. Их результаты основывались на условии Левинсона (см. [5]), установившим, что гарантированно погасить колебания струны можно за время $T = 2l/a$, где l — длина струны, a — скорость распространения возмущений вдоль струны.

В 2006 г. С. Асланов, И. Михайлов и Л. Муравей показали, что если актюатор движется с постоянной скоростью $v \geq a$ на небольшом участке струны, то погасить колебания струны можно за время $T = 2l/v$, т.е. меньшее времени Левинсона.

Отметим, что все ограничения работ [1]–[4] сильно затрудняют использование результатов этих работ в практических приложениях.

Основная идея настоящей работы заключается в использовании нескольких стационарных актюаторов, действующих на небольших участках струны шириной δ . Будем искать функцию $g(t, x)$ в виде

$$g(t, x) = \sum_{i=1}^p W_i(t) f_i(x), \quad (6)$$

где W_i — управляющие функции,

$$f_i(x) = \begin{cases} 1, & |x - x_i| \leq \delta/2, \\ 0, & |x - x_i| > \delta/2, \end{cases}$$

x_i — центры отрезков, в которых помещаются актюаторы.

2. ЧИСЛЕННЫЙ МЕТОД РЕШЕНИЯ ЗАДАЧИ ГАШЕНИЯ КОЛЕБАНИЙ СТРУНЫ

2.1. Численный алгоритм решения начально-краевой задачи для волнового уравнения

Для того чтобы численно решить начально-краевую задачу (1)–(3), аппроксимируем ее явной конечно-разностной схемой второго порядка аппроксимации по обоим переменным. В области решения введем расчетную сетку. Зададим натуральные числа M и N и разобьем рассматриваемые отрезки $\{0 \leq x \leq l, 0 \leq t \leq T\}$ на промежутки точками x_m и t_n соответственно:

$$\omega_h = \{x_m = mh, m = \overline{0, M}\}, \quad \omega_\tau = \{t_n = n\tau, n = \overline{0, N}\},$$

где $h = \frac{l}{M}$ и $\tau = \frac{T}{N}$ — шаги сетки по переменным x и t соответственно. В качестве сетки $D_{h,\tau}$ используем совокупность точек пересечения прямых $\omega_h \times \omega_\tau$. Введем вместо точного решения $u(x_m, t_n)$ сеточную функцию

$\{U_m^n\}_{m=\overline{0,M}, n=\overline{0,N}}$, которая будет удовлетворять следующей конечно-разностной схеме:

$$\frac{U_m^{n+1} - 2U_m^n + U_m^{n-1}}{\tau^2} = a^2 \frac{U_{m+1}^n - 2U_m^n + U_{m-1}^n}{h^2} + g_m^n, \quad m = \overline{1, M-1}, n = \overline{1, N-1}. \quad (7)$$

Она аппроксимирует уравнение (1) со вторым порядком по h и τ .

Выразим U_m^{n+1} :

$$U_m^{n+1} = 2U_m^n - U_m^{n-1} + S(U_{m+1}^n - 2U_m^n + U_{m-1}^n) + \tau^2 g_m^n, \quad \text{где } S = \frac{\tau^2 a^2}{h^2}.$$

Аппроксимируем начальные условия (2). Первое начальное условие запишется в виде

$$U_m^0 = \varphi_m. \quad (8)$$

Для того чтобы аппроксимировать второе начальное условие (2) со вторым порядком по h и τ , введем фиктивный слой по времени $t = -\tau$. Значения сеточной функции на этом слое обозначим U_m^{-1} . Будем полагать, что уравнение (1) справедливо на промежутке $[-\tau; 0]$, т.е. уравнение (7) выполняется при $n = 0$. С использованием “замечательного свойства средней точки” выпишем выражение, аппроксимирующее второе начальное условие (2) со вторым порядком аппроксимации, и выпишем разностную схему (8) на слое $n = 0$:

$$\begin{aligned} \frac{U_m^1 - U_m^{-1}}{2\tau} &= \psi_m, \\ \frac{U_m^1 - 2U_m^0 + U_m^{-1}}{\tau^2} &= a^2 \frac{U_{m+1}^0 - 2U_m^0 + U_{m-1}^0}{h^2} + g_m^0. \end{aligned}$$

Исключим из этих двух уравнений U_m^{-1} .

Тогда получим с учетом (9) выражение для U_m^1 :

$$U_m^1 = \frac{S}{2} (\varphi_{m+1} - 2\varphi_m + \varphi_{m-1}) + \varphi_m + \tau\psi_m + \frac{\tau^2}{2} g_m^0. \quad (9)$$

В результате конечно-разностная аппроксимация уравнения (1) и начальных условий (2) имеет второй порядок аппроксимации по обоим переменным.

Краевые условия имеют вид

$$U_0^n = U_M^n = 0, \quad n = \overline{0, N}. \quad (10)$$

Легко показать, что конечно-разностная схема (8) устойчива по Нейману при условии

$$S \leq 1 \quad \text{или} \quad \frac{a\tau}{h} \leq 1. \quad (11)$$

2.2. Численный алгоритм выбора оптимального управления

Для численного расчета интеграла (5) воспользуемся квадратурной формулой трапеций, которая с учетом краевых условий (10) имеет вид

$$I(T) = \sum_{m=1}^{M-1} \left((U_m^N)^2 + \left(\frac{U_m^N - U_m^{N-1}}{\tau} \right)^2 \right) h. \quad (12)$$

Условием гашения будем полагать выполнение неравенства $I(T) \leq \varepsilon$, где ε — заданная точность вычислений.

Аппроксимируем управляющие функции $W_i(t)$, $i = \overline{1, p}$, кусочно-постоянными функциями такими, что $\forall t \in [t_n, t_{n+1})$, $n = \overline{0, N-1}$, $W_i(t) = W_i^n = \text{const}$.

Тогда интеграл (12) будет являться функцией переменных $W_1^0, \dots, W_1^N, W_2^0, \dots, W_2^N, \dots, W_p^0, \dots, W_p^N$:

$$I(T) = \Phi(W_1^0, \dots, W_1^N, W_2^0, \dots, W_2^N, \dots, W_p^0, \dots, W_p^N). \quad (13)$$

Оптимальные значения W_i^n , минимизирующие (13) с заданной точностью ε и являющиеся искомым решением задачи, будем искать методом градиентного спуска

$$(W_i^j)^{k+1} = (W_i^j)^k - \alpha_k \text{grad } \Phi(W_i^j).$$

Градиент будем находить с помощью метода быстрого автоматического дифференцирования, предложенного Ю.Г. Евтушенко (см. [6]). Опишем идею метода быстрого автоматического вычисления градиента функции $I(U, W)$, где U – решение некоторой начально-краевой задачи $G(U, W) = 0$, а W – управление. Пусть связи $G(U, W) = 0$ неявным образом определяют решение $U = U(W)$.

Тогда

$$G_U \frac{dU}{dW} + G_W = 0. \quad (14)$$

Из уравнения (14) следует

$$G_U \frac{dU}{dW} = -G_W.$$

Градиент $I(U, W)$ по управлению W определяется следующим образом:

$$\frac{dI}{dW} = I_U \frac{dU}{dW} + I_W.$$

Введем функцию Лагранжа

$$L = I(U(W), W) + pG(U(W), W).$$

Стационарные точки L определяются из равенств

$$\begin{aligned} L_U &= I_U + pG_U = 0, \\ L_W &= I_W + pG_W = 0, \\ L_p &= G(U(W), W) = 0. \end{aligned} \quad (15)$$

Из первого уравнения системы (15) с учетом (14) выражение для градиента примет вид

$$\frac{dI}{dW} = I_U \frac{dU}{dW} + I_W = I_W - pG_U \frac{dU}{dW} = I_W + pG_W.$$

А так как $I_W = 0$, то

$$\frac{dI}{dW} = pG_W. \quad (16)$$

Итак, будем искать минимум функционала (12), причем конечно-разностное уравнение, начальные и краевые условия рассматриваются как связи. Выпишем функцию Лагранжа, учитывая связи с помощью переменных множителей Лагранжа P_m^n :

$$\begin{aligned} L = & \sum_{m=1}^{M-1} \left((U_m^N)^2 + \left(\frac{U_m^N - U_m^{N-1}}{\tau} \right)^2 \right) h + \sum_{m=1}^{M-1} \sum_{n=2}^N P_m^n \left(2U_m^{n-1} - U_m^{n-2} + S(U_{m+1}^{n-1} - 2U_m^{n-1} + U_{m-1}^{n-1}) - \right. \\ & \left. - U_m^n + \tau^2 g_m^{n-1} \right) + \sum_{m=0}^M P_m^0 (\varphi_m - U_m^0) + \sum_{m=1}^{M-1} P_m^1 \left(\frac{S}{2} (\varphi_{m+1} - 2\varphi_m + \varphi_{m-1}) + \varphi_m + \tau \psi_m + \frac{\tau^2}{2} g_m^0 - U_m^1 \right) + \\ & + \sum_{n=1}^N P_0^n (0 - U_0^n) + P_M^n (0 - U_M^n). \end{aligned}$$

Стационарные точки функции Лагранжа находятся из соотношений $\frac{\partial L}{\partial U_m^n} = 0$.

Тогда множители Лагранжа P_m^n определяются следующим образом:

$$\begin{aligned} P_m^N &= 2hU_m^N + \frac{2h}{\tau^2} (U_m^N - U_m^{N-1}), \quad m = \overline{1, M-1}, \\ P_m^{N-1} &= 2P_m^N + S(P_{m-1}^N - 2P_m^N + P_{m+1}^N) - \frac{2h}{\tau^2} (U_m^N - U_m^{N-1}), \quad m = \overline{1, M-1}, \\ P_m^n &= 2P_m^{n+1} - P_m^{n+2} + S(P_{m-1}^{n+1} - 2P_m^{n+1} + P_{m+1}^{n+1}), \quad m = \overline{1, M-1}, \quad n = \overline{1, N-2}, \\ P_0^n &= 0; \quad P_M^n = 0, \quad n = \overline{0, N}, \\ P_m^1 &= 0, \quad m = \overline{1, M-1}, \\ P_m^0 &= 0, \quad m = \overline{1, M-1}. \end{aligned} \quad (17)$$

Положим в (6) ширину актьюаторов $\delta = h$. Тогда $\text{grad } \Phi(W_i^j)$ находится по формулам

$$\frac{\partial \Phi}{\partial W_m^0} = \begin{cases} \frac{\tau^2}{2} P_m^1, & m = m_1 \vee \dots \vee m = m_p, \\ 0, & m \neq m_1 \wedge \dots \wedge m \neq m_p, \end{cases} \quad m = \overline{1, M-1}, \quad (18)$$

$$\frac{\partial \Phi}{\partial W_m^n} = \begin{cases} \tau^2 P_m^{n+1}, & m = m_1 \vee \dots \vee m = m_p, \\ 0, & m \neq m_1 \wedge \dots \wedge m \neq m_p, \end{cases} \quad m = \overline{1, M-1}, \quad n = \overline{1, N-1}. \quad (19)$$

Здесь m_i – номера точек отрезка $[0, l]$, в которых помещены центры актьюаторов.

Если, например, $S = 1/4$ и $h \rightarrow 0$, то значения P_m^n являются решениями уравнения

$$P_{tt} = P_{xx} \quad (20)$$

с краевыми условиями

$$P|_{x=0} = 0, P|_{x=l} = 0, 0 \leq t \leq T, \quad (21)$$

и начальными условиями

$$P|_{t=T} = \frac{1}{a} \frac{\partial U}{\partial t}, \quad P_t|_{t=T} = \frac{1}{a} U(T, x), \quad 0 \leq x \leq l. \quad (22)$$

Легко видеть, что эта задача аналогична задаче (1)–(3).

Таким образом, чтобы вычислить, достаточно один раз решить задачу (20)–(22) с использованием разностных уравнений (17) и сразу выписать значение градиента по формулам (18), (19).

3. ПРИМЕРЫ РАСЧЕТОВ

3.1. Сравнение результатов прямого расчета с аналитическим решением

Для проверки правильности прямого расчета был проведен расчет начально-краевой задачи

$$\begin{aligned} u_{tt} &= a^2 u_{xx}, \quad t > 0, \quad 0 < x < l, \quad u > 0 \\ u|_{t=0} &= \sin \frac{\pi x}{l}, \quad u_t|_{t=0} = 0, \quad 0 < x < l, \\ u|_{x=0} &= 0, \quad u|_{x=l} = 0, \end{aligned}$$

которая имеет аналитическое решение $u(t, x) = \cos \frac{\pi a t}{l} \sin \frac{\pi x}{l}$.

Расчет проводился со следующими входными параметрами: $l = 1, T = 2, a = 1, M = 20, N = 80$. Результаты численного расчета совпали с аналитическим решением с точностью $\varepsilon = 10^{-5}$.

3.2. Примеры расчетов для различного числа актьюаторов

Пример 1. Рассматривалась задача гашения колебаний струны с использованием одного актьюатора. В расчете принималось $a = 1, l = 1$, начальное положение центра актьюатора

$$x_1 = \frac{3}{16}, \quad h = \frac{1}{16}, \quad \tau = \frac{1}{32}.$$

Начальные возмущения задавались функциями $\varphi(x) = \sin 2\pi x, \psi(x) = 0$. Точность вычислений $\varepsilon = 10^{-4}$. Процесс гашения колебаний и оптимальное управление $W_1(t)$, позволяющее погасить начальные колебания за время $T = 2.0$, показаны на фиг. 1 и 2 соответственно.

Пример 2. Были проведены расчеты с теми же параметрами, что и в примере 1, но с использованием двух актьюаторов. Начальные положения центров актьюаторов

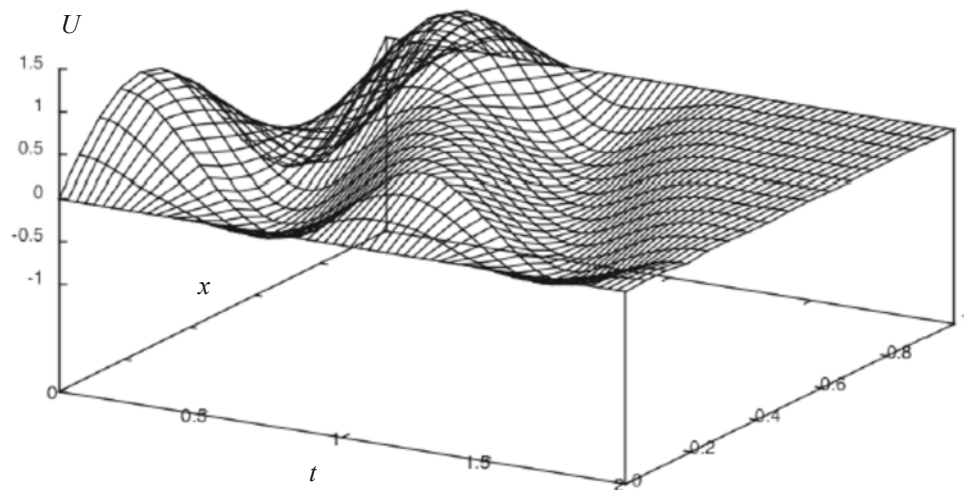
$$x_1 = \frac{3}{16}, \quad x_2 = \frac{11}{16}, \quad h = \frac{1}{16}, \quad \tau = \frac{1}{32}.$$

За время $T = 1.2$ происходит практически полное гашение колебаний. Процесс гашения колебаний и графики оптимальных управлений $W_1(t), W_2(t)$ представлены на фиг. 3 и 4 соответственно.

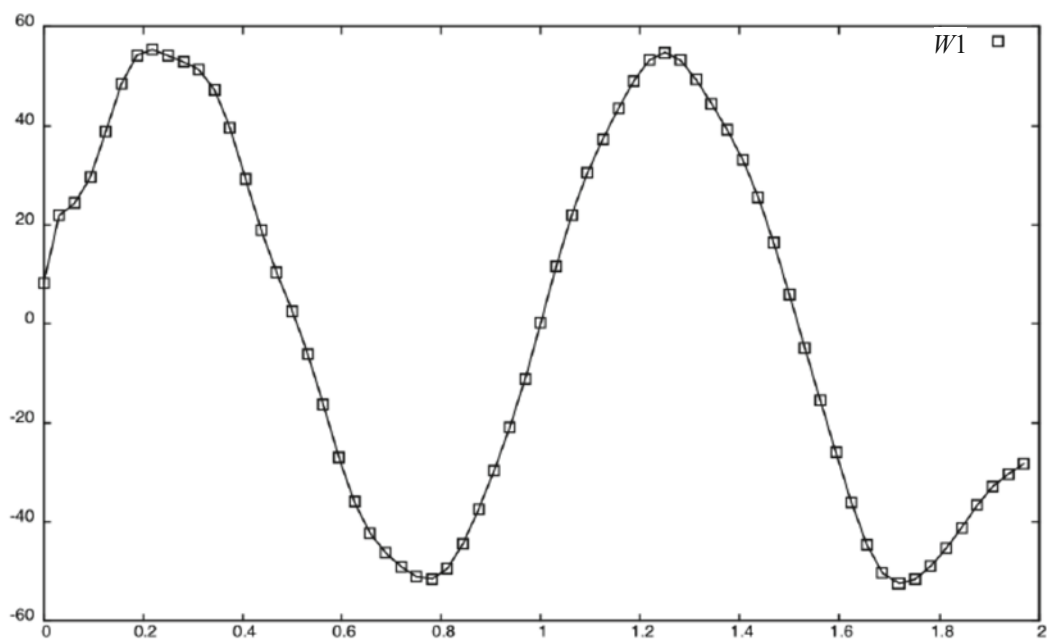
Пример 3. Были проведены расчеты с теми же параметрами, что и в примере 1, но с использованием трех актьюаторов. Начальные положения центров актьюаторов

$$x_1 = \frac{3}{16}, x_2 = \frac{7}{16}, \quad x_3 = \frac{11}{16}, \quad h = \frac{1}{16}, \quad \tau = \frac{1}{32}.$$

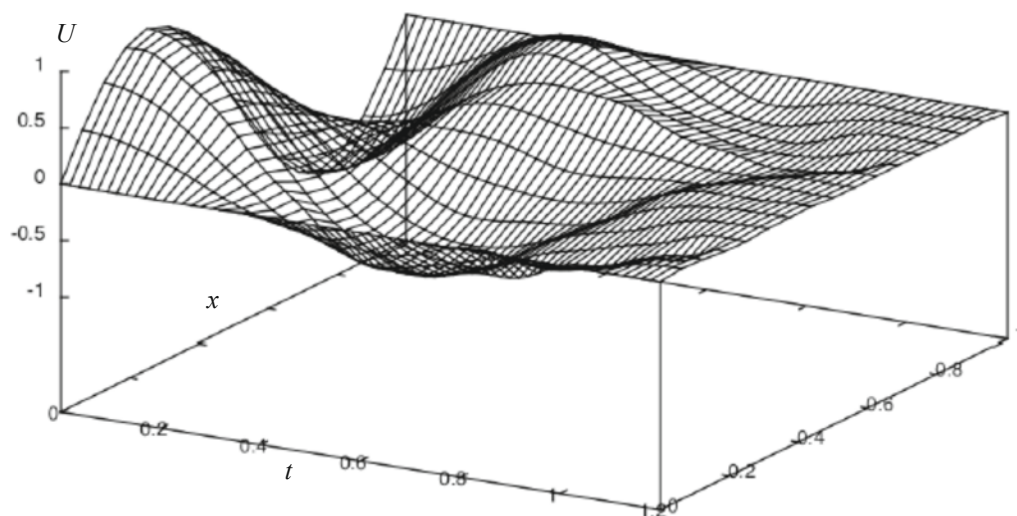
Начальные возмущения удалось погасить за время $T = 1.0$. Процесс гашения колебаний и графики оптимальных управлений $W_1(t), W_2(t), W_3(t)$ представлены на фиг. 5 и 6 соответственно.



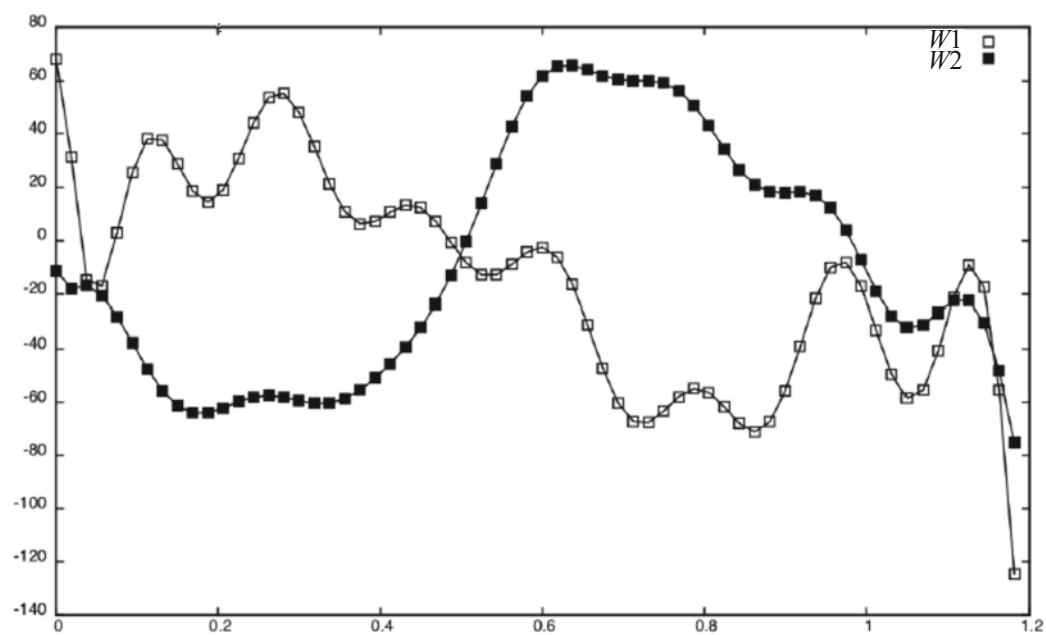
Фиг. 1. Процесс гашения колебаний, $T = 2.0$.



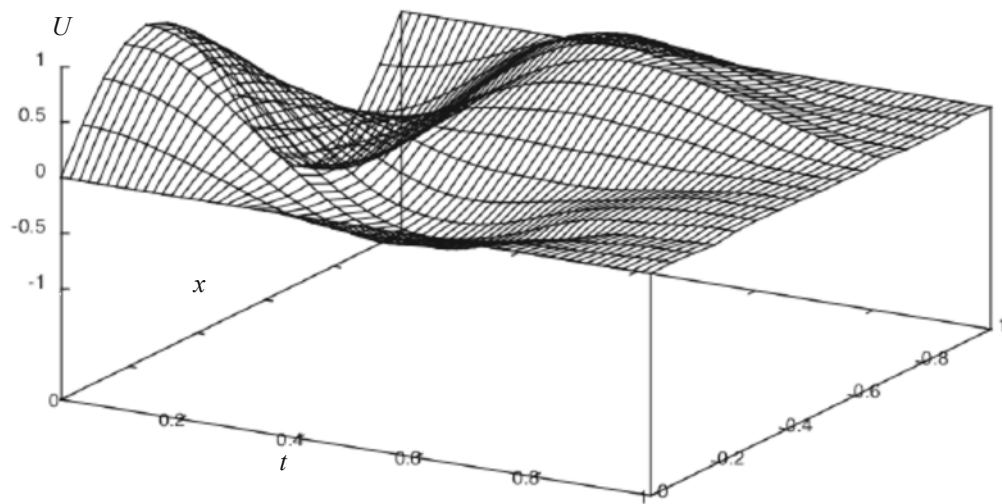
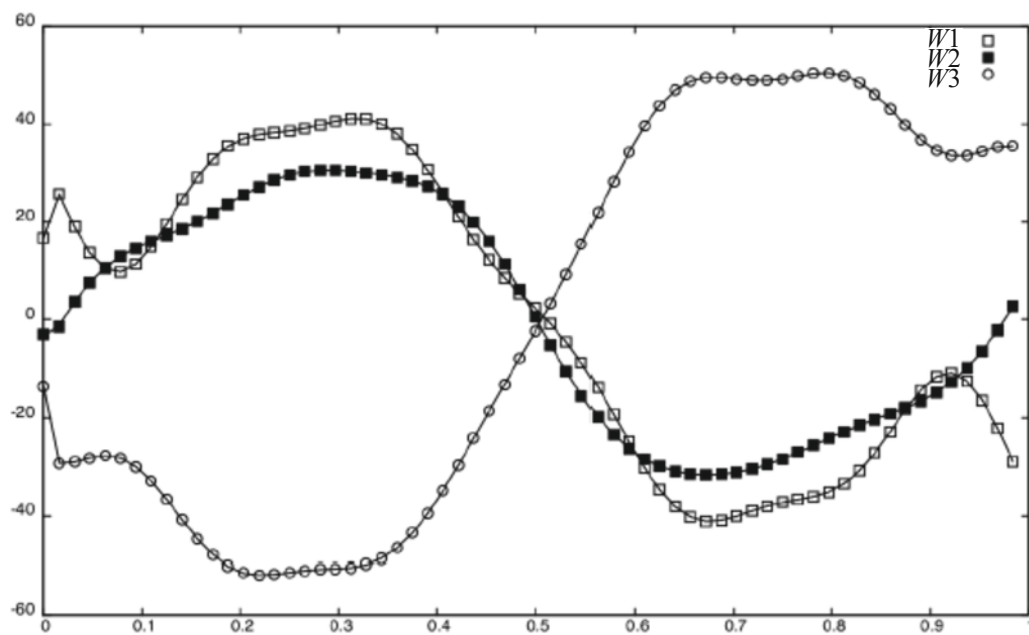
Фиг. 2. Оптимальное управление $W_1(t)$.



Фиг. 3. Процесс гашения колебаний, $T = 1.2$.



Фиг. 4. Оптимальные управления $W_1(t)$, $W_2(t)$.

Фиг. 5. Процесс гашения колебаний, $T=1.0$.Фиг. 6. Оптимальные управления $W_1(t)$, $W_2(t)$, $W_3(t)$.

СПИСОК ЛИТЕРАТУРЫ

1. Бутковский А.Г. Методы управления системами с распределенными параметрами. М.: Наука, 1975. 366 с.
2. Russel D.L. Controllability and stabilization theory for linear partial differential equations // SIAM Rev. 1978. V. 20. № 5. P. 639–739.
3. Lagness J. Control of wave process with distributed controls supported on a subregion // SIAM J. Control and Optim. 1983. V. 1. № 1. P. 68–85.
4. Асланов С.Ж., Михайлов И.Е., Муравей Л.А. Аналитические и численные методы в задаче гашения колебаний струны точечным демпфером // Мехатроника, автоматизация, управление. 2006. № 7. С. 28–35.
5. Levinson N. Gap and density theorem // Am. Math. Soc. Colog. Publ. 1940. V. 26.
6. Евтушенко Ю.Г. Оптимизация и быстрое автоматическое дифференцирование. М.: ВЦ им. А.А. Дородницына РАН, 2013. 144 с.

ON THE NUMERICAL DAMPING OF STRING VIBRATIONS USING SEVERAL STATIONARY ACTUATORS

I. E. Mikhailov^{a,b,*}

^a 119333 Moscow, Vavilov Str., 44, Federal Research Center Computer Science and Control, RAS, Russia

^b 105005 Moscow, 2nd Baumanskaya str., 5, building 1, Bauman Moscow State Technical University, Russia

*e-mail: mikh_igor@mail.ru

Received: 17.03.2024

Revised: 17.03.2024

Accepted: 31.05.2024

Abstract. The task is to transfer the string from the initial disturbed state to a state of rest in the shortest possible time. The damping of the string vibrations is carried out using several stationary actuators. The minimized functional is a certain integral. Vibration damping is controlled using a function included in the right part of the hyperbolic equation describing the transverse vibrations of the string and simulating the actions of the actuators. Computational algorithms for solving the problem based on the grid method and the gradient method of finding the minimum of functions of many variables have been developed, and the gradient is calculated using the method of rapid automatic differentiation proposed by Yu. G. Yevtushenko. Examples of calculations of string vibration damping using a different number of actuators are given.

Keywords: damping of string vibrations, rapid automatic differentiation.

ЯВНО-НЕЯВНАЯ ЧИСЛЕННАЯ СХЕМА ДЛЯ ЗАДАЧ ДИНАМИКИ УПРУГОВЯЗКОПЛАСТИЧЕСКИХ СРЕД С РАЗУПРОЧНЕНИЕМ¹⁾

© 2024 г. А.В. Шевченко^{1,2}, И.С. Никитин², В.И. Голубев^{1,2,*}, И.Б. Петров¹

¹ 141701 Долгопрудный, М.о., Институтский пер., 9, МФТИ (НИУ), Россия

² 123056 Москва, ул. 2-ая Брестская, 19/18, ИАП РАН, Россия

*e-mail: golubev.vi@mpt.ru

Поступила в редакцию 04.03.2024 г.

Переработанный вариант 25.04.2024 г.

Принята к публикации 31.05.2024 г.

Построена явно- неявная схема для численного решения определяющей системы уравнений упруговязкопластической модели сплошной среды с учетом эффекта разупрочнения. Схема включает явную аппроксимацию уравнений движения и неявную аппроксимацию определяющих соотношений, содержащих малый параметр времени релаксации в знаменателе нелинейных свободных членов. Получены точные корректировочные формулы для девиаторов напряжений после упругого шага расчета в случае линейной функции вязкости и нелинейного закона разупрочнения. Полученные решения неявной аппроксимации для девиаторов напряжений рассматриваемой упруговязкопластической системы допускают предельный переход при стремлении времени релаксации к нулю. Корректировочные формулы, полученные таким предельным переходом, можно трактовать как регуляризаторы численных решений некорректных упругопластических систем с разупрочнением. Библ. 25. Фиг. 4.

Ключевые слова: математическое моделирование, упруговязкопластические среды, полулинейные гиперболические системы, явно- неявные схемы повышенного порядка.

DOI: 10.31857/S0044466924090094, EDN: WJSHRR

1. ВВЕДЕНИЕ

Упруговязкопластическая (УВП) модель описывает нестационарное деформирование неупругой сплошной среды, которая проявляет вязкие свойства за пределом упругости (см. [1]–[3]). При фиксированном значении статического предела текучести, который определяет поверхность текучести Мизеса, динамический предел текучести зависит от скорости деформаций (см. [3], [4]). Дифференциальная часть нестационарных систем уравнений упруговязкопластических сред (см. [2], [4]) совпадает с системой уравнений динамической теории упругости, а нелинейные свободные члены определяются вязкопластическими составляющими скорости деформаций и содержат характерное время релаксации τ в знаменателе.

В нестационарных процессах с характерным временем много большим, чем τ упруговязкопластические среды ведут себя как упругопластические (УП) (см. [4], [5]), а при $\tau \rightarrow 0$ УВП системы уравнений переходят в УП системы типа Прандтля–Рейса (см. [5]–[8]).

УП системы уравнений для моделей типа Прандтля–Рейса, к которым сводятся теории пластического течения, не являются дивергентными. Для описания сильных разрывов необходимо строить обобщенные решения, используя формулировки на основе вариационных неравенств (см. [7], [8]). В то же время УВП системы уравнений имеют дивергентную форму, которая позволяет строить консервативные численные методы их решения (см. [9]).

Обе эти модели допускают обобщения для учета эффекта изотропного кинематического упрочнения, который заключается в увеличении статического предела текучести с ростом интенсивности пластических (вязкопластических) деформаций (см. [2], [4]). В этом случае обе системы уравнений остаются гиперболическими, постановки начально-краевых задач динамического нагружения деформируемых тел являются корректными

¹⁾ Работа выполнена при финансовой поддержке РНФ (код проекта 19-71-10060), <https://rscf.ru/project/19-71-10060/>.

(см. [10], [11]). Обзор численных методов решения УП и УВП задач путем проектирования на поверхность текучести для УП и УВП моделей общего вида с упрочнением сделан в работе [12].

Однако некоторые материалы (например, различные геосреды) проявляют эффект разупрочнения, который заключается в уменьшении статического предела текучести с ростом пластических деформаций, что физически связано с увеличением числа микродефектов в единице объема при неупругом деформировании. В этом случае операторы УП теории на шаге по времени теряют свойство положительной определенности. Критерий Адамара корректности краевых задач нарушается, и пошаговые алгоритмы численного решения не работают (см. [13]).

Операторы УВП теории свободны от этого недостатка и допускают корректную постановку начально-краевых задач даже для случая разупрочняющегося материала. Современный обзор УП и УВП моделей, в том числе, включающих эффект разупрочнения, для описания механизмов повреждения и развития полос скольжения в мягких грунтах, а также методов численного моделирования этих эффектов сделан в работе [14]. В работе [15] исследованы неявные алгоритмы интегрирования по времени для УВП системы с упрочнением/разупрочнением, в которых для решения нелинейной системы неявной аппроксимации применяются итерационные методы различного типа (Ньютона–Рафсона, Катона, Ортиц–Симо).

Ранее был предложен явно-неявный численный метод решения динамической системы уравнений для упрочняющейся УВП среды, в том числе, повышенного порядка, и для общего вида функций вязкости (см. [16]–[18]). Этот метод подразумевает аналитическое решение нелинейной алгебраической системы уравнений неявной аппроксимации, содержащих в знаменателе параметр (в том числе, и малый) времени релаксации, без применения каких-либо итерационных процедур. В результате получены корректировочные формулы для девиаторов напряжений в случае идеальной (см. [16], [17]) и упрочняющейся (см. [18]) УВП среды.

В настоящей работе аналогичная явно-неявная схема построена для нестационарной модели разупрочняющейся УВП среды. Этот эффект описывается с помощью динамического уравнения для параметра разупрочнения, который растет с ростом интенсивности вязкопластических деформаций. Предел текучести при этом падает по заданному нелинейному закону в зависимости от параметра разупрочнения. В результате аналитического решения алгебраических уравнений неявной аппроксимации второго порядка получены формулы для корректировки девиаторов напряжений и параметра разупрочнения после упругого шага расчета.

Таким образом, достоинством УВП систем уравнений и полученных на их основе явно-неявных численных схем является то, что они отражают физику динамических процессов деформирования, обеспечивают регуляризацию неединственных УП систем уравнений типа Прандтля–Рейса и допускают построение корректных численных методов для расчета деформирования материала с разупрочнением.

2. ОПРЕДЕЛЯЮЩИЕ УРАВНЕНИЯ ИЗОТРОПНОЙ УВП СПЛОШНОЙ СРЕДЫ С РАЗУПРОЧНЕНИЕМ

Система уравнений изотропной УВП среды с разупрочнением при малых деформациях имеет вид

$$\begin{aligned}\rho \frac{\partial \mathbf{v}}{\partial t} &= \nabla \cdot \boldsymbol{\sigma}, \\ \frac{\partial \boldsymbol{\sigma}}{\partial t} &= \left(\lambda + \frac{2}{3} \mu \right) \mathbf{e} : \mathbf{I}, \\ \frac{\partial \mathbf{s}}{\partial t} &= 2\mu \mathbf{e}' - 2\mu \frac{\mathbf{s}}{\sqrt{\mathbf{s} : \mathbf{s}}} \left\langle F \left(\frac{\sqrt{\mathbf{s} : \mathbf{s}}}{\sigma_s(\kappa)} - 1 \right) \right\rangle / \tau, \\ \frac{\partial \kappa}{\partial t} &= \sqrt{\mathbf{e}'_{vp} : \mathbf{e}'_{vp}}, \sigma_s(\kappa) = \sigma_0 f(\kappa), \\ \mathbf{e}'_{vp} &= \frac{\mathbf{s}}{\sqrt{\mathbf{s} : \mathbf{s}}} \left\langle F \left(\frac{\sqrt{\mathbf{s} : \mathbf{s}}}{\sigma_s(\kappa)} - 1 \right) \right\rangle / \tau, \mathbf{s} = \boldsymbol{\sigma} - \sigma \mathbf{I}, \sigma = (\boldsymbol{\sigma} : \mathbf{I}) / 3, \\ \mathbf{e} &= (\nabla \otimes \mathbf{v} + \mathbf{v} \otimes \nabla) / 2, \mathbf{e}' = \mathbf{e} - (\mathbf{e} : \mathbf{I}) \mathbf{I} / 3,\end{aligned}\tag{1}$$

где $\mathbf{v}$ — вектор скорости, $\boldsymbol{\sigma}$ — тензор напряжений, $\mathbf{s}$ — девиатор напряжений, σ — среднее напряжение, $\mathbf{e}$ — тензор скорости деформации, $\mathbf{e}'$ — девиатор тензора скорости деформации, $\mathbf{e}'_{vp}$ — девиатор тензора скорости вязкопластической деформации, $\sqrt{\mathbf{s} : \mathbf{s}}$ — второй инвариант девиатора напряжений, $\sigma_s(\kappa)$ — переменный предел текучести, зависящий от параметра разупрочнения κ , σ_0 — начальное значение предела текучести (до начала процесса разупрочнения), $F \left(\frac{\sqrt{\mathbf{s} : \mathbf{s}}}{\sigma_s(\kappa)} - 1 \right)$ — в общем случае нелинейная функция вязкости, $F \geq 0$, $F(0) = 0$, $\langle F \rangle = FH(F)$, $H(F)$ — функция Хэвисайда, $f(\kappa)$ — функция разупрочнения, $f(0) = 1$, $f'(\kappa) < 0$, τ — характерное время релаксации компонент девиатора напряжений на поверхность текучести, ρ — плотность среды, λ и μ — модули упругости Ламе.

Система уравнений (1) рассматривается в параллелепипедной области $D : 0 \leq x \leq L_x, 0 \leq y \leq L_y, 0 \leq z \leq L_z$, где L_x, L_y, L_z — ее размеры вдоль соответствующих координатных осей. Дополнительно задаются следующие начальные условия: $\mathbf{v}(x, y, z, t = 0) = \mathbf{0}$, $\boldsymbol{\sigma}(x, y, z, t = 0) = \mathbf{0}$, $\kappa(x, y, z, t = 0) = 0$ и граничные условия $(\boldsymbol{\sigma} \cdot \vec{n})|_{\partial D} = \mathbf{f}$. Здесь $\mathbf{n}$ — вектор внешней нормали, $\mathbf{f}$ — приложенная к границе области ∂D сила.

С учетом выражения для девиатора $\mathbf{e}'_{vp}$ эволюционное уравнение для параметра разупрочнения примет вид

$$\frac{\partial \kappa}{\partial t} = \left\langle F \left(\frac{\sqrt{\mathbf{s} : \mathbf{s}}}{\sigma_s(\kappa)} - 1 \right) \right\rangle / \tau.$$

3. НЕЯВНАЯ АППРОКСИМАЦИЯ ВТОРОГО ПОРЯДКА ДЛЯ ПРОИЗВОЛЬНОЙ ФУНКЦИИ РАЗУПРОЧНЕНИЯ

Построим неявную аппроксимацию второго порядка уравнения для девиаторов тензора напряжений с нелинейным свободным членом из УВП системы с разупрочнением (1) путем аппроксимации правой части полулинейных уравнений в виде полусуммы слагаемых на верхнем и нижнем слоях по времени:

$$\begin{aligned} \frac{\mathbf{s}^{n+1} - \mathbf{s}^n}{\Delta t} &= 2\mu \frac{\mathbf{e}'^{n+1} + \mathbf{e}'^n}{2} - \frac{2\mu}{\tau} \left(\frac{\mathbf{s}^{n+1}}{\sqrt{\mathbf{s}^{n+1} : \mathbf{s}^{n+1}}} \left\langle F \left(\frac{\sqrt{\mathbf{s}^{n+1} : \mathbf{s}^{n+1}}}{\sigma_0 f(\kappa^{n+1})} - 1 \right) \right\rangle + \frac{\mathbf{s}^n}{\sqrt{\mathbf{s}^n : \mathbf{s}^n}} \left\langle F \left(\frac{\sqrt{\mathbf{s}^n : \mathbf{s}^n}}{\sigma_0 f(\kappa^n)} - 1 \right) \right\rangle \right) / 2, \\ \frac{\kappa^{n+1} - \kappa^n}{\Delta t} &= \frac{1}{\tau} \left(\left\langle F \left(\frac{\sqrt{\mathbf{s}^{n+1} : \mathbf{s}^{n+1}}}{\sigma_0 f(\kappa^{n+1})} - 1 \right) \right\rangle + \left\langle F \left(\frac{\sqrt{\mathbf{s}^n : \mathbf{s}^n}}{\sigma_0 f(\kappa^n)} - 1 \right) \right\rangle \right) / 2. \end{aligned} \quad (2)$$

Индексами n и $n+1$ помечены значения искомых величин на верхнем и нижнем слоях разбиения по времени, Δt — шаг по времени. Эту нелинейную систему уравнений для безразмерных $\bar{\mathbf{s}}^{n+1}$ можно записать в виде

$$\begin{aligned} \bar{\mathbf{s}}^{n+1} &= \bar{\mathbf{s}}_e^{n+1} - \frac{1}{\delta} \left(\frac{\bar{\mathbf{s}}^{n+1}}{\sqrt{\bar{\mathbf{s}}^{n+1} : \bar{\mathbf{s}}^{n+1}}} \left\langle F \left(\frac{\sqrt{\bar{\mathbf{s}}^{n+1} : \bar{\mathbf{s}}^{n+1}}}{f(\kappa^{n+1})} - 1 \right) \right\rangle + \frac{\bar{\mathbf{s}}^n}{\sqrt{\bar{\mathbf{s}}^n : \bar{\mathbf{s}}^n}} \left\langle F \left(\frac{\sqrt{\bar{\mathbf{s}}^n : \bar{\mathbf{s}}^n}}{f(\kappa^n)} - 1 \right) \right\rangle \right), \\ \kappa^{n+1} - \kappa^n &= \frac{1}{\beta_2} \left(\left\langle F \left(\frac{\sqrt{\bar{\mathbf{s}}^{n+1} : \bar{\mathbf{s}}^{n+1}}}{f(\kappa^{n+1})} - 1 \right) \right\rangle + \left\langle F \left(\frac{\sqrt{\bar{\mathbf{s}}^n : \bar{\mathbf{s}}^n}}{f(\kappa^n)} - 1 \right) \right\rangle \right). \end{aligned} \quad (3)$$

Здесь введены безразмерные величины $\bar{\mathbf{s}}^{n+1} = \mathbf{s}^{n+1}/\sigma_0$, $\bar{\mathbf{s}}^n = \mathbf{s}^n/\sigma_0$, $\bar{\mathbf{s}}_e^{n+1} = \mathbf{s}_e^{n+1}/\sigma_0$, где $\mathbf{s}_e^{n+1} = \mathbf{s}^n + \mu(\mathbf{e}'^{n+1} + \mathbf{e}'^n)\Delta t$ — компоненты девиатора после упругого шага по времени, $\delta = \frac{\tau}{\Delta t} \frac{\sigma_0}{\mu}$ — безразмерный, возможно малый, параметр системы уравнений, $\beta_2 = \frac{2\tau}{\Delta t}$ — еще один безразмерный, возможно, малый параметр системы уравнений. В данном построении, за исключением особо оговоренных случаев, мы не будем устанавливать ограничений на величины δ и β_2 .

Отметим, что с учетом проведенного ранее явного шага расчета уравнений движения и значений компонент скорости на верхнем временном слое, компоненты «упругого» девиатора $\bar{\mathbf{s}}_e^{n+1}$ также можно считать известными, как и значения $\bar{\mathbf{s}}^n$ на n -ом слое.

Нелинейную систему, из которой необходимо найти неизвестные компоненты девиатора напряжений $\bar{\mathbf{s}}^{n+1}$ и параметр разупрочнения κ^{n+1} на верхнем слое, можно записать следующим образом:

$$\begin{aligned} \bar{\mathbf{s}}^{n+1} \left(\delta + \frac{\left\langle F \left(\frac{\sqrt{\bar{\mathbf{s}}^{n+1} : \bar{\mathbf{s}}^{n+1}}}{f(\kappa^{n+1})} - 1 \right) \right\rangle}{\sqrt{\bar{\mathbf{s}}^{n+1} : \bar{\mathbf{s}}^{n+1}}} \right) + \bar{\mathbf{s}}^n \frac{\left\langle F \left(\frac{\sqrt{\bar{\mathbf{s}}^n : \bar{\mathbf{s}}^n}}{f(\kappa^n)} - 1 \right) \right\rangle}{\sqrt{\bar{\mathbf{s}}^n : \bar{\mathbf{s}}^n}} &= \delta \bar{\mathbf{s}}_e^{n+1}, \\ \kappa^{n+1} &= \kappa^n + \frac{1}{\beta_2} \left\langle F \left(\frac{\sqrt{\bar{\mathbf{s}}^{n+1} : \bar{\mathbf{s}}^{n+1}}}{f(\kappa^{n+1})} - 1 \right) \right\rangle + \frac{1}{\beta_2} \left\langle F \left(\frac{\sqrt{\bar{\mathbf{s}}^n : \bar{\mathbf{s}}^n}}{f(\kappa^n)} - 1 \right) \right\rangle. \end{aligned} \quad (4)$$

Сворачивая уравнение для девиаторов напряжений последовательно с $\bar{\mathbf{s}}^{n+1}$, $\bar{\mathbf{s}}^n$, $\bar{\mathbf{s}}_e^{n+1}$, получим нелинейную алгебраическую систему уравнений для свертков $X = \sqrt{\bar{\mathbf{s}}^{n+1} : \bar{\mathbf{s}}^{n+1}}$, $Y = \sqrt{\bar{\mathbf{s}}^{n+1} : \bar{\mathbf{s}}^n}$ и $Z = \sqrt{\bar{\mathbf{s}}^{n+1} : \bar{\mathbf{s}}_e^{n+1}}$ с уже вычисленными правыми частями $T = \sqrt{\bar{\mathbf{s}}^n : \bar{\mathbf{s}}^n}$, $S = \sqrt{\bar{\mathbf{s}}^n : \bar{\mathbf{s}}_e^{n+1}}$, $\Sigma = \sqrt{\bar{\mathbf{s}}_e^{n+1} : \bar{\mathbf{s}}_e^{n+1}}$.

Из этой системы находим, что

$$\begin{aligned} \delta X + \left\langle F(X/f(\kappa^{n+1}) - 1) \right\rangle &= \delta \tilde{S}, \\ \beta_2 \kappa^{n+1} - \left\langle F(X/f(\kappa^{n+1}) - 1) \right\rangle &= \beta_2 \tilde{\kappa}^n, \end{aligned} \quad (5)$$

где

$$\tilde{s}^{n+1} = \bar{s}_e^{n+1} - \frac{\bar{s}^n}{T} \frac{\langle F(T/f(\kappa^n) - 1) \rangle}{\delta}, \quad \tilde{\kappa}^n = \kappa^n + \frac{1}{\beta_2} \langle F(T/f(\kappa^n) - 1) \rangle, \quad \tilde{S} = \sqrt{\tilde{s}^{n+1} : \bar{s}^{n+1}}, \quad T = \sqrt{\bar{s}^n : \bar{s}^n}.$$

Нетрудно видеть, что промежуточный девиатор $\tilde{s}^{n+1}$ и его свертка $\tilde{S}$ вычисляются по результатам упругого шага по времени. Из (4) и (5) следует формула для искомых компонент девиатора и параметра упрочнения на верхнем временном слое в следующем виде:

$$\begin{aligned} \bar{s}^{n+1} &= \frac{\delta \tilde{s}^{n+1}}{\left(\delta + \frac{\langle F(X/f(\kappa^{n+1}) - 1) \rangle}{X} \right)} = \frac{\tilde{s}^{n+1}}{\tilde{S}} X, \\ \kappa^{n+1} &= \tilde{\kappa}^n + \delta \left(\tilde{S} - X \right) / \beta_2. \end{aligned} \quad (6)$$

Таким образом, для полного определения компонент девиатора $\bar{s}^{n+1}$ и параметра разупрочнения κ^{n+1} следует решить уравнение (5) и подставить результат в (6). После этого можно совершить предельный переход при $\delta \rightarrow 0$ в формуле для промежуточного девиатора $\tilde{s}^{n+1}$, которая содержит малый параметр в знаменателе. Для этого первое уравнение в (5), справедливое для $n + 1$ -го слоя по времени, перепишем для n -го слоя:

$$\delta T + \langle F(T/f(\kappa^n) - 1) \rangle = \delta \sqrt{\bar{s}^n : \bar{s}^n}. \quad (7)$$

Тогда в формуле для промежуточного девиатора, с учетом того, что $\delta/\beta_2 = \sigma_0/(2\mu)$, удастся снять особенность при $\delta \rightarrow 0$:

$$\begin{aligned} \tilde{s}^{n+1} &= \bar{s}_e^{n+1} - \frac{\bar{s}^n}{\sqrt{\bar{s}^n : \bar{s}^n}} \left\langle \sqrt{\tilde{s}^n : \tilde{s}^n} - \sqrt{\bar{s}^n : \bar{s}^n} \right\rangle, \\ \kappa^{n+1} &= \tilde{\kappa}^n + \sigma_0 \left(\tilde{S} - X \right) / (2\mu), \\ \tilde{\kappa}^n &= \kappa^n + \sigma_0 \left\langle \sqrt{\tilde{s}^n : \tilde{s}^n} - \sqrt{\bar{s}^n : \bar{s}^n} \right\rangle / (2\mu). \end{aligned} \quad (8)$$

4. СЛУЧАЙ ЛИНЕЙНОЙ ФУНКЦИИ ВЯЗКОСТИ И НЕЛИНЕЙНОГО РАЗУПРочНЕНИЯ

Для того чтобы решить нелинейную алгебраическую систему (5) и получить явные выражения для неизвестных величин $\bar{s}^{n+1}$ и κ^{n+1} , необходимо конкретизировать функцию вязкости $F(x)$ и функцию разупрочнения $f(\kappa)$.

Для получения аналитического решения примем простую линейную зависимость $F(x) = x$. При этом не будем ограничивать диапазон значений безразмерных параметров $\beta_2 \geq 0$, $\delta \geq 0$. Также рассмотрим случай нелинейной функции $f(\kappa)$, которая фактически (в силу убывания) описывает процесс разупрочнения, так как при $\alpha > 0$ имеем падение предела текучести:

$$f(\kappa) = \frac{1}{1 + \alpha \kappa}, \quad (9)$$

где величина параметра α регулирует режим и скорость изменения предела текучести. В этих предположениях система (5) примет вид

$$\begin{aligned} \delta X + \langle X(1 + \alpha \kappa^{n+1}) - 1 \rangle &= \delta \tilde{S}, \\ \beta_2 \kappa^{n+1} - \langle X(1 + \alpha \kappa^{n+1}) - 1 \rangle &= \beta_2 \tilde{\kappa}^n. \end{aligned} \quad (10)$$

Ее аналитическое решение приводит к формулам для девиатора и параметра разупрочнения на верхнем слое по времени вида ($\gamma = \alpha/\beta_2$)

$$\begin{aligned} \bar{s}^{n+1} &= \frac{\tilde{s}^{n+1}}{\tilde{S}} \left(\tilde{S} - \frac{\sqrt{\left(1 + \alpha \tilde{\kappa}^n + \delta(1 - \gamma \tilde{S})\right)^2 + 4\gamma\delta \left(\tilde{S}(1 + \alpha \tilde{\kappa}^n) - 1\right)} - \left(1 + \alpha \tilde{\kappa}^n + \delta(1 - \gamma \tilde{S})\right)}{2\gamma\delta} \right), \\ \kappa^{n+1} &= \tilde{\kappa}^n + \frac{\delta}{\beta_2} \left(\frac{\sqrt{\left(1 + \alpha \tilde{\kappa}^n + \delta(1 - \gamma \tilde{S})\right)^2 + 4\gamma\delta \left(\tilde{S}(1 + \alpha \tilde{\kappa}^n) - 1\right)} - \left(1 + \alpha \tilde{\kappa}^n + \delta(1 - \gamma \tilde{S})\right)}{2\gamma\delta} \right). \end{aligned} \quad (11)$$

Данные формулы могут быть значительно упрощены, если предположить, что коэффициент разупрочнения α мал, точнее справедливо $\gamma \ll 1$. Тогда итоговые упрощенные формулы будут иметь вид

$$\begin{aligned} X &\approx \tilde{S} - \frac{\tilde{S}(1 + \alpha\tilde{\kappa}^n) - 1}{\left(1 + \alpha\tilde{\kappa}^n + \delta(1 - \gamma\tilde{S})\right)}, \\ \bar{s}^{n+1} &\approx \frac{\tilde{s}^{n+1}}{\tilde{S}} \left(\tilde{S} - \frac{\tilde{S}(1 + \alpha\tilde{\kappa}^n) - 1}{\left(1 + \alpha\tilde{\kappa}^n + \delta(1 - \gamma\tilde{S})\right)} \right), \\ \kappa^{n+1} &\approx \tilde{\kappa}^n + \frac{\delta}{\beta_2} \frac{\tilde{S}(1 + \alpha\tilde{\kappa}^n) - 1}{\left(1 + \alpha\tilde{\kappa}^n + \delta(1 - \gamma\tilde{S})\right)}. \end{aligned} \quad (12)$$

Данные формулы допускают предельный переход при $\delta \rightarrow 0$:

$$\begin{aligned} \bar{s}^{n+1} &\approx \frac{\tilde{s}^{n+1}}{\tilde{S}(1 + \alpha\tilde{\kappa}^n)}, \\ \kappa^{n+1} &\approx \tilde{\kappa}^n + \frac{\sigma_0}{2\mu} \frac{\tilde{S}(1 + \alpha\tilde{\kappa}^n) - 1}{(1 + \alpha\tilde{\kappa}^n)}. \end{aligned} \quad (13)$$

При этом условие реализации УВП режима имеет вид $\tilde{S}(1 + \alpha\tilde{\kappa}^n) - 1 \geq 0$.

5. СХЕМА РАСЧЕТА УПРУГОГО ШАГА

Система уравнений задачи линейной теории упругости для изотропного тела имеет вид

$$\frac{\partial \mathbf{u}}{\partial t} + \sum_{i=1,2,3} A_{x_i}(\rho, \lambda, \mu) \frac{\partial \mathbf{u}}{\partial x_i} = \mathbf{0}. \quad (14)$$

Вектор искомых функций $\mathbf{u}$ включает в себя компоненты скорости смещения и компоненты тензора напряжений. Параметры λ , μ и ρ определяются свойствами описываемой среды. В работе используется сеточно-характеристический метод (см. [19]). Используя метод расщепления по пространственным направлениям (см. [20]), исходная задача разделяется на три одномерные линейные гиперболические системы уравнений

$$\frac{\partial \mathbf{u}}{\partial t} + A_{x_i} \frac{\partial \mathbf{u}}{\partial x_i} = \mathbf{0}, \quad i = 1, 2, 3. \quad (15)$$

Каждая матрица A_{x_i} может быть представлена в виде

$$\Lambda = \Omega_{x_i} A_{x_i} \Omega_{x_i}^{-1}, \quad (16)$$

где матрицы Ω_{x_i} и $\Omega_{x_i}^{-1}$ вычисляются аналитически, а матрица Λ диагональная и состоит из собственных значений матрицы A_{x_i} . Рассмотрим случай однородной среды, когда матрицы A_{x_i} не зависят от координат. В дальнейшем в выкладках опустим индекс x_i у матриц. Переход от неизвестных функций $\mathbf{u}$ к вектору из инвариантов Римана $\boldsymbol{\omega} = \Omega \mathbf{u}$ приводит к системе независимых одномерных линейных уравнений переноса с постоянными коэффициентами:

$$\frac{\partial \boldsymbol{\omega}}{\partial t} + \Lambda \frac{\partial \boldsymbol{\omega}}{\partial x_i} = \mathbf{0}. \quad (17)$$

Решением линейного уравнения переноса с постоянным коэффициентом является произвольная функция $\omega_j(t, x_i, x_k, x_l) = \phi(x_i - \lambda_j t, x_k, x_l)$. Тогда справедливо следующее равенство:

$$\omega_j(t + \Delta t, x_i, x_k, x_l) = \omega_j(t, x_i - \lambda_j \Delta t, x_k, x_l). \quad (18)$$

Пусть $\lambda_j > 0$. Для $\lambda_j < 0$ дальнейший вывод может быть проделан аналогично. Зафиксируем далее набор узлов $(x_{m-2}, x_{m-1}, x_m, x_{m+1})$ в пространстве координат и по известным значениям ω_j в них восстановим однозначно коэффициенты полинома третьей степени $P_3(x)$. Здесь нижний индекс m означает номер узла в равномерной пространственной сетке вдоль координаты x_i с шагом h . Будем вычислять приближенно значения ω_j на новом временном слое по формуле $\omega_j(t + \Delta t, x_i, x_k, x_l) = \omega_j(t, x_i - \lambda_j \Delta t, x_k, x_l) \approx P_3(x_i - \lambda_j \Delta t)$. Тогда конечное

выражение для искоемых функций может быть получено как $\mathbf{u}^{n+1} = \Omega^{-1} \boldsymbol{\omega}^{n+1}$, где компоненты вектора $\boldsymbol{\omega}^{n+1}$ вычислены по формуле

$$\omega_m^{n+1} = \omega_m^n - \frac{K}{2} (\omega_{m+1}^n - \omega_{m-1}^n) + \frac{K^2}{2} (\omega_{m-1}^n - 2\omega_m^n + \omega_{m+1}^n) + \frac{K(K^2 - 1)}{6} (\omega_{m-2}^n - 3\omega_{m-1}^n + 3\omega_m^n - \omega_{m+1}^n), \quad (19)$$

где $K = \lambda_j \Delta t / h$ – число Куранта. Условие устойчивости данного алгоритма $K < 1$. Случай неоднородных сред рассмотрен подробнее в работе [21]. Метод сохранения повышенного порядка аппроксимации схемы вблизи границ расчетной области предложен, например, в работе [22].

Таким образом, вычислительный алгоритм для решения динамической задачи УВП среды может быть сформулирован следующим образом:

- по известным значениям величин $\mathbf{s}^n$, σ^n и $\mathbf{v}^n$ на текущем временном слое выполняется упругий шаг по времени с использованием сеточно-характеристического метода;
- по формулам (11) производится корректировка полученных значений $\mathbf{s}^{n+1}$ и вычисление значения κ^{n+1} на следующем временном слое.

6. РЕЗУЛЬТАТЫ ЧИСЛЕННЫХ РАСЧЕТОВ

Данный вычислительный алгоритм был применен для решения полноволновых трехмерных динамических задач. На первом этапе были выбраны две задачи, для которых в случае УП среды без разупрочнения известны аналитические решения (см. [23]). Рассматривался параллелепипед размерами $50 \times 50 \times 10\,000$ м, заполненный изотропной УВП средой. Параметры среды были следующие: скорость продольных волн $C_p = \sqrt{\frac{\lambda+2\mu}{\rho}} = 4500$ м/с, скорость поперечных волн $C_s = \sqrt{\frac{\mu}{\rho}} = 2250$ м/с, плотность $\rho = 2500$ кг/м³, $\sigma_0 = 112\,500$ Па. Задача решалась на равномерной расчетной сетке с шагом 5 м по пространству и 1 мс по времени, что соответствует выполнению условию устойчивости Куранта.

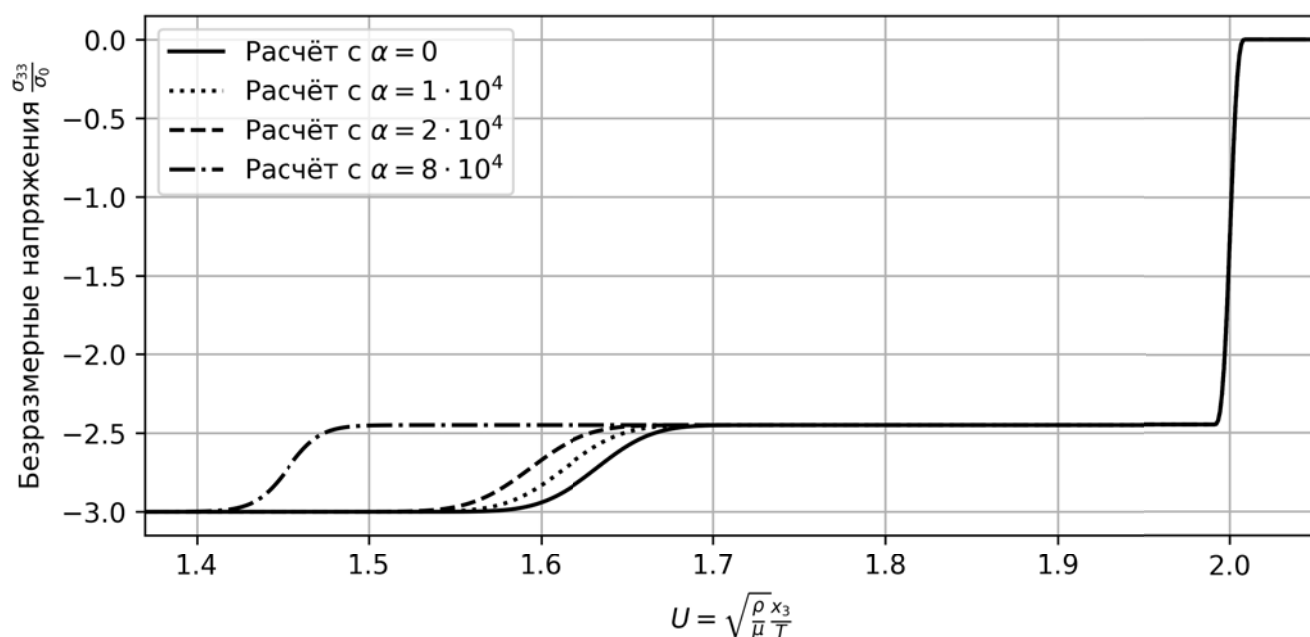
К одной границе расчетной области в начале расчета прикладывалось постоянное сжимающее напряжение величиной 337 500 Па, которое в дальнейшем расчете удерживалось постоянным. Известно аналитическое решение данной квазиодномерной задачи: для упругой среды – распространение продольной волны со скоростью C_p , для упругопластической среды без разупрочнения – распространение продольной волны со скоростью C_p и пластической волны со скоростью $C_f = \sqrt{\frac{\lambda+(2/3)\mu}{\rho}} \approx 3674$ м/с. Картины пространственного распределения напряжения σ_{33} в среде через 2 с от начала расчета приведены на фиг. 1. При этом значение параметра δ было задано равным 1. Достаточно точно разрешается разрыв в решении, соответствующий продольной волне, а также фронт пластической волны. Отметим, что разупрочнение материала приводит к снижению скорости распространения пластической волны, что видно по отставанию ее фронта при увеличении параметра α .

Далее была решена задача о совместной нормальной $\sigma_n = 249\,107$ Па и тангенциальной $\sigma_\tau = -72\,322$ Па нагрузках. Все параметры модели оставались неизменными, кроме обновленного значения $C_s = 2598$ м/с. Результаты расчетов представлены на фиг. 2. Полученные численные решения воспроизводят все типы упругих и пластических волн.

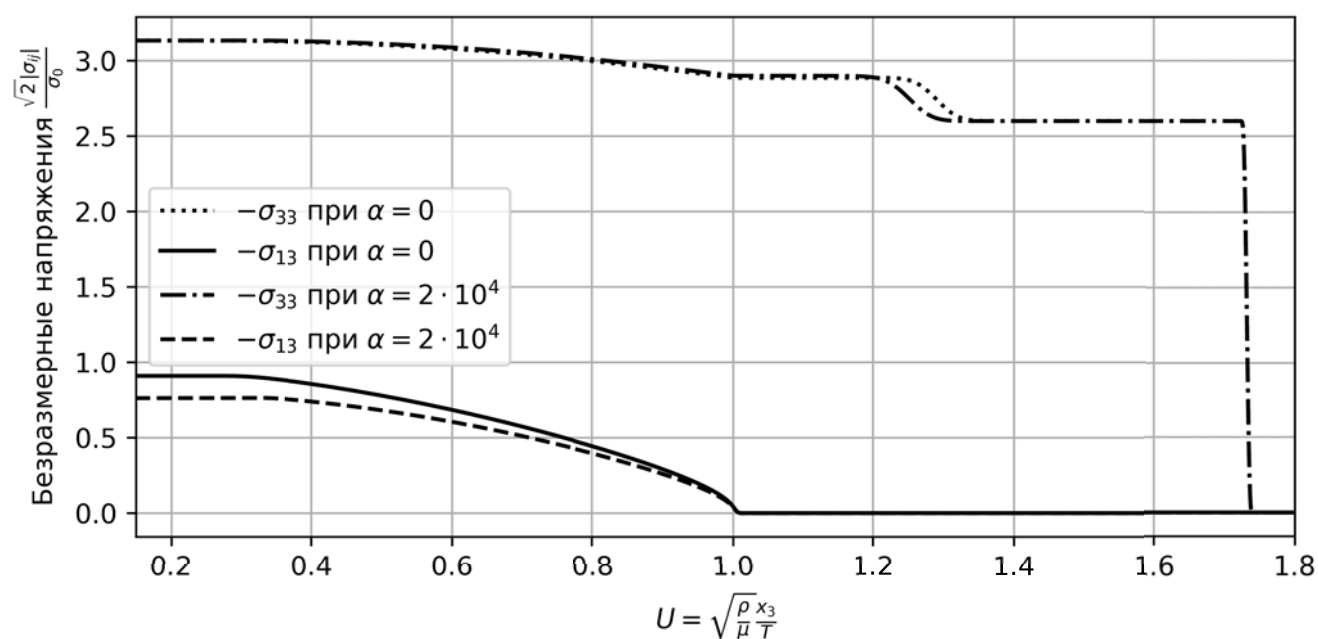
В дальнейшем была решена задача о нагружении нормальной нагрузкой полупространства $z \geq 0$, содержащего включение из отличающегося по упругим характеристикам материала. Выбор данной постановки задачи обусловлен общезначимыми соображениями о наличии концентрации напряжений вблизи контрастных границ, что должно привести к локализованному в пространстве переходу материала в пластическое состояние. Размеры рассматриваемой области были равны $2000 \times 2000 \times 1700$ м. Параметры вмещающего массива $C_p = 4800$ м/с, $C_s = 2400$ м/с, $\rho = 3125$ кг/м³. В центральной части области $x \in [700, 1300]$ м, $y \in [700, 1300]$ м, $z \in [700, 1300]$ м было задано кубическое включение с параметрами $C_p = 300$ м/с, $C_s = 150$ м/с, $\rho = 1000$ кг/м³. Величина приложенной нормальной нагрузки задавалась соотношением

$$\sigma_n = A \times \begin{cases} t/t_0, 0 \leq t \leq t_0, \\ 1, t > t_0, \end{cases} \quad (20)$$

где $t_0 = 51$ мс, амплитуда $A = -250\,000$ Па. Использовалось разбиение по пространству с шагом 5 м и по времени с шагом 1 мс. Всего было рассчитано 800 шагов по времени. Во всей расчетной области считались заданными параметры $\sigma_0 = 112\,500$ Па, $\tau = 112.5$. На фиг. 3 и фиг. 4 представлены пространственные распределения параметра разупрочнения κ на момент окончания расчета. На фиг. 3 размытая серая зона вокруг кубического

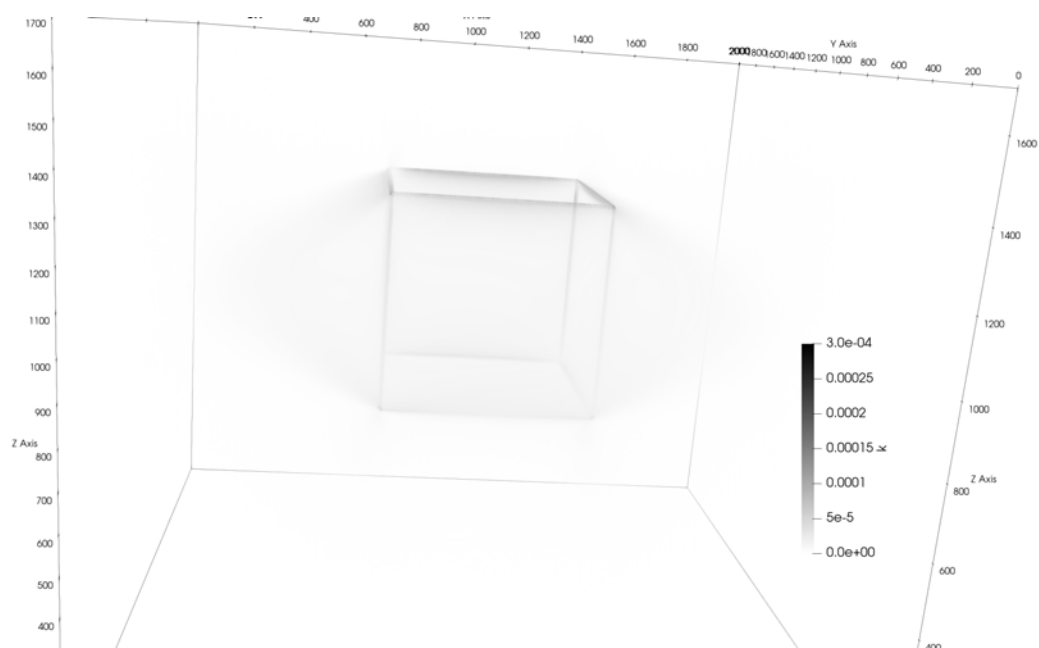


Фиг. 1. Распределение σ_{33} в УВП среде под нормальной нагрузкой через 2 с от момента ее приложения.

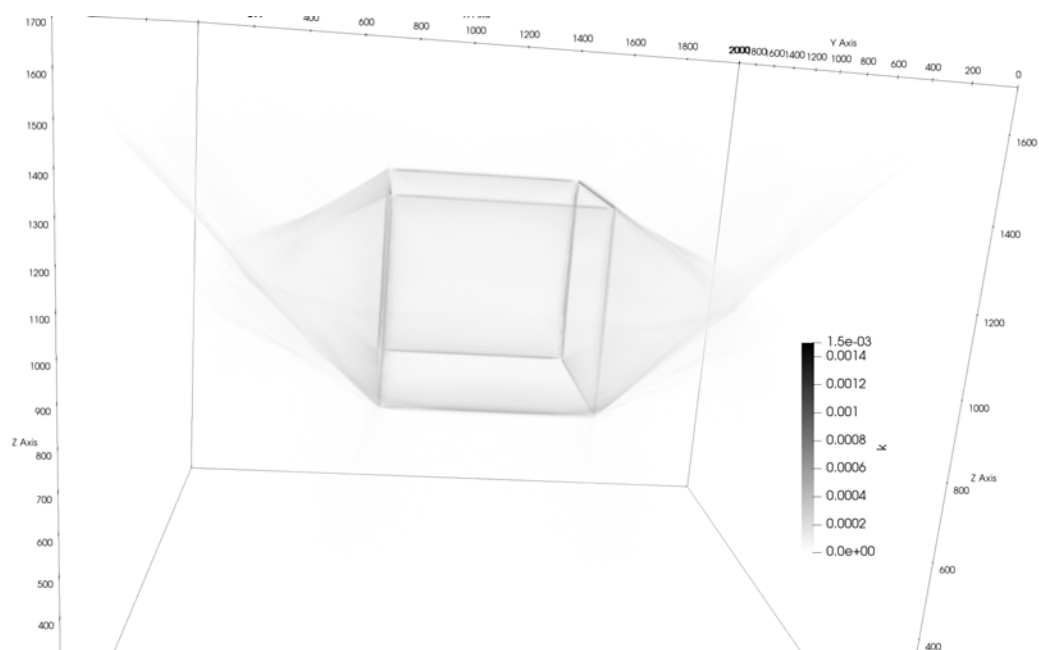


Фиг. 2. Распределение σ_{33} и σ_{13} в УВП среде под смешанной нагрузкой через 2 с от момента ее приложения.

включения в окружающей УВП среде без разупрочнения ($\alpha = 0$) представляет собой область вязкопластических деформаций ($\kappa > 0$). Она связана с релаксацией девиаторов напряжений на поверхность текучести Мизеса в некотором объеме вокруг включения, которое является концентратором напряжений. На фиг. 4 при $\alpha > 0$ (разупрочнение) хорошо видны сформировавшиеся плоскости скольжения, идущие от ребер куба. Этот эффект локализации неупругих деформаций в узких линиях (2D задачи) и плоскостях скольжения (3D задачи) характерен для сред с разупрочнением (см. [24]).



Фиг. 3. Итоговое распределение параметра разупрочнения k ($\alpha = 0$).



Фиг. 4. Итоговое распределение параметра разупрочнения k ($\alpha = 2 \times 10^4$).

7. ЗАКЛЮЧЕНИЕ

В настоящей работе был рассмотрен процесс деформирования УВП среды под действием внешней нагрузки. Отличительной особенностью рассматриваемой задачи являлся учет процесса разупрочнения материала. В общем случае произвольных функций вязкости и разупрочнения предложена процедура построения явно- неявной схемы. Подход, основанный на применении неявной схемы для дискретизации определяющих соотношений, содержащих малый параметр времени релаксации в знаменателе нелинейных свободных членов, позволил свести задачу к решению нелинейной алгебраической системы уравнений на неизвестные значения девиаторов напряжений и параметра разупрочнения.

Для специального случая линейной функции вязкости и нелинейной функции разупрочнения специального вида полученная система уравнений разрешена аналитически, что позволило построить явный расчетный алгоритм. Он опирается на численное решение определяющей системы уравнений линейной изотропной теории упругости. Для этого был использован сеточно-характеристический метод на расширенном шаблоне. Отметим, что последующая корректировочная процедура выполняется независимо в каждом узле расчетной сетки. Данное свойство построенного вычислительного алгоритма, например, не снижает эффективность использования современных CPU (см. [25]) и технологий параллельных вычислений (OpenMP, MPI).

Построенные вычислительные алгоритмы были успешно применены для численного решения трех трехмерных динамических задач: о нормальной нагрузке УВП однородного полупространства, о смешанной нагрузке УВП однородного полупространства и о нормальной нагрузке неоднородного УВП полупространства.

Отметим, что в реальных геоматериалах (грунтах) предел текучести не монотонен, с начальным участком упрочнения, затем разупрочнения и выходом на конечную асимптоту (см. [14]). Поэтому при моделировании реальных материалов необходимо в общем случае использовать комбинацию из корректировочных формул для упрочнения и разупрочнения. Также в используемой нами классической модели не учитываются объемные вязкопластические эффекты, проявляющиеся в реальных грунтах (см. [14]). Это можно отнести к ограничениям разработанного вычислительного алгоритма. В настоящей работе мы хотели остаться в рамках классической УВП модели, при этом обобщение на более сложные УВП модели может быть выполнено аналогично. К недостаткам текущей программной реализации можно отнести тот факт, что расчет упругого решения и его последующая корректировка реализованы отдельными циклами по узлам расчетной сетки, что может незначительно увеличивать расчетное время. Однако данное архитектурное решение позволило напрямую использовать имеющиеся программные наработки по линейно упругой задаче.

СПИСОК ЛИТЕРАТУРЫ

1. Фрейденталь А., Гейрингер Х. Математические теории неупругой сплошной среды. М.: Наука, 1962. 431 с.
2. Кукуджанов В.Н. Вычислительная механика сплошных сред. М.: Физматлит., 2008. 320 с.
3. Perzyna P. *Fundamental problems in viscoplasticity* // Adv. Appl. Mech. 1966. V. 9. С. 243–377.
4. Новацкий В.К. Волновые задачи теории пластичности. М.: Мир, 1978. 310 с.
5. Коларов Д., Балтов А., Бончева Н. Механика пластических сред. М.: Мир. 1979. 302 с.
6. Кукуджанов В.Н. *Распространение волн в упруговязкопластических материалах с диаграммой общего вида* // Механика твердого тела. 2001. № 5. С. 96–111.
7. Дюво Г., Лионс Н. Неравенства в механике и физике. М.: Наука, 1980. 384 с.
8. Садовский В.М. Разрывные решения в задачах динамики упругопластических сред. М., 1997. 208 с.
9. Куликовский А.Г., Погорелов Н.В., Семенов А.Ю. Математические вопросы численного решения гиперболических систем уравнений. М.: Физматлит, 2001. 600 с.
10. Wilkins M.L. *Computer simulation of dynamic phenomena*. Berlin-Heidelberg-New-York, 1999. 264 с.
11. Simo J.C., Hughes T.J. *Elastoplasticity and viscoplasticity – computational aspects*. New York: Springer, 1988.
12. Ortiz M., Simo J.C. *An analysis of a new class of integration algorithms for elastoplastic constitutive relations* // Int. J. Numer. Meth. Eng. 1986. V. 23. № 3. P. 353–366.
13. Бурого Н.Г. *Моделирование разрушения упругопластических тел* // Вычисл. механика сплошных сред. 2008. Т. 1. № 4. С. 5–20.

14. *Urmi Z.A., Saeidi A., Chavali R., Yerro A. Failure mechanism, existing constitutive models and numerical modeling of landslides in sensitive clay: a review // Geoenvironmental Disasters. 2023. V. 10. P. 14.*
15. *Yin Z., Li J., Jin Y., Liu, Feng Y. Estimation of Robustness of Time Integration Algorithms for Elasto-Viscoplastic Modeling of Soils // Inter. J. Geomechanics. 2019. V. 19. № 2. P. 0401819.*
16. *Голубев В.И., Никитин И.С., Бураго Н.Г., Голубева Ю.А. Явно- неявные схемы расчета динамики упруговязко-пластических сред с малым временем релаксации // Дифференц. ур-ния. 2023. Т. 59. № 6. С. 803–813.*
17. *Голубев В.И., Никитин И.С. Уточненные схемы расчета динамики упруговязкопластических сред // Ж. вычисл. матем. и матем. физ. 2023. Т. 63. № 10. С. 1674–1686.*
18. *Golubev V.I., Nikitin I.S., Mi X. Numerical schemes of higher approximation orders for dynamic problems of elastoviscoplastic media // J. Siberian Federal University. Math. and Phys. 2024. V. 17. № 1. P. 8–17.*
19. *Магомедов К.М., Холодов А.С. Сеточно-характеристические численные методы: учебное пособие для вузов. М.: Юрайт, 2023.*
20. *Golubev V.I., Shevchenko A.V., Petrov I.B. Raising convergence order of grid-characteristic schemes for 2D linear elasticity problems using operator splitting // Comput. Res. and Model. 2022. V. 14. Iss. 4. P. 899–910.*
21. *Khokhlov N.I., Petrov I.B. High-order grid-characteristic method for systems of hyperbolic equations with piecewise constant coefficients // Diff. Equat. 2023. V. 59. P. 985–997.*
22. *Shevchenko A.V., Golubev V.I. Boundary and contact conditions of higher order of accuracy for grid-characteristic schemes in acoustic problems // Comput. Math. and Math. Phys. 2023. V. 63. P. 1760–1772.*
23. *Bleich H.H., Nelson I.I. Plane waves in an elastic-plastic half-space due to combined surface pressure and shear // ASME J. Appl. Mech. 1966. V. 33. № 1. P. 149–158.*
24. *Бураго Н.Г., Никитин И.С. Алгоритмы сквозного счета для процессов разрушения // Компьют. исслед. и моделирование. 2018. Т. 10. № 5. С. 645–666.*
25. *Furgailo V., Ivanov A., Khokhlov N. Research of techniques to improve the performance of explicit numerical methods on the CPU // 2019 Ivannikov Memorial Workshop (IVMEM), Velikiy Novgorod, Russia, 2019. P. 79–85.*

An explicit-implicit numerical scheme for problems of dynamics of elastic-viscoplastic media with softening

A. V. Shevchenko^{a,b}, I. S. Nikitin^b, V. I. Golubev^{a,b,*}, I. B. Petrov^a

^a 141701 Dolgoprudny, Moscow region, Institutsky Lane, 9, Moscow Institute of Physics and Technology (NRU), Russia

^b 123056 Moscow, 2nd street, Brestskaya, 19/18, IAP RAS, Russia,

*e-mail: golubev.vi@mipt.ru

Received: 04.03.2024

Revised: 25.04.2024

Accepted: 31.05.2024

Abstract. An explicit-implicit scheme is constructed for the numerical solution of the defining system of equations of the elastic-viscoplastic continuum model, taking into account the softening effect. The scheme includes an explicit approximation of the equations of motion and an implicit approximation of the defining relations containing a small relaxation time parameter in the denominator of the free terms. Precise correction formulas for stress deviators are obtained after the elastic calculation step in the case of the linear viscosity function and the nonlinear law of softening. The obtained solutions of the implicit approximation for the stress deviators of the elastic-viscoplastic system under consideration allow for a limiting transition when the relaxation time tends to zero. Correction formulas, obtained by such a limiting transition, can be interpreted as regularizers of numerical solutions of incorrect elastoplastic systems with a softening effect.

Keywords: mathematical modeling, elastic-viscoplastic media, semi-linear hyperbolic systems, explicit-implicit schemes of increased order.

Посвящается профессору Евгению Михайловичу Шахову в связи с его 90-летием

РЕШЕНИЕ ЗАДАЧ КИНЕТИЧЕСКОГО ВЗАИМОДЕЙСТВИЯ БЫСТРЫХ ГРУПП ЧАСТИЦ С ИСПОЛЬЗОВАНИЕМ АНАЛИТИЧЕСКИХ И ЧИСЛЕННЫХ МЕТОДОВ

© 2024 г. В.В. Аристов^{1,*}, И.В. Воронич^{1,**}

¹119333 Москва, ул. Вавилова, 44, ФИЦ ИУ РАН, Россия

*e-mail: aristovvl@yandex.ru

**e-mail: i.voronich@frcsc.ru

Поступила в редакцию 15.07.2023 г.

Переработанный вариант 10.04.2024 г.

Принята к публикации 31.05.2024 г.

Приводится краткий обзор постановок задач, связанных с высокоскоростными пучками, с акцентом на использование аналитических решений, а также описаны результаты численного решения некоторых задач данного класса. В рамках кинетической теории на основе аналитического метода рассмотрены процессы взаимодействия групп частиц (молекул) в предположении высокой коррелированности скоростей частиц (дельта-функция в качестве плотности распределения). Задачи о взаимодействии пучков с выбыванием и без выбывания частиц изучаются численно с помощью метода прямого статистического моделирования. Для задачи с выбыванием частиц (пересечение и взаимодействие тонких пучков) получено хорошее согласие с аналитическим решением. Для задачи без выбывания частиц (соударение потоков) получено численное решение типа бегущей ударной волны предельного сжатия, формирующейся при соударении потока со стенкой. Показана роль ударных трансформант на начальной стадии процесса. Рассмотрена задача о проникновении пучка в покоящийся газ вплоть до стадии формирования факела, отмечено сходство ее начальной стадии с задачей о тонких пучках. Подчеркивается плодотворность использования аналитических методов на этапе первичного анализа проблемы и при верификации численных решений. Библ. 27. Фиг. 8.

Ключевые слова: кинетическое уравнение, разреженный газ, высокоскоростные течения, столкновение пучков частиц, аналитические решения, статистическое моделирование.

DOI: 10.31857/S0044466924090102, EDN: WJRVWI

1. ВВЕДЕНИЕ

Среди задач кинетической теории и динамики разреженных газов важное место всегда занимали проблемы изучения высокоскоростных течений. В 1970-е годы Е.М. Шахов предложил некоторые эффективные подходы к моделированию высокоскоростных течений разреженного газа на основе процедуры дискретных скоростей. Первый подход заключается в построении адаптивной сетки в пространстве скоростей с целью учета локальной температуры — для точек в физическом пространстве с большей температурой размер скоростной области и шаг по скорости увеличиваются. В определенном смысле это использование аналогии с методами частиц в статистическом моделировании — чем выше температура, тем большую область в пространстве скоростей занимает ансамбль частиц в ячейке и тем больше разница между скоростями частиц. Данный подход применялся, в частности, к задаче о структуре ударной волны с очень большой интенсивностью, характеризуемой числом Маха $M = 100$ [1] (полезно заметить, что эта работа была выполнена под непосредственным руководством Е.М. Шахова). В рамках принятой идеализации не рассматривались процессы, связанные с возбуждением внутренних степеней свободы, диссоциацией и ионизацией молекул. Развитый численный метод позволил существенно сократить время расчетов. В дальнейшем схемы с адаптивными сетками в пространстве скоростей развивались [2–4], в работе [2] делается ссылка на [1] как на первый пример такой схемы. Другой подход основан на допущении, что внутренняя энергия быстрых групп частиц (молекул газа), например, в пучках, пренебрежимо мала по сравнению с их кинетической энергией. Решение такого класса задач потребовало применения иных математических методов, основанных на особенностях в пространстве скоростей в виде дельта-функции Дирака, отвечающей нулевой температуре. Были построены асимптотические приближения и развит новый численный

метод, основанный на использовании обобщенных функций и их аппроксимаций. С помощью этого подхода была решена задача, которая вначале была названа задачей о сильном взрыве в разреженном газе [5], но при уточнении получила более точное определение как задача о выбросе (инжекции) высокоскоростных частиц [6]. К данной тематике можно отнести и задачи о рассеянии импульсных пучков в газе [7, 8]. Некоторые приложения этих задач были описаны в работах [5–8]: рассеяние быстрых частиц на неподвижных молекулах фона, исследование верхних слоев атмосферы с помощью инъекции пучков бария, столкновение групп частиц при очень больших скоростях и т. д.

В дальнейшем оказалось, что такие методы и результаты могут найти применение в некоторых областях, не относящихся к динамике разреженного газа, но допускающих аналогичные математические модели. В кратком обзоре можно отметить постановки задач, выходящие за рамки описанных физических явлений. В последние десятилетия получили развитие математические подходы для описания биологических, экономических, исторических и других явлений [9, 10]. Среди них используются методы статистической физики и кинетической теории [11, 12]. Модели описания распространения эпидемий и других аналогичных процессов традиционно опираются на уравнения параболического типа. В этих моделях, начиная с работ А.Н. Колмогорова, И.Г. Петровского, Н.С. Пискунова и Р. Фишера [13, 14], рассматриваются решения типа бегущей волны. Подобные явления в литературе на русском языке называются автоволновыми процессами [15, 16]. Описанный кинетический подход был применен в [17] для формулировки обобщенного уравнения, которое переходит в решение [7, 8] в частном случае постоянных плотностей фона и пучка. Данная модель была применена в работе [17] для описания процессов быстрого вторжения (блицкрига), результаты которой были отмечены в [18]. Сходная, но более простая линейная модель была применена для изучения распространения пандемии COVID-19. При изучении данной проблематики часто используются диффузионные модели распространения [19], но более адекватной видится кинетическая модель [20, 21].

2. ПОСТАНОВКИ ЗАДАЧ

Ряд рассмотренных в литературе постановок задач связан с описанием движения групп частиц, тепловой энергией (температурой) которых можно пренебречь по сравнению с их кинетической энергией. В таких задачах взаимодействие движущихся частиц происходит с частицами покоящегося фонового газа, для которого также принимается допущение нулевой температуры [6, 7]. Рассматривались одномерная, плоская и сферически-симметричная конфигурации. Первоначально постановка задач прямо соотносилась с вариантами известной задачи газовой динамики о сильном взрыве [22], которые допускают при некоторых упрощениях автомодельные решения, но изучались и численно в более общих постановках [23–25]. Достаточно быстро выяснилось, что изучаемые течения в разреженном газе имеют свою специфику, связанную с характером динамики разреженного газа на временах, сопоставимых со средним временем межмолекулярного столкновения. Поскольку рассматривается движение частиц, имеющих одинаковые по модулю скорости, температура невозмущенного пучка нулевая. В дальнейшем такого рода задачи получили название задач о сильном выбросе (инжекции) частиц газа. Важно, что на начальной стадии развития процесса можно построить асимптотическое решение для функции распределения в виде аналитического выражения, включающего ударную трансформанту [26]. По прошествии нескольких времен столкновения появляется тепловой разброс скоростей частиц и постепенно решение выходит на сплошносредние асимптотики, соответствующие задачам о газодинамическом сильном взрыве.

Приведем кратко постановку такого рода задач и изложим наиболее интересные и значимые результаты. В качестве основы описания принято уравнение Больцмана в виде

$$\frac{\partial f(t, x, \xi)}{\partial t} + \xi \frac{\partial f(t, x, \xi)}{\partial x} = \frac{1}{\text{Kn}} I(f, f). \quad (1)$$

Здесь $f(t, x, \xi)$ — функция распределения, $x = (x, y, z)$ — вектор координат, $\xi = (\xi_x, \xi_y, \xi_z)$ — вектор молекулярной скорости, $I(f, f)$ — интеграл столкновений, Kn — число Кнудсена. Уравнение (1) записано в безразмерном виде, как и все последующие выражения. Для получения безразмерных переменных используются подходящие масштабы длины, числовой плотности, температуры и скорости $L, n_{\text{ref}}, T_{\text{ref}}, u_{\text{ref}} = \sqrt{RT_{\text{ref}}}$, R — газовая постоянная. Масштаб времени является производным от масштабов длины и скорости: $\tau = L/u_{\text{ref}}$. Для описания столкновений принята модель твердых сфер. Число Кнудсена $\text{Kn} = \lambda/L$, где $\lambda = 1/(\sqrt{2}\sigma n_{\text{ref}})$ — длина свободного пробега (σ — полное сечение столкновения), L — масштаб течения.

Для дальнейшего анализа представим функцию распределения в виде суперпозиции членов в виде дельта-функций для пучка, фонового газа и добавки. В случае сферически-симметричной постановки решение ищется в виде [6]:

$$f(t, r, \xi) = \frac{N(t)}{4\pi r^2} \delta(\xi - \mathbf{U}) \delta(r - Ut) + n(t, r) \delta(\xi) + F(t, r, \xi), \quad (2)$$

где первое слагаемое соответствует $N(t)$ частицам пучка (их число убывает по мере столкновений с частицами фона), которые разлетаются от центра с радиальной скоростью $\mathbf{U} = (U, 0, 0)$, r – радиальная координата, $\delta(r)$, $\delta(\xi)$ – дельта-функции Дирака, второе слагаемое соответствует частицам фонового газа с числовой плотностью $n(t, r)$, третье слагаемое F соответствует частицам, выбывшим из первых двух групп за счет столкновений. Такое представление функции распределения наиболее наглядно представляет ситуацию, когда нет столкновений частиц третьей группы с частицами первых двух групп (приближение первых столкновений).

После подстановки в уравнение Больцмана функции распределения, преобразованной к сферическим координатам в пространстве скоростей, уравнение (1) распадается на систему из трех уравнений, которая приведена в [6]. Динамика фракции частиц, образовавшихся в результате столкновений частиц пучка и фона, в [6] описывается с использованием S -модели для интеграла столкновений. Часть интеграла обратных столкновений, отвечающая взаимодействию частиц фронта $r = Ut$ и фона, выписывается точно с использованием ударной трансформанты [26], такого типа решение будет рассмотрено ниже.

Можно переформулировать эту задачу на общий случай столкновения пучков (допуская в качестве одного из них фоновый газ):

$$f(t, x, \xi) = n_1(t, x) \delta(\xi - \mathbf{U}_1) + n_2(t, x) \delta(\xi - \mathbf{U}_2) + F(t, x, \xi), \quad (3)$$

где n_1 и n_2 – числовые плотности сталкивающихся групп молекул со скоростями $\mathbf{U}_1$ и $\mathbf{U}_2$.

Для дальнейшего представляет интерес асимптотическое решение при $t \rightarrow 0$, поскольку количество рассеянных частиц F на начальном этапе достаточно мало, так как при $t = 0$ имеем $F = 0$. Полагая $F = 0$ в правых частях исходной системы [6–8], получим упрощенную систему, пригодную для описания начального этапа. Рассмотрим такую асимптотическую систему для случая столкновения тонкого плоского пучка полуширины y_0 с покоящимся газом ($U_1 = U$, $U_2 = 0$):

$$\frac{\partial n_1}{\partial t} + U \frac{\partial n_1}{\partial x} = -\sigma n_1 n_2; \quad (4a)$$

$$\frac{\partial n_2}{\partial t} = -\sigma n_1 n_2; \quad (4б)$$

$$\frac{\partial F}{\partial t} + \xi_x \frac{\partial F}{\partial x} + \xi_y \frac{\partial F}{\partial y} = \Phi. \quad (4в)$$

Здесь интеграл столкновений Φ , соответствующий взаимодействию частиц пучка и фона, выписывается с использованием ударного трансформанта [26]:

$$\Phi = 2\sigma U n_1 n_2 T(\xi, 0, U) H(y_0 - y), \quad (5)$$

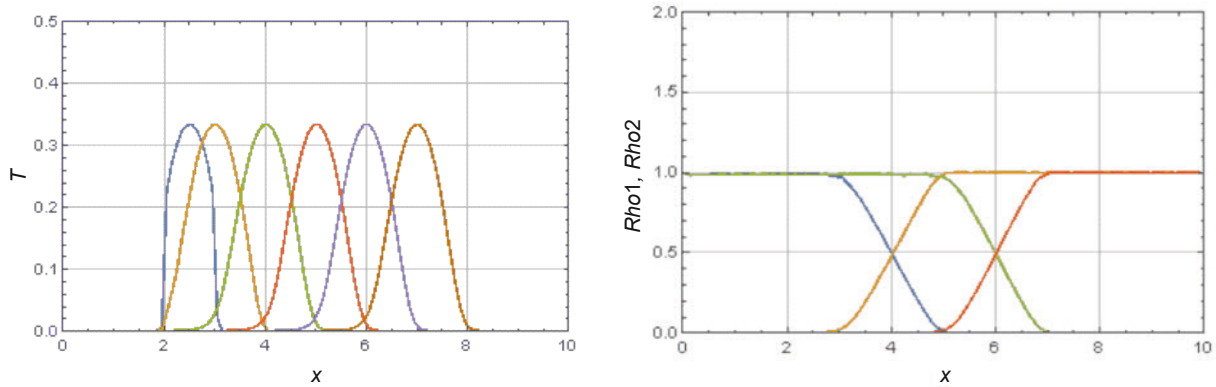
где

$$T(\xi, 0, U) = \frac{1}{(\pi U^2)} \delta \left(\left(\left(\xi_x - \frac{U}{2} \right)^2 + (\xi_y)^2 + (\xi_z)^2 \right)^{\frac{1}{2}} - \frac{U}{2} \right).$$

Можно видеть, что в таком приближении уравнения (4а, б) не зависят от (4в). По истечении достаточного времени ($\sigma U t \gg 1$) решение задачи (4а), (4б) как нелинейной гиперболической системы с характеристиками $x - Ut = \text{const}$ и $x = \text{const}$ устанавливается таким образом, что профили n_1 и n_2 описываются решением вида связанных волн, бегущих со скоростью $D = U n_{10} / (n_{10} + n_{20})$, где n_{10} и n_{20} – невозмущенные плотности пучка и фона. Отсюда становится очевидным, что фронт пучка движется в фоновом газе с меньшей скоростью. При равенстве плотностей пучка и фона n_{10} и n_{20} получаем $D = U/2$. Поскольку толщина тонкого пучка много меньше длины свободного пробега в нем, на начальном этапе не происходит вторичных столкновений рассеянных частиц F с частицами фона и пучка. Точное решение данной задачи [8] позволяет верифицировать более сложные модели взаимодействия пучка с фоновым газом, которые интересны для перспективных приложений. Решение данной задачи используется в качестве тестового для более общей задачи о столкновении пучка и фона.

3. ЧИСЛЕННЫЙ АНАЛИЗ ЗАДАЧ О РАССЕЯНИИ ТОНКОГО ПУЧКА

Для получения численных решений применялся вариант численного метода прямого статистического моделирования DSMC2D [27]. В методе используется схема мажорантной частоты с дополнительными ограничениями на выбор пар сталкивающихся частиц аналогично методу молекулярной динамики. Принята модель



Фиг. 1. Профили температуры и плотностей пучка и фона.

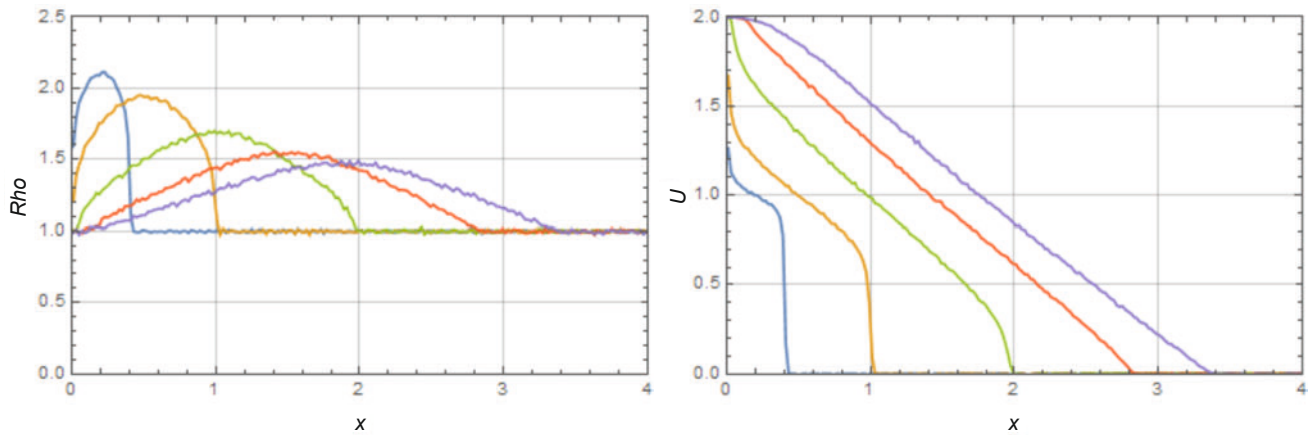
газа из частиц-твердых шаров. Для понимания специфики решений рассматриваемых задач важно, что в модели газа из твердых шаров средняя длина пробега не зависит от средней относительной скорости и тем самым от температуры.

Для сопоставления с аналитическим решением из [7, 8] изучается частный случай этой задачи. Рассматривается входящий слева ($x < 2, y \leq h$) пучок со скоростью $U = 2$, числовой плотностью $n_{10} = 1$ и температурой $T_{10} = 0.001$. Масштабом числовой плотности n_{ref} является значение в невозмущенном пучке, масштаб температуры $T_{\text{ref}} = 1$, масштаб длины $L = \lambda_{10}$, где λ_{10} — средняя длина свободного пробега частиц в пучке. Поперечный размер пучка ограничен возможностью разделения частиц на принадлежащие пучку и выбывшие из него в результате столкновений без учета их обратного влияния на пучок, поэтому ограничение на полуширину пучка $h \ll 1$. Физически это оправдано до тех пор, пока не образуется «атмосфера» из выбывших частиц и начнутся их столкновения между собой и с частицами пучка и фона. Справа ($x \geq 2, y \geq 0$) находится покоящийся газ с плотностью $n_{20} = 1$ и температурой $T_{20} = 0.001$. Несмотря на то, что температуры групп частиц по отдельности малы ($T_{10} \ll 1, T_{20} \ll 1$), температура в области их взаимодействия величина порядка единицы: $T_{\Sigma} = (1/3)(n_{10}(U - U^*)^2 + n_{20}(0 - U^*)^2)/(n_{10} + n_{20}) = 1/3$, где $U^* = (U + 0)/2 = U/2 = 1$. Алгоритм моделирования организован так, что частицы, испытавшие одно столкновение, выбывают из расчета и далее не учитываются. В качестве физической аналогии данного процесса можно привести аннигиляцию частиц и античастиц. Как и для теоретического решения, по прошествии нескольких характерных времен столкновений устанавливается режим бегущей волны.

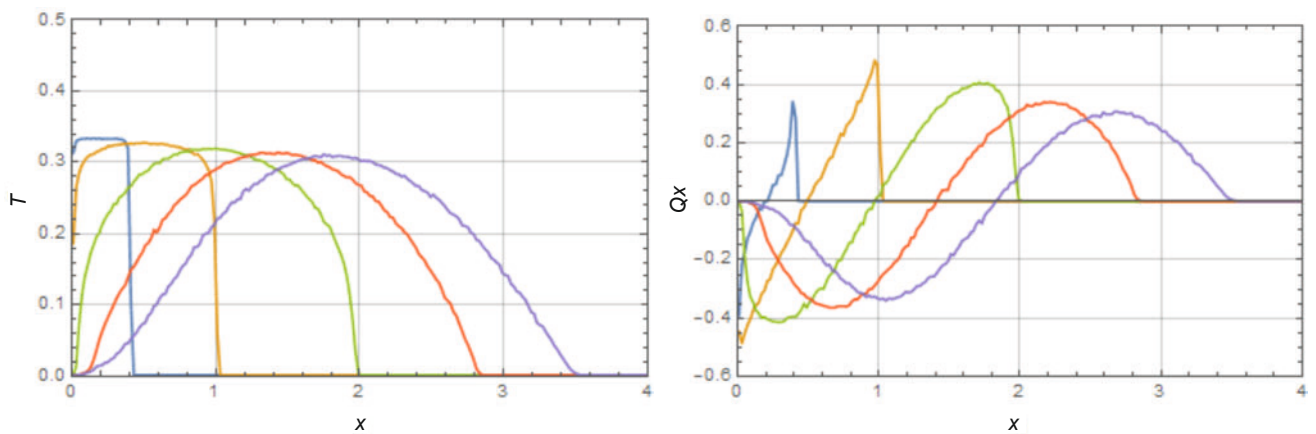
На фиг. 1а показаны распределения температуры в моменты времени $t = 0.5; 1; 2; 3; 4; 5$ (на графике слева направо). На фиг. 1б показаны распределения плотностей пучка (синяя и зеленая линии) и фона (оранжевая и красная линии) в моменты времени $t = 2$ и 4. Скорость движения фронта в расчете оказывается единичной: $D = 1$ (расстояние в 2 единицы длины фронт проходит за 2 единицы времени), что соответствует аналитическому решению. Из фиг. 1а следует, что этап формирования бегущей волны заканчивается к моменту $t \cong 2$. Полученное согласие расчетных и теоретических результатов позволяет верифицировать численный метод и алгоритм.

Для более общей задачи о проникании тонкого пучка в покоящийся газ рассматривается входящий слева ($x < 0, y \leq h$) пучок со скоростью $U = 2$, плотностью $n_{10} = 1$ и температурой $T_{10} = 0.001$. Справа ($x \geq 0, y \geq 0$) находится покоящийся газ с плотностью $n_{20} = 1$ и температурой $T_{20} = 0.001$. Поперечный размер плоского пучка: $y \leq h = 1/10$. Выбор масштабов такой же, как принят ранее. Граница $x = 0$ при $y > h$ моделировалась зеркально отражающей стенкой (ситуация проникновения пучка в резервуар). В данной задаче частицы пучка и фона, испытавшие столкновения, не выбывают и участвуют в последующих столкновениях с формированием глобального течения, которое вовлекает все больше частиц. На фиг. 2 и 3 показаны распределения основных переменных (числовая плотность n на фиг. 2а, продольная скорость u на фиг. 2б, температура T на фиг. 3а, тепловой поток Q_x на фиг. 3б) на плоскости симметрии $y = 0$ в моменты времени $t = 0.2; 0.5; 1; 2$. На фиг. 4а, б показаны поля температуры T в моменты времени $t = 1$ и 2. На фиг. 5а, б показаны поля компонент теплового потока Q_x и Q_y в момент $t = 2$.

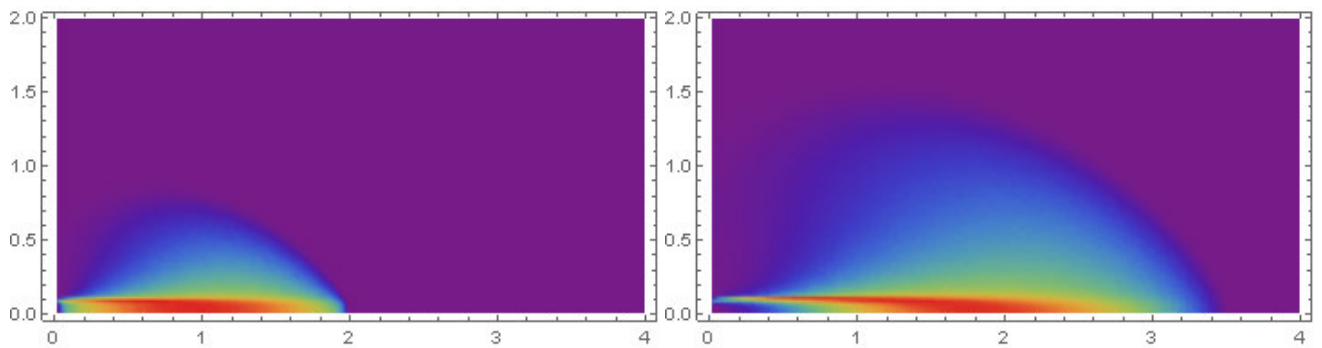
Из фиг. 2–5 видно, что первоначально происходит проникновение пучка в покоящийся газ (плотности частиц пучка и фона суммируются). Затем по мере продвижения пучка и столкновений происходит выбывание частиц из зоны взаимодействия, возмущение плотности в пучке снижается, снижается и скорость фронта пучка. После прохождения пучком дистанции $l \sim 3 \div 5$ длин пробега, образуется конфигурация факела, состоящая из зоны удва, где уже нет частиц фонового газа, и зоны взаимодействия с поперечно расширяющимся фронтом.



Фиг. 2. Профили плотности ρ (а) и скорости U (б) на плоскости симметрии в моменты времени $t = 0.2, 0.5, 1, 2$.



Фиг. 3. Профили температуры T (а) и теплового потока Q_x (б) на плоскости симметрии в моменты времени $t = 0.2, 0.5, 1, 2$.

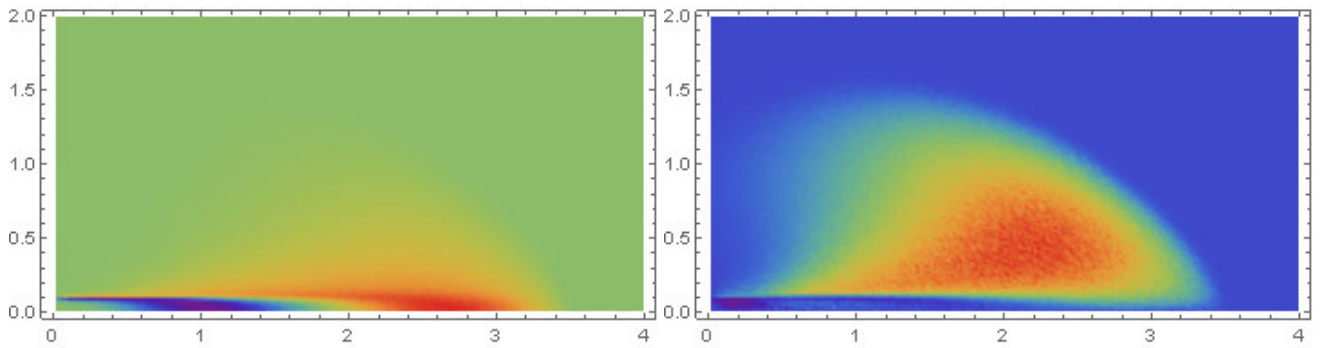


Фиг. 4. Поля температуры T в моменты $t = 1$ (а) и 2 (б).

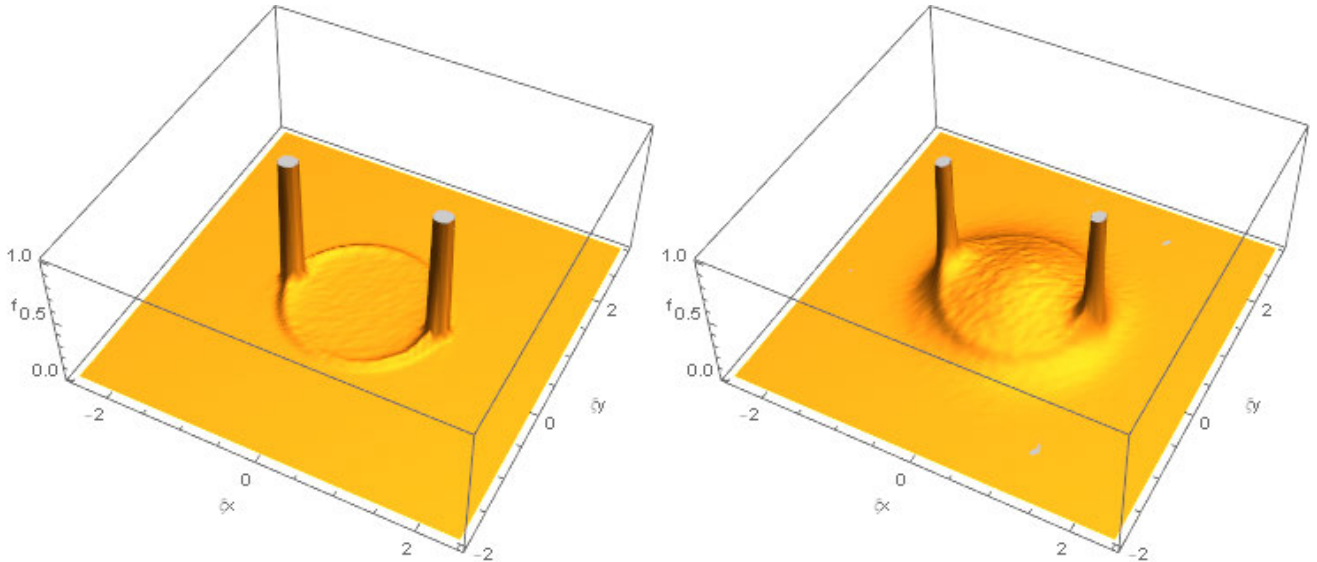
4. ЗАДАЧА О СТОЛКНОВЕНИИ ПОТОКОВ

Рассмотрим задачу, логически дополняющую предыдущую. Пусть два разреженных потока (пучка большой ширины) приведены в начальный момент в соприкосновение. В рамках использования известных асимптотических решений рассмотрим лобовое столкновение потоков, такая задача является пространственно одномерной.

Входящий слева ($x < 0$) поток имеет скорость $U_{10} = 1$, плотность $n_{10} = 1$ и температуру $T_{10} = 0.001$, входящий справа ($x > 0$) поток имеет скорость $U_{20} = -1$, плотность $n_{20} = 1$ и температуру $T_{20} = 0.001$. Температура системы $T_{\Sigma} = 1/3$. Выбор масштабов аналогичен предыдущему разделу.



Фиг. 5. Поля компонент теплового потока Q_x (а) и Q_y (б) в момент $t = 2$.



Фиг. 6. Функция распределения в окрестности точки $x = 0$ при $t = 0.2, 2$.

Для сопоставления аналитического и численного решений на начальном этапе процесса удобно рассмотреть эволюцию функции распределения частиц в окрестности точки столкновения потоков $x = 0$. Хотя рассматриваемая задача представляет пространственно-неоднородное решение, процессы в зоне столкновения потоков могут быть описаны на начальном этапе с помощью модели однородной релаксации. Для этого можно использовать простое асимптотическое при $t \rightarrow 0$ решение уравнения (4в), которое в пространственно-однородном случае примет вид

$$\frac{\partial F(t)}{\partial t} = \Phi_{\infty}.$$

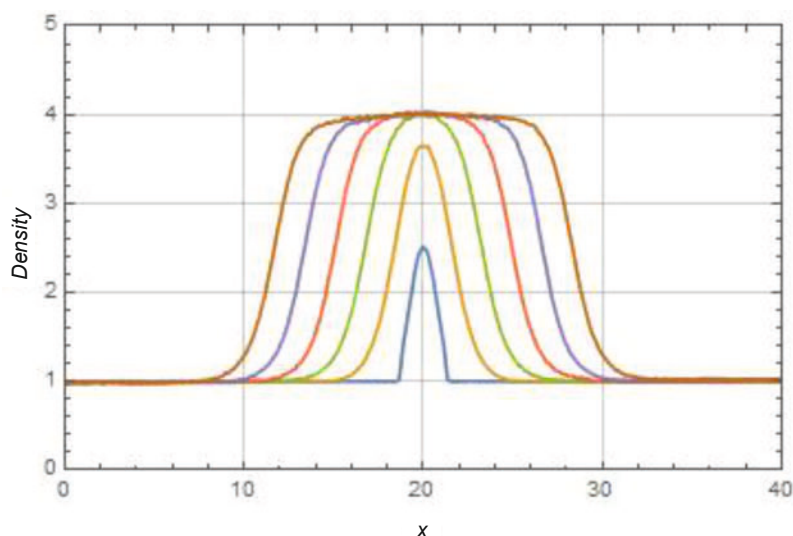
Для малых времен $t \rightarrow 0$ получаем решение [6]:

$$F(t) = 3n^2 \frac{\sigma}{(2\pi)} \delta(\xi - U/2) t.$$

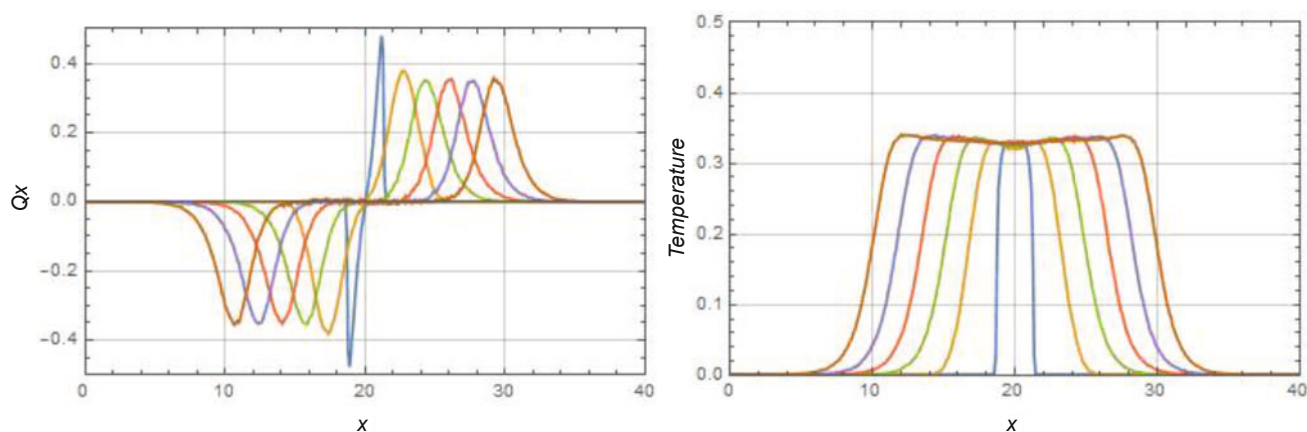
Данное выражение показывает, как выглядит функция распределения в результате первых столкновений частиц потоков — скорости рассеянных частиц находятся на сфере (окружности) с радиусом, следующим из ударной трансформанты. По мере релаксации выбывшие из пучков частицы за несколько столкновений приобретают скорости, близкие к равновесным. Процесс формирования равновесной плотности распределения идет до момента $t \sim 10$, когда достигается предельное сжатие газа в зоне столкновения потоков и формируются бегущие влево и вправо ударные волны.

На фиг. 6 представлен вид функции распределения (для удобства анализа здесь проведено суммирование по компоненте молекулярной скорости ξ_z) в окрестности точки $x = 0$ в моменты времени $t = 0.2$ и 2 .

Происходящие при формировании бегущих ударных волн процессы могут быть хорошо проиллюстрированы распределениями плотности, температуры и теплового потока в различные моменты времени, фиг. 7 и 8а, б.



Фиг. 7. Профили плотности n в моменты времени $t = 1.25; 5; 10; 15; 20; 25$.



Фиг. 8. Профили теплового потока Q_x (а) и температуры T (б) в моменты времени $t = 1.25; 5; 10; 15; 20; 25$.

Как можно видеть, в рассматриваемой задаче формируются ударные волны предельной интенсивности (предельного сжатия), так как из-за малой температуры пучков число Маха велико. Как известно, для одноатомного газа предел повышения плотности при ударном сжатии равен 4, что можно увидеть на фиг. 7. Образовавшиеся ударные волны бегут влево и вправо по газу с малой температурой, фиг. 8б. После формирования бегущих ударных волн скорость их движения выходит на теоретическую величину, которую можно вычислить из закона сохранения массы [22].

Из профилей теплового потока и температуры на фиг. 8а, б видно, что после формирования сильных ударных волн имеется зона небольшого локального минимума температуры за счет поступательной неравновесности, в этой зоне знаки градиента температуры и теплового потока совпадают. В дальнейшем это явление сохраняется, с поправкой на то, что возмущение теплового потока передвигается вместе с ударной волной. Это свидетельствует о существовании зоны неклассического переноса на расстоянии $l \sim 5 \div 10$ длин пробега за бегущей ударной волной.

5. ЗАКЛЮЧЕНИЕ

Модели и методы динамики разреженного газа для задач описания высокоскоростных течений позволяют исследовать практически значимые неравновесные явления — столкновение пучков с учетом и без учета выбывания частиц. Для задачи с безвозвратным выбыванием частиц (тонкие пучки) получено хорошее согласие с аналитическим решением. Для задачи без выбывания частиц (соударение потоков) получено численное решение типа бегущей ударной волны предельного сжатия, формирующейся, например, при соударении вы-

сокоскоростного потока со стенкой. Для этого случая показана роль ударных трансформант на начальной стадии процесса. Рассмотрена задача о проникновении пучка в покоящийся газ вплоть до формирования факела, отмечено сходство ее начальной стадии с первой задачей. Отметим, что развитие созданной Е.М. Шаховым методологии дает возможность проводить моделирование различных явлений в областях, не относящихся прямо к кинетической теории, что говорит об определенной общности рассматриваемых моделей.

СПИСОК ЛИТЕРАТУРЫ

1. *Аристов В.В.* Метод переменных сеток в скоростном пространстве в задаче о сильной ударной волне // Ж. вычисл. матем. и матем. физ. 1977. Т. 17. № 4. С. 261–267.
2. *Brull S., Mieussens L.* Local discrete velocity grids for deterministic rarefied flow simulations // J. of Comput. Phys. 2014. V. 266. P. 22–46.
3. *Arslanbekov R.R., Kolobov V.I., Frolova A.A.* Kinetic Solvers with Adaptive Mesh in Phase Space // Phys. Rev. E. 2013. V. 88. 063301.
4. *Xiao T., Liu C., Xu K., Cai Q.* A velocity-space adaptive unified gas kinetic scheme for continuum and rarefied flows // J. of Comput. Phys. 2020. V. 415. 109535.
5. *Аристов В.В., Шахов Е.М.* Задача о сильном взрыве в разреженном газе // Ж. вычисл. матем. и матем. физ. 1979. Т. 19. № 5. С. 1276–1287.
6. *Аристов В.В., Шахов Е.М.* Течение разреженного газа, вызванное сильным точечным выбросом конечной массы // Ж. вычисл. матем. и матем. физ. 1985. Т. 25. № 7. С. 1066–1077.
7. *Aristov V.V., Shakhov E.M.* Scattering of impulsive molecular beam in a rarefied gas // Proc. of 15th Int. Symp. on Rarefied Gas Dyn. Ed. B.G. Teubner. 1986. V. 1. P. 266–275.
8. *Аристов В.В., Шахов Е.М.* Нелинейное рассеяние импульсного молекулярного пучка в разреженном газе // Ж. вычисл. матем. и матем. физ. 1987. Т. 27. № 2. С. 159–164.
9. *Мюррей Д.* Математическая биология. Т. II / Ред. Г.Ю. Ризниченко. М.: – Ижевск, 2011.
10. История и Математика: Проблемы периодизации исторических макропроцессов. М.: КомКнига, 2006.
11. Special Issue: Statistical Mechanics and Social Sciences I // J. Stat. Phys. 2013. V. 151. N 1–2. Special Issue: Statistical Mechanics and Social Sciences II // J. Stat. Phys. 2013. V. 151. N 3–4.
12. *Prigogine I., Herman R.* Kinetic Theory of Vehicular Traffic. New York: American Elsevier, 1971.
13. *Колмогоров А.Н., Петровский И.Г., Пискунов Н.С.* Исследование уравнения диффузии, соединенной с возрастанием количества вещества и его применение к одной биологической проблеме // Бюллетень МГУ. Сер. А. Матем. и механ. 1937. Т. 1. С. 1–25.
14. *Fischer R. A.* The wave of advance of advantageous genes // Ann. Eugenics. 1937. V. 7. P. 355–369.
15. *Васильев В.А., Романовский Ю.М., Яхно В.Г.* Автоволновые процессы. М.: Наука, 1987.
16. *Елькин Ю.Е.* Автоволновые процессы // Матем. биология и биоинформатика. 2006. Т. 1. № 1. С. 27–40.
17. *Aristov V.V., Ilyin O.V.* Kinetic Models for Historical Processes of Fast Invasion and Aggression // Phys. Rev. E. 2015. V. 91. 04286.
18. Blitzkrieg basics // Scientific American. 2015. N 6. P. 22.
19. *Aciole P.H.* Diffusion as a first model of spread of viral infection // Am. J. Phys. 2020. V. 88(8). P. 600–604.
20. *Aristov V.V., Stroganov A.V., Yastrebov A.D.* Simulation of Spatial Spread of the COVID-19 Pandemic on the Basis of the Kinetic-Advection Model // Physics. 2021. V. 3. P. 85–102.
21. *Аристов В.В., Строганов А.В., Ястребов А.Д.* Применение модели кинетического типа для изучения пространственного распространения COVID21 // Докл. АН. Физика. Техн. науки. 2021. Т. 498. С. 27–32.

22. *Седов Л.И.* Распространение сильных взрывных волн // Прикл. матем.и механ. 1946. Т. 10. Вып. 2. С. 241–250.
23. *Taylor G.* The formation of blast wave by a very intense explosion // Rept RC–210, 27 June 1941. Civil Defence Research Committee, 1941.
24. *von Neumann J.* The point source solution // Bethe H.A., Fuchs K., Hirschfelder J.O. et al. Blast wave. Los-Alamos Scientific Laboratory Rept LA-2000. 1958. P. 27–55.
25. *Коробейников В.П.* Задачи точечного взрыва. М.: Наука, 1988.
26. *Баранцев Р.Г.* Об ударных трансформантах кинетического уравнения аэродинамики разреженных газов // Аэродинамика разреженных газов. Л.: Изд-во ЛГУ., 1963. Вып. 1. С. 80–91.
27. *Аристов В.В., Воронич И.В., Забелок С.А.* Исследование неклассического переноса с применением численных методов решения уравнения Больцмана // Ж.вычисл.матем. и матем. физ. 2023. №12. С. 2025–2034.

SOLVING PROBLEMS OF KINETIC INTERACTION OF FAST GROUPS OF PARTICLES USING ANALYTICAL AND NUMERICAL METHODS

V. V. Aristov*, I. V. Voronich**

119333 Moscow, Vavilov Str., 44, Federal Research Center Computer Science and Control, RAS, Russia

** e-mail: aristovvl@yandex.ru*

***e-mail: i.voronich@frcscs.ru*

Received: 15.07.2023

Revised: 10.04.2024

Accepted: 31.05.2024

Abstract. A brief overview of the problem statements related to high-speed beams is given, with an emphasis on the use of analytical solutions, and the results of numerical solution of some problems of this class are described. Within the framework of kinetic theory, the processes of interaction of groups of particles (molecules) are considered on the basis of the analytical method, assuming a high correlation of particle velocities (delta function as the distribution density). The problems of the interaction of beams with and without the elimination of particles are studied numerically using the method of direct statistical modeling. For the problem with the elimination of particles (intersection and interaction of thin beams), a good agreement with the analytical solution was obtained. For the problem without the elimination of particles (collision of flows), a numerical solution of the type of traveling shock wave of extreme compression formed when the flow collides with the wall is obtained. The role of shock transformants at the initial stage of the process is shown. The problem of beam penetration into a stationary gas up to the stage of plume formation is considered, and the similarity of its initial stage with the problem of thin beams is noted. The fruitfulness of using analytical methods at the stage of primary analysis of the problem and verification of numerical solutions is emphasized.

Keywords: kinetic equation, rarefied gas, high-speed flows, collision of particle beams, analytical solutions, statistical modeling.

ЧИСЛЕННЫЙ РАСЧЕТ ВЛИЯНИЯ ПРОЦЕССА МИГРАЦИИ МЕТАНА НА РЕЗУЛЬТАТЫ СЕЙСМИЧЕСКОЙ РАЗВЕДКИ В ЗОНАХ ВЕЧНОЙ МЕРЗЛОТЫ¹⁾

© 2024 г. Е. К. Гусева^{1,*}, В. И. Голубев^{1,**}, И. Б. Петров^{1,***}

¹ 141701 Долгопрудный, М.о., Институтский пер., 9, МФТИ (НИУ), Россия

*e-mail: guseva.ek@phystech.edu

**e-mail: golubev.vi@mipt.ru

***e-mail: petrov@mipt.ru

Поступила в редакцию 23.02.2024 г.

Переработанный вариант 27.04.2024 г.

Принята к публикации 31.05.2024 г.

На протяжении длительного времени в шельфовых и прибрежных зонах Арктического региона наблюдается интенсивная эмиссия метана в атмосферу из недр вечной мерзлоты. Из-за потенциальной опасности этого явления для окружающей среды и инфраструктуры появляется необходимость периодического мониторинга газовых карманов, в том числе, с помощью проведения наземной сейсморазведки. Настоящая работа содержит результаты исследования данного процесса методами численного моделирования. Построена модель слоистого многолетнемерзлого песчаного грунта с криволинейными границами между пластами, отражающая основную специфику региона. Исследован процесс миграции газа в вертикальном и горизонтальном направлениях с помощью увеличения количества метановых резервуаров. Используется определяющая гиперболическая система линейной теории упругости. Задача решается сеточно-характеристическим методом в двумерной постановке на прямоугольной сетке, в каждой ячейке которой задаются упругие характеристики слоев. Подробно изучаются образующиеся волновые структуры на полученных синтетических волновых картинах и сейсмограммах, позволяющие определить направление распространения газа. Полученные результаты можно использовать для трактовки подобных натурных экспериментов. Библ. 22. Фиг. 5. Табл. 1.

Ключевые слова: численное моделирование, прямая задача сейсморазведки, вечная мерзлота, миграция газа, сеточно-характеристический метод.

DOI: 10.31857/S0044466924090119, EDN: WJPVOQ

1. ВВЕДЕНИЕ

Численное моделирование является эффективным инструментом изучения волновых процессов, возникающих в различных динамических задачах (см. [1]). Одним из подобных процессов является сейсморазведка, широко используемая для изучения структуры геологических сред (см. [2]), в том числе, с целью обнаружения полезных ископаемых, таких как нефть и газ. Значительное внимание к себе привлекают прибрежные и шельфовые зоны Арктического региона, богатые запасами природных ресурсов. Однако низкие температуры создают ряд специфических особенностей, влияющих на результаты полевых измерений. Одной из главных отличительных черт области является наличие пластов вечной мерзлоты (см. [3]) — слоев мерзлого грунта, находящегося под воздействием отрицательных температур на протяжении длительного времени. Структура грунта неоднородна, в слоях присутствует много пустот, ледовых и газовых включений, границы между пластами оказываются существенно криволинейны. Моделирование сейсморазведки при таких условиях позволяет оценить влияние подобных особенностей на результаты измерений. Таким образом, одной из целей настоящей работы является отражение основных характеристик региона и создание модели грунта, приближенной к реальным условиям многолетнемерзлых песчаных пород в полярных и приполярных прибрежных зонах.

В дополнение к этому, работа сосредоточена на изучении явления миграции газа. Арктический регион богат своими запасами газа, часто в форме метангидрата, занимающего образующиеся в грунте пустоты. В стандартных условиях данные образования окружены замёрзшим грунтом и не представляют угрозы. Однако в связи с глобальным потеплением лед начинает оттаивать, и на протяжении многих лет стабильно наблюдается эмиссия метана в атмосферу в шельфовых и прибрежных зонах (см. [4], [5]). Подобные явления представляет собой

¹⁾ Работа выполнена при финансовой поддержке РФФИ (проект № 21-71-10015).

проблему для окружающей среды и инфраструктуры региона (см. [6]) и могут вызвать дополнительной рост температуры из-за нарастающего парникового эффекта и лавинообразного усиления процесса. В отдельных ситуациях газ выходит на поверхность вместе с сильным взрывом, образующим воронку (см. [7], [8]). Существует гипотеза о том, что подобные метановые бомбы в случае всеобщего взрыва могут привести к глобальным негативным последствиям (см. [9]). Дополнительную угрозу подобные образования представляют в случае разработки месторождений. Бурение может растопить сдерживающие газ мерзлые пласты и привести к детонации и разрушению всего бурового оборудования и окружающих строений. Таким образом, обнаружение и мониторинг метановых карманов является актуальной и необходимой задачей.

Для решения данной проблемы используют различные методики проведения измерений, в том числе, и сейсморазведку, позволяющую детектировать отражения исходного волнового фронта от интересующих включений, которыми в рассматриваемом случае являются метановые резервуары. С одной стороны, выделение таких полезных волн (см. [10]) позволяет обнаружить подобные потенциально опасные объекты. С другой стороны, многочисленные отклики, возникающие в слоистых грунтах, затрудняют трактовку полученных результатов измерений. Проведение периодических натурных измерений – дорогой и длительный процесс, поэтому численное моделирование позволяет более экономно провести детальный анализ синтезированных волновых картин и сейсмограмм в процессе миграции газа. Таким образом, в настоящей работе предлагается использовать сеточно-характеристический метод из [11] для решения этой динамической задачи. Данный метод нашел широкое применение в подобных проблемах из-за возможности точного воспроизведения волновых структур в объектах сложной формы (см. [12], [13]). В работе используются определяющие уравнения линейной теории упругости (см. [14]). Главным отличием от предыдущих подобных исследований [15], [16] является использование усовершенствованной двумерной постановки задачи. Вводятся криволинейные границы между слоями, выбираются типичные для региона упругие и геометрические характеристики пластов. Миграция газа моделируется как последовательное увеличение числа метановых включений в вертикальном и горизонтальном направлениях. Значительная часть работы посвящена детальному анализу итоговых волновых структур и возможности обнаружения распространения газа.

2. ОПРЕДЕЛЯЮЩАЯ СИСТЕМА УРАВНЕНИЙ И МЕТОД РАСЧЕТА

Для описания динамического поведения грунта использовались уравнения линейной теории упругости (см. [14]). Неизвестными выступали скорость точек среды $\mathbf{v}$ и тензор напряжений σ :

$$\rho \dot{\mathbf{v}} = \nabla \cdot \sigma + \mathbf{f}, \quad (1)$$

$$\dot{\sigma} = \lambda(\nabla \cdot \mathbf{v})I + \mu(\nabla \otimes \mathbf{v} + (\nabla \otimes \mathbf{v})^T). \quad (2)$$

Здесь ρ – это плотность среды, $\mathbf{f}$ – вектор внешней силы, λ и μ – параметры Ламе. В рамках данной модели возникают продольные и поперечные волны со скоростями $c_p = \sqrt{\frac{\lambda+2\mu}{\rho}}$ и $c_s = \sqrt{\frac{\mu}{\rho}}$. Плотность и скорости упругих волн полностью задают поведение упругого материала.

В работе динамическая задача рассматривалась в двумерной постановке. Как было показано ранее в [17], преимуществами данного подхода является значительное сокращение времени расчетов при сохранении основных свойств получаемых волновых картин. Для решения определяющей системы уравнений гиперболического типа (1)–(2) использовался сеточно-характеристический метод [11]. Согласно методу сначала формируется вектор неизвестных $\mathbf{U} = (v_x, v_y, \sigma_{xx}, \sigma_{yy}, \sigma_{xy})^T$ и система преобразуется к каноническому виду

$$\frac{\partial \mathbf{U}}{\partial t} + A_x \frac{\partial \mathbf{U}}{\partial x} + A_y \frac{\partial \mathbf{U}}{\partial y} = \mathbf{g}. \quad (3)$$

Здесь вводятся матрицы A_i вида

$$A_x = - \begin{pmatrix} 0 & 0 & 1/\rho & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/\rho \\ \lambda + 2\mu & 0 & 0 & 0 & 0 \\ \lambda & 0 & 0 & 0 & 0 \\ 0 & \mu & 0 & 0 & 0 \end{pmatrix}, \quad A_y = - \begin{pmatrix} 0 & 0 & 0 & 0 & 1/\rho \\ 0 & 0 & 0 & 1/\rho & 0 \\ 0 & \lambda & 0 & 0 & 0 \\ 0 & \lambda + 2\mu & 0 & 0 & 0 \\ \mu & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Затем проводится расщепление по координатам и физическим процессам. Для начала делается шаг алгоритма вдоль оси x : $\frac{\partial \mathbf{U}}{\partial t} + A_x \frac{\partial \mathbf{U}}{\partial x} = 0$, потом вдоль y : $\frac{\partial \mathbf{U}}{\partial t} + A_y \frac{\partial \mathbf{U}}{\partial y} = 0$, а в конце проводится необходимая коррекция решения согласно $\frac{\partial \mathbf{U}}{\partial t} = \mathbf{g}$. Для получившихся уравнений используется условие гиперболичности системы (3)

для представления матриц A_i , $i = \{x, y\}$, в форме $A_i = L_i^{-1} \Lambda_i L_i$, где L_i — матрица собственных векторов, $\Lambda_i = \text{diag}\{0, -c_s, c_s, -c_p, c_p\}$ — матрица собственных значений. Таким образом, появляется возможность ввести новые переменные — инварианты Римана: $\mathbf{w}_i = L_i \mathbf{U} = (w_i^1, w_i^2, w_i^3, w_i^4, w_i^5)^T$. В результате система (3) вдоль каждого направления преобразуется к системе одномерных независимых линейных уравнений переноса вида

$$\frac{\partial \mathbf{w}_i}{\partial t} + \Lambda_i \frac{\partial \mathbf{w}_i}{\partial \xi_i} = 0. \quad (4)$$

На данном этапе для решения каждого уравнения в (4) использовалась сеточно-характеристическая схема третьего порядка аппроксимации [18], монотонизированная с помощью сеточно-характеристического критерия монотонности [19]:

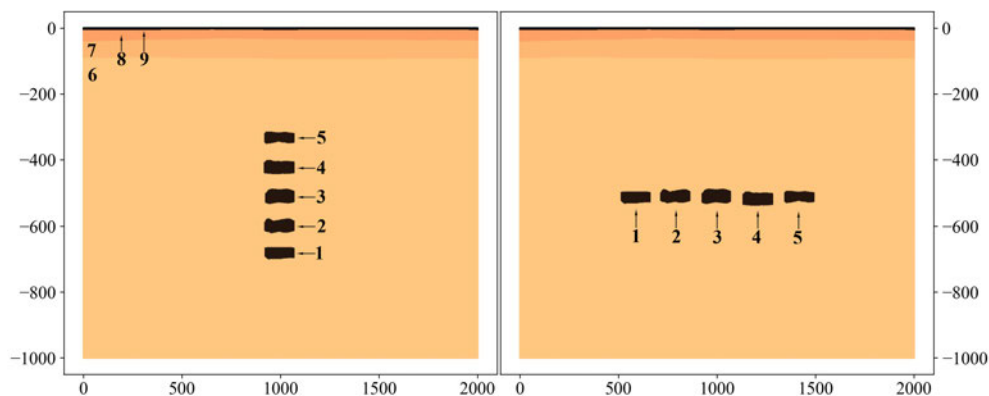
$$[w_i^j]_m^{n+1} = [w_i^j]_m^n + \frac{c_i^j}{2} ([w_i^j]_{m-1}^n - [w_i^j]_{m+1}^n) + \frac{c_i^{j^2}}{2} ([w_i^j]_{m-1}^n - 2[w_i^j]_m^n + [w_i^j]_{m+1}^n) - \frac{c_i^j(1 - c_i^{j^2})}{6} ([w_i^j]_{m-2}^n - 3[w_i^j]_{m-1}^n + 3[w_i^j]_m^n - [w_i^j]_{m+1}^n).$$

Здесь вводится число Куранта $c_i^j = \frac{\lambda_i^j \tau}{h_i}$, где λ_i^j — соответствующее собственное значение, h_i — пространственный шаг, $j = \{1, 5\}$. При расчетах значение данного числа было задано меньше единицы, таким образом, выполнялось условие устойчивости расчетной схемы. Согласно критерию монотонности, если $[w_i^j]_m^{n+1} \leq \min [w_i^j]_1^n, [w_i^j]_2^n$, то проводилась коррекция $[w_i^j]_m^{n+1} = \min [w_i^j]_1^n, [w_i^j]_2^n$. При $[w_i^j]_m^{n+1} \geq \max [w_i^j]_1^n, [w_i^j]_2^n$ считалось, что $[w_i^j]_m^{n+1} = \max [w_i^j]_1^n, [w_i^j]_2^n$. В данном случае $[w_i^j]_1^n, [w_i^j]_2^n$ — это значения сеточной функции при $t = t_n$ в двух ближайших к характеристике точках.

3. РАСЧЕТНАЯ ОБЛАСТЬ

Для проведения расчетов рассматривался прибрежный участок вечной мерзлоты, размером 2 км по ширине и 1 км по глубине, изображенный на фиг. 1. В качестве основы рассматривалась модель многолетнемерзлого песчаного грунта, сформулированная и обобщенная в работе [20] на основе результатов ряда полевых исследований, проведенных различными коллективами. Было создано 4 слоя, границы между которыми задавались существенно криволинейными. Верхний слой представлял собой узкий пласт талого грунта, под которым располагались мерзлые слои с постепенно увеличивающейся акустической жесткостью и плотностью. Упругие и геометрические характеристики пластов приведены в табл. 1.

В нижнем слое грунта на характерных глубинах залегания располагались резервуары с метангидратом типичного размера с упругими параметрами, взятыми из [15]. В процессе моделирования число включений варьировалось. Предполагалась миграция газа в вертикальном направлении от 1 резервуара (в обозначениях фиг. 1) до 5 и в горизонтальном направлении от центрального резервуара 3. Форма резервуаров с одинаковыми номерами задавалась одинаковой, отличающейся от прямоугольной криволинейными верхней и нижней границами. Для сравнения была протестирована постановка в отсутствии включений, что позволило оценить как влияние слоистого грунта, так и вклад каждого резервуара в полученные результаты расчетов.



Фиг. 1. Макет расчетной области.

Таблица 1. Упругие и геометрические характеристики слоев, номера соответствуют фиг. 1

№	c_1 , м/с	c_2 , м/с	ρ , кг/м ³	Глубина слоя, м	Толщина слоя, м	Ширина слоя, м
1	1000	707	900	662–664	29–37	150
2	1000	707	900	575–586	27–43	152
3	1000	707	900	487–500	29–46	148
4	1000	707	900	399–406	34–42	154
5	1000	707	900	313–318	28–37	152
6	4100	2300	2750	87–93	907–913	1000
7	3900	2150	2650	30–40	50–60	1000
8	3500	1950	2650	3–5	27–25	1000
9	900	490	2100	0	3–5	1000

Для моделирования была создана одна прямоугольная сетка с пространственными шагами вдоль горизонтальной и вертикальной осей, равным 2 м и 1 м соответственно. Шаг по времени был задан равным 0.1 мс. Источник сигнала в форме импульса Рикера с частотой 30 Гц помещался в центр области на глубину 1 м. На верхней дневной поверхности через каждые 4 м размещались приемники сигнала для построения сейсмограмм. В каждом узле расчетной области хранились упругие параметры соответствующих слоев. Дневная поверхность, считалась свободной, на остальных границах ставилось характеристическое условие поглощения волн. Использовалась форма граничных условий, описанная в [11].

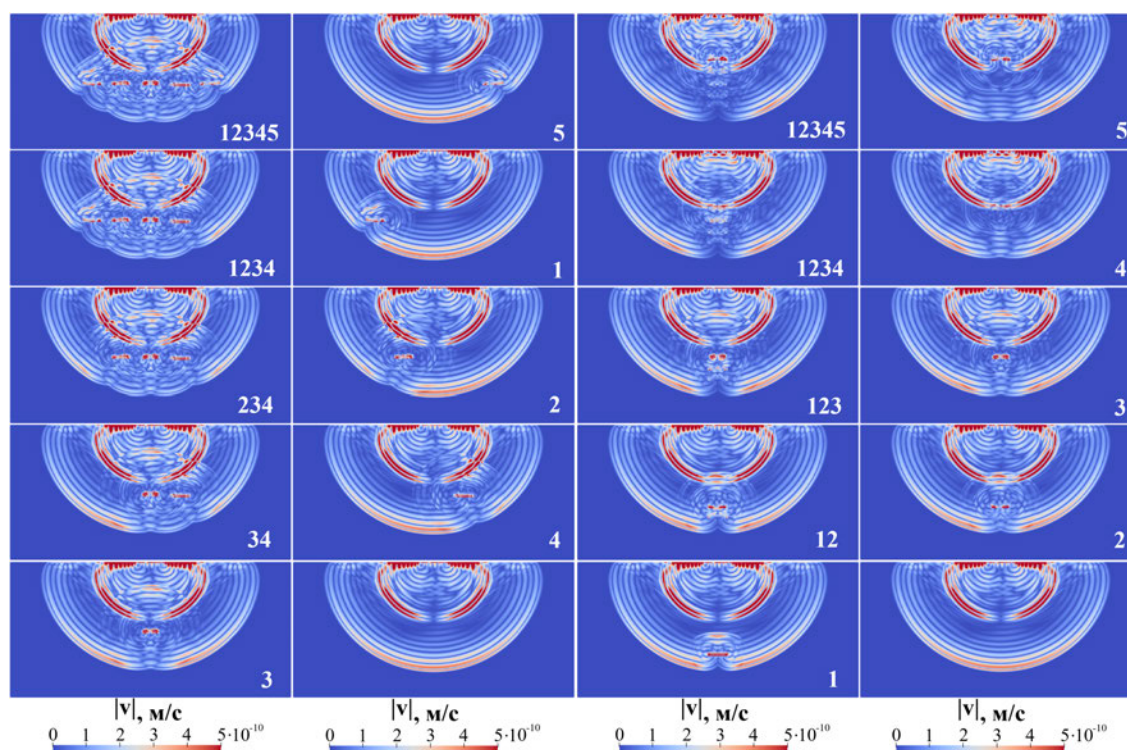
Отметим, что использование прямоугольной расчетной сетки вносит дополнительную ошибку аппроксимации при описании геометрии криволинейных границ раздела слоев модели. Однако, как было показано ранее [16], данный подход сохраняет все особенности сейсмического сигнала, незначительно повышая амплитуды отдельных волн. Его преимуществами является значительное повышение скорости расчета и снижением объема используемой оперативной памяти.

4. РЕЗУЛЬТАТЫ РАСЧЕТОВ

4.1. Волновые картины

Первым результатом проведенных расчетов являются волновые картины, изображенные на фиг. 2. Постановка без включений позволяет проанализировать волновой фронт, распространяющийся от источника, и оценить влияние структуры грунта на приходящий к дневной поверхности сигнал. Исходный сигнал имеет стандартную форму, состоящую из продольной и поперечной волн. Из-за наличия криволинейных границ картина оказывается слегка несимметричной. Наличие узких приповерхностных пластов приводит к возникновению кратных волн (см. [21]), многократных переотражений исходного фронта, которые приносят дополнительные осцилляции в его структуру. К тому же, на дневной поверхности вместе с изначальной волной распространяются волны Рэлея (см. [22]), вызывающие круговые структуры на картинах. Подобные волны значительно зашумляют интересные отклики от различных включений и затрудняют анализ результатов натурных измерений.

Согласно фиг. 2 выделенные волны вызывают осцилляции также в отраженных от включений волнах, принимающих характерную сферическую форму. В случае миграции газа в горизонтальном направлении отклики от каждого включения формируют треугольную структуру совокупного отклика. При рассмотрении волновых картин в постановке со всеми резервуарами и с последовательным удалением боковых включений можно наблюдать вклад каждого метанового слоя в общее отражение. Резервуар 3 формирует центральную часть фронта, которая первая возвращается на дневную поверхность. Отклики от остальных включений оказываются несимметричными, до них исходная волна доходит позднее, что приводит к формированию наклонных структур в совокупном отражении. Таким образом, можно сделать вывод о том, что распространение газа в горизонтальном направлении, приводит к постепенному расширению отраженного от резервуаров волнового фронта. В дополнение к этому можно ожидать прихода данного сигнала на большее количество приемников на дневной поверхности.

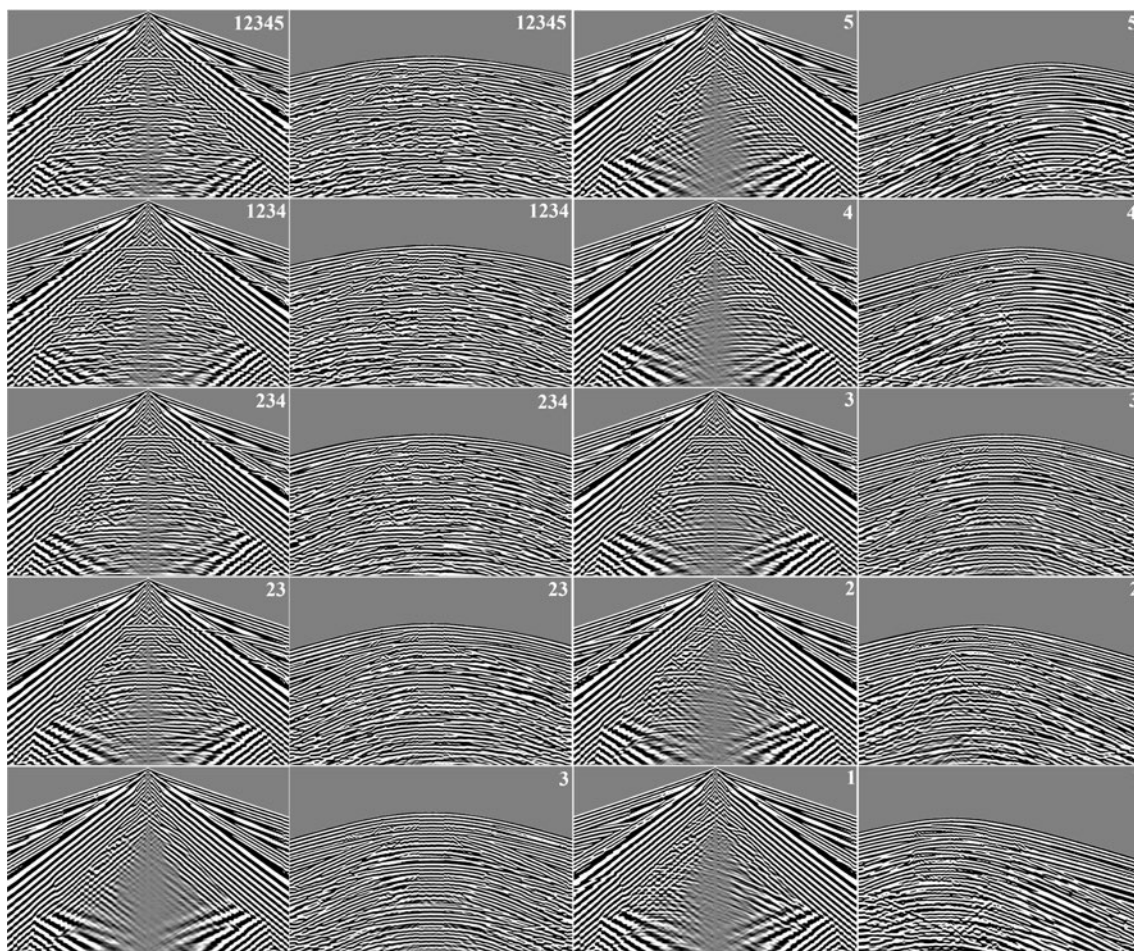


Фиг. 2. Волновые картины в момент времени 0.2 мс. Случай с горизонтальным распространением резервуаров слева, с вертикальным — справа. Цифрами указаны номера резервуаров, включенных в расчетную область.

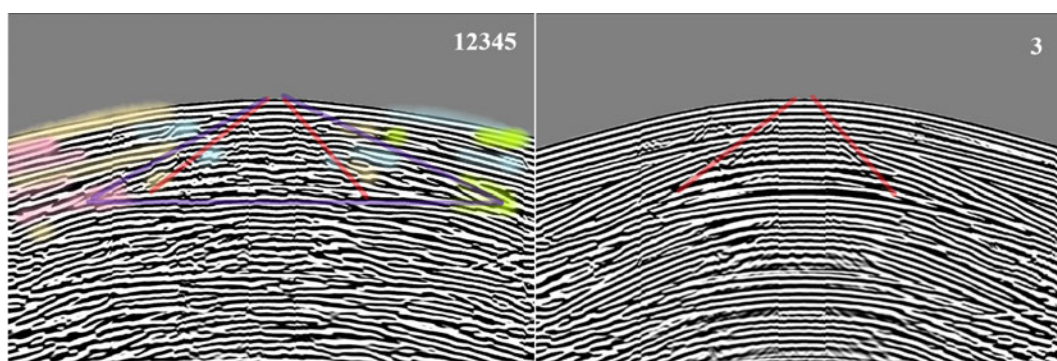
Другая картина формируется в случае миграции газа в вертикальном направлении. При подобных условиях форма основного оклика зависит преимущественно от наиболее близкого к поверхности резервуара. На фиг. 2 это проявляется в сходстве картин в постановках с несколькими включениями и при наличии только верхнего из них. Нижние резервуары привносят лишь дополнительные осцилляции в волновой фронт, которые приходят к дневной поверхности значительно позднее. Фиг. 2 также демонстрирует, как при увеличении числа резервуаров и приближении газа к поверхности первое отражение происходит раньше. В результате, хотя по приходящему к дневной поверхности сигналу точное число резервуаров может быть сложно определить, уменьшение времени прихода откликов гарантирует распространение газа в вертикальном направлении. Таким образом, благодаря периодическим измерениям и последовательным их сравнениям появляется возможность обнаружить миграцию метана и предотвратить потенциальный урон, связанный с выбросом газа и образованием воронок.

4.2. Синтетические сейсмограммы

Выводы, полученные при рассмотрении волновых картин, подтверждаются синтезированными сейсмограммами, изображенными на фиг. 3, 4 для случая с горизонтальным распространением метана и на фиг. 5 для миграции газа в вертикальном направлении. Для начала, так как волновой фронт имеет большое число осцилляций, многократные отклики также проявляются на сейсмограммах. Таким образом, наглядно демонстрируется негативное влияние кратных волн на возможность интерпретации результатов. Сходно с волновыми картинками на фиг. 2, несимметричность отражений из-за криволинейных границ областей также видна на сейсмограммах. К тому же, на необработанных картинах наблюдается распространение исходного фронта и видны поверхностные волны, сильно скрывающие отклики от резервуаров. В дополнение к этому во второй половине расчета появляются отражения от боковых границ расчетной области, артефакты, связанные с несовершенством выбранных граничных условий. С одной стороны, данные волны нефизичны и поэтому представляют собой проблему. Однако, с другой стороны, они появляются ближе к концу расчетов, значительно позднее того, как начинают регистрироваться интересные отражения. В связи с подобными эффектами возникает необходимость вычисления разности между сигналами в постановке с включениями и без них. В итоге появляется возможность анализировать только полезные волны, а именно, отклики от резервуаров.

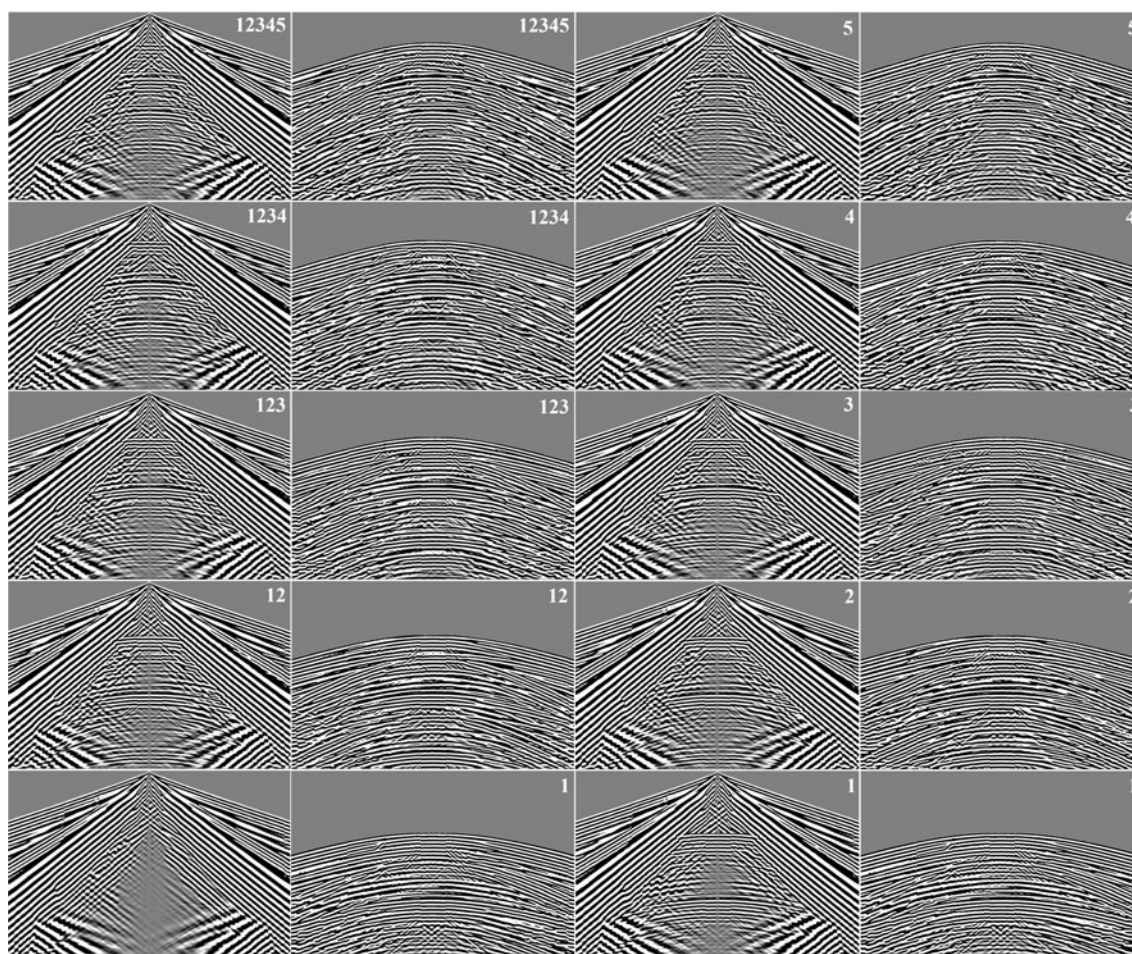


Фиг. 3. Сейсмограммы в случае с горизонтальным распространением резервуаров. Первый, третий столбцы — изначальные сейсмограммы, второй и четвертый — разница между постановками с резервуарами и без них. Цифрами указаны номера резервуаров, включенных в расчетную область.



Фиг. 4. Обработанные сейсмограммы в случае с горизонтальным распространением резервуаров. Слева — постановка с 5 резервуарами, справа — только с центральным. Цветами обозначены части картины, вызванные определенными включениями, линии соответствуют визуально выделяемым структурам.

Согласно нижним двум рядам на фиг. 3 отражение от каждого резервуара в отдельности соответствует их расположению по отношению к источнику сигнала. Отклик от центрального резервуара оказывается более симметричным по сравнению с остальными. Отраженные волны от включений, располагающихся ближе к ис-



Фиг. 5. Сейсмограммы в случае с вертикальным распространением резервуаров. Расположение аналогично фиг. 3.

точнику, приходят к дневной поверхности раньше остальных и фиксируются вначале приемниками, находящимся непосредственно над ними. Данные отклики наблюдаются как на обработанных сейсмограммах, так и на исходных. В случае наличия нескольких резервуаров отражения от каждого включения накладываются друг на друга, что затрудняет их идентифицирование. Тем не менее, при рассмотрении обработанных сейсмограмм последовательно, начиная со случая с одним включением и заканчивая постановкой со всеми, появляется возможность проследить возникновение дополнительных волн. Таким образом, удастся заметить постепенное нарушение исходной картины и появление новой структуры. На исходных сейсмограммах отличия видны гораздо хуже. Однако при миграции газа амплитуда откликов постепенно становится больше и уже при наличии трех включений она оказывается сопоставима с поверхностными волнами. В итоге совокупный отклик получается легче идентифицировать.

Фигура 4 позволяет оценить сейсмограммы до и после миграции газа в горизонтальном направлении. Цветами дополнительно показаны вклады откликов от отдельных включений в общее отражение. Светло-розовым цветом указано влияние отклика от 1 резервуара, голубым — от 2, светло-желтым — от 4, ярко-зеленым — от 5. Резервуар 3 создает базу, на которую накладываются остальные волны. Отклики только от центрального включения формируют треугольную структуру, выделенную красными линиями. При миграции газа данная структура расширяется до фиолетового треугольника, аналогично эффектам на волновых картинах на фиг. 2. Таким образом, детектирование распространения газа в горизонтальном направлении требует последовательных измерений и детального анализа изменения структур откликов на сейсмограммах.

В отличие от случая миграции метана в горизонтальном направлении, отслеживание его перемещения к поверхности значительно упрощается за счет возможности фиксирования постепенного уменьшения времени прихода первого отклика. Данный эффект наблюдается на исходных и обработанных сейсмограммах на фиг. 5. К тому же, аналогично волнам на фиг. 2 в случае наличия нескольких включений структуры откликов, в ос-

новном, определяются отражениями от наиболее близкого к верхней границе резервуара. При этом вклад в регистрируемый сигнал остальных резервуаров начинает проявляться позднее, что особенно заметно в трассах из отдаленных от центра приемников сигнала. Таким образом, хотя в каждой сейсмограмме может быть затруднительно определить конкретное число резервуаров, последовательное сравнение измерений позволяет эффективно регистрировать миграцию газа.

5. ЗАКЛЮЧЕНИЕ

В результате настоящей работы методами численного моделирования были решены прямые задачи сейсморазведки в условиях вечной мерзлоты при миграции метана в горизонтальном и вертикальном направлениях. Была создана постановка задачи, отражающие реальные упругие и геометрические характеристики многолетнемерзлых песчаных пород в полярных и приполярных прибрежных зонах Арктического региона. Грунт разделялся на узкий верхний талый слой, под которым располагались более плотные пласты. Все границы между слоями задавались криволинейными. Распространение газа моделировалось с помощью постепенного увеличения числа метановых включений. Использование сеточно-характеристического метода позволило с достаточной точностью воспроизвести сложные волновые структуры и отражения от различных областей расчетной постановки.

Подробно изучено волновое многообразие на полученных волновых картинах и сейсмограммах. Таким образом, были обнаружены волны Рэлея на свободной дневной поверхности. К тому же наличие слоистой структуры привело к возникновению кратных волн, привносящих осцилляции в изначальный фронт и значительно затрудняющих трактовку построенных сейсмограмм. Для облегчения анализа результатов были дополнительно рассмотрены сейсмограммы разности между измерениями с включениями и без них. В случае миграции газа в горизонтальном направлении было обнаружено образование треугольных структур на волновых картинах и сейсмограммах, с увеличивающимся основанием при постепенном добавлении в постановку большего числа включений. При распространении метана к дневной поверхности оказалось, что основной вклад в результаты измерений вносит только верхний резервуар. В итоге появляется возможность связать появление новых газовых карманов с уменьшением времени регистрации первого отклика. Таким образом, были зафиксированы характерные структуры на волновых картинах и сейсмограммах, позволяющие определить направление распространения газа.

Предложенные постановки задачи могут использоваться при решении обратных задач сейсморазведки. Представленный анализ результатов может помочь в трактовке подобных измерений. Дальнейшие улучшения построенных моделей заключаются в рассмотрении дополнительных ледовых включений, возможных полостей и нефтенасыщенных слоев. К тому же, опора на конкретные геологические разрезы и сравнение с натурными данными могут значительно повысить качество проводимого анализа. Для получения результатов, наиболее приближенных к реальным условиям, представляется необходимым проведение трехмерных расчетов. Все указанные задачи являются направлениями предстоящей работы.

СПИСОК ЛИТЕРАТУРЫ

1. *Li L.* Special issue on numerical modeling in civil and mining geotechnical engineering // *Processes*. 2022. V. 10. № 8. 1571.
2. *Bath M.* Introduction to seismology. V. 27. Birkhauser, 2013.
3. *Romanovsky V.E., Drozdov D.S., Oberman N.G., Malkova G.V., Kholodov A.L., Marchenko S.S., Moskalenko N.G., Sergeev D.O., Ukraintseva N.G., Abramov A.A., Gilichinsky D.A., Vasiliev A.A.* Thermal state of permafrost in Russia // *Permafrost Periglac.* 2010. V. 21. P. 136–155.
4. *Bogoyavlensky V.I., Kishankov A., Yanchevskaya A., Bogoyavlensky I.V.* Forecast of gas hydrates distribution zones in the Arctic Ocean and adjacent offshore areas // *Geosciences*. 2018. V. 8. № 12. P. 453.
5. *Anisimov O., Zimov S.* Thawing permafrost and methane emission in Siberia: Synthesis of observations, reanalysis, and predictive modeling // *Ambio*. 2021. V. 50. P. 2050–2059.
6. *Streletskiy D.A., Clemens S., Lanckman J.-P., Shiklomanov N.I.* The costs of Arctic infrastructure damages due to permafrost degradation // *Environmental Res. Lett.* 2023. V. 18. № 1. 015006.
7. *Bogoyavlensky V.I., Bogoyavlensky I.V., Nikonov R.A., Kargina T.N., Chuvilin E., Bukhanov B., Umnikov A.* New catastrophic gas blowout and giant crater on the Yamal Peninsula in 2020: Results of the expedition and data processing // *Geosciences*. 2021. V. 11. № 2. P. 71.

8. *Богоявленский В.И., Богоявленский И.В., Никонов Р.А., Каргина Т.Н.* Мониторинг развития Мордыяхского объекта взрыва газа на Ямале на основе данных дистанционного зондирования Земли // Арктика: экология и экономика. 2022. Т. 12. № 4. С. 513–523.
9. *Glikson A.* The methane time bomb // *Energy Procedia*. 2018. V. 146. P. 23–29.
10. *Favorskaya A.V., Petrov I.B.* Wave responses from oil reservoirs in the Arctic shelf zone // *Dokl. Earth Sc.* 2016. V. 466. P. 214–217.
11. *Favorskaya A.V., Petrov I.B.* Grid-Characteristic Method // In: *Favorskaya, A., Petrov, I. (eds) Innovations in Wave Processes Modelling and Decision Making. Smart Innovation, Systems and Technologies*. 2018. V. 90. P. 117–160.
12. *Petrov I.B., Kozhemyachenko A.A., Favorskaya A.V.* Simulation of Vibrations of Railway Structures by the Grid-Characteristic Method // *Dokl. Phys.* 2023. V. 68. P. 30–34.
13. *Golubev V.I., Nikitin I.S., Beklemysheva K.A.* Model of fractured medium and nondestructive control of composite materials // *Chinese J. Aeronautics*. 2023. V. 37. № 2. P. 93–99.
14. *Новацкий В.* Теория упругости. М.: Мир, 1975.
15. *Stognii P.V., Petrov I.B., Khokhlov N.I.* Numerical modelling of the reflected waves from gas cavities and methane bombs in the heterogeneous media of the Northern regions // *IOP Conf. Ser.: Materials Science and Engineering*. 2020. V. 927. № 1. 012006.
16. *Stognii P.V., Khokhlov N.I., Petrov I.B., Favorskaya A.V.* The comparison of two approaches to modeling the seismic waves spread in the heterogeneous 2D medium with gas cavities // In: *Favorskaya, M.N., Favorskaya, A.V., Petrov, I.B., Jain, L.C. (eds) Smart Modelling For Engineering Systems. Smart Innovation, Systems and Technologies*. 2021. V. 214 P. 101–114.
17. *Stognii P.V., Khokhlov N.I., Petrov I.B.* Numerical modeling of wave processes in multilayered media with gas-containing layers: comparison of 2D and 3D models // *Dokl. Math.* 2019. V. 100. P. 586–588.
18. *Rusanov V.V.* Difference schemes of the third order of accuracy for the forward calculation of discontinuous solutions // *Doklady Akademii Nauk*. 1968. V. 180. P. 1303–1305.
19. *Guseva E.K., Golubev V.I., Petrov I.B.* Linear quasi-monotone and hybrid grid-characteristic schemes for the numerical solution of linear acoustic problems // *Numer. Analys. Appl.* 2023. V. 16. P. 112–122.
20. *Трифонов Б.А., Милановский С.Ю., Несынов В.В.* Оценка сейсмических воздействий в условиях деградации мерзлоты // *Вестник КРАУНЦ. Сер.: Науки о Земле*. 2022. Т. 56. № 4. С. 59–74.
21. *Fatyanov A.G.* A wave method of suppressing multiple waves for any complex subsurface geometry // *Numer. Analys. Appl.* 2012. V. 5. P. 187–190.
22. *Nkemzi D.* A new formula for the velocity of Rayleigh waves // *Wave Motion*. 1997. V. 6. № 2. P. 199–205.

NUMERICAL CALCULATION OF THE EFFECT OF THE METHANE MIGRATION PROCESS ON THE RESULTS OF SEISMIC EXPLORATION IN PERMAFROST ZONES

E. K. Guseva*, V. I. Golubev**, I. B. Petrov***

141701 Dolgoprudny, Institutsky Lane, 9, MIPT (National Research University), Russia

** e-mail: guseva.ek@phystech.edu*

*** e-mail: golubev.vi@mipt.ru*

****e-mail: petrov@mipt.ru*

Received: 23.02.2024

Revised: 27.04.2024

Accepted: 31.05.2024

Abstract. For a long time, intense methane emissions into the atmosphere from the subsurface of permafrost have been observed in the shelf and coastal zones of the Arctic region. Due to the potential danger of this phenomenon to the environment and infrastructure, there is a need for periodic monitoring of gas pockets, including through ground-based seismic exploration. This paper contains the results of a study of this process using numerical modeling. A model of layered frozen sandy soil with curvilinear boundaries between the sheets is constructed, reflecting the main specifics of the region. The process of gas migration in vertical and horizontal directions is reconstructed by increasing the number of methane reservoirs. The defining system of hyperbolic equations of the linear theory of elasticity is used. The problem is solved by the grid-characteristic method in a two-dimensional formulation on a rectangular grid, in each cell of which the elastic characteristics of the layers are set. The resulting wave structures are studied in detail on the obtained synthetic wave patterns and seismograms, which make it possible to determine the direction of gas propagation. The results obtained can be used to interpret similar field experiments.

Keywords: numerical modeling, direct seismic survey, permafrost, gas migration, grid-characteristic method.

СПУТНИК НА ЭЛЛИПТИЧЕСКОЙ ОРБИТЕ: О ЧИСЛЕННОМ ОБНАРУЖЕНИИ ПЕРИОДИЧЕСКИХ ДВИЖЕНИЙ И ИССЛЕДОВАНИИ ИХ УСТОЙЧИВОСТИ

© 2024 г. А.А. Буров^{1,*}, В.И. Никонов^{1,**}

¹ 119333 Москва, ул. Вавилова, 44, ФИЦ ИУ РАН, Россия

*e-mail: jtm@narod.ru

**e-mail: nikon_v@list.ru

Поступила в редакцию 09.10.2023 г.

Переработанный вариант 09.10.2023 г.

Принята к публикации 31.05.2024 г.

Рассматриваются уравнения плоских колебаний спутника на эллиптической орбите. Для численного обнаружения периодических решений применяется сочетание метода сечений Пуанкаре и предложенного ранее подхода, опирающийся на аналог принципа сжимающих отображений. Численно выявлен ряд классов периодических решений и исследованы необходимые условия их устойчивости. Этим движениям уделяется особое внимание, поскольку в общем случае они трудно поддаются аналитическому изучению. Библ. 47. Фиг. 6.

Ключевые слова: плоские движения спутника на эллиптической орбите, отображение Пуанкаре, инвариантные торы, хаотическая динамика, уравнение Белецкого, периодические движения, необходимые условия устойчивости, теория Ляпунова–Флоке.

DOI: 10.31857/S0044466924090124, EDN: WJLLDO

ВВЕДЕНИЕ

Уравнение колебаний спутника на эллиптической орбите, известное в литературе как уравнение В. В. Белецкого (см. [1, 2]), можно отнести к уравнениям маятникового типа с периодическим возбуждением. В случаях когда орбита спутника отлична от круговой, уравнения движения оказываются неинтегрируемыми, и оказывается актуальной задача выявления устойчивых периодических решений, особенно для больших значений эксцентриситета. Как было подмечено (см. [3–5]), существенную роль при отыскании таких решений играет “стробоскопический метод”, по существу совпадающий с методом отображения Пуанкаре за период.

Сочетание этого метода с обсуждавшимся авторами ранее подходом, опирающимся на визуальное выявление областей, и дальнейшим итерационным численно-аналитическим, данный метод предоставляет определенные возможности для отыскания периодических решений. В настоящей работе этот подход применяется для выявления периодических движений спутника при значениях эксцентриситета, существенно отличных от нуля. Необходимые условия устойчивости выявленных движений проверяются с помощью теории Ляпунова–Флоке.

Работа посвящена 60-летию публикаций [3–5].

1. СИСТЕМЫ МАЯТНИКОВОГО ТИПА

Рассмотрим систему вида

$$\dot{q} = v, \quad \dot{v} = f(v, q, t), \quad (1.1)$$

где $f : S^1(q) \times R^1(v) \times R^1(t) \rightarrow R$ — функция, гладкая по q и v , периодическая по q и t с периодом 2π . В общем случае функция f также зависит от параметров. Системы вида (1.1) называют системами маятникового типа.

Для таких уравнений, как правило численно, строится отображение Пуанкаре за период, т.е. отображения плоскости $t = 0$ на себя за период T вдоль потока в расширенном фазовом пространстве (q, v, t) . Наблюдения за следами траекторий системы на этой плоскости для типичных систем вида (1.1), не близких к интегрируемым, в общем случае позволяет выявить области визуально “регулярного” поведения, окруженные областями “нерегулярного” поведения¹.

¹ Речь идет об “островах” и целых “архипелагах” в “море хаоса” в терминологии, предложенной В.В. Белецким [9].

С точки зрения выявления устойчивых периодических движений перспективными выглядят области фазового пространства, близкие на сечениях Пуанкаре к “островам регулярности”, являющихся, вообще говоря, следами не до конца разрушенных или деформированных инвариантных торов. Для отыскания самих периодических решений предлагается воспользоваться подходом, предложенным авторами в [6]. Этот способ состоит в следующем. Прежде всего фиксируются параметры задачи. Потом численно строится отображение Пуанкаре за период и визуально (“эмпирически”) выделяется “кандидат” на звание “острова регулярности” или, иначе говоря, “следа разрушенного или деформированного тора”. Выбирается начальное значение, принадлежащее этому следу, по нему строится некоторая часть траектории, представленная в силу использования конечно-разностного численного моделирования массивом точек, принадлежащих плоскости (q, v) . Эти точки, расположенные вблизи “острова регулярности”, с помощью метода наименьших квадратов приближаются кривой второго порядка. Как правило, такая кривая оказывается эллипсом. Далее, аналитически вычисляется центр этой кривой, после чего итерация повторяется: в качестве приближенного решения на этом шаге итераций выбирается решение с начальными условиями, отвечающими центру кривой второго порядка, найденному на предыдущем шаге. Как показали рассмотренные ранее примеры [6, 7], такая сходимость оказывается достаточно быстрой, при том, что в рассмотренных примерах правые части уравнений вида (1.1) оказываются разрывными функциями времени и претерпевают разрывы первого рода.

Естественно ожидать, что найденное таким образом периодическое решение будет устойчиво, по крайней мере в первом приближении. Для проверки этого утверждения выполняется следующее. Найденное периодическое решение, представленное, вообще говоря, конечным набором точек, приближается тригонометрическим полиномом высокого порядка. Далее, осуществляется линеаризация уравнений движения в окрестности найденного решения. Затем численно определяется матрица монодромии. Наконец, на основе анализа ее следа с помощью теории Ляпунова–Флоке (см., например, [8]) делается вывод об устойчивости в первом приближении (или неустойчивости) изучаемого периодического решения.

2. ПЛОСКИЕ КОЛЕБАНИЯ И ВРАЩЕНИЯ СПУТНИКА НА ЭЛЛИПТИЧЕСКОЙ ОРБИТЕ

Применим описанный подход к задаче о плоских колебаниях спутника на эллиптической орбите. Как известно, в рамках такой постановки движения описываются уравнением Белецкого, представимом в виде [1] (см. также [2])

$$(1 + e \cos v) \ddot{\delta} + n^2 \sin \delta = 2e\dot{\delta} \sin v + 4e \sin v, \quad 0 \leq n^2 = 3 \frac{A - C}{B} \leq 3. \quad (2.1)$$

Здесь v — истинная аномалия центра масс спутника, используемая в качестве независимой переменной, δ — угол, характеризующий поворот спутника относительно орбитальной системы отсчета, A , B и C — главные центральные моменты инерции спутника, причем момент инерции B отвечает оси, перпендикулярной к плоскости орбиты, $0 \leq e < 1$ — эксцентриситет орбиты центра масс.

То же уравнение в нормальной форме Коши имеет вид системы второго порядка

$$\dot{\delta} = v, \quad \dot{v} = \frac{1}{1 + e \cos v} (2ev \sin v - n^2 \sin \delta + 4e \sin v), \quad (2.2)$$

которое и составит предмет дальнейшего обсуждения².

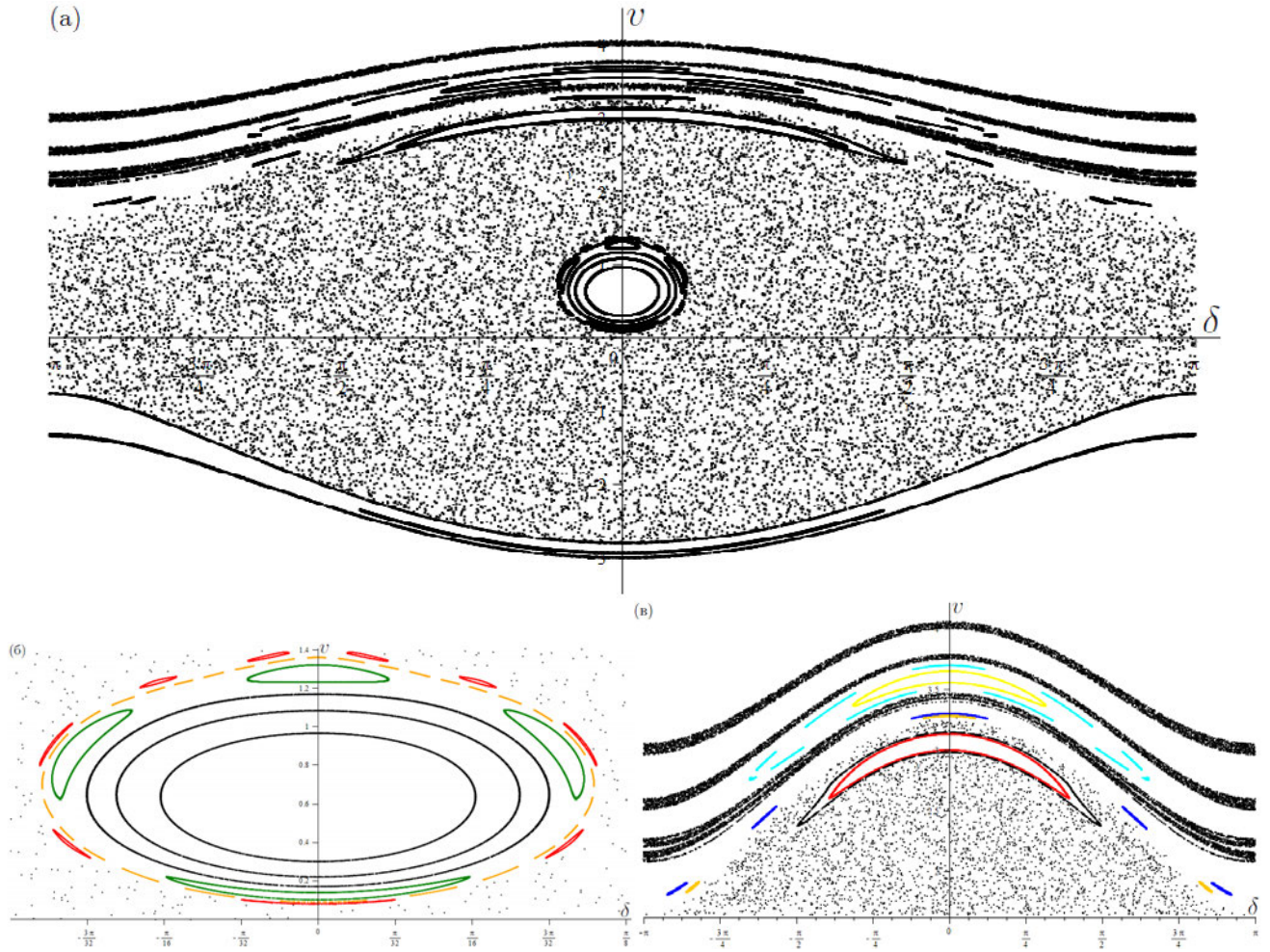
2.1. Периодические решения при существенно отличных от нуля значениях эксцентриситета

Как известно (см. [10]), при достаточно малых значениях эксцентриситета $e \neq 0$ сепаратрисы расщепляются, и в их окрестности образуется тонкий хаотический слой. Также разрушаются резонансные торы и слегка деформируются нерезонансные торы, что обеспечивает, благодаря теореме Мозера о повороте кольца, существование большого числа периодических движений.

Применим предлагаемую технику к случаю, когда $e = 0.16$, $n^2 = 2$. Общая картина представлена на фиг. 1а. Из нее с целью дальнейшего исследования на плоскости $[\delta, v]$ были выделены области A : $[-\pi/8, \pi/8] \times [0, 1.4]$ (фиг. 1б) и B : $[-3\pi/4, 3\pi/4] \times [1.5, 4]$ (фиг. 1в).

Область А. Прежде всего обратим внимание на область A . Черному острову предположительно отвечает периодическое решение A_I — колебания периода 2π . Применение к нему изложенного метода за 16 итераций позволяет с точностью 10^{-11} выявить это периодическое решение: оно отвечает начальным условиям $\delta(0) = 0.0000002735$, $v(0) = 0.6094296495$. Это движение изображено на фиг. 2. Вычисления для этого решения матрицы монодромии показывают, что на этом решение ее след составляет -0.787747701 , и поэтому выполнены необходимые условия устойчивости.

²Предъявленное в [10] гамильтоново представление уравнений движения в работе не используется.



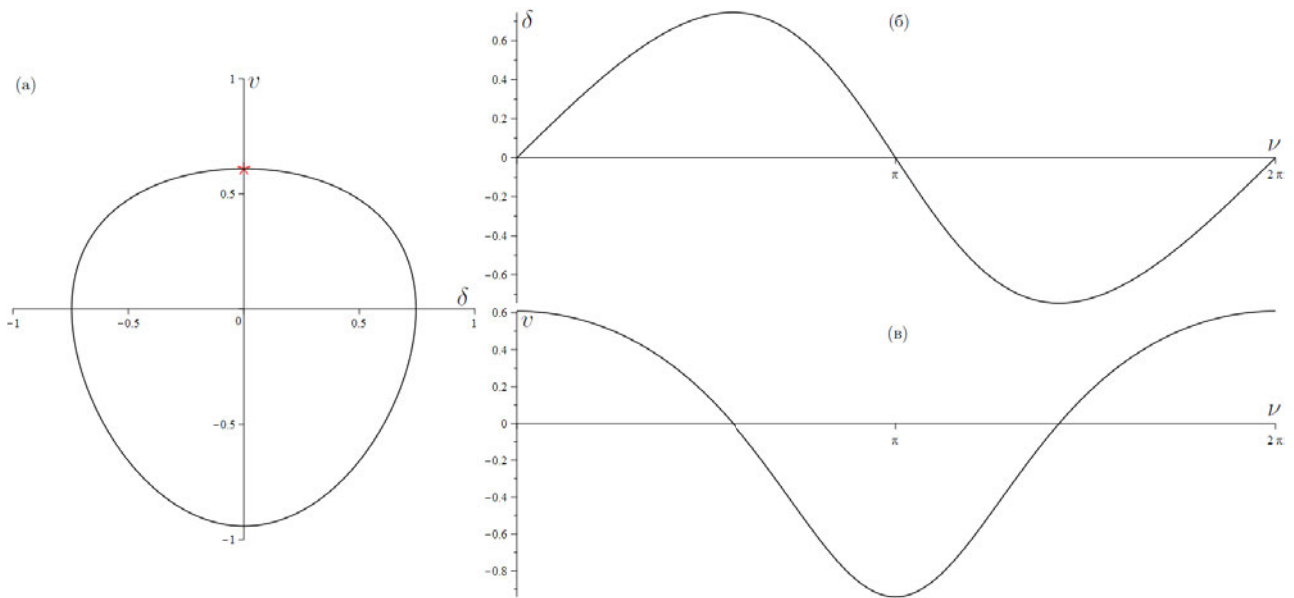
Фиг. 1. Сечение Пуанкаре при $e = 0.16$, $n^2 = 2$: (а) — общий вид, (б), (в) — выделенные области А и В.

Наблюдаемый архипелаг из четырех зеленых островов предположительно отвечает периодическому решению A_{II} — колебания периода 8π . Применение к ним изложенного метода за 15 итераций позволяет с точностью 10^{-9} выявить это периодическое решение: оно отвечает начальным условиям $\delta(0) = -0.00001348$, $v(0) = 1.27508013$. Это движение изображено на фиг. 3. Вычисления для этого решения матрицы монодромии показывают, что на этом решение ее след составляет 1.577671089, и поэтому выполнены необходимые условия устойчивости.

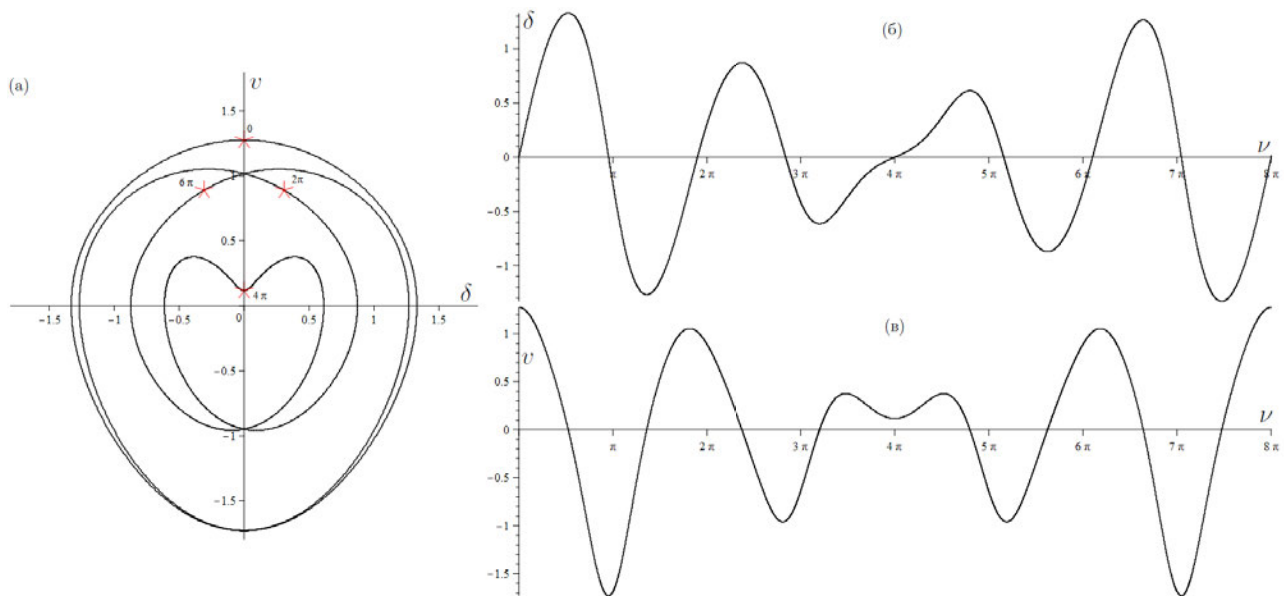
Острова, изображенные красным на фиг. 1б, отвечают периодическому решению A_{III} — колебания периода 18π . Применение к ним изложенного метода за 14 итераций позволяет с точностью 10^{-9} выявить это периодическое решение: оно отвечает начальным условиям $\delta(0) = 0.00013070$, $v(0) = 0.07810302$. Это движение изображено на фиг. 4. Вычисления для этого решения матрицы монодромии показывают, что на этом решение ее след составляет 0.03202327, и поэтому выполнены необходимые условия устойчивости.

Пунктирная кривая, изображенная желтым, на первый взгляд представляет собой остатки “разрушенного тора”. Однако увеличение картины позволяет установить, что за каждым пунктирным штрихом стоит “остров регулярности”, продольные размеры которого существенно превосходят поперечные размеры. Применение к ним изложенного метода за 14 итераций позволяет с точностью 10^{-9} выявить это 60π периодическое решение A_{IV} : оно отвечает начальным условиям $\delta(0) = -0.00074051$, $v(0) = 1.35920192$. Это движение изображено на фиг. 5. Вычисления для этого решения матрицы монодромии показывают, что на этом решение ее след составляет -1.99980036 , и поэтому выполнены необходимые условия устойчивости.

Замечание 1. Представляет интерес ответ на вопрос, принадлежит ли найденное решение хотя бы одному из классов периодических решений, найденных недавно в [11].



Фиг. 2. Колебание A_I . Период 2π . Начальные условия: $(\delta(0), v(0)) = (0.0000002735, 0.6094296495)$.



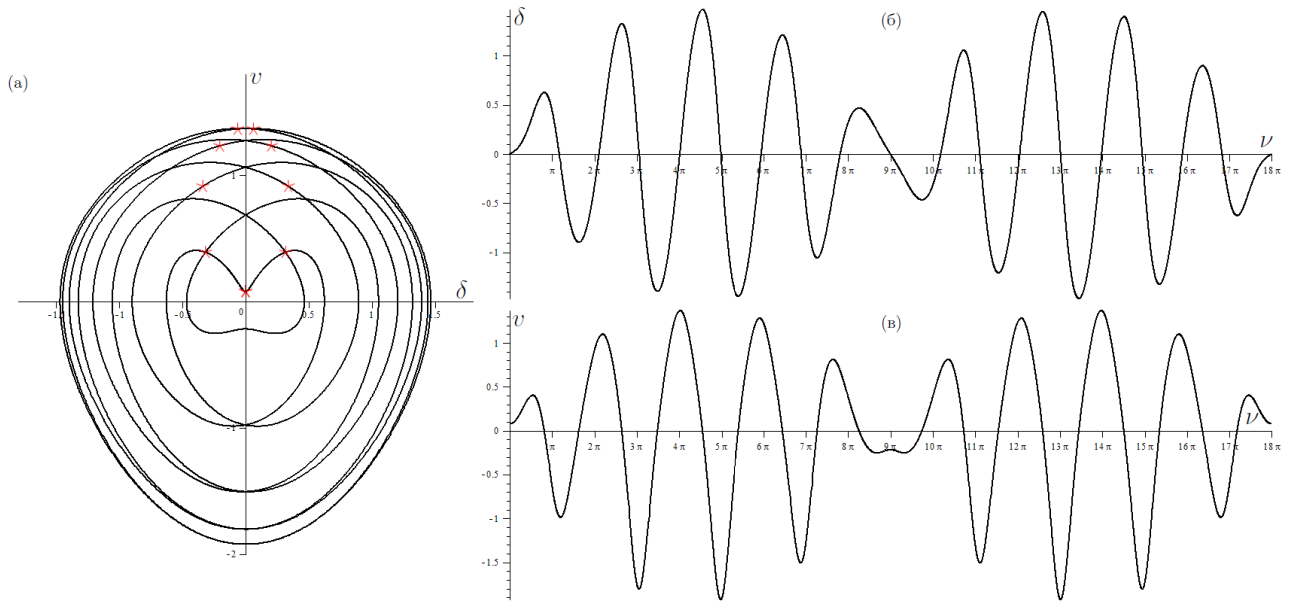
Фиг. 3. Колебание A_{II} . Период 8π . Начальные условия: $(\delta(0), v(0)) = (-0.00001348, 1.27508013)$.

Область В. Обратим теперь внимание на область В. Островам из этой области отвечают вращения, на которых скорости являются положительными периодическими функциями.

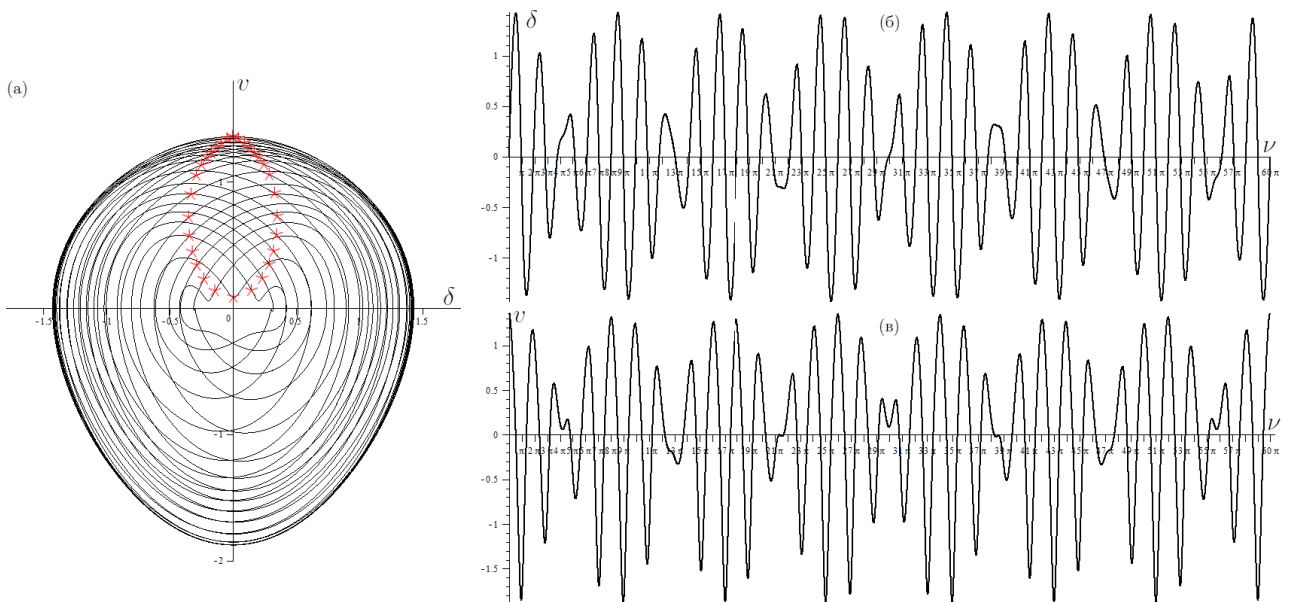
Так, острову, изображенному красным на фиг. 1в, отвечает вращение B_I с 2π периодическим изменением скорости. Применяя изложенный выше метод, за 13 итераций удалось с точностью 10^{-9} выявить это решение: оно отвечает начальным условиям $\delta(0) = 0.00007554$, $v(0) = 3.06266105$.

Рыжим островам (фиг. 1в) отвечает вращение B_{II} с 6π периодическим изменением скорости. Применяя изложенный выше метод, за 16 итераций удалось с точностью 10^{-9} выявить это периодическое решение: оно отвечает начальным условиям $\delta(0) = 0.00130513$, $v(0) = 3.28182441$.

Островам, изображенным синим (фиг. 1в), отвечает вращение B_{III} с 10π периодическим изменением скорости. Применяя изложенный выше метод, за 15 итераций удалось с точностью 10^{-9} выявить это периодическое решение: оно отвечает начальным условиям $\delta(0) = 0.00471541$, $v(0) = 3.30139051$.



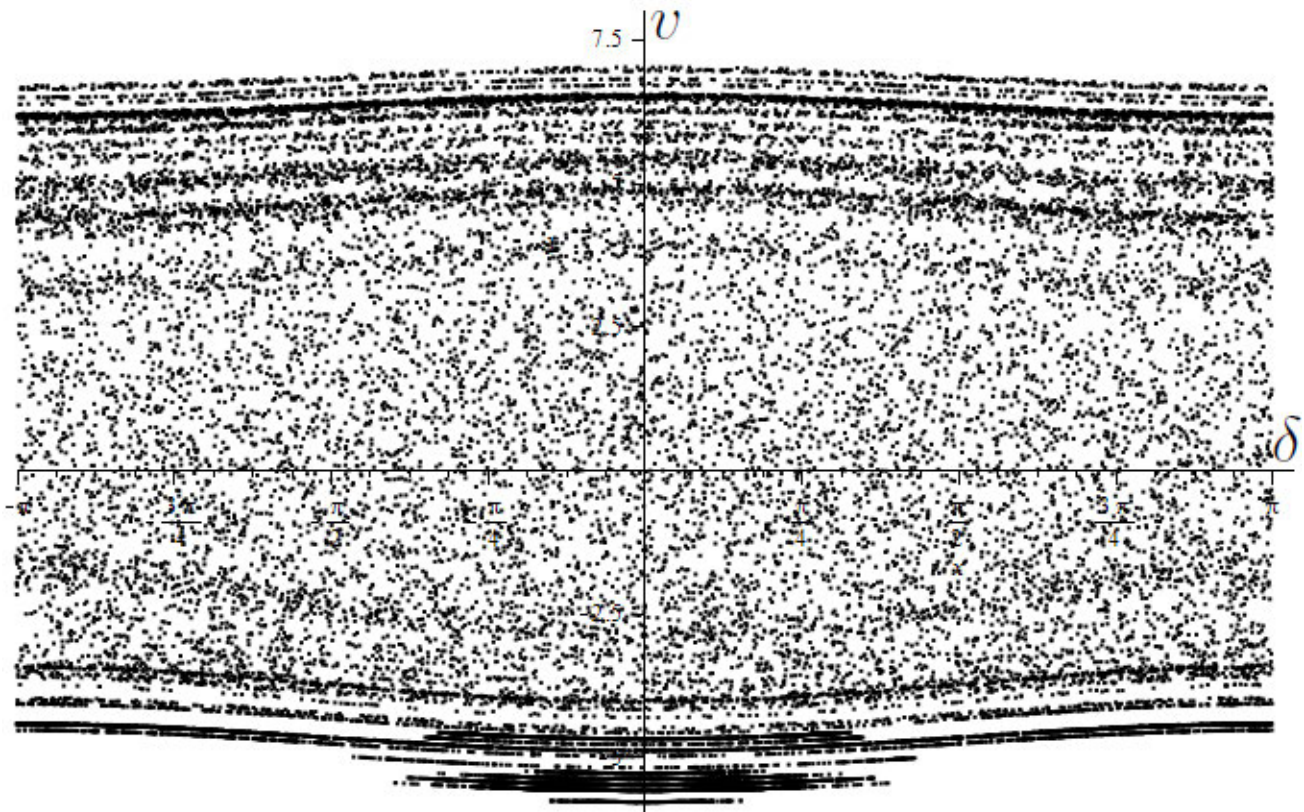
Фиг. 4. Колебание A_{III} . Период 18π . Начальные условия: $(\delta(0), v(0)) = (-0.00074051, 1.35920192)$.



Фиг. 5. Колебание A_{IV} . Период 60π . Начальные условия: $(\delta(0), v(0)) = (0.00013070, 0.07810302)$.

Голубым островам (фиг. 1в) отвечает вращение B_{IV} с 22π периодическим изменением скорости. Применяя изложенный выше метод, за 14 итераций удалось с точностью 10^{-9} выявить это периодическое решение: оно отвечает начальным условиям $\delta(0) = -0.00265124$, $v(0) = 3.69945572$.

Наконец, желтому острову (фиг. 1в) отвечает вращение B_V с 2π периодическим изменением скорости. Применяя изложенный выше метод, за 16 итераций удалось с точностью 10^{-9} выявить это периодическое решение: оно отвечает начальным условиям $\delta(0) = 0.00024514$, $v(0) = 3.60198732$.



Фиг. 6. Сечение Пуанкаре при $e = 0.95$, $n^2 = 2$.

2.2. Развитой хаос при больших значениях $e > 0$

Как было показано в предыдущем разделе, при $e = 0.16$ развитой хаос имеет место не всюду в фазовом пространстве. Картина существенно меняется для достаточно больших значений $e > 0$. Так при $e = 0.95$ сплошное “море хаоса”, см. фиг. 6. Отсутствие наблюдаемых следов разрушенных или деформированных торов делает непригодным применение описанного выше алгоритма по выявлению периодических решений.

3. ИСТОРИЧЕСКИЕ ЗАМЕЧАНИЯ

В настоящей работе основное внимание уделяется применению метода нахождения периодических движений к уравнению плоских колебаний спутника на эллиптической орбите. Это уравнение, выведенное в [1] (см. также [2]), является предметом многочисленных исследований как точными, так и приближенными, численными методами. Так исследование периодических решений в зависимости от значений эксцентриситета, восходящее к работам [3–5, 12, 13], продолженное в многочисленных публикациях, потребовало, в частности, разработки, адаптации и применения весьма тонких методов анализа нелинейных дифференциальных уравнений (см. [14–24]). Для регуляризации уравнений движения при значениях эксцентриситета, близких к единице, применялась теория степени Лере-Шаудера (см. [25–27]).

Из многочисленных работ, посвященных различным свойствам решений уравнения Белецкого, в частности, случаям быстрым вращениям и резонансным явлениям, выделим публикации [14, 28–38].

Неинтегрируемость уравнений движения, обусловленная расщеплением сепаратрис, при значениях эксцентриситета, близких к нулю, доказана в [10] (см. также [39–42]). При построении отображений Пуанкаре для уравнений Белецкого, посчитанных при разных значениях параметров в [9] (см., также, например, [43]), отчетливо наблюдаются “острова регулярности”, составляющие “архипелаги регулярности” в “море хаоса”, являющиеся проявлением резонансных явлений (см., например, [44]). Выделим также публикацию [45], посвященную численному исследованию изменения ориентации малых спутников планет при хаотическом вращении. Численному исследованию бифуркаций периодических решений и их устойчивости посвящена работа [11] (см. также [46]). Полное решение вопроса о соответствии обнаруженных численно бифуркаций бифуркациям, обнаруженных ранее аналитически в работах А.Д. Бруно и соавторов остается открытым.

СПИСОК ЛИТЕРАТУРЫ

1. Белецкий В. В. О либрации спутника // Сборник "Искусственные спутники Земли". 1959. № 3. М.: АН СССР. С. 13–31.
2. Белецкий В. В. Движение искусственного спутника относительно центра масс. М.: Наука, 1965. 416 с.
3. Златоустов В. А., Охоцимский Д. Е., Сарычев В. А., Торжеский А. П. Исследование колебаний спутника в плоскости эллиптической орбиты // Космические исследования. 1964. Т. 2. № 5. С. 657–666.
4. Торжеский А. П. Периодические решения уравнения плоских колебаний спутника на эллиптической орбите // Космические исследования. 1964. Т. 2. № 5. С. 667–678.
5. Zlatoustov V. A., Okhotsimsky D. E., Sarychev V. A., Torzhesky A. P. Investigation of a Satellite Oscillations in the Plane of an Elliptic Orbit // The Eleventh International Congress of Applied Mechanics, Munich, Germany. 1964. P. 436–439.
6. Burov A. A., Nikonov V. I. On the nonlinear Meissner equation // International Journal of Non-Linear Mechanics. 2019. V. 110. P. 26–32.
7. Burov A. A., Nikonov V. I. On the motion of the pendulum in an alternating, sawtooth force field // International Journal of Bifurcation and Chaos in Applied Sciences and Engineering. 2020. V. 30. Art No. 2050135.
8. Арнольд В. И. Математические методы классической механики. М.: Наука. 1974.
9. Beletsky V. V. Regulare und Chaotische Bewegung Starrer Korper, Teubner Studienbucher Mechanik, B. G. Teubner, Stuttgart. 1995.
10. Буров А. А. Неинтегрируемость уравнения плоских колебаний спутника на эллиптической орбите // Вестник Московского университета. Серия 1: Математика. Механика. 1984. № 1. С. 71–73.
11. Koh D., Flashner H. Global Analysis of Gravity Gradient Satellite's Pitch Motion in an Elliptic Orbit // Journal of Computational and Nonlinear Dynamics. 2015. V. 10. № 6. Art No. 061020.
12. Baker R. M. L. Librations on a Slightly Eccentric Orbit // ARS Journal. 1960. V. 30. № 1. P. 124–126.
13. Schechter H. B. Dumbbell Librations in Elliptic Orbits // AIAA Journal. 1964. V. 2. № 6. P. 1000–1003.
14. Vinh, Nguyen X. Sur les Solutions Periodiques du Mouvement Plan de Libration des Satellites et des Planetes // Celestial Mechanics. 1973. V. 8. № 3. P. 371–403.
15. Mawhin J. On a differential equation for the periodic motions of a satellite around its center of mass // Асимптотические методы в задачах математической физики. Киев: Издательство Института математики. 1988. С. 150–157.
16. Bruno A. D., Varin V. P. The Limit Problems for the Equation of Oscillations of a Satellite // Celestial Mechanics and Dynamical Astronomy. 1997. V. 67. № 1. P. 1–40.
17. Bruno A. D., Varin V. P. Singularities of Oscillations of a Satellite on Highly Eccentric Elliptic Orbits // Nonlinear Analysis: Theory, Methods & Applications. 1997. V. 30. № 4. P. 2541–2546.
18. Брюно А. Д., Петрович В. Ю. Вычисление периодических колебаний спутника // Математическое моделирование. 1997. Т. 9. № 6. С. 82–94.
19. Bruno A. D., Varin V. P. Generalized Periodic Solutions to the Equation of Oscillations of a Satellite // ZAMM. 1999. V. 79. № Suppl. 2. P. 283–284.
20. Varin V. P. Degeneracies of Periodic Solutions to the Beletsky Equation // Regular and Chaotic Dynamics. 2000. V. 5. № 3. P. 313–328.
21. Bruno A. D., Varin V. P. Classes of families of generalized periodic solutions to the Beletsky equation // Celestial Mechanics and Dynamical Astronomy. 2004. V. 88. P. 325–341.
22. Nuñez D., Torres P. J. Stable odd solutions of some periodic equations modeling satellite motion // Journal of Mathematical Analysis and Applications. 2003. Vol. 279. Issue 2. P. 700–709.

23. *Cabada A., Cid J. A.* On a class of singular Sturm–Liouville periodic boundary value problems // *Nonlinear Analysis: Real World Applications*. 2011. Vol. 12. Issue 4. P. 2378–2384.
24. *Брюно А. Д.* Семейства периодических решений уравнения Белецкого // *Космические исследования*. 2002. Т. 40. № 3. С. 295–316.
25. *Косенко И. И.* Регуляризация предельной задачи о колебаниях спутника на кеплеровой орбите с учетом сил светового давления // *Космические исследования*. 2002. Т. 40. № 6. С. 626–632.
26. *Косенко И. И.* Применение теории степени Лере–Шаудера для аппроксимации колебаний спутника на эллиптической орбите // *Докл. АН*. 2005. Т. 404. № 5. С. 625–627.
27. *Косенко И. И.* Топологическая степень и аппроксимация решений нерегулярных задач механики. Колебания спутника на эллиптической орбите // *Современная математика. Фундаментальные направления*. 2006. Т. 16. С. 68–95.
28. *Черноусько Ф. Л.* Резонансные явления при движении спутника относительно центра масс // *Ж. вычисл. матем. и матем. физ.* 1963. Т. 3. Вып. 3. С. 528–538.
29. *Zlatoustov V. A., Markeev A. P.* Stability of Planar Oscillations of a Satellite in an Elliptic Orbit // *Celestial Mechanics*. 1973. V. 7. № 1. P. 31–45.
30. *Сарычев В. А., Сазонов В. В., Златоустов В. А.* Периодические колебания спутника в плоскости эллиптической орбиты // *Космические исследования*. 1977. Т. 15. № 6. С. 809–834.
31. *Сарычев В. А., Сазонов В. В., Златоустов В. А.* Периодические вращения спутника в плоскости эллиптической орбиты // *Космические исследования*. 1979. Т. 17. № 2. С. 190–207.
32. *Сарычев В. А., Сазонов В. В., Златоустов В. А.* Несимметричные периодические колебания спутника в плоскости эллиптической орбиты // *Космические исследования*. 1980. Т. 18. № 1. С. 3–10.
33. *Садов С. Ю.* Нормальная форма уравнения колебаний спутника в сингулярном случае // *Матем. заметки*. 1995. Т. 58. Вып. 5. С. 785–789.
34. *Sadov S.* Lissajous Solutions of the Satellite Oscillation Equation: Stability and Bifurcations via Higher Order Averaging // *Nonlinear Phenomena in Complex Systems*. 1999. V. 2. P. 96–101.
35. *Садов С. Ю.* Об устойчивости резонансного вращения спутника относительно центра масс в плоскости орбиты // *Космические исследования*. 2006. Т. 44. № 2. С. 170–181.
36. *Садов С. Ю.* Усреднение уравнения движения спутника относительно центра масс в вырожденных случаях. I. Колебания в абсолютной системе координат // *Космические исследования*. 2008. Т. 46. 2. С. 174–182.
37. *Маркеев А. П.* Об устойчивости колебаний спутника в плоскости эллиптической орбиты // *Доклады академии наук*. 2007. Т. 413. № 3. С. 340–344.
38. *Маркеев А. П.* К задаче о плоских периодических вращениях спутника на эллиптической орбите // *Известия РАН. Механика твердого тела*. 2008. № 3. С. 102–115.
39. *Koch B.-P., Bruhn B.* Chaotic and Periodic Motions of Satellites in Elliptic Orbits // *Zeitschrift fur Naturforschung A*. 1989. V. 44. P. 1155–1162.
40. *Maciejewski A. J.* Non-Integrability of the Planar Oscillations of a Satellite // *Acta Astronomica*. 1995. V. 45. P. 327–344.
41. *Teofilatto P., Graziani F.* On Librational Motion of Spacecraft // *Chaos, Solitons & Fractals*. 1996. V. 7. P. 1721–1744.
42. *Kirchgraber U., Manz U., Stoffer D.* Rigorous Proof of Chaotic Behaviour in a Dumbbell Satellite Model // *J. of Mathematical Analysis and Applications*. 2000. V. 251. № 2. P. 897–911.
43. *Tong X., Rimrott F.* Numerical Studies on Chaotic Planar Motion of Satellite in an Elliptic Orbit // *Chaos, Solitons & Fractals*. 1991. V. 1. № 2. P. 179–186.
44. *Gkolias I., Celletti A., Efthymiopoulos Ch., Pucacco G.* The theory of secondary resonances in the spin–orbit problem // *Monthly Notices of the Royal Astronomical Society*. 2016. V. 459. P. 1327–1339.

45. Мельников А. В. Ориентация фигур малых спутников планет при хаотическом вращении // *Астрономический вестник*. 2020. Т. 54. № 5. С. 458–467.
46. Earpen R. T., Majji M., Alfriend K. T. Attitude dynamics of a rigid body in Keplerian motion // *Celestial Mechanics and Dynamical Astronomy*. 2021. V. 133. Art No. 2.

A SATELLITE IN AN ELLIPTICAL ORBIT: ON THE NUMERICAL DETECTION OF PERIODIC MOTIONS AND THE STUDY OF THEIR STABILITY

A. A. Burov*, V. I. Nikonov**

119333 Moscow, Vavilov Str., 44, Federal Research Center Computer Science and Control RAS, Russia

**e-mail: jtm@narod.ru,*

***e-mail: nikon_v@list.ru*

Received: 09.10.2023

Revised: 09.10.2023

Accepted: 31.05.2024

Abstract. The equations of plane oscillations of a satellite in an elliptical orbit are considered. For the numerical detection of periodic solutions, a combination of the Poincaré cross-section method and the previously proposed approach based on an analogue of the principle of contracting maps is used. A number of classes of periodic solutions have been numerically identified and the necessary conditions for their stability have been investigated. Special attention is paid to these movements, since in general they are difficult to study analytically.

Keywords: flat motions of a satellite in an elliptical orbit, Poincaré mapping, invariant tori, chaotic dynamics, Beletsky equation, periodic motions, necessary stability conditions, Lyapunov—Floquet theory.

ЧИСЛЕННОЕ ИССЛЕДОВАНИЕ ПЕРЕХОДНЫХ РЕЖИМОВ ТЕЧЕНИЯ КОЛМОГОРОВА В КВАДРАТНОЙ ЯЧЕЙКЕ¹⁾

© 2024 г. А. О. Посудневская^{1,2,*}, С. В. Фортова², А. Н. Долуденко³, И. В. Колоколов¹,
В. В. Лебедев¹

¹ 142432 Черногловка, М.о., пр-т Акад. Семенова, 1А,

Институт теоретической физики им. Л.Д. Ландау РАН, Россия

² 123056 Москва, ул. 2-я Брестская, 19/18, Институт автоматизации проектирования РАН, Россия

³ 125412 Москва, ул. Ижорская, 13, стр.2, Объединенный институт высоких температур РАН, Россия

*e-mail: posudnevskaja.ao@phystech.edu

Поступила в редакцию 04.03.2024 г.

Переработанный вариант 04.04.2024 г.

Принята к публикации 31.05.2024 г.

Рассматривается задача двумерного течения вязкой слабосжимаемой жидкости в квадратной ячейке при возбуждении периодической в пространстве статической внешней силой (течение Колмогорова). Представлен новый метод определения структуры течения, основанный на анализе поля завихренности в различные моменты времени. Данный метод использован для классификации типов течений, характеристики которых получены при численном моделировании. Выделены основные режимы течения в зависимости от величин коэффициента трения о дно и силы накачки: ламинарный, хаотический и вихревой режимы. Отдельно исследуются переходные типы течения: квазипериодический режим, который возникает через последовательность бифуркаций при смене ламинарного и хаотического режимов течения, и режим сменяемости, возникающий при переходе от хаотического к вихревому течению. Построены фазовые диаграммы в пространстве амплитуда внешней силы - коэффициент трения о дно, позволяющие по значениям величин коэффициента трения о дно и силы накачки классифицировать тип течения. Библ. 31. Фиг. 8.

Ключевые слова: турбулентность, численное моделирование, течение Колмогорова, вязкая жидкость, вихрь.

DOI: 10.31857/S0044466924090136, **EDN:** WIXOKI

1. ВВЕДЕНИЕ

Данная работа посвящена изучению широко распространенного в природе явления — турбулентного течения вязкой жидкости при наличии действия внешней силы (накачки).

В трехмерной турбулентности движение возникает на масштабах, меньших масштаба накачки [1, 2]. Это связано с возникновением прямого каскада энергии, который за счет нелинейного взаимодействия переносит энергию от интегрального масштаба турбулентности (где она создается накачкой) на все более мелкие масштабы вплоть до диссипативного масштаба, где включается вязкость, преобразующая кинетическую энергию в тепло.

Однако в описании многообразных явлений природы, таких, например, как циклоны и антициклоны, преобладают двумерные свойства системы. Для них создана теория квазидвумерной турбулентности, которая применима в случае, когда горизонтальные масштабы исследуемой системы много больше вертикального и можно считать, что основное движение жидкости происходит в двух горизонтальных направлениях. Двумерное описание турбулентных потоков широко применяется в атмосферных, океанических и астрофизических исследованиях.

Первые теоретические работы [3–7], посвященные двумерной турбулентности, выявляют ее принципиальное отличие от трехмерной. С аналитической точки зрения отличие связано с существованием двух квадратичных положительных величин (энергии и энтропии), сохраняющихся двумерными уравнениями Навье–Стокса. Это приводит к образованию двух разных каскадов, возникающих в результате нелинейного взаимодействия: энтропия переносится от уровня накачки к меньшим масштабам (прямой каскад), тогда как энергия

¹⁾ Работа выполнена при финансовой поддержке Минобрнауки РФ (госзадания 075-00270-24-00, 075-01129-23-00, 075-15-2022-1099, FFWR-2024-0014, 124022400174-3 и АААА-А-19-119041590048-0) и РФФИ (грант РФФИ 23-72-30006). Численные эксперименты выполнены на кластере “Парма” в ИТФ им. Ландау РАН.

переносится к более крупным масштабам (обратный каскад). Энтрофия рассеивается за счет вязкости на малых масштабах, а энергия рассеивается за счет трения о дно.

Накопление энергии на масштабе размера системы приводит к возникновению интенсивного крупномасштабного движения, в том числе крупных вихрей [4, 6, 8]. Тенденция к образованию больших вихрей отмечалась в работах, посвященных двумерной турбулентности, как экспериментальных [9], так и численных [10–12]. Большие когерентные вихри наблюдались при численном моделировании [13–16], основанном на решении двумерного уравнения Навье–Стокса с граничными условиями прилипания. Аналогичная когерентная вихревая структура была создана в лабораторных экспериментах в квадратной ювете [17–19]. Крупные вихри были получены при численном моделировании статической накачки с различными типами крупномасштабной диссипации, о чем сообщалось в работах [20, 21].

Изучению свойств когерентного вихря посвящена целая серия вычислительных работ [22–25]. Первая попытка установить профиль средней скорости когерентного вихря была предпринята в работе [22], где использовалось влияние периодических граничных условий и использовалась кратковременная коррелированная во времени накачка. Авторами показано возникновение устойчивого вихревого диполя. В [23] в аналогичной постановке численно найден профиль средней скорости когерентного вихря, который проявляет свойства изотропии в диапазоне расстояний до центра вихря. В работе [24] показано, что плоский профиль скорости соответствует пассивному (квазилинейному) режиму турбулентных пульсаций. Заметим, что существенное влияние на формирование подобных течений оказывает наличие трения о дно, которое вводится в систему исследуемых уравнений путем добавления слагаемого с коэффициентом, называемым коэффициентом трения о дно. В работе [25] были исследованы условия существования когерентного вихря и показано, что он не может существовать при достаточно большом значении коэффициента трения о дно. В [26] исследованы различные режимы течения в квадратной ячейке.

Данная работа является продолжением исследований, проведенных в работах [27–29], посвященных изучению двумерного течения вязкой слабосжимаемой жидкости в квадратной ячейке, возбуждаемого внешней силой с характерным масштабом накачки меньше размера ячейки. В [29] путем прямого численного моделирования системы уравнений Навье–Стокса численным методом Мак–Кормака получены различные основные типы течений (ламинарный, турбулентный и вихревой) в зависимости от величины коэффициента трения о дно и амплитуды силы накачки.

Целью данной работы является рассмотрение переходных режимов течения вязкой слабосжимаемой жидкости в квадратной ячейке с твердыми стенками, возникающих при смене основных режимов течения, а также построение фазовой диаграммы режимов в зависимости от значения вязкости жидкости, коэффициента трения о дно и амплитуды силы накачки. Особое внимание уделено анализу различных критериев определения различных режимов течения.

2. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ И ЧИСЛЕННЫЕ МЕТОДЫ

Численное моделирование основано на системе уравнения Навье–Стокса. Для замыкания системы используется уравнение слабой сжимаемости, которое хорошо описывает течения в слабосжимаемых средах. Ниже представлен конкретный вид используемых уравнений: неразрывности, движения (в проекции на 2 перпендикулярных направления) и слабой сжимаемости:

$$\begin{aligned}\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \bar{V}) &= 0, \\ \frac{\partial \rho u}{\partial t} + \nabla \cdot (\rho u \bar{V}) &= -\frac{\partial p}{\partial x} + \rho G \sin ky + \mu \Delta u - \zeta u, \\ \frac{\partial \rho v}{\partial t} + \nabla \cdot (\rho v \bar{V}) &= -\frac{\partial p}{\partial y} - \rho G \sin kx + \mu \Delta v - \zeta v, \\ dp &= c^2 \rho_0 \frac{d\rho}{\rho}.\end{aligned}$$

Здесь: ρ — плотность жидкости; $\bar{V} = (u, v)^T$ — вектор скорости, компоненты которого равны, соответственно, u и v ; μ — динамическая вязкость жидкости; p — давление; ζ — коэффициент трения о дно; G — амплитуда внешней силы; k — пространственная частота внешней силы; c — скорость звука.

При численном моделировании в данной работе было выбрано значение $k = 5 \frac{1}{M}$.

В правой части уравнений движения присутствуют слагаемые $\rho G \sin ky$ и $-\rho G \sin kx$, которые моделируют действие силы накачки. Также к правой части уравнений движения добавлены $-\zeta u$ и $-\zeta v$, равные силе трения о дно юветы. Эта сила пропорциональна скорости жидкости и направлена против ее движения. Таким образом, мы учитываем трение о дно в двумерном случае. В качестве граничных условий выбрано прилипание.

Начальные условия и физические свойства жидкости:

$$p(t = 0) = p_0 = 10^5 \text{ Па},$$

$$\rho_0 = 1000 \frac{\text{кг}}{\text{м}^3}.$$

Моделирование проводилось с тремя различными значениями вязкости жидкости: $\mu_1 = 0.006 \text{ Па} \cdot \text{с}$, $\mu_2 = 0.01 \text{ Па} \cdot \text{с}$, $\mu_3 = 0.1 \text{ Па} \cdot \text{с}$.

Размер расчетной области $2\pi \times 2\pi \text{ м}^2$. Моделирование проводилось на разных сетках: 128×128 , 256×256 , 512×512 , 1024×1024 ячеек.

Моделирование течения вязкой жидкости требует больших вычислительных мощностей и занимает много времени, поэтому необходимо оптимизировать расчеты. Для уменьшения времени расчетов в используемой компьютерной программе реализовано распараллеливание вычислительного кода при помощи стандарта OpenMP. Размер вычислительной сетки в большинстве экспериментов подобран таким образом, чтобы результаты качественно описывали исследуемое течение, а время расчетов было минимально.

Численное решение уравнений Навье–Стокса основано на методе искусственной сжимаемости [30]. При этом гиперболическая часть уравнений решается явным методом Мак–Кормака [31], а параболическая часть стандартным методом конечных разностей. Схема Мак–Кормака имеет второй порядок точности по пространству и времени. Каждый этап расчета на каждом временном шаге разделен на 4 шага: разности вперед и разности назад у предиктора вдоль направления X , а также разности вперед и назад у предиктора по направлению Y . На этапе корректора аналогично, за исключением того, что шаг “вперед” меняется на “назад” и, наоборот. Эти шаги циклически сменяют друг друга с каждым временным шагом. Метод Мак–Кормака хорошо зарекомендовал себя при решении гиперболических уравнений газо- и гидродинамики.

3. РЕЗУЛЬТАТЫ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ

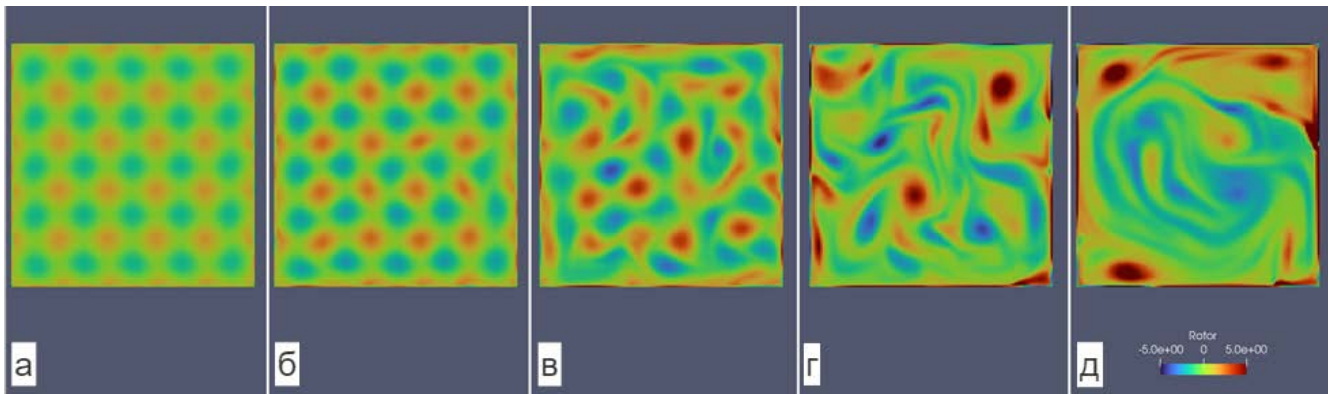
В работе [29] были получены три основных режима течения (фиг. 1а, г, д): ламинарный, при котором картина движения жидкости сохраняет свою начальную форму и не меняется со временем; хаотический (или турбулентный режим), характеризующийся наличием в течении хаотически перемещающихся вихрей различного размера и времени жизни; и вихревой, при котором возникает один крупный когерентный вихрь, занимающий практически всю расчетную область и существующий продолжительное время. Предметом исследования данной работы являются переходные режимы течения, которые возникают при смене основных режимов. Ниже приведены основные результаты численных экспериментов: построены фазовые диаграммы режимов течения в зависимости от коэффициентов трения о дно и величины накачки с учетом переходных состояний, проанализированы различные критерии возникновения режимов исследуемых течений.

3.1. Режимы течения и фазовая диаграмма

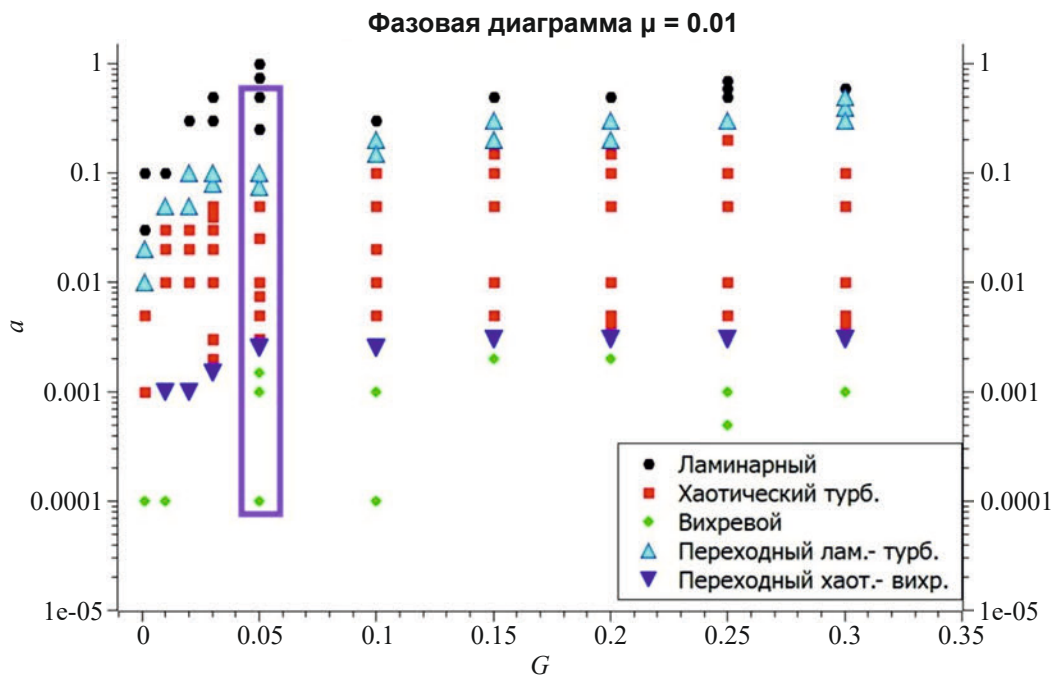
Известно, что наличие трения значительно изменяет структуру течения. Рассмотрим влияние величины коэффициента трения о дно на формирование режимов течения. Для удобства рассмотрим нормированный на плотность коэффициент трения:

$$\alpha = \frac{\xi}{\rho_0}.$$

На фиг. 1 представлены различные типы формирующегося течения в зависимости от величины этого коэффициента, все другие параметры моделирования одинаковы. При малых величинах коэффициента ($\alpha = 0.1 \dots 0.3$), между ламинарным (фиг. 1а) и хаотическим (фиг. 1г) режимами течений существует переход, для которого характерно хаотически флуктуирующее течение с определенным средним видом, характерным для начального поля течения (фиг. 1б, в). Наблюдается система из вихрей, имеющих положительную и отрицательную завихренности (фиг. 1б). Их положение и форма имеют небольшие отклонения от начального ламинарного распределения вихрей. В целом картина течения носит колебательный характер относительно начального распределения и сохраняется на протяжении всего численного эксперимента. Развитие течения (при уменьшении коэффициента трения о дно до $\alpha = 0.1$) приводит к увеличению амплитуды колебаний вихрей около своего начального положения. При дальнейшем уменьшении коэффициента трения размеры и форма вихрей изменяется (фиг. 1в). Критерием перехода к хаотическому режиму (фиг. 1г) является отрыв вихрей от исходного расположения и их хаотическое движение по ячейке. Для этого режима характерно случайное объединение и разрушение вихрей различных размеров, их короткое время жизни. Хаотическое течение на рисунке соответствует значению коэффициента трения $\alpha = 0.01$. Уменьшение величины коэффициента трения ($\alpha = 0.0001$) приводит к возникновению вихревого режима (фиг. 1д), для которого характерно формирование одной крупной структуры, занимающей всю расчетную ячейку. Этот режим подробно численно исследован в работе [28].



Фиг. 1. Поля завихренности при фиксированных параметрах $G = 0.05$, $\mu = 0.01$, и различных α в ламинарном (а) — для $\alpha = 0.2$, $Re = 125 \cdot 10^3$, переходных (б) — для $\alpha = 0.2$, $Re = 220 \cdot 10^3$, (в) — для $\alpha = 0.1$, $Re = 378 \cdot 10^3$, турбулентном (г) — для $\alpha = 0.01$, $Re = 754 \cdot 10^3$ и вихревом (д) — для $\alpha = 0.0001$, $Re = 1257 \cdot 10^3$ режимах течения.



Фиг. 2. Фазовая диаграмма течений в зависимости от величины коэффициента трения о дно и амплитуды накачки.

Важным результатом работы являются фазовые диаграммы режимов течений жидкости, построенные для трех различных значений вязкости: $\mu_1 = 0.006$ Па · с, $\mu_2 = 0.01$ Па · с, $\mu_3 = 0.1$ Па · с. Так как для разных значений вязкости диаграммы оказались аналогичными, делаем вывод, что вязкость не оказывает существенного влияния на режим течения. Приведем фазовую диаграмму при $\mu_2 = 0.01$ Па · с (фиг. 2). По оси X отложена амплитуда силы накачки G , по оси Y — нормированный на начальную плотность коэффициент трения о дно α . Вертикальным прямоугольником выделена область, рассмотренная ранее и представленная на (фиг. 1).

Каждой точке на графике соответствует длительный расчет с определенными параметрами амплитуды вынуждающей силы и коэффициента трения о дно. Черными точками обозначены расчеты, соответствующие ламинарному режиму. Красные квадраты описывают хаотический режим течения. Зеленые ромбы соответствуют вихревому режиму. Отдельно выделены переходные режимы: синим обозначен переход между ламинарным и турбулентным течениями, а голубым — между хаотическим режимом и когерентными вихрями. При малой силе трения наблюдается вихревой режим, при большой — течение ламинарно, а между данными состояниями наблюдается хаотическая турбулентность.

Наибольшее влияние на режим течения оказывает именно трение о дно. Вязкость и величина вынуждающей силы заметно влияют на скорость течения: чем больше вязкость, тем медленнее течение; чем больше сила накачки, тем быстрее течение.

Для подробного рассмотрения режимов течения использовались два различных подхода.

3.2. Анализ максимума завихренности

Первый метод основан на рассмотрении поля завихренности, которое в нашем случае является полем Z -компоненты ротора скорости Ω_z . Значение завихренности в ячейке с номерами i, j определяется следующим соотношением:

$$\Omega_{zij} = \frac{\partial v_{ij}}{\partial x} - \frac{\partial u_{ij}}{\partial y},$$

где u_{ij}, v_{ij} суть X - и Y -компоненты скорости движения жидкости в ячейке с номерами i и j по горизонтали и вертикали соответственно; i, j меняются от 1 до N , где N — размер сетки в одном направлении.

X - и Y -компоненты ротора скорости нас не интересуют, так как в двумерном случае они тождественно равны нулю.

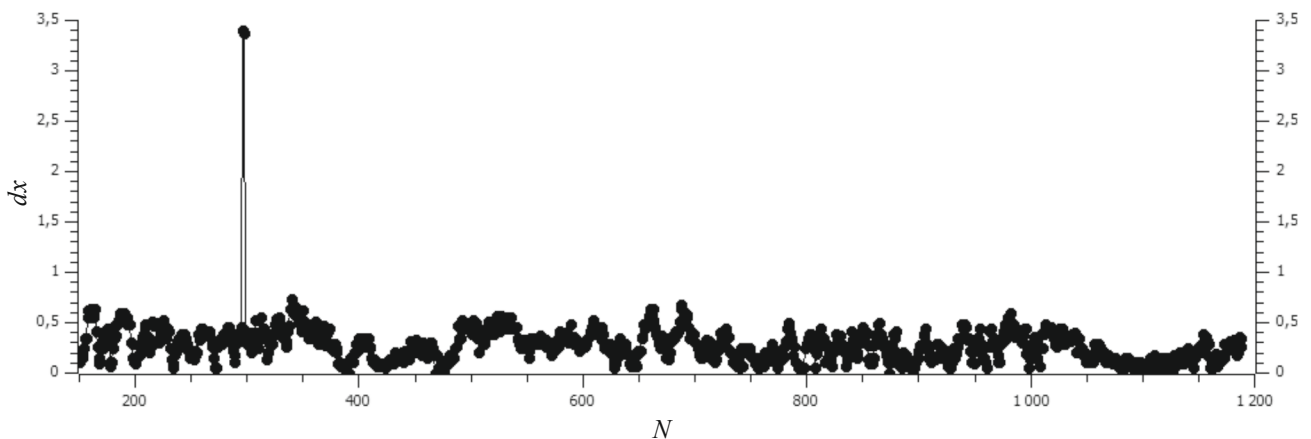
Предлагаемая методика основана на определении максимума поля завихренности в каждый момент времени. Приведем последовательность вычислений, необходимых для определения критерия определения режима течения.

1. На каждом временном шаге (обозначим выдачу индексом N) определяются координаты (x_{mN}, y_{mN}) точки с максимальной завихренностью. Из рассмотрения исключаются области вблизи стенок в связи с тем, что около них продуцируется большое количество мелких вихрей с большой величиной завихренности.
2. Далее вычисляется изменение координаты максимальной завихренности за один шаг по времени.

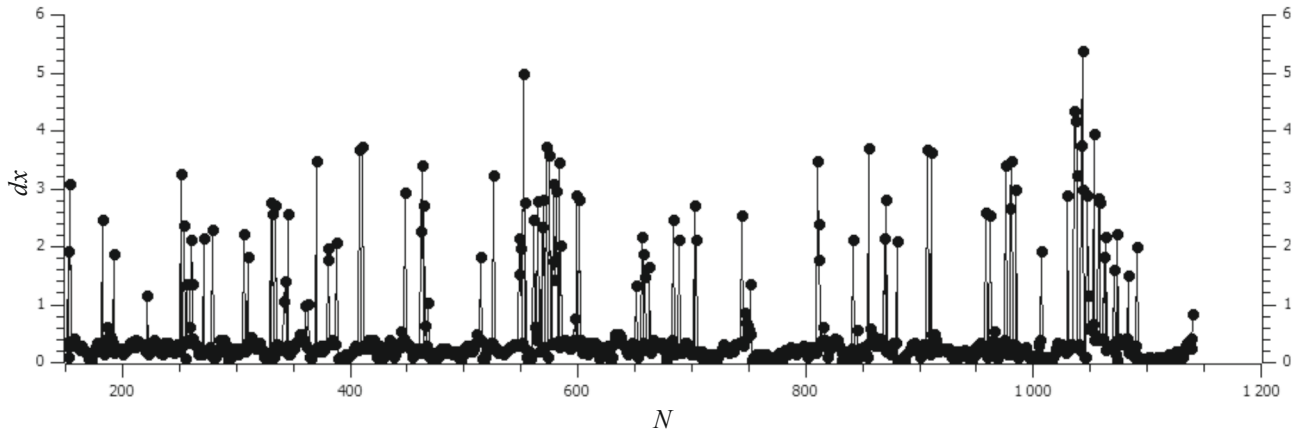
$$dx_N = \sqrt{(x_{mN+1} - x_{mN})^2 + (y_{mN+1} - y_{mN})^2}.$$

3. Строится график зависимости изменения координаты максимальной завихренности в зависимости от шага по времени.
4. Если точка максимальной завихренности меняется без скачков (фиг. 3), то это является критерием возникновения одного вихря с центром в данной точке, который совершает вращение вокруг центра ячейки. Если есть много скачкообразных изменений координаты максимальной завихренности (фиг. 4), то, по нашему мнению, этот режим можно назвать хаотическим. Одиночный пик на графике для вихревого режима (фиг. 3) связан со случайным попаданием мелкого пристеночного вихря с большим значением завихренности в рассматриваемую центральную область.

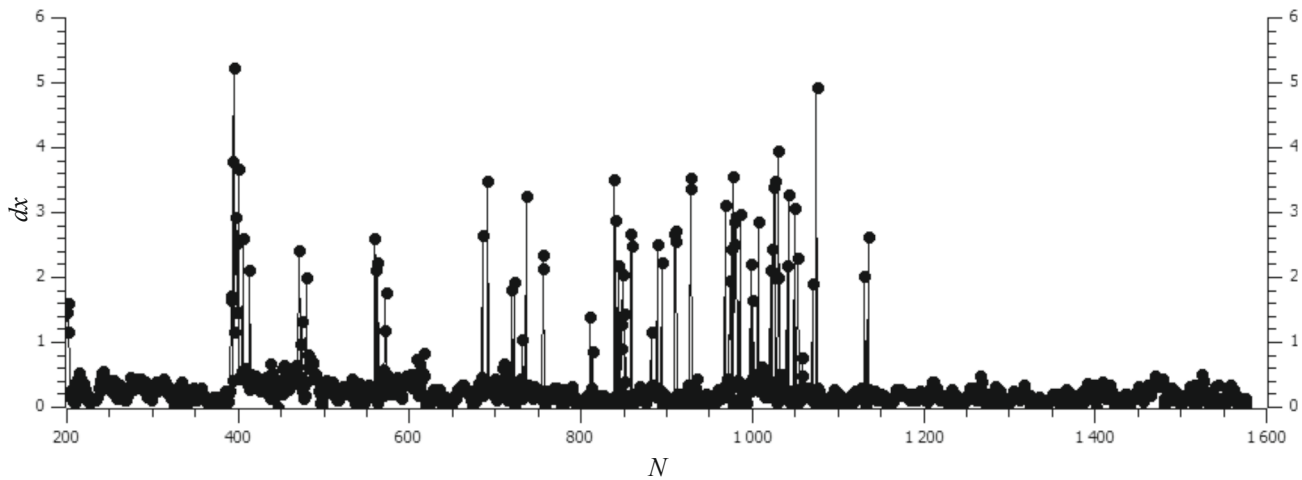
Между хаотическим и вихревым типами течения существует режим сменяемости, когда при одних и одинаковых начальных параметрах моделирования существуют разные режимы течений. Для этого режима течения характерна смена во времени вихревого и хаотического режимов (фиг. 5).



Фиг. 3. Зависимость изменения координаты точки с максимальной завихренностью от выдачи по времени, вихревой режим.



Фиг. 4. Зависимость изменения координаты точки с максимальной завихренностью от выдачи по времени, хаотический турбулентный режим.



Фиг. 5. График зависимости изменения координаты точки с максимальной завихренностью от выдачи по времени, режим сменяемости.

3.3. Анализ фурье-компонент скорости

Другой подход, использованный для рассмотрения переходных режимов течения, основан на анализе фурье-разложения поля скорости:

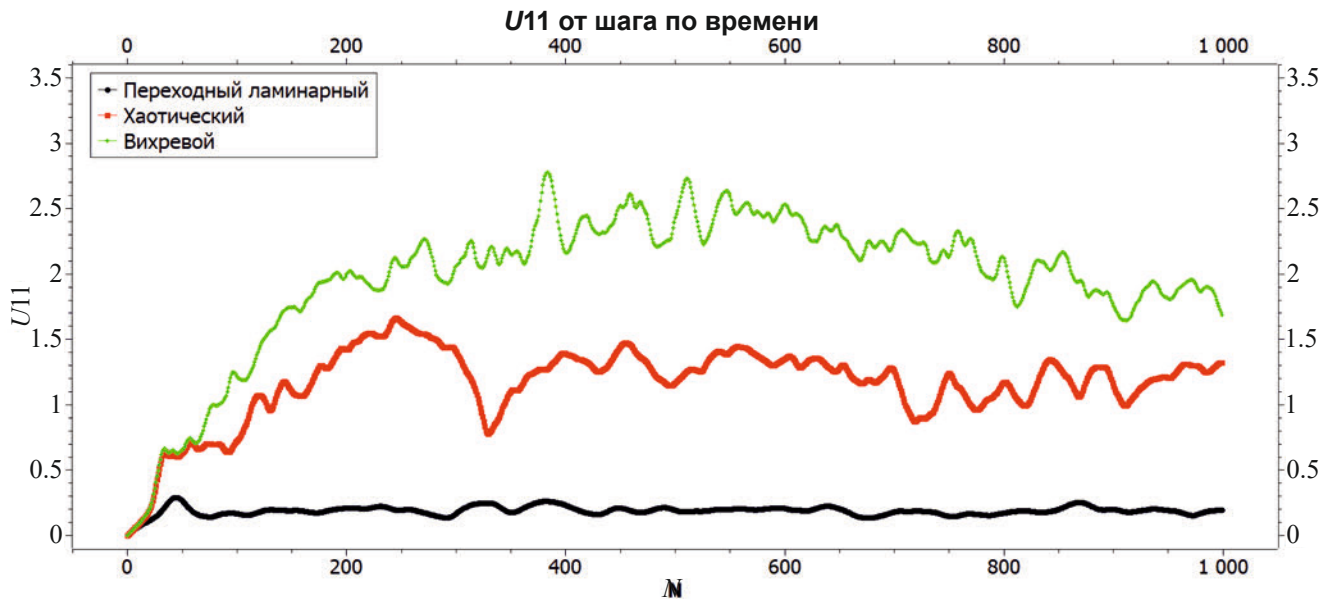
$$u(x, y) = \sum_{l, m} u_{lm}^* \sin \pi l x \cos \pi m y,$$

$$v(x, y) = \sum_{l, m} v_{lm}^* \cos \pi l x \sin \pi m y.$$

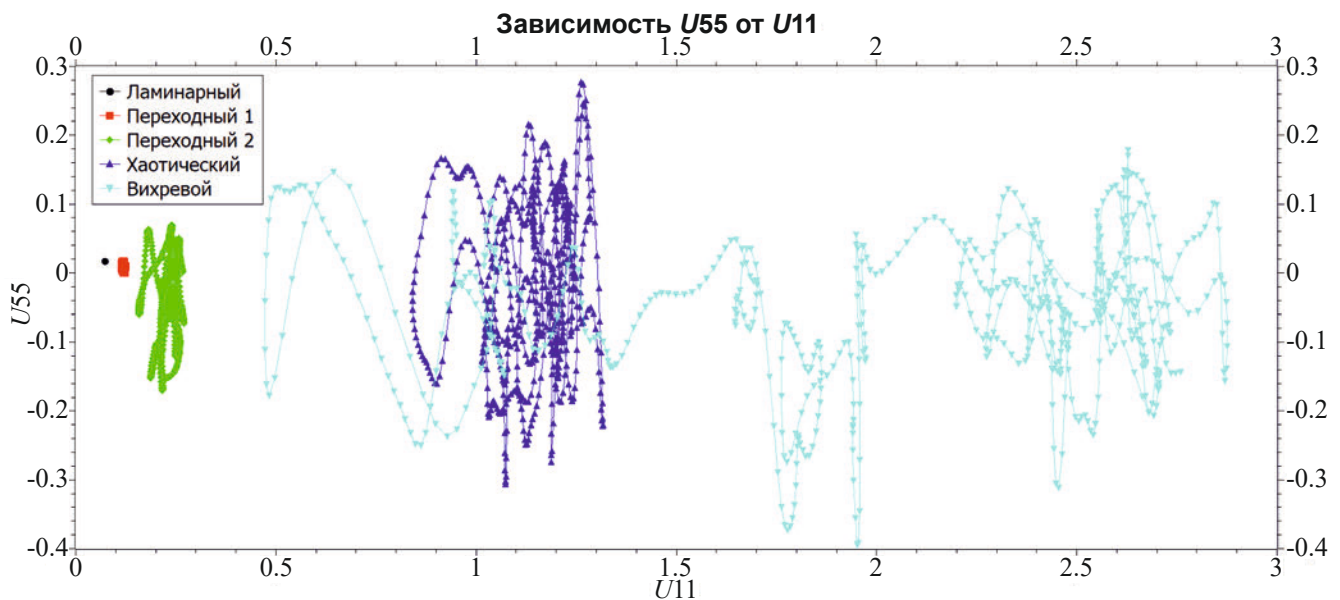
Рассмотрим фурье-компоненты горизонтальной составляющей скорости на масштабе всей ячейки u_{11}^* и на масштабе накачки u_{55}^* . Вертикальную компоненту скорости можем не рассматривать, так как ее свойства будут аналогичны.

Для удобства и универсальности компоненты Фурье-разложения рассматривались в безразмерной форме U_{lm} . Обезразмеривание проведено по формуле $U_{lm} = \frac{u_{lm}^*}{\sqrt{G/k}}$.

Для вихревого, хаотического и ламинарного режимов течения построены графики зависимости коэффициента Фурье U_{11} от номера выдачи по времени N . На фиг. 6 зеленой линией обозначен вихревой режим, красной –



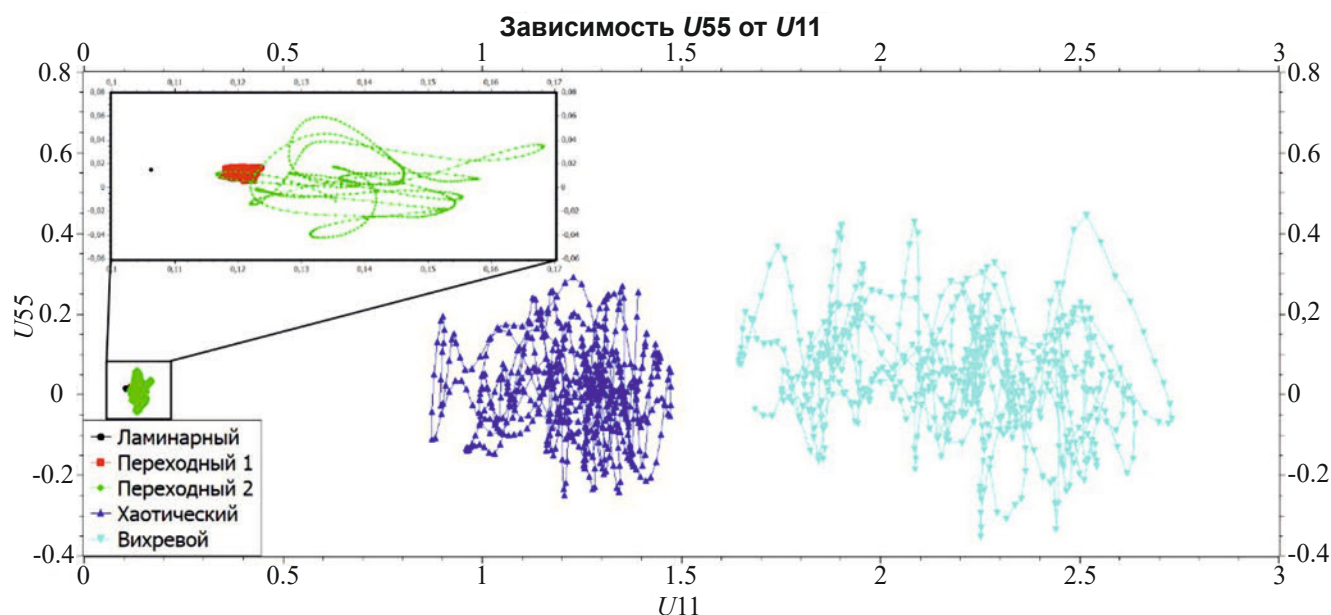
Фиг. 6. Зависимость коэффициента U_{11} от шага по времени для разных типов течения.



Фиг. 7. Зависимость U_{55} от U_{11} для разных режимов течения при фиксированных параметрах $G = 0.05$, $\mu = 0.01$, и различных α (ламинарный $\alpha = 0.3$, $Re = 125 \cdot 10^3$, переходный 1 $\alpha = 0.2$, $Re = 220 \cdot 10^3$, переходный 2 $\alpha = 0.1$, $Re = 378 \cdot 10^3$, хаотический $\alpha = 0.01$, $Re = 754 \cdot 10^3$, вихревой $\alpha = 0.0001$, $Re = 1257 \cdot 10^3$).

хаотический и черной — ламинарный. Наибольшее значение U_{11} имеет для вихревого режима, что показывает его лидирующую роль в этом типе течения.

Для анализа перехода течения от ламинарного к хаотическому и от хаотического к вихревому построим график зависимости фурье-компонент скорости U_{55} от U_{11} (фиг. 7, 8) для каждого из исследуемых типов течения. Каждой точке на графике соответствует одна выдача по времени, соседние выдачи соединены линией. Графики построены для установившихся режимов течения (100–250 с). В ламинарном случае скорость течения стабильна и соотношение между различными гармониками разложения скорости постоянно во времени. На графике (фиг. 7, 8) это представлено точкой черного цвета. Это состояние соответствует ламинарному режиму течения на фиг. 1а. В переходном режиме возникают квазипериодические колебания скорости, что показано красной линией (фиг. 7, 8) (соответствует режиму на фиг. 1б). При приближении к турбулентному режиму зависимость



Фиг. 8. Зависимость U_{55} от U_{11} для разных режимов течения при фиксированных параметрах $G = 0.15$, $\mu = 0.01$, и различных α (ламинарный $\alpha = 0.4$, $Re = 315 \cdot 10^3$, переходный 1 $\alpha = 0.35$, $Re = 315 \cdot 10^3$, переходный 2 $\alpha = 0.3$, $Re = 378 \cdot 10^3$, хаотический $\alpha = 0.01$, $Re = 1570 \cdot 10^3$, вихревой $\alpha = 0.0001$, $Re = 1885 \cdot 10^3$).

между U_{55} и U_{11} становится хаотической, представлена зеленой линией (фиг. 7, 8). Это соответствует переходному режиму с сильными флуктуациями (фиг. 1в). При уменьшении трения о дно течение становится более хаотическим, на графике зависимости U_{55} от U_{11} ему соответствует синий цвет. Голубым цветом на графике (фиг. 7, 8) обозначен вихревой режим (соответствует типу течения на фиг. 1д). Для него характерно большое значение величины гармоники U_{11} . Это означает, что она играет решающую роль в образовании вихревого режима течения. Отметим также, что при уменьшении донного трения значение гармоники U_{11} систематично возрастает. При большем значении амплитуды вынуждающей силы G вихревой режим получился более стабильным, чем при меньшем значении G , что можно видеть по диапазону значений U_{11} на графиках (фиг. 7, 8) для вихревого режима. Данные результаты согласуются с данными из работы [26].

4. ВЫВОДЫ

Рассмотрена задача двумерного течения вязкой жидкости в квадратной ячейке при воздействии внешней силы (течение Колмогорова) и наличии донного трения. Анализ возникающих типов течения при переходе от ламинарного типа течения к вихревому происходит через ряд бифуркаций, включающих квазипериодический и турбулентный типы течения. Для исследуемых течений построены фазовые диаграммы в пространстве амплитуда внешней силы - коэффициент трения о дно. Анализ фурье-гармоник разложения скорости для рассматриваемых типов течений показал ведущую роль гармоники U_{11} в вихревом течении.

СПИСОК ЛИТЕРАТУРЫ

1. Monin A.S., Yaglom A.M. Statistical Fluid Mechanics. V. I, MIT Press, 1971. V. II, Diver, 2007.
2. Белоцерковский О.М., Опарин А.М. Численный эксперимент: от порядка к хаосу. Изд. 2-е доп. М.: Наука, 2000.
3. Бетчелор Дж. К. Теория однородной турбулентности. М.: Изд-во иностр. лит., 1955.
4. Kraichnan R.H. Inertial Ranges in Two-Dimensional Turbulence // Phys. Fluids. 1967. V. 10. P. 1417–1423.
5. Leith C. E. Diffusion Approximation for Two-Dimensional Turbulence // Phys. Fluids. 1968. V. 11. P. 671–672.
6. Batchelor G. K. Computation of the Energy Spectrum in Homogeneous Two-Dimensional Turbulence // Phys. Fluids. 1969. V. 12 P. 233–239.
7. Kolmogorov A. N. Proceedings of the USSR Academy of Sciences // Doklady Akademii Nauk SSSR 1941 30 299303.

8. *Kolokolov I. V., Lebedev V. V.* Large-scale flow in two-dimensional turbulence at static pumping // JETP Lett. 2017. V. 106, № 10. P. 659–661.
9. *Sommeria J.* Experimental study of the two-dimensional inverse energy cascade in a square box // Fluid Mech. 1986 V. 170, P. 139–168.
10. *Smith L. M., Yakhot V.* Bose condensation and small-scale structure generation in a random force driven 2D turbulence // Phys. Rev. Lett. 1993 V. 71, P. 352.
11. *Smith L. M., Yakhot V.* Finite-size effects in forced two-dimensional turbulence // Fluid Mech. 1994 V. 274, P. 115–138.
12. *Borue V.* Inverse energy cascade in stationary two-dimensional homogeneous turbulence // Phys. Rev. Lett. 1994 V. 72, 1475.
13. *Clercx H.J.H.* A spectral solver for the Navier–Stokes equations in velocity-vorticity formulation for flows with two nonperiodic directions // Comp. Phys. 1997 V. 137, P. 186.
14. *Clercx H.J.H., Maassen S.R., van Heijst G.J.F.* Spontaneous spin-up during the decay of 2D turbulence in a square container with rigid boundaries // Phys. Rev. Lett. 1998 V. 80, 5129.
15. *Clercx H.J.H., Nielsen A.H., Torres D.J., Coutsias E.A.* Two-dimensional turbulence in square and circular domains with no-slip walls // Eur. J. Mech. B-Fluids 2001 V. 20 P. 557–576.
16. *Molenaar D., Clercx H.J.H., van Heijst G.J.F.* Angular momentum of forced 2D turbulence in a square no-slip domain // Physica D 2004 V.196, P. 329–340.
17. *Xia H., Punzmann H., Falkovich G., Shats M. G.* Turbulence-Condensate Interaction in Two Dimensions // Phys. Rev. 2008 Lett. 101 194504.
18. *Xia H., Shats M., Falkovich G.* Spectrally condensed turbulence in thin layers // Phys. Fluids 2009 V. 21 125101.
19. *Francois N., Xia Y., Punzmann H., Ramsden S., Shats M.* Three-Dimensional Fluid Motion in Faraday Waves: Creation of Vorticity and Generation of Two-Dimensional Turbulence // Phys. Rev. 2014 X4 021021.
20. *Yue-Kin Tsang, Young W.R.* Forced-dissipative two-dimensional turbulence: a scaling regime controlled by drag // Phys. Rev. 2009 E 79, 045308.
21. *Yue-Kin Tsang.* Nonuniversal velocity probability densities in two-dimensional turbulence: The effect of large-scale dissipation // Phys. Fluids 2010 22, 115102.
22. *Chertkov M., Connaughton C., Kolokolov I., Lebedev V.* Dynamics of Energy Condensation in Two-Dimensional Turbulence // Phys. Rev. 2007 Lett 99 084501.
23. *Laurie J., Boffetta G., Falkovich G., Kolokolov I., Lebedev V.* название // Phys. Rev. 2014 Lett. 113 254593.
24. *Kolokolov I.V., Lebedev V.V.* Structure of coherent vortices generated by the inverse cascade of two-dimensional turbulence in a finite box Phys. Rev. 2016 E 93 033104.
25. *Kolokolov I.V., Lebedev V.V.* Coherent vortex in two-dimensional turbulence: Interplay of viscosity and bottom friction // Phys. Rev. E 102 023108 (2020).
26. *Mishra P. K., Herault J., Fauve S., Verma M. K.* Dynamics of reversals and condensates in two-dimensional Kolmogorov flows // Phys. Rev. 2015 E 91 053005.
27. *Denisenko V.V., Doludenko A.N., Fortova S.V., Kolokolov I.V., Lebedev V.V.* Numerical modeling of the Kolmogorov flow in a viscous media, forced by the periodic on space static force // Computer Research and Modeling. 2022. V. 14. T. Issue 4. P. 741–753.
28. *Doludenko A.N., Fortova S.V., Kolokolov I.V., Lebedev V.V.* Coherent vortex in a spatially restricted two-dimensional turbulent flow in absence of bottom friction // Physics of Fluids 2021 V. 33 011704.
29. *Doludenko A.N., Fortova S.V., Kolokolov I.V., Lebedev V.V.* Coherent vortex versus chaotic state in two-dimension turbulence // Annals of Physics 2022 V. 447. part 2 169072.

30. *Anderson D., Tannehill J.C., Pletcher R.H.* Computational Fluid Mechanics and Heat Transfer, third ed., Taylor and Francis, 2016.
31. *Флетчер К.* Вычислительные методы в динамике жидкостей. Т. 1–2. М.: Мир, 1991.

NUMERICAL STUDY OF THE TRANSIENT MODES OF THE KOLMOGOROV FLOW IN A SQUARE CELL

A. O. Posudnevskaya^{a,b,*}, S. V. Fortova^b, A. N. Doludenko^c, I. V. Kolokolov^a, V. V. Lebedev^a

^a 142432 Chernogolovka, Moscow region, Akad. Semenova Ave., 1A,

L.D. Landau Institute of Theoretical Physics of the Russian Academy of Sciences, Russia

^b 123056, 19/18 2nd Brestskaya str., Moscow, Institute of Design Automation of the Russian Academy of Sciences, Russia

^c 125412 Moscow, Izhorskaya str., 13, bldg 2, Joint Institute of High Temperatures

of the Russian Academy of Sciences, Russia

**e-mail: posudnevskaya.ao@phystech.edu*

Received: 04.03.2024

Revised: 04.04.2024

Accepted: 31.05.2024

Abstract. The problem of two-dimensional flow of a viscous slightly compressible liquid in a square cell under excitation by a static external force periodic in space (Kolmogorov flow) is considered. A new method for determining the flow structure based on the analysis of the vorticity field at various points in time is presented. This method is used to classify the types of flows whose characteristics are obtained by numerical modeling. The main flow modes are distinguished depending on the values of the bottom friction coefficient and the pumping force: laminar, chaotic and vortex modes. Transitional flow types are studied separately: the quasi-periodic regime, which arises through a sequence of bifurcations during the change of laminar and chaotic flow modes, and the regime of alternation, which occurs during the transition from chaotic to vortex flow. Phase diagrams in the space of the amplitude of the external force — the bottom friction coefficient are constructed, making it possible to classify the type of flow according to the values of the bottom friction coefficient and the pumping force.

Keywords: turbulence, numerical modeling, Kolmogorov flow, viscous liquid, vortex.

*Посвящается доктору физико-математических наук, академику РАН
Юрию Гавриловичу Евтушенко в связи с его 85-летием.*

АЛГОРИТМ АДАПТАЦИИ РАСЧЕТНОЙ СЕТКИ К ПОЛЮ ТЕЧЕНИЯ, СОДЕРЖАЩЕМУ ГОЛОВНОЙ СКАЧОК УПЛОТНЕНИЯ¹⁾

© 2024 г. И.В. Воронич^{1,*}, Н.С. Смирнова^{1,**}, В.А. Титарев^{1,***}

¹ 119333 Москва, ул. Вавилова, 44, ФИЦ ИУ РАН, Россия

*e-mail: i.voronich@frccsc.ru

**e-mail: nsmirnova@frccsc.ru

***e-mail: vladimir.titarev@frccsc.ru

Поступила в редакцию 08.04.2024 г.

Переработанный вариант 13.05.2024 г.

Принята к публикации 31.05.2024 г.

Построение качественных расчетных сеток играет существенную роль в получении высокоточных результатов в задачах расчета внешнего обтекания тел высокоскоростным потоком. Приоритетной задачей является адаптация расчетной сетки к разрывам, в первую очередь к головным скачкам уплотнения. В настоящей работе рассмотрен вариант алгоритма адаптации расчетной сетки как механической системы с упругими связями к полю течения, содержащему головной скачок уплотнения. В результате применения алгоритма к типичной структурированной расчетной сетке происходит ее перестройка в областях большого градиента поля путем втягивания в область скачка сеточных линий с сохранением качества элементов сетки. Рассмотренные задачи показывают возможность применения описанного алгоритма к реальным задачам обтекания тел. Библ. 24. Фиг. 6.

Ключевые слова: адаптация сетки, вычислительная аэродинамика, сверхзвуковые течения, параллельные вычисления.

DOI: 10.31857/S0044466924090145, EDN: WIQSBX

ВВЕДЕНИЕ

Реализованные к настоящему времени модели и алгоритмы генерации расчетных сеток позволяют успешно проводить расчеты обтекания геометрических форм различной сложности [1, 2]. Задача повышения точности численных решений предъявляет в первую очередь требования к расчетной сетке — по качеству элементов, по количеству элементов на масштаб особенности течения, по соответствию сеточных линий линиям тока [3]. Для повышения эффективности расчета сверхзвуковых течений приоритетной задачей является адаптация расчетной сетки к разрывам, в первую очередь к головным скачкам уплотнения [4, 6].

Неоднократно было показано, что адаптация расчетной сетки эквивалентна увеличению количества ее элементов, порой весьма существенному, в зависимости от способа адаптации [7–10]. Традиционно большое внимание уделяется методам построения и начальной адаптации расчетных сеток к форме обтекаемых поверхностей и полям течения. Эти методы делятся на алгебраические (трансфинитная интерполяция), методы конформных преобразований, вариационные методы, методы решения уравнений эллиптического типа [1, 4, 6, 7, 11–15]. Наряду с указанными методами развивался второй подход, основанный на измельчении начальной расчетной сетки путем деления элементов по ребрам, на которых достигаются наибольшие перепады целевой функции [8–10]. Такой подход сейчас широко представлен в коммерческом программном обеспечении.

Преимущество первой группы методов заключается в неизменном количестве узлов расчетной сетки, сохранении ее топологии и непрерывности, соответствии семейств сеточных линий общей структуре течения и

¹⁾ Работа выполнялась с использованием инфраструктуры Центра коллективного пользования “Высокопроизводительные вычисления и большие данные” (ЦКП “Информатика”) ФИЦ ИУ РАН и МСЦ РАН (г. Москва).

форме обтекаемых поверхностей. Недостатки заключаются в неоднородном качестве элементов для алгебраических методов и ограниченных возможностях контроля адаптации расчетной сетки к деталям поля течения для остальных методов.

Вторая группа методов в качестве преимущества дает алгоритмическую простоту, но имеет серьезные недостатки: потерю непрерывности расчетной сетки, более низкое качество производных элементов, необходимость контроля количества добавляемых элементов. В рамках подхода встраиваемых расчетных сеток задача заключается в построении функции-сигнализатора и разработке алгоритма встраивания новых элементов путем деления исходных. Такой подход обладает свойством локальности, затрагивая только некоторую подобласть расчетной сетки, но не обеспечивает ее непрерывности. Последнее обстоятельство частично обесценивает результат адаптации из-за нарушения аппроксимации численного решения. Подход применяется как для структурированных расчетных сеток, так и для неструктурированных расчетных сеток общего типа.

В недавнее время в связи с решением сопряженных задач на подвижных сетках развились методы адаптации, основанные на силовых физических аналогиях [5, 6, 16–22]. Эти аналогии рассматривают элементы сетки как упругие деформируемые объекты и наделяют их механическими связями, определяемыми с учетом решаемой задачи. Преимущество этой категории методов заключается в том, что расчетная сетка представляется как механическая система узлов (масс), соединенных связями, жесткость которых определяется свойствами поля. На сегодняшний день считается, что задача адаптации при надлежащем выборе коэффициента жесткости имеет корректное решение, при этом на перемещения узлов могут быть наложены дополнительные ограничения, и к полученной сетке могут быть применены процедуры сглаживания. Данная модель адаптации в силу физической универсальности может быть применена к сеткам произвольной размерности и топологии.

В настоящей работе рассмотрен вариант алгоритма адаптации структурированной расчетной сетки как механической системы с упругими связями к полю течения, содержащему головной скачок уплотнения.

1. МОДЕЛЬ РАСЧЕТНОЙ СЕТКИ КАК МЕХАНИЧЕСКОЙ СИСТЕМЫ

Расчетная сетка рассматривается как механическая система точечных масс, в которой ребра элементов представлены упругими связями со свойствами, определяемыми полем течения. Поле течения в виде выбранной функции течения $P(x)$ определено в расчетной области, где задана сетка. Функция $P(x)$ может быть непрерывной или дискретной $\{P_i\}$, заданной в узлах сетки $\{x_i\}$. Во втором случае предполагается, что существует непрерывное продолжение дискретной сеточной функции на всю расчетную область в виде той или иной интерполяционной функции так, чтобы можно было вычислить значения функции в новых положениях узлов. В качестве функции $P(x)$ можно использовать поле плотности или поле давления, которые везде положительны и имеют заведомо ограниченные вариации. Удобно нормировать значения функции P с использованием ее характерного значения в поле течения.

Под адаптацией расчетной сетки понимается изменение координат части ее узлов в соответствии с функцией $P(x)$ и выбранным алгоритмом адаптации. Координаты узлов на границе адаптируемой части расчетной сетки могут фиксироваться (например, на стенке) либо их движение может быть ограничено (например, условием принадлежности узла кривой или поверхности либо условием близости к границе).

Уравнения динамики механической системы с заданными начальными и граничными условиями и ограничениями запишем в форме:

$$\begin{aligned} \ddot{x}_i &= \sum_{j(i)} K(P(x_{j(i)}), P(x_i)) (x_{j(i)} - x_i) - 2\gamma_i \dot{x}_i, \quad i = 1, \dots, N, \\ x_{iw} &= x_{iw}^0, \quad G(x_{is}) = 0, \\ x_i(t=0) &= x_i^0, \quad \dot{x}_i(t=0) = 0. \end{aligned} \quad (1)$$

Здесь x_i — координаты узлов, массы узлов приняты единичными (хотя может быть сделан и иной выбор), $K(P(x_j), P(x_i))$ представляет собой функцию жесткости, зависящую от взаимного положения узлов и поля $P(x)$, $\gamma_i > 0$ — локальный коэффициент демпфирования, нижний индекс w указывает неподвижные узлы на границе, s — ограниченно подвижные узлы на границе, G — функция, задающая ограничение движения узлов на границе, индекс $j(i)$ нумерует ближайших соседей узла с номером i . Таким образом, система (1) представляет собой динамическую систему из N упруго связанных масс с демпфированием при заданных начальных и граничных условиях [23]. Размерность задачи и структура связей могут быть произвольными. В настоящей работе рассматриваются структурированные расчетные сетки как наиболее экономичные для расчета задач сверхзвукового обтекания.

Уравнения (1) могут быть сопоставлены с механическими колебательными системами [23]. Параметр демпфирования γ_i служит для обеспечения сходимости к стационарному решению системы (1), соответствующему

положению равновесия. Для обеспечения оптимальной сходимости целесообразно выбирать значение параметра γ_i близким к границе колебательных режимов локального осциллятора:

$$\gamma_i \approx \omega_{0i} = \sqrt{\sum_{j(i)} K(P(x_{j(i)}), P(x_i))}.$$

В такой модели расчетная сетка является напряженной — силы упругости на ребрах исчезают только в случае сетки, стянутой в точку, и конечное состояние расчетной сетки определяется достижением баланса сил упругости на ребрах элементов в адаптируемой части расчетной области с учетом наложенных граничных условий. Это обстоятельство отличает настоящий подход от широко принятых, которые рассматривают начальную сетку как ненапряженную [18–20], в которой начальные положения узлов считаются равновесными и силы упругости возникают только при отклонении узлов от начальных значений, что адекватно для описания малых отклонений узлов. Так как адаптация расчетной сетки к полю течения со скачком уплотнения требует больших деформаций начальной сетки для обеспечения разрешения разрывов, в настоящей работе принята модель напряженной сетки.

В качестве определения функции $K(P(x_j), P(x_i))$ примем выражение следующего вида:

$$K(P(x_j), P(x_i)) = F(a, b, c, \delta, z) = \frac{a+b}{2} + \frac{b-a}{2} \tanh\left(\frac{z-c}{\delta}\right), \quad (2)$$

$$z = |\Delta_{ij} P(x) / \Delta_{ij}(x)|, \quad \Delta_{ij} P = P(x_j) - P(x_i), \quad \Delta_{ij}(x) = \|x_j - x_i\|.$$

Данное выражение связывает коэффициент жесткости упругой связи на ребре расчетной сетки с градиентом поля $P(x)$ в направлении ребра между узлами i и j . Параметры a, b, c, δ выбираются таким образом, чтобы обеспечить нужный минимальный (a) и максимальный (b) уровень функции F , центрированной в точке c и имеющей ширину переходной зоны δ . Практика показывает, что в определении функции F важны прежде всего отношение параметров b/a и правильный выбор параметра c в соответствии с разрешаемым градиентом, при этом параметр δ может быть выбран достаточно малым по сравнению с уровнем порога c .

2. ЧИСЛЕННЫЙ АЛГОРИТМ

Предложенная система уравнений (1) для динамической системы из N упруго связанных масс с демпфированием решается численно, при заданных начальных и граничных условиях. Дифференциальное уравнение для узла с номером i аппроксимируется разностной схемой второго порядка точности по времени:

$$\frac{x_i^{n+1} - 2x_i^n + x_i^{n-1}}{\Delta t^2} = \sum_{j(i)} K(P(x_{j(i)}^n), P(x_i^n)) (x_{j(i)}^n - x_i^n) - 2\gamma_i^n \frac{x_i^{n+1} - x_i^{n-1}}{2\Delta t}. \quad (3)$$

Положение узлов на новом временном шаге определяется решением системы:

$$x_i^{n+1} = \frac{2x_i^n + (\gamma_i^n \Delta t - 1)x_i^{n-1} + \Delta t^2 \sum_{j(i)} K(P(x_{j(i)}^n), P(x_i^n)) (x_{j(i)}^n - x_i^n)}{1 + \gamma_i^n \Delta t} \quad (4)$$

с учетом ограничений, наложенных граничными условиями.

Шаг по времени для схемы (4) в настоящей работе выбирается из условия разрешения периода колебаний для самых жестких связей:

$$\Delta t < \frac{2\pi}{\max_i \omega_{0i}}.$$

Если поле $P(x)$ является результатом интерполяции сеточной функции, в процессе адаптации необходимо проводить повторную интерполяцию после перерасчета поля с учетом новых положений узлов.

Критериями остановки итерационного процесса могут служить установление положений узлов или достижение некоторого ограничения. Следует отметить, что наложение избыточных ограничений на расчетную сетку в процессе адаптации не требуется, так как механическая модель практически не допускает пересечений ребер элементов, обеспечивая выстраивание и сгущение сеточных линий в зонах больших градиентов.

Улучшение качества полученной адаптированной расчетной сетки может быть достигнуто с помощью процедуры взвешенного сглаживания (здесь W — собственный вес узла):

$$x_i = \frac{1}{W + \sum_{j_i} 1} \left(W x_i^{n+1} + \sum_{j_i} x_{j_i}^{n+1} \right). \quad (5)$$

Данная процедура может применяться для улучшения конечной расчетной сетки избирательно в области элементов большой скошенности или скорости роста размеров.

3. РАСЧЕТНЫЕ ПРИМЕРЫ

В настоящем разделе рассмотрены примеры применения описанного алгоритма к аналитически заданным и дискретным полям в двухмерном случае. В первом случае это позволяет выявить поведение алгоритма и влияние его параметров при адаптации расчетной сетки к гладкому полю, содержащему заданный перепад параметров и переходную зону. Во втором случае технология применяется к дискретным вычислительным полям, формирующимся при обтекании тела сверхзвуковым потоком газа. В этих примерах выполняется адаптация и перерасчет поля на новой сетке, чтобы эффект мог быть оценен с практической стороны.

В расчетах применялась процедура сглаживания (5) с весом $W = 1$. У стенки перемещения узлов при адаптации и сглаживании ограничиваются путем умножения на экспоненциальную функцию вида

$$f(s, s_*) = 1 - \exp(-(s/s_*)^3), \quad (6)$$

так что на поверхности перемещение узла равно нулю, тогда как на внешней границе оно полностью определяется алгоритмом адаптации. Здесь s — расстояние до стенки, s_* — характерный масштаб, определяемый с учетом расстояния от стенки до фронта ударной волны для конкретной задачи. Применение (6) дает возможность избежать ухудшения точности моделирования пограничного слоя на стенке.

3.1. Адаптация к аналитически заданному полю

Рассмотрим приложения алгоритма адаптации расчетной сетки к аналитически заданному полю около цилиндра с имитацией головного скачка уплотнения. Поле задается в полярных координатах (r, θ) выражением вида

$$P(r, \theta) = F(\psi(\theta), 1, r_s(\theta), \Delta(\theta), r), \quad r \geq 1, \quad (7)$$

где общий вид функции F определен в (2). Здесь функция $\psi(\theta)$ задает изменение по полярному углу интенсивности поля за скачком, функция $r_s(\theta)$ задает форму скачка, функция $\Delta(\theta)$ задает изменение ширины скачка.

Форму скачка положим заданной в виде

$$r_s(\theta) = 1 + 0.5 \exp(0.5\theta^2).$$

Первое поле представляет ситуацию с постоянным перепадом параметров на скачке $\psi(\theta) = 2$ и постоянной шириной скачка $\Delta(\theta) = 0.02$.

Второе поле представляет ситуацию с уменьшающимся перепадом параметров на скачке

$$\psi(\theta) = 2(1 - 0.001 \exp(0.5\theta^2))$$

и увеличивающейся шириной скачка

$$\Delta(\theta) = 0.02 \exp(0.5\theta^2).$$

Результаты для аналитически заданных полей получены на расчетной сетке размерности 50×50 ячеек с квадратичным распределением узлов по радиусу [1] в области $1 \leq r \leq 4$, $0 \leq \theta \leq \pi$; пристенная ячейка высотой 10^{-4} . На фиг. 1 приведены начальная, адаптированная и сглаженная расчетные сетки для первого поля. Функция F , используемая в (1)–(2), построена при параметрах:

$$a = 1, \quad b = 3, \quad c = \maxgrad/10, \quad \delta = 0.3.$$

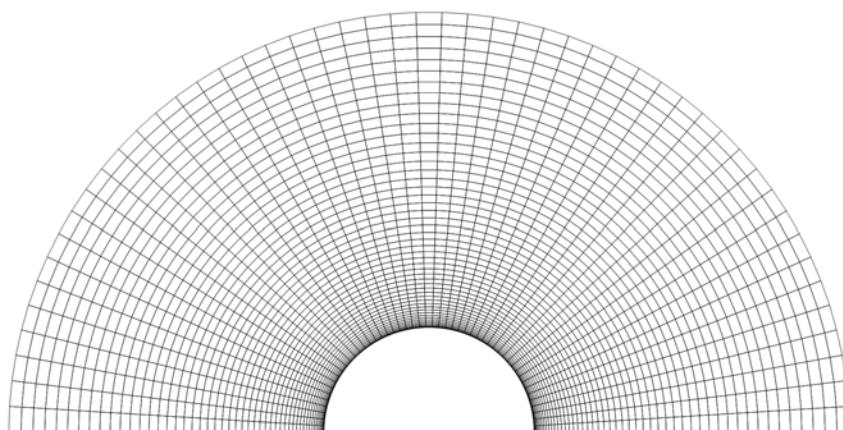
Здесь $\maxgrad$ — максимальное значение дискретного градиента величины P вдоль ребра во всей расчетной области:

$$\maxgrad = \max_{ij} |\Delta_{ij} P(\mathbf{x}) / \Delta_{ij}(\mathbf{x})|.$$

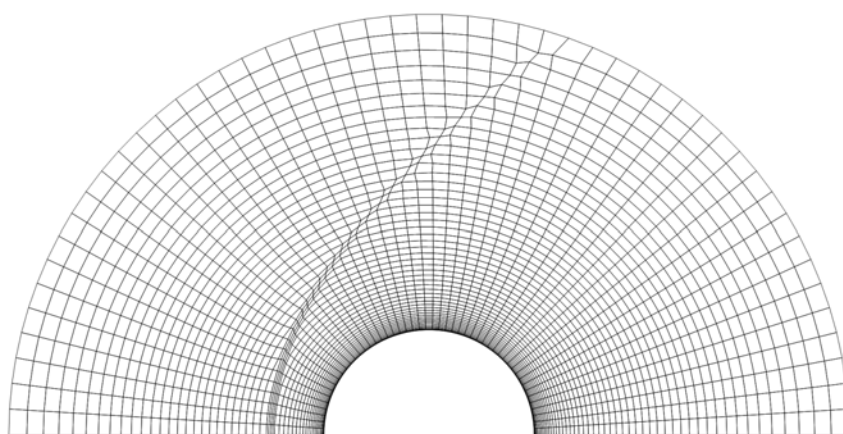
На фиг. 2 приведены начальная, адаптированная и сглаженная расчетные сетки для второго поля с параметрами функции F

$$a = 1, \quad b = 5, \quad c = \maxgrad/10, \quad \delta = 0.5.$$

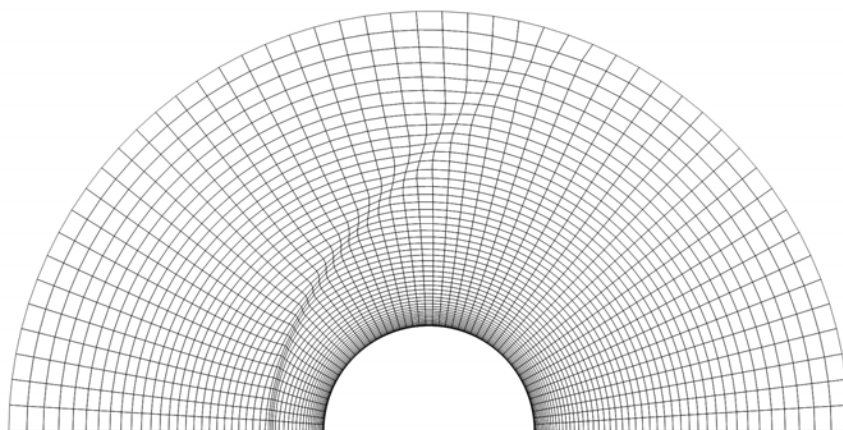
В рассматриваемых примерах $\maxgrad = 23$. В обоих случаях адаптация происходит адекватно — линии сетки выстраиваются вдоль скачка, происходит втягивание линий в скачок. Финальная сетка после сглаживания обладает достаточной для практики ортогональностью (минимальный угол при вершине элемента равен 30 градусам) и ограниченностью скорости роста элементов вдоль сеточных линий.



(а) начальная сетка



(б) адаптированная сетка без сглаживания

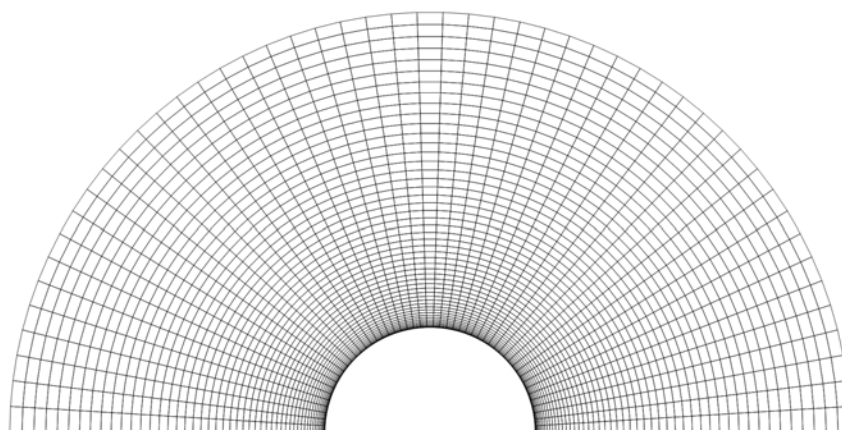


(в) адаптированная сетка после сглаживания

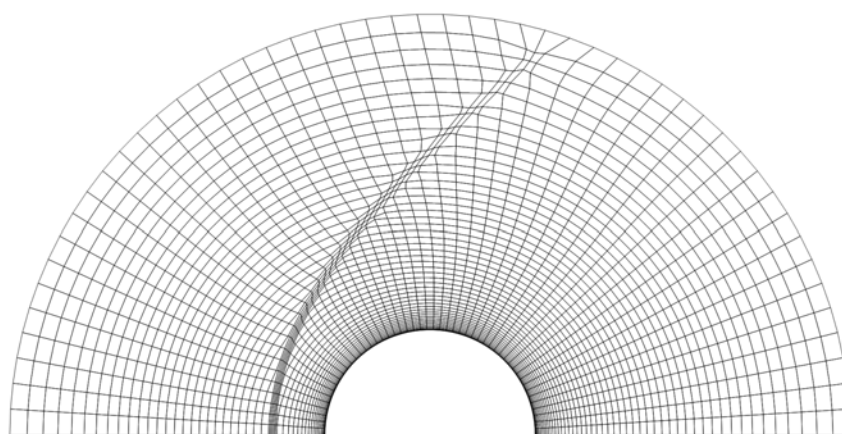
Фиг. 1. Адаптация расчетной сетки к первому аналитически заданному полю около цилиндра.

3.2. Адаптация к дискретному полю

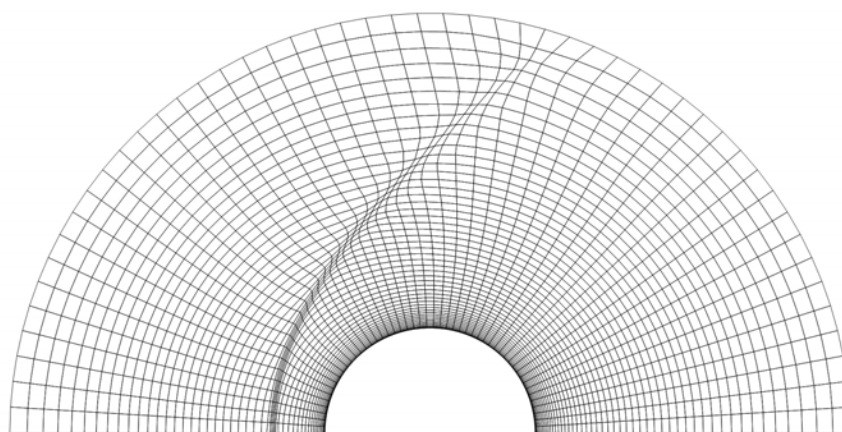
Рассмотрим задачи об адаптации расчетной сетки к двумерному полю течения двухатомного совершенного вязкого газа около цилиндра и эллипса. Во всех тестовых примерах число Маха набегающего потока $M_\infty=10$, число Рейнольдса $Re = 13.5 \times 10^6$ в расчете на размер 1 метр, температурный фактор $T_w/T_0 = 0.12$. Поле те-



(а) начальная сетка



(б) адаптированная сетка без сглаживания



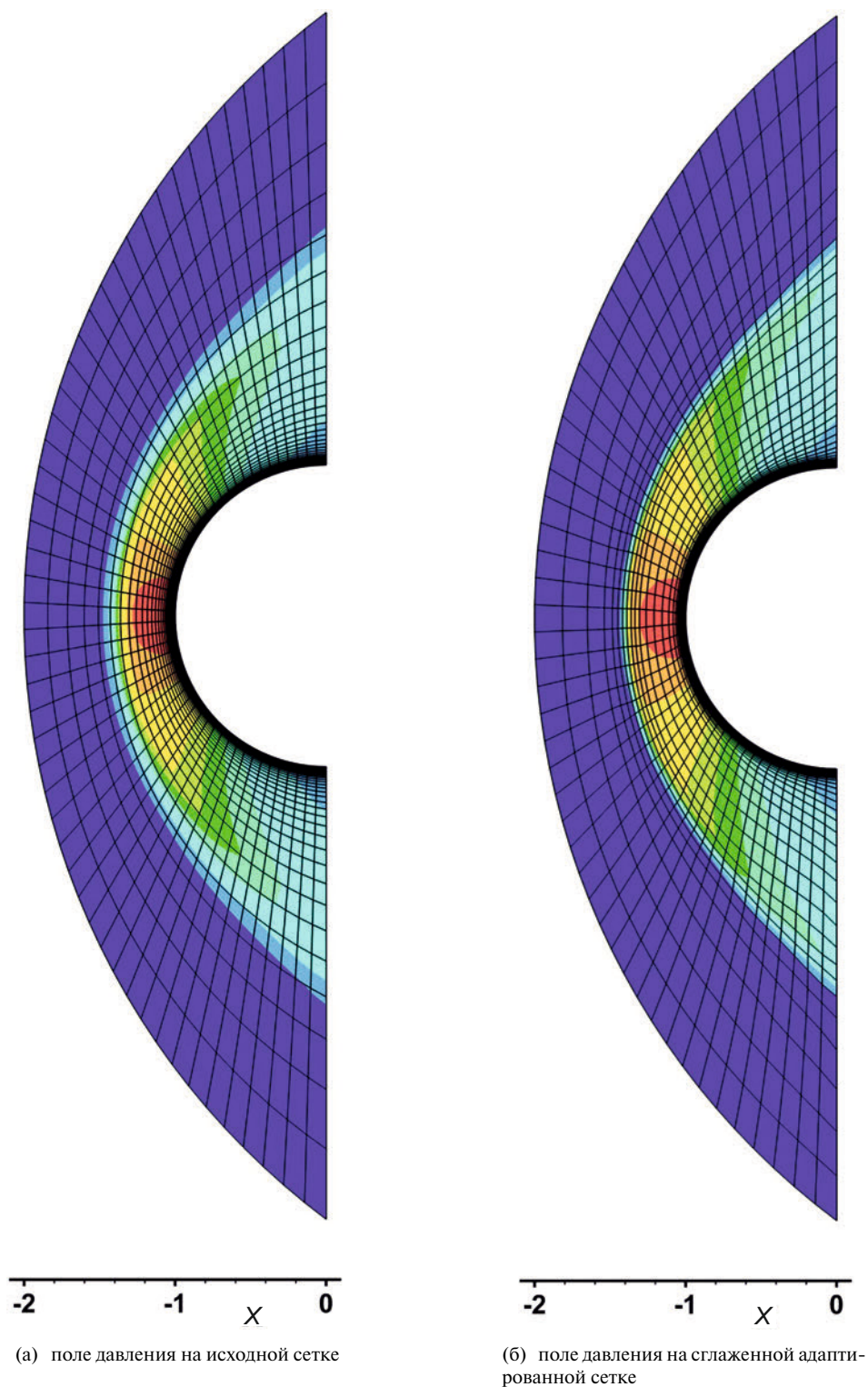
(в) адаптированная сетка после сглаживания

Фиг. 2. Адаптация расчетной сетки ко второму аналитически заданному полю около цилиндра.

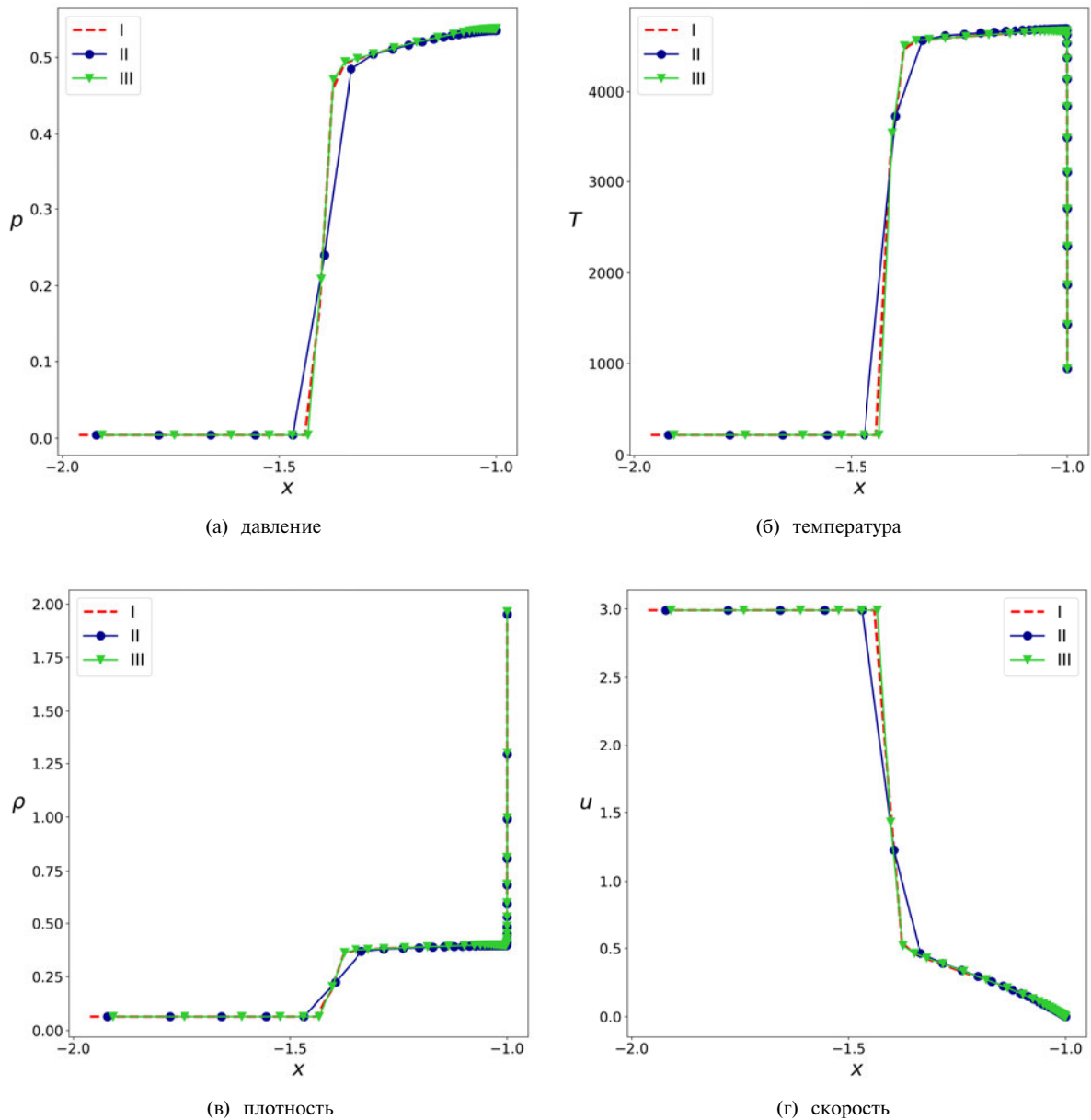
числения на исходной и адаптированной сетках получено путем решения трехмерных уравнений Навье-Стокса совершенного газа с помощью неявного численного метода второго порядка аппроксимации [24]. Вычисления проводились на одном сервере ЦКП “Информатика” ФИЦ ИУ РАН и МСЦ РАН с использованием до 32 физических ядер. Для построения функции $P(x)$ использовалось нормированное поле давления, что обусловлено

его слабыми изменениями в пограничном слое.

Исходная расчетная сетка для половины цилиндра радиуса 1 метр содержала 50×50 ячеек. Поскольку детальное разрешение пограничного слоя не являлось целью расчетов, высота первой ячейки принималась равной $h \approx 10^{-4}$ м. На фиг. 3 приведены результаты применения алгоритма адаптации и изменение поля давления.



Фиг. 3. Адаптация расчетной сетки к дискретному полю около половины цилиндра.



Фиг. 4. Распределения давления (МПа), температуры (К), плотности (кг/м³) и скорости (км/с) вдоль линии торможения перед цилиндром. I – 100 ячеек по радиусу; II – 50 ячеек по радиусу; III – 50 ячеек по радиусу после адаптации.

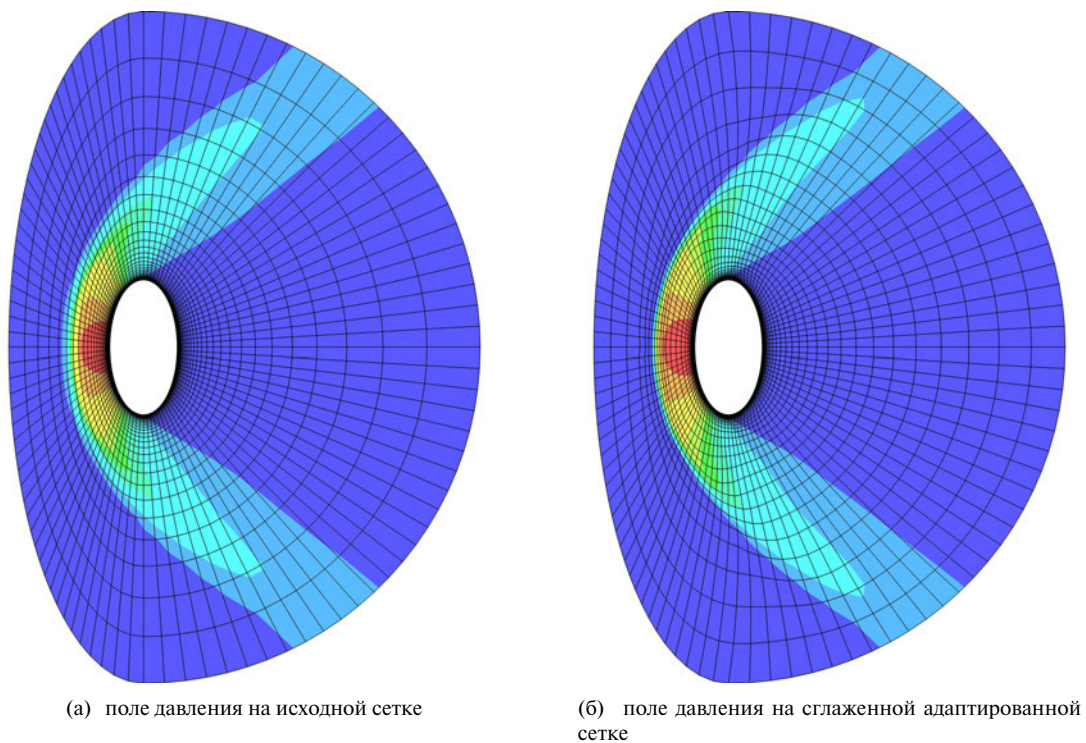
На фиг. 4 приведены распределения размерных величин вдоль линии торможения. В дополнение к основной серии расчетов на фиг. 4 приведены также результаты вычислений на неадаптированной сетке, содержащей 100 ячеек по радиусу с сохранением высоты первой ячейки. Параметры функции F задавались следующие:

$$a = 1, \quad b = 5, \quad c = \maxgrad/10, \quad \delta = 0.3.$$

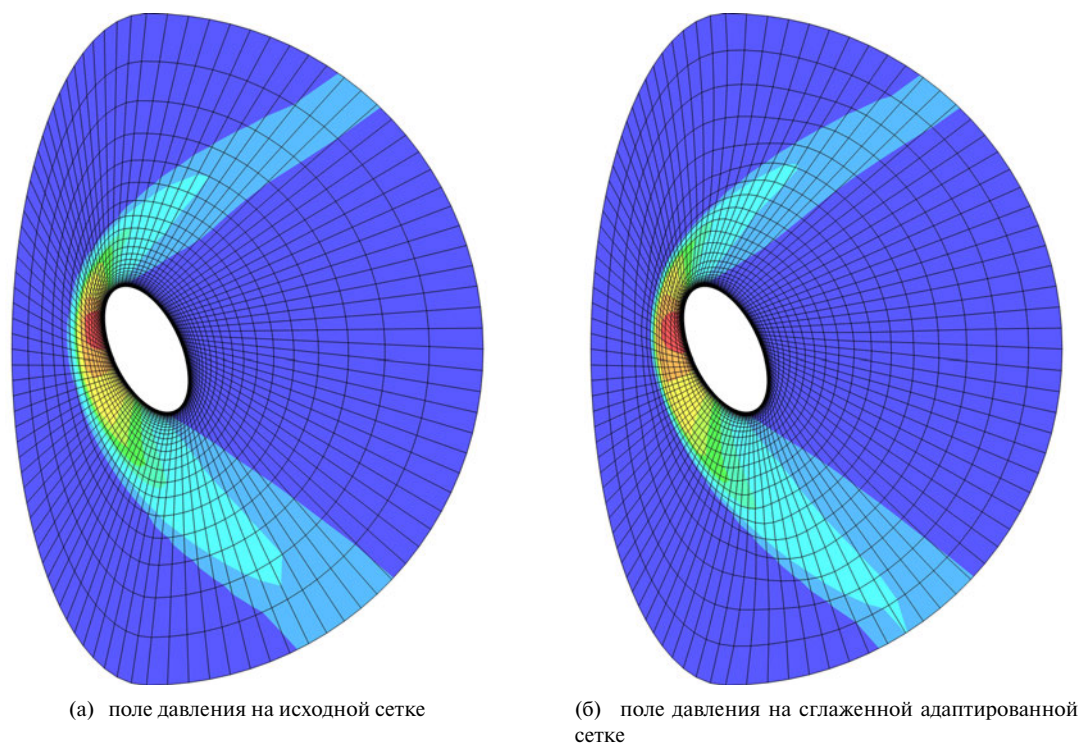
Результаты демонстрируют сгущение узлов к области скачка уплотнения, а также показывают значительное количественное уточнение профилей в окрестности скачка: использование адаптации позволяет проводить расчеты на приблизительно вдвое менее подробной сетке.

Исходные расчетные сетки для обтекания полного эллипса с полуосями 1 и 2 метра состояли из 50 ячеек по радиальному направлению и 100 ячеек по углу, высота первой ячейки составляла $h \approx 10^{-4}$ м. Параметры функции F задавались следующие:

$$a = 1, \quad b = 10, \quad c = \maxgrad/10, \quad \delta = 0.22.$$



Фиг. 5. Адаптация расчетной сетки к дискретному полю около эллипса.



Фиг. 6. Адаптация расчетной сетки к дискретному полю около эллипса с повернутой на 25 градусов большой полуосью.

На фиг. 5 представлены результаты применения алгоритма адаптации и изменение поля давления для эллипса с вертикально расположенной большой полуосью. На фиг. 6 изображены результаты адаптации расчетной сетки для обтекания эллипса с большой полуосью, повернутой на угол 25° относительно начального положения. В процессе адаптации сеточные линии стягиваются к скачку, повторяя фронт ударной волны.

ЗАКЛЮЧЕНИЕ

Рассмотренная модель адаптации расчетной сетки как механической системы с упругими связями показывает работоспособность: происходит перестройка расчетной сетки в областях с большими значениями градиента поля с воспроизведением формы головной части скачка и втягиванием в область скачка дополнительных сеточных линий. Качество элементов структурированной расчетной сетки в отношении ортогональности сохраняется на приемлемом для практических расчетов уровне с минимальным углом при вершине элемента $30\text{--}50$ градусов. Дополнительное повышение качества расчетной сетки производится взвешенным сглаживанием. Полученные численные решения демонстрируют уточнение профилей параметров в окрестности головного скачка уплотнения. Это приводит к заметному снижению уровня численной ошибки в зоне скачка и возможности использовать примерно в 2 раза более экономные сетки по направлению к фронту ударной волны. Рассмотренные тестовые задачи показывают возможность применения алгоритма к более общим задачам внешнего обтекания.

СПИСОК ЛИТЕРАТУРЫ

1. *Thompson J.F., Warsi Z.U.A., Mastin C.W.* Numerical grid generation: Foundations and Applications. North Holland, 1985.
2. *Молчанов А.М., Щербаков М.А., Янышев Д.С., Куприков М.Ю., Быков Л.В.* Построение сеток в задачах авиационной и космической техники. Учебное пособие. М.: МАИ, 2013.
3. *Chawner J.R., Michal T., Slotnick J.P., Rumsey C.L.* Summary of the 1st AIAA Geometry and Mesh Generation Workshop (GMGW-1) and Future Plans. AIAA Paper 2018-0128.
4. *Ворожцов Е.В., Яненко Н.Н.* Методы локализации особенностей в вычислительной газодинамике. Новосибирск, 1985.
5. *Garanzha V.A.* Variational principles in grid generation and geometric modelling: theoretical justifications and open problems // Numerical Linear Algebra with Applications. 2004. V. 11, № 5–6. P. 535–563.
6. *Sheshadri A., Crabill J., Jameson A.* Mesh deformation and shock capturing techniques for high-order simulation of unsteady compressible flows on dynamic meshes. AIAA Paper 2015-1741, 2015.
7. *Суржиков С.Т.* Аналитические методы построения конечно-разностных сеток для расчета аэротермодинамики спускаемых космических аппаратов // Вестник Московского государственного технического университета им. Н. Э. Баумана. Серия “Машиностроение”. 2004. № 2.
8. *Афендииков А.Л., Меркулов К.Д., Пленкин А.В.* Динамическая локальная адаптация сеток на основе вейвлет-анализа в задачах газовой динамики // Препринты ИПМ им. М.В. Келдыша. 2014. № 99. С. 26.
9. *Меркулов К.Д., Меньшов И.С.* Динамически перестраиваемые декартовы сетки с локальным измельчением для расчета задач газовой динамики // Физико-химическая кинетика в газовой динамике. 2015. Т. 16, № 4. С. 1–11.
10. *Стручков А.В., Козелков А.С., Жучков Р.Н., Уткина А.А., Саразов А.В.* Численное моделирование задач аэродинамики со статической адаптацией сетки под особенности решения // Вопросы атомной науки и техники, сер. Математическое моделирование физических процессов. 2019. Т. 2. С. 55–67.
11. *Иваненко С.А., Чарахчян А.А.* Криволинейные сетки из выпуклых четырехугольников // Ж. вычисл. матем. и матем. физ. 1988. Т. 28, № 4. С. 503–514.
12. *Лебедев А.С., Лисейкин В.Д., Хакимзянов Г.С.* Разработка методов построения адаптивных сеток // Вычисл. технологии. 2002. Т. 7, № 3. С. 29–43.
13. *Лысенков А.В.* Разработка метода построения многоблочной адаптивной сетки для расчета сложных двумерных течений на базе полной системы уравнений Эйлера // Ученые записки ЦАГИ. 2002. Т. 33. № 3, 4.

14. *Вальгер С.А.* Создание вычислительных технологий для расчета ветровых и ударно-волновых воздействий на конструкции. Дисс. на соискание к.ф.-м.н. Новосибирск. 2015. 220 С.
15. *Huang T-H, Chen J-S, Tupek M.R., Beckwith F.N., Fan H.E.* A variational multiscale immersed meshfree method for fluid structure interactive systems involving shock waves // *Comput. Methods Appl. Mech. Eng.* 2022. № 389. 114396.
16. *Selim M.M., Koomullil R.P.* Mesh deformation approaches – a survey // *J. of Physical Mathematics*. 2016. V. 7. № 2. 1000181.
17. *Garanzha V.A., Kudryavtseva L.N.* Hyperelastic springback technique for construction of prismatic mesh layers // 26th International Meshing Roundtable (IMR26), 18-21 September 2017, Barcelona, Spain. *Procedia Engineering*. 2017. V. 203.
18. *Mehta A.* Mesh deformation algorithms for fluid-structure interaction studies. 2020.
19. *Takizawa K., Tezduyar T.E., Avsar R.* A low-distortion mesh moving method based on fiber-reinforced hyperelasticity and optimized zero-stress state // *Computational Mechanics*. 2020. № 65. P. 1567–1591.
20. *Shamanskiy A., Simeon B.* Mesh moving techniques in fluid-structure interaction: robustness, accumulated distortion and computational efficiency // *Computational Mechanics*. 2021. V. 67. P. 583–600.
21. *Вершков В.А.* Математическое моделирование процесса обтекания шарнирного несущего винта вертолета методом деформируемых неструктурированных сеток. Дисс. на соискание к.т.н. Москва. 2021. 116 С.
22. *Garanzha V., Kaporin I., Kudryavtseva L., Protais F., Sokolov D.* In the Quest for Scale-optimal Mappings // *ACM Transactions on Graphics*. 2024. V. 43, № 1. P. 1–16.
23. *Карлов Н.В., Кириченко Н.А.* Колебания, волны, структуры. М.: Физматлит, 2003.
24. *Петров М.Н., Тамбова А.А., Титарев В.А., Утюжников С.В., Чикиткин А.В.* Программный комплекс FlowModellium для расчета высокоскоростных течений сжимаемого газа // *Ж. вычисл. матем. и матем. физ.* 2018. Т. 58. № 11. С. 1932–1954.

ALGORITHM FOR MESH ADAPTATION TO A FLOW FIELD WITH A BOW SHOCK WAVE

I. V. Voronich*, N. S. Smirnova**, V. A. Titarev***

119333 Moscow, FRC CSC RAS, Russia

**e-mail: i.voronich@frcsc.ru*

***e-mail: nsmirnova@frcsc.ru*

****e-mail: vladimir.titarev@frcsc.ru*

Received: 08.04.2024

Revised: 13.05.2024

Accepted: 31.05.2024

Abstract. The construction of high-quality computational grids plays an essential role in obtaining high-precision results in the problems of calculating the external high-speed flow around bodies. The priority task is to adapt the computational grid to discontinuities, primarily to leading edge shock waves. In this paper, a variant of the algorithm for adapting the computational grid as a mechanical system with elastic connections to the flow field containing the bow shock wave is considered. As a result of applying the algorithm to a typical structured computational grid, it is rebuilt in areas of a large field gradient by drawing grid lines into the shock area while maintaining the quality of grid elements. The considered problems show the possibility of applying the described algorithm to real problems of external flow.

Keywords: grid adaptation, computational aerodynamics, supersonic flows, parallel computing.

ЧИСЛЕННОЕ ИССЛЕДОВАНИЕ УДАРНО-ВОЛНОВЫХ ПРОЦЕССОВ В МАТЕРИАЛАХ ПОД ВОЗДЕЙСТВИЕМ СВЕРХКОРОТКИХ ЛАЗЕРНЫХ ИМПУЛЬСОВ С ИСПОЛЬЗОВАНИЕМ МОДЕЛИ БАЕРА–НУНЦИАТО¹⁾

© 2024 г. П. А. Чупров^{1,*}, С. В. Фортова¹, В. В. Шепелев¹

¹ 123056 Москва, 2-ая Брестская ул., 19/18, Институт автоматизации проектирования РАН, Россия

*e-mail: petchu@mail.ru

Поступила в редакцию 29.02.2024 г.

Переработанный вариант 08.04.2024 г.

Принята к публикации 31.05.2024 г.

Представлена математическая модель, основанная на многофазной модели Баера–Нунциато. Эффективность модели продемонстрирована на численном решении ударно-волновых задач в конденсированных средах при наличии явной контактной границы с вакуумом. Рассматриваются результаты численного моделирования задач о взаимодействии фемтосекундного лазерного излучения с мишенью из алюминия. Показано преимущество применения модели Баера–Нунциато по сравнению с однофазной гидродинамической моделью при расчете динамики контактной границы. Простота реализации и возможность легкого введения дополнительных субмоделей, таких как горение, делает этот подход привлекательным для моделирования высокоэнергетических процессов в многофазных средах. Библ. 16. Фиг. 7.

Ключевые слова: фемтосекундное лазерное излучение, гидродинамика, металлы.

DOI: 10.31857/S0044466924090159, EDN: WIQNKT

1. ВВЕДЕНИЕ

Сверхкороткие лазерные импульсы (< 1 нс) — это одно из новых и активно развивающихся направлений техники, находящее широкое применение в промышленности. С их помощью возможно добиться модификации поверхностной структуры материала с необычайной точностью, недоступной механическому воздействию. Одним из ярких примеров такой технологии является поверхностное упрочнение материала LSP (laser shock peening) (см. [1]– [3]). Поверхность материала, упрочненного таким образом, показывает отличную сопротивляемость усталостным нагрузкам и малое количество дефектов поверхности, что обуславливает применение этой технологии в авиастроении, например, для упрочнения лопаток двигателей. Применяются лазеры и для создания поверхностных маломасштабных структур, например LIPSS (laser induced periodic surface structures) (см. [4], [5]). С помощью таких периодических структур можно значительно изменить свойства поверхности, например, ее коэффициент трения или отражательную характеристику. Таким образом, существует большой интерес к исследованию, особенно численному, процессов, вызванных воздействием сверхкороткого лазерного импульса. Существенным свойством таких коротких импульсов энергии является тот факт, что поглощение энергии лазерного импульса происходит целиком внутри небольшого слоя толщиной d_T , причем время нагрева существенно меньше характерного акустического времени. Это приводит к тому, что сразу после облучения в материале присутствует нагретый слой малой толщины с огромным давлением (см. [6]). Этот замечательный факт позволяет не создавать сложные численные алгоритмы, напрямую моделирующие лазерное излучение и его взаимодействие с веществом, а использовать теоретические оценки для задания корректных начальных условий. Особый интерес представляет исследование волновой динамики процесса, особенно на малых временах. При существенном затухании волн в материале они больше не способны ни вызывать пластические деформации, ни приводить к абляции вещества, что и обуславливает интерес к начальной стадии процесса. Большое распространение получили такие методы, как молекулярно-динамическое моделирование (см. [7]– [9]). Молекулярно-динамическое моделирование позволяет получать невероятно точные результаты, в некоторых

¹⁾ Работа выполнена в рамках государственного задания АААА-А19-124022400174-3.

случаях максимально приближенные к полной симуляции реально протекающих процессов. Однако попытки исследования крупномасштабных многомерных явлений с помощью такого метода в настоящее время встречают трудности, связанные с недостаточной мощностью доступных вычислительных ресурсов. Таким образом, возникает потребность в более экономном методе, основанном, например, на принципах механики сплошной среды. Известны успешные попытки моделирования начальных этапов процесса с помощью численного алгоритма, основанного на гидродинамической системе уравнений Эйлера. Однако подобный подход приводит к появлению нефизичных эффектов на границе между материалами, например, границе металл-вакуум (см. [10]). В настоящей работе предлагается метод, основанный на двухфазной гидродинамической системе уравнений Баера–Нунциато, в которой одна из фаз представляет собой конденсированное вещество мишени, а другая – идеальный газ. Представленный метод верифицируется на модельных задачах, а полученные результаты сравниваются с результатами моделирования другими методами.

2. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

Математическая модель основана на системе уравнений Баера–Нунциато (см. [11]):

$$\mathbf{U}_t + \mathbf{F}_x(\mathbf{U}) + \mathbf{G}_y(\mathbf{U}) = \mathbf{H}(\mathbf{U}) \bar{\alpha}_x + \mathbf{S}(u) \bar{\alpha}_y$$

$$\mathbf{U} = \begin{bmatrix} \bar{\alpha} \\ \bar{\alpha}\bar{q} \\ \bar{\alpha}\bar{q}\bar{v} \\ \bar{\alpha}\bar{q}\bar{u} \\ \bar{\alpha}\bar{q}\bar{E} \\ \alpha q \\ \alpha q v \\ \alpha q u \\ \alpha q E \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} 0 \\ \bar{\alpha}\bar{q}\bar{v} \\ \bar{\alpha}(\bar{q}\bar{v}^2 + \bar{p}) \\ \bar{\alpha}\bar{u}\bar{v} \\ \bar{\alpha}\bar{v}(\bar{q}\bar{E} + \bar{p}) \\ \alpha q v \\ \alpha(qv^2 + p) \\ \alpha uv \\ \alpha v(qE + p) \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} 0 \\ \bar{\alpha}\bar{q}\bar{u} \\ \bar{\alpha}\bar{v}\bar{u} \\ \bar{\alpha}(\bar{q}\bar{u}^2 + \bar{p}) \\ \bar{\alpha}\bar{u}(\bar{q}\bar{E} + \bar{p}) \\ \alpha q u \\ \alpha v u \\ \alpha(qu^2 + p) \\ \alpha u(qE + p) \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} -\bar{v} \\ 0 \\ \bar{p} \\ \bar{p}\bar{v} \\ 0 \\ -\bar{p} \\ -\bar{p}\bar{v} \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} -\bar{u} \\ \bar{p} \\ 0 \\ \bar{p}\bar{u} \\ -\bar{p} \\ 0 \\ -\bar{p}\bar{u} \end{bmatrix}.$$

Здесь t — время, x, y — пространственные координаты, α — объемная доля, v — x -компонента скорости, u — y -компонента скорости, q — плотность, p — давление, E — полная удельная энергия, e — удельная внутренняя энергия. Параметры с чертой сверху соответствуют твердой фазе, параметры без черты — газовой фазе. Параметры $\bar{p}$, $\bar{v}$ и $\bar{u}$ — давление и скорости на межфазной границе, которые были выбраны следующим образом: $\bar{v} = \bar{v}$, $\bar{u} = \bar{u}$ и $\bar{p} = \bar{p}$.

Модель Баера–Нунциато можно отнести к классу *diffuse interface*, т.е. моделей со взаимопроникновением фаз. Это означает, что в модели предполагается наличие в каждой ячейке всех фаз. В случае, когда некоторый объем вычислительной области должен соответствовать физической области, занимаемой лишь одной фазой, например, металлом, объемная доля остальных фаз полагается равной некоторой заранее определенной малой константе. Таким образом, и граница между различными фазами представляет собой не резкий перепад, а плавное изменение соотношения объемных долей. Таким образом, наличие в каждой ячейке нескольких фаз приводит, с одной стороны, к тому, что решаемые уравнения не меняются во всей расчетной области, а с другой стороны, к тому, что в каждой ячейке разные фазы обладают разными давлениями и скоростями. Для стабильной работы алгоритма и наилучшему моделированию реальных процессов эти давления и скорости должны стремиться к равновесию между фазами, что приводит к необходимости некоторых алгоритмов релаксации давления и скорости. К настоящему времени в литературе присутствуют хорошо обоснованные и проверенные алгоритмы для задач динамики пористых и гранулированных сред (см., например, [12]), которые используют уравнения состояния обеих фаз для нахождения равновесного давления. Применительно к задачам описания ударно-волновых процессов в твердых телах применить классические механизмы релаксации затруднительно по причине возникновения отрицательных давлений, что делает невозможным использование уравнения состояния идеального газа для моделирования области вакуума. В связи с этим для настоящей работы был разработан следующий прототип алгоритма релаксации давления и скорости:

- Для начала определяется «эффективное» давление $\hat{p} = \bar{p}\bar{\alpha}$.
- Решается нелинейное уравнение на $\hat{q}$ — плотность после релаксации: $\bar{e}(\hat{p}, \hat{q}) - \bar{e}(\bar{p}, \bar{q}) + \frac{1}{2}(\bar{p} + \hat{p})\left(\frac{1}{\bar{q}} - \frac{1}{\hat{q}}\right) = 0$, где $\bar{e}$ — внутренняя энергия твердой фазы.
- С помощью найденной плотности обновляется значение объемной доли $\hat{\alpha} = \bar{\alpha}\frac{\bar{q}}{\hat{q}}$.

- С помощью подстановки найденных равновесных значений давления и плотности в закон сохранения энергии находятся равновесное значение скорости $\varepsilon(\hat{p}, \hat{q}) + \hat{v}^2/2 + \hat{u}^2/2 = \varepsilon(p, q) + v^2/2 + u^2/2$. В предположении, что кинетическая энергия равномерно распределяется по направлениям, вышеуказанные соотношения дополняются равенством $\hat{v}^2 - v^2 = \hat{u}^2 - u^2$.

Важным свойством является отсутствие параметров второй (газовой) фазы в этих выражениях, что позволяет использовать идеальный газ для моделирования области, физически соответствующей области вакуума. Это позволяет существенно упростить вычислительный процесс, так как нет необходимости явным образом разрешать границу между областями с разным наполнением, т.е. возможен сквозной счет. В настоящей работе будет продемонстрирована работоспособность такого подхода. Здесь и далее при упоминании моделирования вакуума будет подразумеваться именно использование идеального газа при нормальных условиях в начальный момент времени.

Необходимо отдельно упомянуть, что используемая в работе математическая модель не учитывает упруго-пластические эффекты, наблюдаемые в твердых телах. В силу этого факта, модель не будет способна количественно точно предсказывать итоговые значения формы зоны упрочнения и массы выброшенного вещества. Однако для исследования начальных этапов высокоэнергетических процессов, при которых материалы ведут себя как слабосжимаемые жидкости, настоящая модель может быть применена с большим успехом, как будет показано далее.

3. УРАВНЕНИЯ СОСТОЯНИЯ

Для замыкания представленной в предыдущем разделе системы уравнений необходимо задаться уравнением состояния (УРС) вещества. В настоящей работе исследуются свойства твердых тел при экстремальных условиях, что требует аккуратного подхода к данному вопросу. Для всех исследуемых веществ было использовано уравнение состояния в форме Ми-Грюнайзена. В работе используется два вещества, соответственно и два разных УРС — для фторида лития и для алюминия. УРС для фторида лития выглядит следующим образом:

$$p = p_r(q) + \Gamma q(\varepsilon - \varepsilon_r),$$

$$p_r = q_0 c^2 \frac{1-x}{(1-s(1-x))^2}, \varepsilon_r = \frac{p_r}{q_0} \frac{1-x}{2}, x = \frac{q_0}{q}.$$

Здесь $q_0 = 2650 \text{ кг/м}^3$ — референсная плотность вещества, $c = 5150 \text{ м/с}$ — референсная скорость звука, $s = 1.35$ — показатель адиабаты, $\Gamma = 0.71$ — параметр Грюнайзена. Для моделирования алюминия было использовано уравнение состояния, подробно описанное в [10]:

$$p = p_c(q) + \Gamma q(\varepsilon - \varepsilon_c),$$

$$p_c = p_0 y(y^a - y^b), y > 1,$$

$$p_c = p_g(y^n - 1), y < 1,$$

$$\varepsilon_c = \frac{p_0}{q_0} \left(\frac{y^a}{a} - \frac{y^b}{b} \right), y > 1,$$

$$\varepsilon_c = \frac{p_0}{q_0} (a - b) \left[\frac{1}{n} \left(\frac{y^{n-1}}{n-1} + \frac{1}{y} \right) - \frac{1}{n-1} - \frac{1}{ab} \right], y < 1,$$

$$n = p_0(a - b)/p_g.$$

Здесь $q_0 = 2750 \text{ кг/м}^3$ — холодная плотность вещества, $\Gamma = 1.2$ — параметр Грюнайзена, $p_0 = 560.964 \text{ ГПа}$, $a = 1.1266$, $b = 0.9775$, $p_g = 15 \text{ ГПа}$ — параметры холодной кривой, $y = \frac{q}{q_0}$.

4. ЧИСЛЕННЫЙ МЕТОД

Для численного решения системы 2 используется конечно-объемный метод. Для этого система разделяется на две части: первое уравнение системы и остальные уравнения. Такой выбор обусловлен необходимостью отдельной аппроксимации уравнения переноса объемной доли для предотвращения численных осцилляций. В общем виде решение оставшейся системы записывается как

$$\mathbf{U}_{i,j}^{n+1} = \mathbf{U}_{i,j}^n - \frac{\Delta t^n}{\Delta x} [\mathbf{F}_{i+1/2,j} - \mathbf{F}_{i-1/2,j}] - \frac{\Delta t^n}{\Delta y} [\mathbf{G}_{i,j+1/2} - \mathbf{G}_{i,j-1/2}] + \mathbf{H}_{i,j}^n \bar{\alpha}_x + \mathbf{S}_{i,j}^n \bar{\alpha}_y.$$

Здесь $\mathbf{F}_{i+1/2,j}$, $\mathbf{G}_{i,j+1/2}$, $\mathbf{F}_{i-1/2,j}$, $\mathbf{G}_{i,j-1/2}$ — численные аппроксимации потоков сквозь соответствующую стенку ячейки. Для их расчета используется HLLC (Harten-Lax-van Leer Contact) — подобный метод для двумерной системы уравнений Баера—Нунциато. Для повышения порядка аппроксимации использовалась minmod-реконструкция решения. Данный метод подробно описан и верифицирован, например, в [13] и основан на теоретической работе [14], а также [15], [16]. Выбор численного метода особенно важен для задач с явной границей между областями, занятыми только одним из нескольких веществ. Известно, что метод HLL (Harten-Lax-van Leer) для уравнений Баера—Нунциато не способен сохранять стационарный разрыв объемной доли, т.е. даже в отсутствии скоростей и перепадов давлений разрыв объемной доли будет "расползаться". Такое явление неприемлемо в исследуемом типе задач. Повышение порядка пространственной аппроксимации также служит в основном для более точного передания положения границы между веществами.

Ниже приведены формулы для численной аппроксимации потока в x -направлении $\mathbf{F}_{i+1/2,j}$ (выражения для потока в y -направлении полностью аналогичны):

$$\mathbf{F}_{i+1/2,j} = \begin{bmatrix} \Phi_{i+1/2,j} \\ \Phi_{i+1/2,j} \end{bmatrix}, \quad \mathbf{U}_{i,j} = \begin{bmatrix} \mathbf{W}_{i,j} \\ \overline{\mathbf{W}}_{i,j} \end{bmatrix},$$

$$\Phi_{i+1/2,j} = \begin{cases} \Phi_{i,j}, & \text{если } S_{i+1/2,j}^- \geq 0, \\ \Phi_{i,j} + S_{i+1/2,j}^- (\mathbf{Q}_*^- - \mathbf{W}_{i,j}), & \text{если } S_{i+1/2,j}^- < 0, S_{i+1/2,j}^* \geq 0, \\ \Phi_{i+1,j} + S_{i+1/2,j}^+ (\mathbf{Q}_*^+ - \mathbf{W}_{i+1,j}), & \text{если } S_{i+1/2,j}^* < 0, S_{i+1/2,j}^+ \geq 0, \\ \Phi_{i+1,j}, & \text{если } S_{i+1/2,j}^+ \leq 0, \end{cases}$$

$$S_{i+1/2,j}^+ = \max(v_{i,j} + c_{i,j}, v_{i+1,j} + c_{i+1,j}), \quad S_{i+1/2,j}^- = \min(v_{i,j} - c_{i,j}, v_{i+1,j} - c_{i+1,j}),$$

$$S_{i+1/2,j}^* = \frac{p_{i+1,j} - p_{i,j} + \varrho_{i,j} u_{i,j} (S_{i+1/2,j}^- - u_{i,j}) - \varrho_{i+1,j} u_{i+1,j} (S_{i+1/2,j}^+ - u_{i+1,j})}{\varrho_{i,j} (S_{i+1/2,j}^- - u_{i,j}) - \varrho_{i+1,j} (S_{i+1/2,j}^+ - u_{i+1,j})},$$

$$\mathbf{Q}_*^- = \begin{bmatrix} C^- \\ C^- S_{i+1/2,j}^* \\ C^- v_{i,j} \\ C^- \left(\frac{p_{i,j}}{\varrho_{i,j}} + (S_{i+1/2,j}^* - u_{i,j}) \left(S_{i+1/2,j}^* + \frac{p_{i,j}}{\varrho_{i,j} (S_{i+1/2,j}^- - u_{i,j})} \right) \right) \end{bmatrix},$$

$$C^- = \frac{\alpha_{i,j} \varrho_{i,j} (S_{i+1/2,j}^- - u_{i,j})}{(S_{i+1/2,j}^- - S_{i+1/2,j}^*)},$$

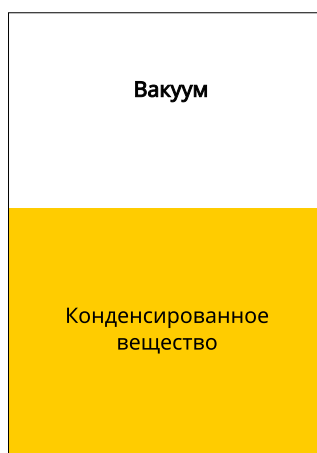
$$\mathbf{Q}_*^+ = \begin{bmatrix} C^+ \\ C^+ S_{i+1/2,j}^* \\ C^+ v_{i+1,j} \\ C^+ \left(\frac{p_{i+1,j}}{\varrho_{i+1,j}} + (S_{i+1/2,j}^* - u_{i+1,j}) \left(S_{i+1/2,j}^* + \frac{p_{i+1,j}}{\varrho_{i+1,j} (S_{i+1/2,j}^+ - u_{i+1,j})} \right) \right) \end{bmatrix},$$

$$C^+ = \frac{\alpha_{i+1,j} \varrho_{i+1,j} (S_{i+1/2,j}^+ - u_{i+1,j})}{(S_{i+1/2,j}^+ - S_{i+1/2,j}^*)}.$$

Здесь c — скорость звука, определяемая с помощью уравнения состояния соответствующей фазы. Записанные выше выражения для расчета численного потока подсистемы газовой фазы обобщаются и на подсистему конденсированной фазы с помощью элементарной замены величин без черты на величины с чертой. В этом смысле используемый солвер является симметричным относительно фаз.

5. ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ

В настоящей работе рассматриваются результаты численного моделирования задач о взаимодействии фемтосекундного лазерного излучения с мишенями из твердых материалов. Важно подчеркнуть, что используемая математическая модель не подразумевает прямого моделирования лазерного пучка и первоначальной стадии



Фиг. 1. Схематическая постановка первой модельной задачи.

поглощения энергии излучения мишенью. Эти процессы происходят на существенно меньших временных масштабах, поэтому вполне обоснованным является рассмотрение волновых процессов в системе в предположении уже полностью установившегося в начальный момент времени распределения температуры (давления и плотности) внутри мишени.

5.1. Центрированная волна разрежения

Для первоначальной верификации модели была рассмотрена одномерная задача разгрузки нагретого слоя конденсированной фазы фторида лития (мишень) на границе с вакуумом. Данная задача изучена и представляет собой хороший способ верификации модели. Известно, что для конденсированного вещества распределения плотности и давления при расширении в вакуум имеют четкую ступеньку, что отличает этот случай от расширения газа в вакуум, где распределения давления и плотности в волне гладкие. Геометрическая постановка задачи показана на фиг. 1.

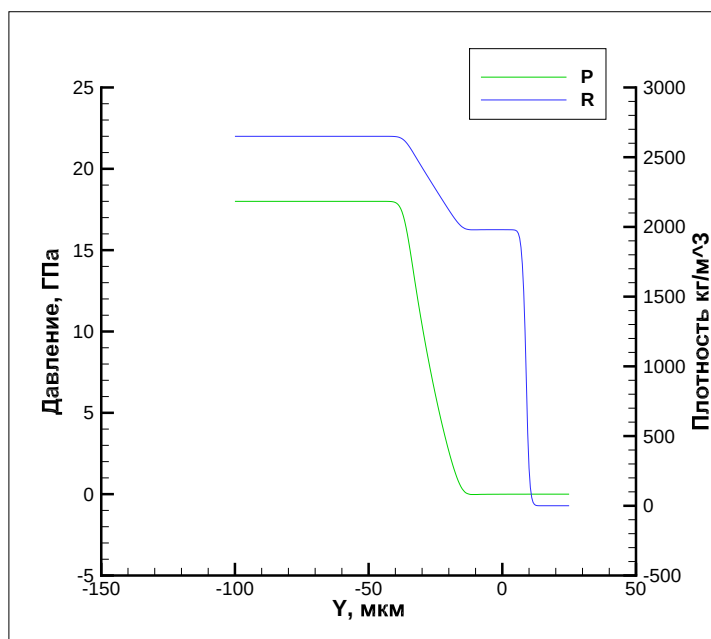
Начальное давление конденсированной фазы — 18 ГПа, в начальный момент граница между мишенью и вакуумом находится в нуле координат. Плотность вещества в начальный момент времени равна 2650 кг/м³. Толщина мишени — 100 мкм, слой вакуума над мишенью — 25 мкм. С обеих сторон поставлены неотражающие граничные условия. На фиг. 2 показаны пространственные распределения давления и плотности на момент времени 7 нс. Отчетливо видна характерная для ЦВР в конденсированных средах "ступенька" плотности, образовавшаяся на месте первоначальной границы. В полном соответствии с известными результатами находится и тот факт, что на этой "ступеньке" давление обращается в нуль. Таким образом, можно утверждать, что распространение волны разрежения в вакуум моделируется корректно, что позволяет перейти к более сложным верификационным задачам.

5.2. Динамика тонкого нагретого слоя

Особый интерес в контексте исследуемых задач представляет динамика тонкого нагретого слоя вещества (фторида лития), который, с одной стороны, контактирует с вакуумом, а с другой — с холодным веществом. В такой постановке внутрь нагретого слоя будут распространяться две волны разрежения: со стороны свободной поверхности и со стороны холодного вещества. При взаимодействии двух волн разгрузки в веществе появится зона растягивающих напряжений, которая при должной амплитуде может привести к потере сплошности веществом — отрыву. Для моделирования отрыва в настоящей работе использовался простой алгоритм. При падении скорости звука вещества до значений ниже 1 км/с в ячейке предполагалось нарушение сплошности, что выражалось в уменьшении объемной доли вещества с одновременным увеличением плотности для сохранения массы в ячейке:

$$\begin{aligned}\bar{\alpha}_i^s &= 0.9\bar{\alpha}_i, \text{ если } \bar{c}_i < 1 \text{ км/с}, \\ \bar{\rho}_i^s &= \bar{\alpha}_i\bar{\rho}_i/\bar{\alpha}_i^s.\end{aligned}\tag{1}$$

Здесь величины с верхним индексом s соответствуют значениям в ячейке после процедуры моделирования откола. Вследствие уменьшения объемной доли в дальнейшем механизм релаксации давления приводит давление в такой ячейке к нулевому. Как будет видно из результатов моделирования, такой алгоритм, хоть и не



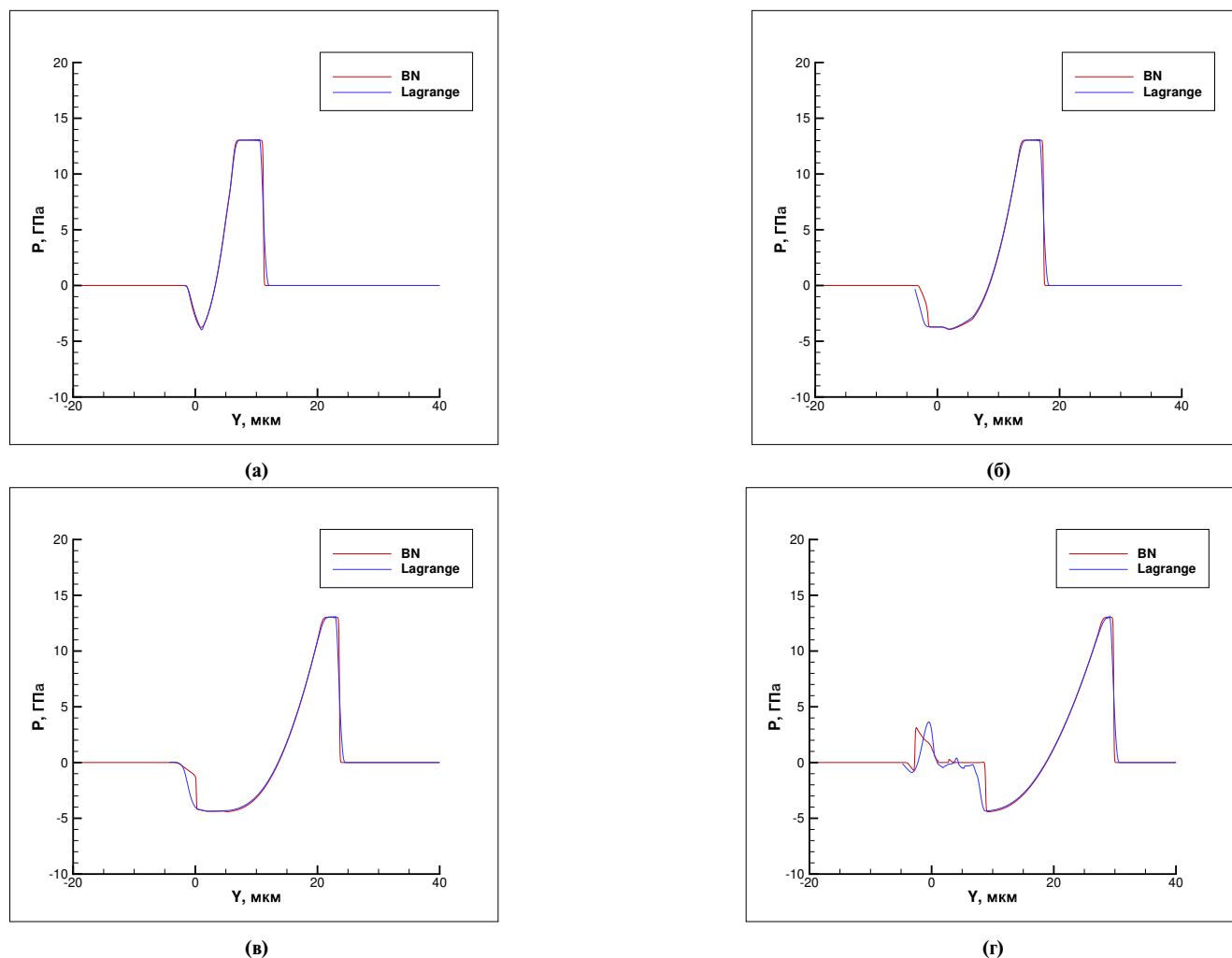
Фиг. 2. Результаты верификационного расчета.



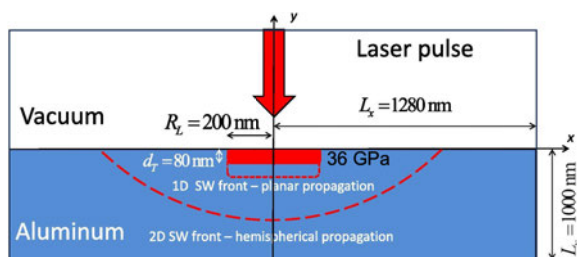
Фиг. 3. Схематичная постановка второй модельной задачи.

претендует на абсолютную корректность с физической точки зрения, показывает качественно и количественно соответствующие другим моделям результаты.

Таким образом, была рассмотрена одномерная модельная задача, схематичная постановка которой изображена на фиг. 3. В качестве референса использовались результаты численного эксперимента, проведенного с помощью модели, основанной на лагранжевом формализме. Начальное давление горячей области — 25 ГПа, плотность вещества в начальный момент времени равна 2650 кг/м^3 во всем объеме фторида лития, толщина горячей области — 5 мкм. Результаты численного моделирования представлены на фиг. 4. Представленные результаты показывают отличное количественное совпадение уходящей вглубь вещества ударной волны. Небольшие различия вызваны разными численными методами, используемыми для решения задач. На контакте с вакуумом различия более заметны, однако не меняют картину течения качественно. Важным элементом этого теста является обработка откольных явлений. В момент времени приблизительно 3.1 нс в области отрицательных давлений вещество достигает состояния откола. Как видно из фиг. 4г, оба метода показывают качественно и количественно совпадающие результаты. Полученные результаты позволяют сделать вывод о возможности представленной модели для исследования многофазных задач с присутствием границы с вакуумом.



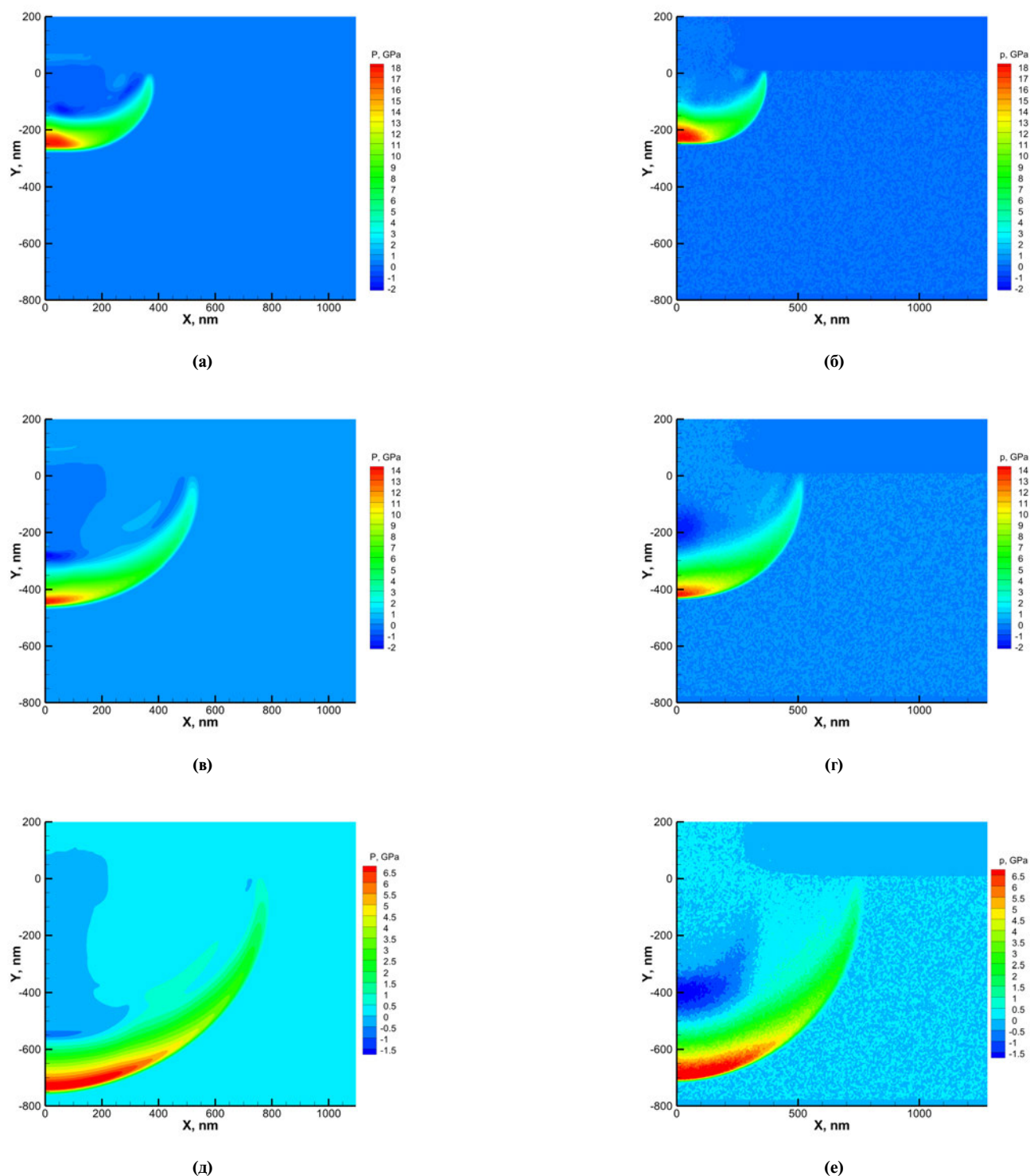
Фиг. 4. Пространственные распределения давления в моменты времени 1 нс (а), 2 нс (б), 3 нс (в), 4 нс (г).



Фиг. 5. Постановка плоской задачи о взаимодействии лазерного импульса с алюминиевой мишенью.

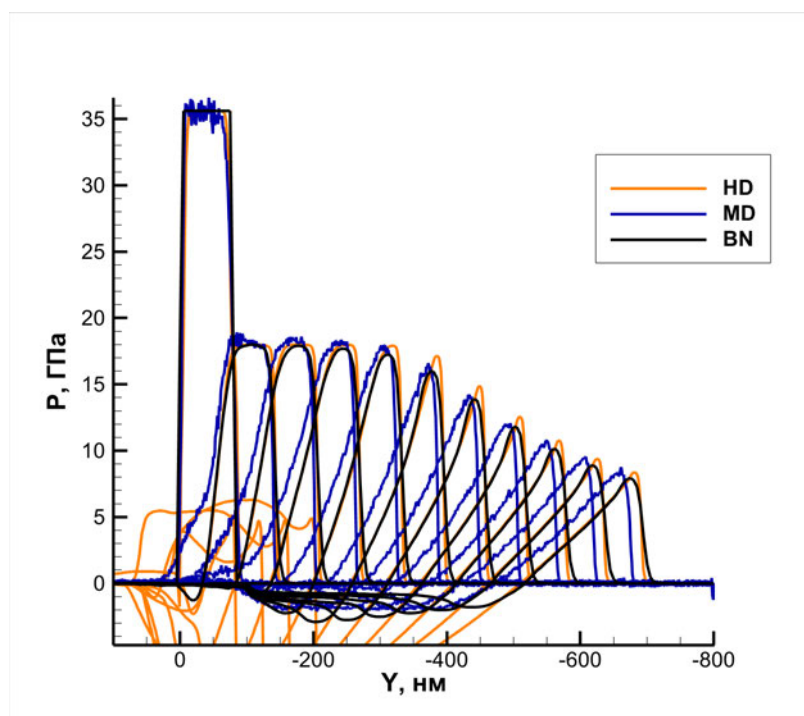
5.3. Объемная алюминиевая мишень

В настоящем разделе рассматривается численное исследование плоской задачи о взаимодействии фемтосекундного лазерного импульса с толстой алюминиевой мишенью. Постановка задачи отображена на фиг. 5. В начальный момент времени ширина нагретой области составляет 200 нм, глубина прогрева — 40 нм. Давление горячей зоны — 35.6 ГПа. Эта же задача была численно исследована с помощью алгоритма, основанного на молекулярно-динамическом подходе в приближении жидкого алюминия в работе [10]. Результаты сравнения представлены на фиг. 6. Изначальная область повышенной температуры порождает ударную волну (УВ), распространяющуюся по веществу. На протяжении первых пикосекунд процесса эта волна практически плоская, распространяющаяся параллельно оси X . Однако, с течением времени ее фронт изгибается и к моменту време-



Фиг. 6. Пространственные распределения давления в моменты времени: (а), (б) – 28.8 нс, (в), (г) – 57.6 нс, (д), (е) – 105.6 нс. Слева – BN, справа – MD.

ни 105 пс уже выглядит как дуга окружности. Амплитуда и положение фронта УВ показывают хорошее совпадение с результатами молекулярно-динамического моделирования на всех этапах процесса. Не менее важным процессом является распространение волн разгрузки (ВР). В прогретый участок входят три ВР: одна с поверхности, одна из холодного алюминия снизу и одна из холодного алюминия справа. После взаимодействия эти



Фиг. 7. Профили давления на оси канала для разных моделей.

волны создают область растягивающих напряжений, которая в дальнейшем распространяется за фронтом УВ. Как будет показано далее, характер этой области не будет влиять на верное отображение УВ, однако неточности в моделировании этой зоны могут быть критичными при исследовании процессов образования кратеров и полостей, а также в целом мешать стабильности решения. Во все три момента времени, представленные на фиг. 6, можно заметить, что представленная в настоящей работе модель недостаточно точно передает характеристики факела — области, заполненной выброшенным из основной массы мишени веществом. Факел расположен непосредственно над зоной изначального прогрева и в молекулярно-динамическом расчете имеет более выраженную форму. Динамика в этой области в основном зависит от испарения и конденсации вещества мишени. Разработка и включение субмодели, позволяющей адекватно описывать указанные процессы в представленной модели является предметом дальнейшего исследования.

Для более точного сравнения результатов рассмотрим одномерные профили давления на оси симметрии. К имеющемуся набору данных были добавлены результаты расчета однофазным гидродинамическим кодом. Эти результаты представлены на фиг. 7. Относительно описания треугольной волны, бегущей вправо, оба гидродинамических кода показывают очень близкий результат, лишь немного отличаясь от молекулярно-динамического моделирования. Особый интерес представляет область за этой волной, в которой возникает “яма”, отрицательного давления. Описываемый в этой статье метод показывает удовлетворительное качественное и количественное соответствие с MD моделированием, что отличает его от однофазной модели, в которой возникают отрицательные давления во много раз большие. Более того, в результатах предлагаемого метода отсутствуют и нефизичные значительные колебания давления на месте первоначального контакта с вакуумом. Отметим еще одно преимущество рассматриваемой модели. В однофазной модели используется одно уравнение состояния на всей расчетной области. Это приводит к некорректному моделированию ударно-волновых процессов на межфазной границе. Используемая многофазная модель позволяет применять свое уравнение состояния для каждой из фаз, что приводит к получению физически верной волновой структуры решения.

6. ЗАКЛЮЧЕНИЕ

В настоящей работе представлена математическая модель, основанная на многофазной модели Баера–Нунциато и использованная для решения ударно-волновых задач в конденсированных средах при наличии явной контактной границы с вакуумом. Эта математическая модель, совместно с конечно-объемным методом ее решения, были верифицированы на различных задачах динамики экстремальных состояний вещества и по-

казали качественное и количественное совпадение результатов как с теоретическими предсказаниями, так и с результатами других апробированных моделей. В отличие от однофазной гидродинамической модели, динамика контактной границы намного более приближена к реальности. Существенным преимуществом предложенной модели является простота реализации, возможность легкого введения дополнительных моделей, таких как горение, а также сравнительно малые вычислительные затраты.

СПИСОК ЛИТЕРАТУРЫ

1. *Kaczmarek A., Denis P., Krajewski M., Moscicki T., Malolepszyand Hoffman M.* Improved Laser Ablation Method for the Production of Luminescent Carbon Particles in Liquids // *Materials*. 2021. V. 14. Iss. 9.
2. *Liao Yiliang, Ye Chang, Cheng G.J.* A review: Warm laser shock peening and related laser processing technique // *Optics & Laser Tech.* 2016. V. 78. Iss. 2. P. 15–24.
3. *Walsh N., Costello J.T., Kelly T.J.* Optical diagnostics of laser-produced aluminium plasmas under water // *Appl. Phys. B*. 2017. V. 123. Iss. 6. P. 179.
4. *Haiying Song, Liu Shi-Bing, Liu H., Wang Yang, Chen Tao, Dong Xiang-Ming* Mechanism for femtosecond laser-induced periodic subwavelength structures on solid surface: surface two-plasmon resonance //
5. *Petrovic S., Gakovic B., Perusko D., Stratakis E., Radovic I., Cekada M., Fotakis C., Jelenkovic B.* Femtosecond laser-induced periodic surface structure on the Ti-based nanolayered thin films // *J. Appl. Phys.* 2013. V. 114. P 233108.
6. *Inogamov N.A., Zhakhovskii V.V., Ashitkov S.I., Petrov Yu.V., Agranat M.B., Anisimov S.I., Nishihara K., Fortov V.E.* Nanospallation induced by an ultrashort laser pulse // *J. Exp. Theor. Phys.* 2008. V. 107. Iss. 1. P 1–19.
7. *Иногамов Н.А., Перов Е.А., Жаховский В.В., Шепелев В.В., Петров Ю.В., Фортובה С.В.* Лазерная ударная волна: пластичность, толщина слоя остаточных деформаций и переход из упругопластического в упругий режим распространения // *Письма в ЖЭТФ*. 2022. Т. 115. № 2. С 80–88.
8. *Xu J., Xue D., Gaidai O., Wang Y., Shaolin Xu* Molecular dynamics simulation of femtosecond laser ablation of Cu50Zr50 metallic glass based on two-temperature model // *Processes*. 2023. V. 11. Iss. 6.
9. *Bauerhenne B., Zijlstra E.S., Garcia M.E.* Molecular dynamics simulations of a femtosecond-laser-induced solid-to-solid transition in antimony // *Appl. Phys. A*. 2017. V. 123. Iss. 9. P 608.
10. *Shepelev V.V., Petrov Yu.V., Inogamov N.A., Zhakhovskii V.V., Perov E.A., Fortova S.V.* Attenuation and inflection of initially planar shock wave generated by femtosecond laser pulse // *Opt. Las. Tech.* 2022. V. 152. P 108100.
11. *Baer M.R., Nunziato J.W.* A two-phase mixture theory for the deflagration-to-detonation transition (ddt) in reactive granular materials // *Int. J. Multiphase Flow*. 1986. V. 12. Iss. 6. P. 861–889.
12. *Saurel R., Favrie N., Petitpas F., Lallemand M.-H., Gavriluk S.L.* Modelling dynamic and irreversible powder compaction // *J. Fl. Mech.* 2010. V. 664. P 348–396.
13. *Chuprov P., Poroshyna Ya., Utkin P.* Numerical simulation of the propagation of a shock wave above the dense layer of particles using the Baer-Nunziato system of equations. // *Comb. and Expl.* 2022. V. 15. Iss. 2. P. 67–74.
14. *Tokareva S.A., Toro E.F.* HLLC-type Riemann solver for the Baer–Nunziato equations of compressible two-phase flow // *J. Comp. Phys.* 2010. V. 229. Iss. 10. P. 3573–3604.
15. *Li Q.* Difference scheme for two-phase flow // *Appl. Math. Mech.* 2004. V. 25. P 536–545.
16. *Liang Shan, Liu Wei, Yuan Li* Solving seven-equation model for compressible two-phase flow using multiple GPUs // *Comp. Fluids*. 2014. V. 99. P 156–171.

NUMERICAL STUDY OF SHOCK WAVE PROCESSES IN MATERIALS UNDER THE INFLUENCE OF ULTRASHORT LASER PULSES USING THE BAER–NUNZIATO MODEL

P. A. Chuprov*, S. V. Fortova, V. V. Shepelev

133056, 19/18 2nd Brestskaya str., Moscow, Institute of Design Automation of the Russian Academy of Sciences, Russia

**e-mail: petchu@mail.ru*

Received: 29.02.2024

Revised: 08.04.2024

Accepted: 31.05.2024

Abstract. A mathematical model based on the Baer–Nunziato multiphase model is presented. The effectiveness of the model is demonstrated by the numerical solution of shock wave problems in condensed media in the presence of an explicit contact boundary with vacuum. The results of numerical simulation of problems of interaction of femtosecond laser radiation with an aluminum target are considered. The advantage of using the Baer–Nunziato model in comparison with the single-phase hydrodynamic model in calculating the dynamics of the contact boundary is shown. The simplicity of implementation and the possibility of easy introduction of additional submodels, such as ignition, makes this approach attractive for modeling high-energy processes in multiphase media.

Keywords: femtosecond laser radiation, hydrodynamics, metals.

О ТОЧНОСТИ РЕШЕТОЧНЫХ УРАВНЕНИЙ БОЛЬЦМАНА НИЗКОГО И ВЫСОКОГО ПОРЯДКОВ В ПРИЛОЖЕНИЯХ К МЕДЛЕННЫМ ИЗОТЕРМИЧЕСКИМ МИКРОТЕЧЕНИЯМ

© 2024 г. О. В. Ильин^{1,*}¹ 119333 Москва, ул. Вавилова, 44, ФИЦ ИУ РАН, Россия

*e-mail: oilyin@gmail.com

Поступила в редакцию 12.12.2023 г.

Переработанный вариант 14.04.2024 г.

Принята к публикации 31.05.2024 г.

Рассмотрен вопрос применимости двумерных решеточных уравнений Больцмана разного порядка на стандартных решетках к описанию медленных изотермических разреженных течений. В качестве эталонной задачи используется двумерное течение Пуазейля при разных числах Кнудсена. Данная задача численно решается с использованием нескольких решеточных уравнений Больцмана низкого и высокого порядков, имеющих от 9-ти до 53-х дискретных скоростей. Результаты сравниваются с решениями линеаризованного уравнения Больцмана, Бхатнагара–Гросса–Крука, которые используются в качестве эталонных. В ходе численных экспериментов показано, что увеличение порядка решеточного уравнения Больцмана (т.е. числа первых моментов локально-максвелловского распределения, воспроизводимых дискретным локальным равновесием решеточного уравнения Больцмана) не всегда приводит к повышению точности. В частности, предложена новая модель низкого порядка для 16-ти скоростей, правильно на качественном уровне описывающая диффузное отражение на твердых границах. Показано, что для данной модели можно получить достаточно точные значения объемного расхода скоростей скольжения для широкого интервала чисел Кнудсена по сравнению с другими рассматриваемыми моделями. Библ. 48. Фиг. 2.

Ключевые слова: решеточные уравнения Больцмана, разреженный газ, микротечения.

DOI: 10.31857/S0044466924090168, EDN: WIPGOR

1. ВВЕДЕНИЕ

Метод решеточных уравнений Больцмана (РУБ), являющийся дискретизацией по пространственным и временным переменным кинетического уравнения Бхатнагара–Гросса–Крука (БГК) (см. [1]), изначально разрабатывался для описания уравнений вязкой сплошной среды (см. [2]). Однако основным объектом описания данного метода является функция распределения жидкости или газа, поэтому РУБ определяют динамику не только гидродинамических моментов, но и моментов высокого порядка. Следовательно, естественным образом возникает задача построения РУБ, которые способны правильно описывать неравновесные эффекты, заключенные в негидродинамических моментах функции распределения. Важность задачи применения РУБ к описанию медленных разреженных течений обусловлена рядом потенциально интересных приложений, таких как моделирование микроэлектромеханических систем (МЭМС), например, микро-сопел (см. [3]) или течений сланцевого газа (см. [4]).

Одним из наиболее естественных способов расширения метода РУБ для описания неравновесных разреженных эффектов является повышение порядка решеточных моделей. Далее для наглядности будет удобно рассмотреть уравнение БГК, дискретизированное только по скоростям, но без дискретизации по пространству и времени:

$$\frac{df_i}{dt} = \frac{1}{\tau}(f_i^{\text{eq}} - f), \quad i = 1, \dots, N, \quad (1)$$

где N — число дискретных скоростей, f_i — функция распределения, τ — время релаксации. Отметим, что классические РУБ есть по сути метод решения уравнения (1) с применением удобной дискретизации по пространственным и временным переменным. Состояние локального равновесия f_i^{eq} для РУБ обычно выбирается в виде (см. [5])

$$f_i^{\text{eq}} = w_i \sum_{n=0}^M a_n H^{(n)}(c_i), \quad i = 1, \dots, N, \quad (2)$$

где $M \geq 2$, $w_i > 0$ — веса решетки (аналоги абсолютно максвелловского распределения), c_i — дискретные скорости, $H^{(m)}$ — полиномы Эрмита порядка m , a_m — коэффициенты, определяющие моменты локально-равновесной функции распределения (в общем случае они могут зависеть от пространственных и временной переменных). Отметим, что для дискретного набора скоростей c_i функция распределения (2) положительна для широкого интервала коэффициентов a_m , в частности, для девятискоростной модели $D2Q9$ состояние локального равновесия неотрицательно для чисел Маха меньших единицы (см. [1]). Так как в текущем исследовании рассматриваются медленные течения, то $f_i^{\text{eq}} \geq 0$. Дискретное по скоростям локальное распределение в виде (2) удобно тем, что коэффициенты a_m выражаются аналитически через гидродинамические величины.

Рассмотрим стандартное локально-максвелловское распределение, вычислим его моменты до порядка M включительно, это очевидно ρ , pu и т.д. Пусть удастся построить набор весов и дискретных скоростей w_i , c_i так, что дискретное состояние равновесия (2) имеет совпадающие с локально-максвелловским моменты до порядка M включительно. Рассмотрим гидродинамические уравнения, получающиеся при разложении дискретной по скоростям модели по методу Чепмена—Энскога до членов порядка Kn^k (Kn — число Кнудсена) включительно. Можно показать, что первые моменты до порядка $M - k$, $k \leq M$, включительно, вычисленные из уравнений гидродинамики, которые получены на основе дискретно-скоростных уравнений (1) будут совпадать с решениями уравнений гидродинамики, полученными из разложения Чепмена—Энскога уравнения БГК (см. [5]). Таким образом, дискретная модель (1)–(2) способна описывать моменты функции распределения конечного порядка ровно с такой же точностью, как и уравнение БГК.

Например, для описания вязкой среды, т.е. моментов до второго порядка включительно (плотность, импульс, тензор напряжений) на уровне уравнений Навье—Стокса (Kn^1) необходимо, чтобы f_i^{eq} имела моменты, равные максвелловским до третьего порядка включительно. В случае, если f_i^{eq} описывает правильно максвелловские моменты до второго порядка включительно, то в соответствующих уравнениях Навье—Стокса при моделировании изотермических течений содержатся ошибки порядка $O(Ma^3)$ (Ma — число Маха), однако важно отметить, что для случая медленных течений эти ошибки несутся и такие РУБ (называемые РУБ *низкого порядка*) часто используют в приложениях. Для РУБ *высокого порядка* соответствующее состояние локального равновесия правильно воспроизводит максвелловские моменты порядка 3 и выше. Отметим, что уравнения термо-гидродинамики (Kn^1) и полные уравнения Барнетта (Kn^2) будут воспроизводиться в случае, если f_i^{eq} обладает максвелловскими моментами соответственно до четвертого и пятого порядков включительно.

Для выполнения указанных условий для моментов следует брать коэффициенты a_m , совпадающими с максвелловскими, а также должны выполняться условия

$$\sum_i w_i H^{(m)}(c_i) H^{(m')}(c_i) = \delta_{mm'}, \quad 0 \leq m, m' \leq M, \quad (3)$$

из которых надо определить w_i , c_i . Один из возможных подходов нахождения весов и дискретных скоростей для РУБ основан на квадратурах Гаусса—Эрмита (см. [6]). В данном подходе веса w_i и скорости c_i имеют аналитические выражения. Проблема этого метода в том, что теории многомерных квадратур Гаусса—Эрмита нет, т.е. данным способом удобно строить одномерные РУБ. Модели РУБ для нескольких пространственных переменных получаются как тензорное произведение одномерных моделей, поэтому число дискретных скоростей будет расти квадратично (для двух пространственных переменных) или кубично (для трех пространственных переменных), что не очень желательно. Другой особенностью данного подхода является то, что получаемые дискретные скорости таковы, что отношения между ними в общем случае есть иррациональное число, поэтому стандартный шаг адвекции (из узла в узел) неприменим. В настоящей работе рассматриваются только РУБ, для которых множество дискретных скоростей имеет вид $(n_x c, n_y c)$, где n_x, n_y — целые числа, и c — постоянная величина (модели “на решетке”), соответствующие решетки будем называть *стандартными*. Построение моделей разного порядка на стандартных решетках (которые могут не иметь представления в виде произведения одномерных моделей) основано на решении системы (3), этой задаче посвящено большое число работ (см. [7]–[13]).

Естественно предположить, что повышение порядка РУБ повышает точность расчетов в задачах моделирования разреженного газа. Исследования показывают, что для правильного описания неравновесных эффектов в режимах, близких к свободномолекулярному $Kn \sim 1$, требуются модели очень высокого порядка (см. [14], [15]). Относительно медленная сходимость решений РУБ к эталонным решениям связана с тем, что для сдвиговых течений в областях с твердыми границами функция распределения не является непрерывной по скоростной координате, направленной перпендикулярно твердой стенке (см. [16]). Таким образом, для разреженных течений возникает необходимость (как в численном счете, так и аналитически) рассматривать отдельно две части функции распределения для скоростей, направленных в стенку и от стенки. Однако разложение по полиномам Эрмита плохо аппроксимирует разрывные функции распределения. Таким образом, возникает задача разложения по ортогональной системе полиномов (в полупространстве) двух частей функции распределения

(для скоростей, направленных в стенку и от стенки). Коэффициенты в разложении, стоящие при данных полиномах, очевидно выражаются через полумоменты функции распределения. В этом случае также правильно описываются полумоменты максвелловской функции распределения (т.е. интегралы берутся по полупространству скоростей). Очевидно, что моменты (как суммы полумоментов) также определены правильно (см. [17]–[19]). Данный метод дает очень точные результаты в силу того, что диффузное отражение частиц правильно воспроизводится, но продолжение этих моделей для нескольких пространственных координат требует большого числа дискретных скоростей, а также данные РУБ не относятся к классу моделей “на решетке”. В недавних исследованиях было предложено находить дискретные скорости и веса на основе квадратур Гаусса–Лежандра (и других квадратур) в полупространстве, такие модели способны описывать режимы, близкие к свободномолекулярным для тестовых течений (см. [20], [21]). Выбор квадратур типа Гаусса–Лежандра связан с тем, что в отличие от квадратур Гаусса–Эрмита в полупространстве при таком выборе удастся гораздо лучше аппроксимировать функции Абрамовитца, которые входят в решение задачи Куэтта для уравнения БГК (см. [20], [21]).

Существует другой подход к описанию неравновесных эффектов на основе РУБ (Гауссов подход), в котором дополнительные веса и решеточные скорости используются не для увеличения порядка модели, а для улучшения аппроксимации локального максвелловского распределения дискретно-скоростным распределением (см. [22], [23]). Также возможно применение гибридного подхода, в котором приграничные области моделируются методом Монте-Карло или уравнением БГК, а для внутренних областей используются РУБ (см. [24]–[28]).

В настоящей работе рассмотрено применение стандартных (“на решетке”) двумерных решеточных уравнений Больцмана разного порядка в приложениях к моделированию течения Пуазейля для разных чисел Кнудсена. Численные эксперименты, направленные на оценку возможности моделирования разреженных течений на основе нерегуляризованных РУБ разного порядка с одним временем релаксации, проводились ранее для разных эталонных течений (см. [29]–[32], [14], [15]). В настоящей работе эти исследования продолжены. Показано, что повышение порядка и числа скоростей РУБ не ведет автоматически к повышению точности, на результаты расчетов большее влияние оказывает правильность воспроизведения условий диффузного отражения. Рассмотрены РУБ для 9-и скоростей, новая модель низкого порядка для 16-и скоростей, две 17-и скоростные модели высокого порядка и модель для 53-х скоростей высокого порядка. Показано, что модели одного порядка с одинаковым числом скоростей (17 скоростей) дают существенно разные результаты, а новая 16-и скоростная модель низкого порядка показывает лучшую точность для переходного режима и режима скольжения в сравнении с остальными моделями. Это связано с тем, что все рассмотренные модели имеют сильно отличающиеся ошибки (формула (11)) для первого полумомента максвелловского распределения. Например, две 17-и скоростные модели имеют ошибки, равные 0.25 и 0.03, точность численных расчетов (табл. 2, фиг. 1, фиг. 2.) коррелирует с данной ошибкой (чем она больше – тем больше отклонение от эталонных значений.)

2. РЕШЕТОЧНЫЕ УРАВНЕНИЯ БОЛЬЦМАНА НА СТАНДАРТНЫХ РЕШЕТКАХ

В настоящей работе рассмотрены стандартные модели (“на решетке”), т.е. набор скоростей для таких моделей имеет вид $\mathbf{c}_i \equiv (n_{i,x}c, n_{i,y}c)$, где $n_{i,x}, n_{i,y}$ – целые числа и $c > 0$ – постоянная величина. Данные РУБ являются дискретизацией уравнения (1) по пространству и времени, решаемые в два шага (столкновение и адвекция)

$$f_i^*(t, \mathbf{x}) = f_i(t, \mathbf{x}) + \frac{\Delta t}{\tau + \frac{\Delta t}{2}} (f_i^{\text{eq}} - f_i)(t, \mathbf{x}), \quad i = 1, \dots, N, \quad (4)$$

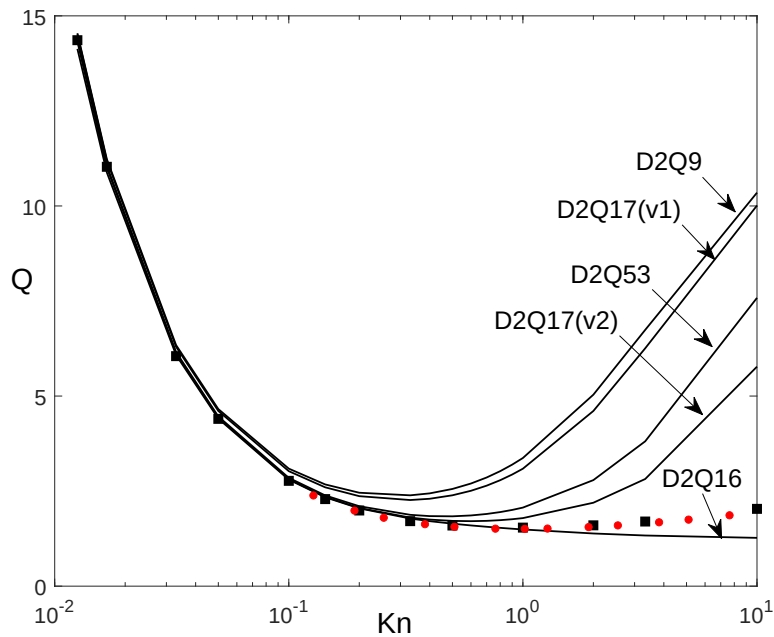
$$f_i(t + \Delta t, \mathbf{x} + \mathbf{c}_i \Delta t) = f_i^*(t, \mathbf{x}), \quad i = 1, \dots, N, \quad (5)$$

где f^* – промежуточная функция распределения после столкновения. В работе $\Delta t = 1$, следовательно, расстояние между узлами решетки есть $\Delta x = c\Delta t = c$, τ – время релаксации. Функция распределения локального равновесия имеет вид

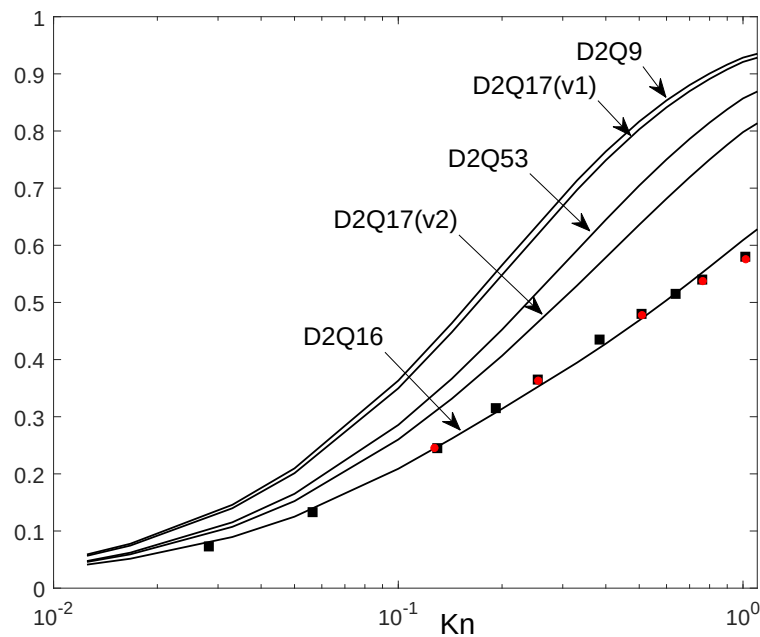
$$f_i^{\text{eq}} = w_i \sum_{m=0}^M \mathbf{a}^{(m)} \mathbf{H}^{(m)}(\mathbf{c}_i). \quad (6)$$

Здесь

$$\begin{aligned} a^{(0)} &= \rho, \quad a_x^{(1)} = \rho u_x, \quad a_y^{(1)} = \rho u_y, \\ a_{xx}^{(2)} &= \frac{1}{2}\rho(\Theta + u_x^2), \quad a_{yy}^{(2)} = \frac{1}{2}\rho(\Theta + u_y^2), \quad a_{xy}^{(2)} = \rho u_x u_y, \\ a_{xxx}^{(3)} &= \frac{1}{6}\rho(u_x^3 + 3u_x\Theta), \quad a_{xxy}^{(3)} = \frac{1}{2}\rho(\Theta + u_x^2)u_y, \quad a_{xyy}^{(3)} = \frac{1}{2}\rho(\Theta + u_y^2)u_x, \quad a_{yyy}^{(3)} = \frac{1}{6}\rho(u_y^3 + 3u_y\Theta), \end{aligned}$$



Фиг. 1. Двумерное течение Пуазейля. Представлены значения объемного расхода Q (формула (10)) для разных чисел Кнудсена). Красные кружки обозначают решение линейаризованного уравнения Больцмана для твердых сфер (см. [39]), черные квадраты соответствуют решению линейаризованного уравнения БГК (см. [38]).



Фиг. 2. Двумерное течение Пуазейля. Представлены значения нормализованных скоростей скольжения на стенках для разных чисел Кнудсена. Красные кружки обозначают решение линейаризованного уравнения Больцмана для твердых сфер (см. [39]), черные квадраты соответствуют решению на основе метода Монте-Карло (см. [12]).

$$\begin{aligned}
a_{xxxx}^{(4)} &= \frac{1}{24}\rho(u_x^4 + 6u_x^2\Theta + 3\Theta^2), \quad a_{yyyy}^{(4)} = \frac{1}{24}\rho(u_y^4 + 6u_y^2\Theta + 3\Theta^2), \\
a_{xxyy}^{(4)} &= \frac{1}{6}\rho(u_x^3 + 3u_x\Theta)u_y, \quad a_{xyyy}^{(4)} = \frac{1}{6}\rho(u_y^3 + 3u_y\Theta)u_x, \quad a_{xxyy}^{(4)} = \frac{1}{4}\rho(\Theta + u_x^2)(\Theta + u_y^2), \\
a_{xxxxx}^{(4)} &= \frac{1}{120}\rho(u_x^5 + 15u_x\Theta^2 + 10u_x^3\Theta), \quad a_{yyyyy}^{(4)} = \frac{1}{120}\rho(u_y^5 + 15u_y\Theta^2 + 10u_y^3\Theta), \\
a_{xxxxy}^{(4)} &= \frac{1}{24}\rho(u_x^4 + 6u_x^2\Theta + 3\Theta^2)u_y, \quad a_{yyyxy}^{(4)} = \frac{1}{24}\rho(u_y^4 + 6u_y^2\Theta + 3\Theta^2)u_x, \\
a_{xxxxy}^{(4)} &= \frac{1}{12}\rho(u_x^3 + 3u_x\Theta)(\Theta + u_y^2), \quad a_{xyyyx}^{(4)} = \frac{1}{12}\rho(u_y^3 + 3u_y\Theta)(\Theta + u_x^2),
\end{aligned}$$

где ρ, u_x, u_y — плотность и скорости потока, также $\Theta \equiv T - 1$ есть отклонение температуры от температуры абсолютно максвелловского распределения (которая равна единице для всех рассматриваемых моделей), T — температура газа.

Таблица 1. Приведены веса w_i для дискретных скоростей c_i , а также скорости c , рассматриваемых в работе решеточных моделей уравнения Больцмана: $D2Q9$ [1], $D2Q16$ [35], два варианта (обозначаемые $v1$ и $v2$) моделей $D2Q17$ [11], $D2Q53$ [8], N_i — число дискретных скоростей в соответствующей скоростной группе. Для модели $D2Q16$ веса равны: $w_1 = \frac{79-\sqrt{97}}{180}, w_2 = 1/2 - w_1 = \frac{11+\sqrt{97}}{180}$. Для краткости для модели $D2Q53$ значения w_i, c приведены с тремя знаками после запятой, значения соответствующих весов с машинной точностью могут быть найдены в [8]

c_i	N_i	$D2Q9$	$D2Q16$	$D2Q17(v1)$	$D2Q17(v2)$	$D2Q53$
0	1	$\frac{16}{36}$	—	$\frac{190-8\sqrt{10}}{405}$	$\frac{35}{288}$	2.053×10^{-1}
$(\pm c, 0)$	4	$\frac{4}{36}$	—	$\frac{12\sqrt{10}-15}{200}$	—	1.048×10^{-1}
$(\pm c, \pm c)$	4	$\frac{1}{36}$	w_1^2	$\frac{150-39\sqrt{10}}{800}$	$\frac{45}{256}$	5.869×10^{-2}
$(\pm 2c, 0)$	4	—	—	—	—	1.679×10^{-2}
$(\pm 2c, \pm c)$	8	—	—	—	—	7.657×10^{-3}
$(\pm 2c, \pm 2c)$	4	—	—	—	$\frac{9}{640}$	1.549×10^{-3}
$(\pm 3c, 0)$	4	—	—	$\frac{295-92\sqrt{10}}{16200}$	$\frac{1}{36}$	5.25×10^{-4}
$(\pm 3c, \pm c)$	8	—	—	—	—	4.779×10^{-4}
$(\pm 3c, \pm 3c)$	4	—	—	$\frac{130-41\sqrt{10}}{64800}$	$\frac{23}{11520}$	1.196×10^{-5}
$(\pm 4c, \pm c)$	8	—	$w_1 w_2$	—	—	6.074×10^{-6}
$(\pm 4c, \pm 4c)$	4	—	w_2^2	—	—	—
$(\pm 5c, 0)$	4	—	—	—	—	3.306×10^{-7}
c		$\sqrt{3}$	$\sqrt{\frac{1}{(16-30w_1)}}$	$\sqrt{\frac{5+\sqrt{10}}{3}}$	$\sqrt{\frac{2}{3}}$	1.128

Скорости c_i, c и веса w_i моделей приведены в табл. 1. РУБ $D2Q9, D2Q16$ являются моделями низкого порядка, для них берется $M = 2$ в формуле (6). Для моделей $D2Q17v1, D2Q17v2$ берется $M = 3, \Theta = 0$; для $D2Q53$ имеем $M = 5$, и Θ может быть отлично от нуля. Подчеркнем, что $D2Q53$ может описывать течения на уровне уравнений Барнетта.

Рассмотрим подробнее модель $D2Q16$. Используемое в работе РУБ $D2Q16$ имеет низкий порядок $M = 2$ и относится к классу моделей “на решетке”, но отметим, что для 16-ти скоростей можно получить модель высокого порядка ($M = 3$), однако в этом случае решетка не является стандартной и имеет скорости $(\sqrt{3} - \sqrt{6}c, \sqrt{3} + \sqrt{6}c)$ (отношение скоростей равно 0.318) (см. [33], [34], [9]). Данная модель выводится из следующих соображений (см. также [35]). Для того чтобы моменты f_i^{eq} совпадали с максвелловскими моментами до второго порядка включительно необходимо, чтобы выполнялись равенства

$$w_1 + w_2 = \frac{1}{2}, \quad w_1 c_1^2 + w_2 c_2^2 = \frac{1}{2}, \quad w_1 c_1^4 + w_2 c_2^4 = \frac{3}{2}, \quad (7)$$

данные соотношения получаются после умножения функции f_i^{eq} , $M = 2$ на 1, c_i , c_i^2 , суммирования по i и последующего приравнивания результата к выражениям для максвелловских моментов. Веса w_i и скорость c вычисляются из уравнений (7) при условии $c_1 = c$, $c_2 = 4c$. В итоге получаем $w_1 = \frac{79-\sqrt{97}}{180} \approx 0.3842$, $w_2 = 1/2 - w_1 = \frac{11+\sqrt{97}}{180} \approx 0.1158$ и $c = \sqrt{\frac{1}{(16-30w_1)}} \approx 0.4727$.

3. ЧИСЛЕННОЕ РЕШЕНИЕ ЗАДАЧИ ПУАЗЕЙЛЯ ДЛЯ РАЗНЫХ ЧИСЕЛ КНУДСЕНА

В качестве тестовой задачи рассмотрим двумерное течение Пуазейля газа между параллельными твердыми стенками под действием внешней силы (направленной вдоль стенок) для разных чисел Кнудсена.

Считаем, что ось y параллельна стенкам канала, переменная x — перпендикулярна стенкам. Так как рассматривается медленное течение, то амплитуда внешней силы мала и в ходе численных расчетов соответствующие члены в РУБ можно брать в линеаризованной форме (см. [1])

$$F_i = w_i c_{i,y} F, \quad (8)$$

где F — безразмерная амплитуда внешней силы, в численных расчетах она берется равной 10^{-6} , максимальное число Маха составляет 10^{-4} .

На открытых концах применяются периодические граничные условия. На твердых стенках задаются кинетические граничные условия (диффузное отражение газа) (см. [36], [37]). Количество пространственных узлов между стенками равно 100. Кинетические граничные условия реализуются следующим образом. Считается, что стенка расположена на расстоянии $0.5c$ от ближайшего узла решетки, в котором моделируется движение газа (приграничный узел). Если для вычисления функции распределения РУБ f_i , соответствующей скорости c_i , необходимо выполнить шаг адвекции из узла, расположенного за стенкой (вне моделируемой области), то используются кинетические граничные условия

$$f_i(t, x, y) = \frac{\sum_{c_j n < -L} |c_j n| f_j(t, x, y)}{\sum_{c_j n > L} |c_j n| w_j} w_i, \quad (9)$$

где L — расстояние между приграничным узлом (x, y) и стенкой, n — нормаль к стенке, направленная внутрь моделируемой области.

Скорости скольжения для течения определяются как $u(0)/u_{\text{max}}$, где $u(0)$ — скорость газа у стенки, u_{max} — максимальная величина продольной скорости (в центре моделируемого канала).

Объемный расход определяется следующим образом (см. [29]):

$$Q = \frac{1}{4U_0 H K n} \int_0^H u(s) ds, \quad (10)$$

где $U_0 = FH^2/(8\rho\nu)$ — средняя скорость для решения уравнения Навье—Стокса с граничными условиями прилипания, H — расстояние между стенками, ρ — плотность газа, ν — вязкость, равная τ для всех моделей, u — продольная скорость.

Для сравнения точности РУБ в качестве эталонных значений объемного расхода используются данные, полученные Черчиньяни с соавторами для линеаризованного уравнения БГК (см. [38]) и Овады с соавторами для линеаризованного уравнения Больцмана (для твердых сфер); для скоростей скольжения используются результаты из статьи Овады с соавторами (см. [39]) и решения на основе метода Монте-Карло (см. [12]).

Отметим, что все рассматриваемые РУБ, кроме $D2Q16$, имеют дискретные скорости, направленные параллельно стенкам. Поэтому при увеличении числа Кнудсена для таких моделей будут наблюдаться эффекты “убегания”, (см. [40], [41]), т.е. частицы, движущиеся параллельно стенкам, практически не будут испытывать столкновений. На практике это приводит к завышению объемного расхода Q . Также большое влияние на результаты расчетов оказывает точность, с которой рассматриваемые модели описывают диффузное отражение. В первую очередь, интерес представляет первый полумомент максвелловского распределения, т.е.

$$m_x^+ = \int_0^\infty \frac{1}{2\pi} \exp\left(-\frac{c_x^2 + c_y^2}{2}\right) c_x dc_x dc_y = \frac{1}{\sqrt{2\pi}}.$$

Для рассматриваемых в настоящей работе РУБ относительные ошибки первого полумомента существенно различаются. Определим относительную ошибку формулой

$$err \equiv \frac{|\sum_{i: c_{i,x} > 0} w_i c_{i,x} - \frac{1}{\sqrt{2\pi}}|}{\frac{1}{\sqrt{2\pi}}}. \quad (11)$$

Таблица 2. Приведены значения объемного расхода для течения Пуазейля при разных числах Кнудсена. *BGK* обозначает решения линеаризованного уравнения Больцмана [38]

Kn	<i>BGK</i>	<i>D2Q9</i>	<i>D2Q16</i>	<i>D2Q17(v1)</i>	<i>D2Q17(v2)</i>	<i>D2Q53</i>
0.0125	14.36	14.54	14.13	14.51	14.33	14.38
0.0167	11.03	11.23	10.88	11.19	11.01	11.06
0.033	6.05	6.35	6.08	6.30	6.13	6.16
0.05	4.40	4.66	4.40	4.60	4.43	4.46
0.1	2.77	3.09	2.80	3.02	2.82	2.85
0.143	2.29	2.68	2.34	2.60	2.35	2.39
0.2	1.99	2.46	2.05	2.37	2.06	2.11
0.33	1.71	2.39	1.78	2.27	1.80	1.88
0.4	—	2.44	1.71	2.30	1.75	1.84
0.5	1.595	2.56	1.64	2.39	1.72	1.84
0.6	—	2.70	1.60	2.51	1.71	1.86
0.7	—	2.86	1.56	2.64	1.72	1.90
0.8	—	3.02	1.53	2.79	1.74	1.95
0.9	—	3.20	1.51	2.94	1.76	2.01
1	1.54	3.37	1.49	3.09	1.79	2.07
2	1.60	5.03	1.39	4.61	2.19	2.79
3.33	1.70	6.74	1.33	6.26	2.82	3.81
10	2.03	10.35	1.27	10.01	5.77	7.59

Наименьшую ошибку (11), равную 0.004, имеет модель *D2Q16*. Для решеточных моделей *D2Q9*, *D2Q17(v1)*, *D2Q53*, *D2Q17(v2)* ошибки *err* равны 0.28, 0.25, 0.11, 0.03 соответственно.

Полученные значения объемного расхода представлены в табл. 2 и на фиг. 1. Решеточные уравнения Больцмана *D2Q53*, *D2Q17(v2)* при $Kn < 0.05$ имеют примерно одинаковую точность, модель *D2Q16* дает значения немного более отличающиеся от эталонного. При $0.05 \leq Kn \leq 1$ модель *D2Q16* воспроизводит значения объемного расхода, которые отличаются от решения линеаризованного уравнения БГК всего на несколько процентов. Это связано с тем, что остальные РУБ менее точно, чем *D2Q16*, описывают условия диффузного отражения, а также подвержены эффектам “убегания”. Если исключить из рассмотрения *D2Q16*, то наименее подвержена этому эффекту модель *D2Q17(v2)*, которая наиболее правильно предсказывает значения *Q*. В свободномолекулярном режиме точность всех моделей падает, *D2Q16* не воспроизводит кнудсеновский минимум, остальные РУБ завышают объемный расход. Аналогичным образом модель *D2Q16* дает наиболее правильные скорости скольжения в режиме скольжения и переходном режиме, фиг. 2.

4. ЗАКЛЮЧЕНИЕ

В работе рассматривается задача описания медленных изотермических разреженных течений на основе решеточных уравнений Больцмана разного порядка на стандартных решетках. Исследуются РУБ с количеством дискретных скоростей от 9-ти до 53-х. Основной интерес настоящего исследования представляет вопрос влияния структуры дискретных скоростей на точность моделирования разреженного газа, поэтому включение дополнительных времен релаксации, модификация времени релаксации в зависимости от числа Кнудсена и расстояния до стенки, дополнительная регуляризация РУБ в работе не рассматриваются. В качестве тестовой задачи используется медленное двумерное течение Пуазейля для разных чисел Кнудсена.

Для течения Пуазейля показано, что увеличение порядка для РУБ и числа дискретных скоростей не обязательно ведет к увеличению точности. Ранее аналогичный результат был подтвержден для течения Куэтта

(см. [31]). На результаты расчетов основное влияние оказывает точность воспроизведения диффузного отражения, а также появление эффектов “убегания” (при наличии у модели скоростей, параллельных стенкам, частицы с этими скоростями почти не испытывают столкновений). Условия диффузного отражения можно качественно описать полумоментами максвелловского распределения. Результаты численного счета показывают, что, чем хуже модель воспроизводит первый максвелловский полумомент, тем большие отличия она дает в сравнении с эталонными решениями.

Например, модель $D2Q53$ описывает гидродинамику на уровне уравнений Барнетта, но имеет худшую точность, чем модели $D2Q16$, $D2Q17(v2)$, в силу того что последние две модели лучше воспроизводят первый полумомент распределения Максвелла (условия диффузного отражения). Среди двух моделей одинакового порядка и одинаковым числом скоростей ($D2Q17(v1)$, $D2Q17(v2)$) существенно более близкие к эталонным решения имеет модель с меньшей относительной ошибкой первого полумомента распределения Максвелла ($D2Q17(v2)$).

Построена новая модель для 16-ти дискретных скоростей ($D2Q16$), отличающаяся от предложенной ранее 16-ти скоростной модели (см. [29]) тем, что новая модель имеет низкий порядок, но относится к классу моделей “на решетке”. Несмотря на понижение порядка, для РУБ $D2Q16$ значения объемного расхода и скоростей скольжения на стенках при $Kn \leq 1$ имеют отклонения от эталонных (уравнения Больцмана и БГК, метод Монте-Карло) всего на несколько процентов. Преимущество данной модели относительно остальных во многом обусловлено постановкой задачи: рассмотрено медленное течение, нет эффектов переноса тепла, нет дискретных скоростей, которые параллельны стенкам. Отметим, что недостатком модели $D2Q16$ является то, что она не может полноценно описать пограничный слой и не предсказывает кнудсеновский минимум. Также, в силу того, что новая $D2Q16$ есть РУБ низкого порядка, она вносит кубические ошибки по числу Маха в уравнения гидродинамики (которые несущественны для медленных течений), тогда как остальные рассмотренные модели лишены этого недостатка. Кроме того, в отличие от всех остальных рассматриваемых моделей решеточное уравнение $D2Q53$ может описывать процессы переноса тепла. Отдельно стоит отметить, что на результаты моделирования течений с применением решеточных уравнений Больцмана может оказывать влияние появление ложных инвариантов (см. [42]). В теории дискретных уравнений Больцмана задаче построения моделей с правильным числом инвариантов посвящено много исследований (см. [43]–[46]). В численных расчетах на основе РУБ эффекты, связанные с лишними инвариантами наблюдаются при симуляции многофазных течений (см. [47]).

Для течений разреженного газа в сложных геометриях углы между дискретными скоростями и стенкой непостоянны, таким образом, точность расчетов для течений в пористых средах может отличаться от результатов для течения Пуазейля. В частности, также представляет интерес задача моделирования течения Пуазейля для разных чисел Кнудсена решеточными моделями Больцмана при разных углах поворота решетки. Также важной является задача оценки точности моделей для течений в микроканалах с шероховатостями и течений через пористые среды с разной геометрий пор (см. [48]).

СПИСОК ЛИТЕРАТУРЫ

1. Krüger T., Kusumaatmaja H., Kuzmin A., Shardt O., Silva G., Vigen E. The Lattice Boltzmann Method. Principles and Practice. Springer, 2017.
2. McNamara G., Zanetti G. Use of the Boltzmann equation to simulate lattice gas automata // Phys. Rev. Lett. 1988. V. 61. P. 2332–2335.
3. Karniadakis G., Beskok A., Aluru N. Microflows and Nanoflows. Fundamentals and Simulation. Springer, 2005.
4. Wang J., Chen L., Kang Q., Rahman S. The lattice Boltzmann method for isothermal micro-gaseous flow and its application in shale gas flow: A review // Int. J. Heat Mass Transf. 2016. V. 95. P. 94–108.
5. Shan X., Yuan X., Chen H. Kinetic theory representation of hydrodynamics: a way beyond the Navier–Stokes equation // J. Fluid Mech. 2006. V. 550. P. 413–441.
6. Shan X., He X. Discretization of the Velocity Space in the Solution of the Boltzmann Equation // Phys. Rev. Lett. 1998. V. 80. P. 65–67.
7. Philippi P., Hegele Jr. L., dos Santos L., Surmas R. From the continuous to the lattice Boltzmann equation: The discretization problem and thermal models // Phys. Rev. E. 2006. V. 73. P. 056702.
8. Surmas R., Ortiz C., Philippi P. Simulating thermohydrodynamics by finite difference solutions of the Boltzmann equation // Eur. Phys. J. 2009. V. 90. P. 81–90.

9. *Chikatamarla S., Karlin I.* Lattices for the lattice Boltzmann method // *Phys. Rev. E.* 2009. V. 79. P. 046701.
10. *Shan X.* General solution of lattices for Cartesian lattice Bhatnagar–Gross–Krook models // *Phys. Rev. E.* 2010. V. 81. P. 036702.
11. *Shan X.* The mathematical structure of the lattices of the lattice Boltzmann method // *J. Comput. Sci.* V. 17. P. 475–481.
12. *Feuchter C., Schleifenbaum W.* High-order lattice Boltzmann models for wall-bounded flows at finite Knudsen numbers // *Phys. Rev. E.* 2016. V. 94. P. 013304.
13. *Spiller D., Dünweg B.* Semiautomatic construction of lattice Boltzmann models // *Phys. Rev. E.* 2020. V. 101. P. 043310.
14. *Meng J.-P., Zhang Y.* Gauss–Hermite quadratures and accuracy of lattice Boltzmann models for non-equilibrium gas flows // *Phys. Rev. E.* 2011. V. 83. P. 036704.
15. *Shi Y., Wu L., Shan X.* Accuracy of high-order lattice Boltzmann method for non-equilibrium gas flow // *J. Fluid Mech.* 2020. V. 907. P. A25.
16. *Gross E., Ziering S.* Kinetic theory of Linear Shear flow // *Phys. Fluid.* 1958. V. 1. P. 215–223.
17. *Ambrus V., Sofonea V.* Implementation of diffuse-reflection boundary conditions using lattice Boltzmann models based on half-space Gauss–Laguerre quadratures // *Phys. Rev. E.* 2014. V. 89. P. 041301(R).
18. *Ambrus V., Sofonea V.* Application of mixed quadrature lattice Boltzmann models for the simulation of Poiseuille flow at non-negligible values of the Knudsen number // *J. Comput. Sci.* 2016. V. 17. P. 403.
19. *Ambrus V., Sofonea V.* Lattice Boltzmann models based on half-range Gauss–Hermite quadratures // *J. Comput. Phys.* 2016. V. 316. P. 760.
20. *Shi Y.* Velocity discretization for lattice Boltzmann method for noncontinuum bounded gas flows at the micro- and nanoscale // *Phys. Fluid.* 2022. V. 34. P. 082013.
21. *Shi Y.* Comparison of different Gaussian quadrature rules for lattice Boltzmann simulations of noncontinuum Couette flows: From the slip to free molecular flow regimes // *Phys. Fluid.* 2023. V. 35. P. 072015.
22. *Ilyin O.* Intermediate Lattice Boltzmann–BGK method and its application to micro-flows // *J. Phys.: Conf. Ser.* 2019. V. 1163. P. 012030.
23. *Ilyin O.* Gaussian Lattice Boltzmann method and its applications to rarefied flows // *Phys. Fluid.* 2020. V. 32. P. 012007.
24. *Di Staso G., Clercx H., Succi S., Toschi F.* Lattice Boltzmann accelerated direct simulation Monte Carlo for dilute gas flow simulations // *Phil. Trans. R. Soc. A.* 2016. V. 374. P. 20160226.
25. *Di Staso G., Clercx H., Succi S., Toschi F.* DSMC–LBM mapping scheme for rarefied and non-rarefied gas flows // *J. Comput. Sci.* 2016. V. 17. P. 357–369.
26. *Di Staso G., Srivastava S., Arlemark E., Clercx H., Toschi F.* Hybrid lattice Boltzmann–direct simulation Monte Carlo approach for flows in three-dimensional geometries // *Comput. Fluid.* 2018. V. 172. P. 492–509.
27. *Aristov V., Ilyin O., Rogozin O.* A hybrid numerical scheme based on coupling discrete-velocities models for the BGK and LBGK equations // *AIP Conf. Proc.* 2019. V. 2132. P. 060007.
28. *Aristov V., Ilyin O., Rogozin O.* Kinetic multiscale scheme based on the discrete-velocity and lattice-Boltzmann methods // *J. Comput. Sci.* 2020. V. 40. P. 101064.
29. *Kim S., Pitsch H., Boyd I.* Accuracy of higher-order lattice Boltzmann methods for microscale flows with finite Knudsen numbers // *J. Comput. Phys.* 2008. V. 227. P. 8655–8671.
30. *Tang G., Zhang Y., Emerson D.* Lattice Boltzmann models for nonequilibrium gas flows // *Phys. Rev. E.* 2008. V. 77. P. 046701.

31. *Shi Y., Brookes P., Yap Y., Sader J.* Accuracy of the lattice Boltzmann method for low-speed noncontinuum flows // *Phys. Rev. E.* 2011. V. 83. P. 045701(R).
32. *de Izarra L., Rouet J.-L., Izrar B.* High-order lattice Boltzmann models for gas flow for a wide range of Knudsen numbers // *Phys. Rev. E.* 2011. V. 84. P. 066705.
33. *Chikatamarla S., Karlin I.* Entropy and Galilean invariance of lattice Boltzmann theories // *Phys. Rev. Lett.* 2006. V. 97. P. 190601.
34. *Bardow A., Karlin I., Gusev A.* Multispeed models in off-lattice Boltzmann simulations // *Phys. Rev. E.* 2008. V. 77. P. 025701(R).
35. *Ilyin O.* Nonclassical Heat Transfer in a Microchannel and a Problem for Lattice Boltzmann Equations // *Comp. Math. Math. Phys.* 2023. V. 63. P. 2297–2305.
36. *Ansumali S., Karlin I.* Kinetic boundary conditions in the lattice Boltzmann method // *Phys. Rev. E.* 2002. V. 66. P. 026311.
37. *Meng J., Zhang Y.* Diffuse reflection boundary condition for high-order lattice Boltzmann models with streaming-collision mechanism // *J. Comput. Phys.* 2014. V. 258. P. 601–612.
38. *Cercignani C., Lampis M., Lorenzani S.* Variational approach to gas flows in microchannels // *Phys. Fluid.* 2004. V. 16. P. 3426–3437.
39. *Ohwada T., Sone Y., Aoki K.* Numerical analysis of the Poiseuille and thermal transpiration flows between two parallel plates on the basis of the Boltzmann equation for hard-sphere molecules // *Phys. Fluid.* 1989. V. 1. P. 2042–2049.
40. *Toschi F., Succi S.* Lattice Boltzmann method at finite Knudsen numbers // *Europhys. Lett.* 2005. V. 69. P. 549.
41. *Montessori A., Prestininzi P., La Rocca M., Succi S.* Lattice Boltzmann approach for complex nonequilibrium flows // *Phys. Rev. E.* 2015. V. 92. P. 043308.
42. *Kandhai D., Koponen A., Hoekstra A., Kataja M., Timonen J., Sloot P.M.A.* Implementation Aspects of 3D Lattice-BGK: Boundaries, Accuracy, and a New Fast Relaxation Method // *J. Comput. Phys.* 1999. V. 150. P. 482–501.
43. *Bobylev A., Cercignani C.* Discrete Velocity Models Without Nonphysical Invariants // *J. Stat. Phys.* 1999. V. 97. P. 677–686.
44. *Веденяпин В., Орлов Ю.* О законах сохранения для полиномиальных гамильтонианов и для дискретных моделей уравнения Больцмана // *ТМФ.* 1999. Т. 121. С. 307–315.
45. *Vedenyapin V.* Velocity inductive construction for mixtures // *Transp. Theor. Stat. Phys.* 1999. V. 28. P. 727–742.
46. *Bobylev A., Vinerean M.* Construction of Discrete Kinetic Models with Given Invariants // *J. Stat. Phys.* 2008. V. 132. P. 153–170.
47. *Guo Z., Shi B., Zheng C.* Chequerboard effects on spurious currents in the lattice Boltzmann equation for two-phase flows // *Phil. Trans. R. Soc. A.* 2011. V. 369. P. 2283–2291.
48. *Su W., Lindsay S., Liu H., Wu L.* Comparative study of the discrete velocity and lattice Boltzmann methods for rarefied gas flows through irregular channels // *Phys. Rev. E.* 2017. V. 96. P. 023309.

ON THE ACCURACY OF LOW- AND HIGH-ORDER LATTICE BOLTZMANN EQUATIONS IN APPLICATIONS TO SLOW ISOTHERMAL MICROFLOWS

O. V. Ilyin*

119333 Moscow, Vavilov Str., 44, Federal Research Center Computer Science and Control, RAS, Russia

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

Received: 12.12.2023

Revised: 14.04.2024

Accepted: 31.05.2024

Abstract. The question of the applicability of two-dimensional lattice Boltzmann equations of different orders in standard lattices to the description of slow isothermal sparse flows is considered. The two-dimensional Poiseuille flow at different Knudsen numbers is used as a reference problem. This problem is numerically solved using several lattice Boltzmann equations of low and high orders having from 9 to 53 discrete velocities. The results are compared with solutions of the linearized Boltzmann, Bhatnagar-Gross-Krook equations, which are used as reference ones. Numerical experiments have shown that an increase in the order of the Boltzmann lattice equation (i.e., the number of first moments of the local-Maxwell distribution reproduced by the discrete local equilibrium of the Boltzmann lattice equation) does not always lead to an increase in accuracy. In particular, a new low-order model for 16 velocities is proposed, which correctly describes the diffuse reflection at solid boundaries at a qualitative level. It is shown that for this model, it is possible to obtain sufficiently accurate values of the volumetric flow rate of sliding velocities for a wide range of Knudsen numbers in comparison with other models under consideration.

Keywords: lattice Boltzmann equations, rarefied gas, microflows.

ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ ТЕЧЕНИЯ ПОЛИМЕРНОГО РАСТВОРА ДЛЯ КОЛМОГОРОВСКОГО ПОТОКА¹⁾

© 2024 г. В.В. Денисенко^{1,*}, С.В. Фортова¹, В.В. Лебедев², И.В. Колоколов²

¹ 123056 Москва, ул. 2-Я Брестская, 19/18, Институт автоматизации проектирования РАН, Россия

² 142432 Черноголовка, М.о., пр-т Акад. Семенова, 1А, Институт теоретической физики им. Л.Д. Ландау РАН, Россия

*e-mail: ned13@rambler.ru

Поступила в редакцию 22.02.2024 г.

Переработанный вариант 22.02.2024 г.

Принята к публикации 31.05.2024 г.

Предложен численный метод, аппроксимирующий уравнения динамики течения полимерного раствора. Построенная методика основывается на гибридном подходе. Гидродинамическая составляющая течения описывается системой уравнений Навье–Стокса и численно аппроксимируются линейаризованным методом Годунова. Полимерная составляющая течения описывается системой уравнений для вектора растяжения полимеров $\mathbf{R}$ и численно аппроксимируются методом Курганова–Тедмора. С использованием построенной схемы, исследована задача об устойчивости течения полимерного раствора при низких числах Рейнольдса $Re \sim 10$ в квадратной периодической ячейке под действием внешней периодической силы. Путем численного эксперимента изучена неустойчивость данного вида течения, характеризующаяся нарушением его ламинарности. Построены спектральные характеристики полимерного раствора при низких числах Рейнольдса. Библ. 38. Фиг. 6.

Ключевые слова: численное моделирование, эластическая турбулентность, гидродинамическая неустойчивость, течение Колмогорова.

DOI: 10.31857/S0044466924090179, EDN: WICNKA

ВВЕДЕНИЕ

Взаимодействие полимерных молекул с потоками жидкости приводит ко множеству интересных явлений, наблюдаемых в природе и лабораторных экспериментах, а именно к уменьшению турбулентного сопротивления, эластической турбулентности и изменению процесса теплопереноса при естественной конвекции [1, 2]. Лабораторные эксперименты показывают, что даже незначительная концентрация полимеров может значительно изменить свойства ламинарных низкорейнольдсовых ($Re \sim 1$) потоков, и породить новую форму турбулентности, получившую название эластической(полимерной) турбулентности. Для ньютоновского потока жидкости переход к турбулентности определяется нелинейным инерционным членом в уравнении импульса, точнее отношением инерционных сил, описываемых этим членом к диссипативным силам, обусловленным вязкостью среды. Это отношение характеризуется безразмерным параметром — числом Рейнольдса Re . Чем выше Re , тем более неустойчивым становится поток, что приводит к возникновению и развитию турбулентного режима течения. Однако в жидкости, в которой присутствуют полимерные молекулы, возникает так называемый вязкоупругий эффект, вследствие наличия другого нелинейного члена, характеризующего влияние полимерных молекул на поток. Влияние упругих сил на поток, которые вызваны наличием полимеров, описывается еще одним безразмерным параметром — числом Вайсенберга $Wi = \frac{U}{\gamma_0 L}$, определяемым отношением характерного градиента скорости потока жидкости к темпу релаксации полимерной молекулы к своему равновесному состоянию. Здесь γ_0 — коэффициент релаксации полимерной молекулы к своему равновесному состоянию. Как показывают физические эксперименты, возникновение и развитие неустойчивости течения при малых числах Re происходит именно благодаря наличию этих упругих сил [1–4].

Развитие неустойчивостей, обусловленных действием только полимерных сил при низких числах Рейнольдса наблюдается в большом разнообразии экспериментов для криволинейных геометрий потоков. Примерами

¹⁾ Работа выполнена при финансовой поддержке Минобрнауки И.В. Колоколова (госзадание FFWR-2024-0014) и РНФ В.В. Лебедева (код проекта 23-72-30006) в рамках госзадания №124022400174-3 “Вычислительный эксперимент на суперкомпьютерах в задачах механики сплошных сред”.

служат течения Куэтта между двумя цилиндрами, где был предложен механизм развития полимерной неустойчивости [5] и количественно подтвержден экспериментально [6–8], а так же закрученный поток между двумя вращающимися параллельными дисками [9–11], коническо-плоские геометрии [12], криволинейные каналы или поток Дина [13–18], поток в канале [19–22] и неустойчивости следа в потоках, обтекающих различные препятствия [23–26]. Так же результаты по эластической турбулентности были получены в работах [27, 28].

Несмотря на большой интерес, проявляемый к полимерным течениям и широкий спектр потенциальных применений в различных технологических процессах, численное моделирование упругой неустойчивости и эластической турбулентности встречает большие сложности и мало представлена в литературе. При постановке физических экспериментов в настоящее время довольно трудно получить подробную информацию о полимерном растворе, например, поле тензора упругих напряжений полимерных молекул. При этом эти данные имеют большое значение при построении численного исследования механизма течения полимерного раствора и изучения его свойств. Таким образом, актуальной задачей является разработка качественных методов численного моделирования и позволяет путем непосредственного решения системы уравнений, описывающих динамику течения раствора, получить наиболее полную информацию о характеристиках потока. Однако численное моделирование турбулентных потоков полимерных растворов далеко не тривиально [29–32]. Это обусловлено наличием больших градиентов поля растяжений полимерных молекул, возникающих на режимах течений с высоким числом Wi , что приводит к развитию неустойчивостей численной схемы. Для регуляризации численного решения уравнений, описывающих динамику поля растяжений полимеров, обычно вводят дополнительную искусственную диффузию. Искусственная диффузия преобразует уравнения полимерной модели в параболическую форму за счет сглаживания градиентов растяжения полимеров [31].

Эластическая турбулентность была впервые изучена численно в работе [29] в двумерном колмогоровском потоке полимерного раствора. Основные явления, наблюдаемые в лабораторных экспериментах, были воспроизведены с помощью численного моделирования, например, увеличение лобового сопротивления с увеличением Wi и наблюдение степенного энергетического спектра. Трехмерное (3D) моделирование колмогоровского течения вязкоупругой жидкости было выполнено в работе [34]. Авторы показали, что эластическая турбулентность возникает при меньшем числе Вайсенберга Wi по сравнению с двумерным случаем.

В настоящей работе представлены результаты прямого численного моделирования неустойчивости течения вязкой жидкости, содержащей полимерную примесь под действием внешней периодической силы в периодической квадратной двумерной ячейке. Получены зависимости кинетической энергии течения от времени, показывающие переход от ламинарного течения ко вторичному периодическому режиму. Построены спектры скорости и кинетической энергии для гидродинамической составляющей течения, показывающие наличие степенной зависимости для энергетического каскада с коэффициентом $k \sim -3$. Аналогичные спектры растяжений полимерных молекул демонстрируют зависимость $k \sim -1$.

1. ПОСТАНОВКА ЗАДАЧИ. МОДЕЛЬНЫЕ УРАВНЕНИЯ

Система уравнений, описывающих динамику полимерного раствора вязкой жидкости, включает в себя уравнения Навье–Стокса и уравнения, описывающие эволюцию состояния полимерной составляющей [1]. Эта составляющая описывается векторным полем растяжения полимерной молекулы $\mathbf{R}$ — вектором, соединяющим начало и конец молекулы (end-to-end vector). Запишем нашу модель в виде

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{V}) &= 0, \\ \frac{\partial \rho u}{\partial t} + \nabla \cdot (\rho u \mathbf{V}) &= -\frac{\partial p}{\partial x} + \rho G \sin ky + \mu \Delta u + A \frac{\partial}{\partial x} [\gamma(R) \{R^x\}^2] + A \frac{\partial}{\partial y} (\gamma(R) R^x R^y), \\ \frac{\partial \rho v}{\partial t} + \nabla \cdot (\rho v \mathbf{V}) &= -\frac{\partial p}{\partial y} - \rho G \sin kx + \mu \Delta v + A \frac{\partial}{\partial y} [\gamma(R) \{R^y\}^2] + A \frac{\partial}{\partial x} (\gamma(R) R^x R^y), \\ \frac{\partial (\frac{\rho \mathbf{V}^2}{2} + e)}{\partial t} + \nabla \cdot \left(\mathbf{V} \left(\frac{\rho \mathbf{V}^2}{2} + p + e \right) \right) &= \frac{\partial}{\partial x} (\mu \rho \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) + A u \gamma(R) (R^x)^2 + \\ &+ v \mu \rho \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right) + A v \gamma(R) R^x R^y) + \frac{\partial}{\partial y} (\mu \rho \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) + A u \gamma(R) R^x R^y + \\ &+ \mu v \rho \left(\frac{\partial v}{\partial y} - \frac{\partial u}{\partial x} \right) + A v \gamma(R) (R^y)^2) + u \rho G \sin ky - v \rho G \sin kx, \\ \frac{\partial R^x}{\partial t} + u \frac{\partial R^x}{\partial x} + v \frac{\partial R^x}{\partial y} - R^x \frac{\partial u}{\partial x} - R^y \frac{\partial u}{\partial y} + \gamma(R) R^x &= C_d \Delta R^x, \end{aligned} \quad (1.1)$$

$$\frac{\partial R^y}{\partial t} + u \frac{\partial R^y}{\partial x} + v \frac{\partial R^y}{\partial y} - R^x \frac{\partial v}{\partial x} - R^y \frac{\partial v}{\partial y} + \gamma(R) R^y = C_d \Delta R^y,$$

$$p = \sigma e, \quad \sigma = \frac{2}{3}, \quad \gamma(R) = \gamma_0 \left(1 + \frac{R^2}{R_m^2}\right), \quad \mathbf{V} = (u, v)^T.$$

Здесь A — коэффициент концентрации полимеров в растворе, C_d — коэффициент искусственной диффузии полимеров, введенный для стабилизации численного решения, $\gamma(R)$ — модель релаксации полимерной молекулы, G — интенсивность внешней силы. В качестве модели упругости полимеров использовано приближение модели FENE-P [1], когда $R \ll R_m$: $\gamma(R) = \gamma_0 \left(1 + \frac{R^2}{R_m^2}\right)$. Здесь R_m — максимальная величина растяжения полимерной молекулы.

В начале расчетов раствор находился в состоянии покоя: $u(x, y, t = 0) = 0 \frac{\text{м}}{\text{с}}$; $v(x, y, t = 0) = 0 \frac{\text{м}}{\text{с}}$. Плотность и давление принимались равными $\rho(x, y, t = 0) = 10 \frac{\text{кг}}{\text{м}^3}$; $p(x, y, t = 0) = 10^3$ Па. Начальные значения растяжения полимерных молекул $R^x(x, y, t = 0) = 10^{-2}$ м; $R^y(x, y, t = 0) = 10^{-2}$ м. Коэффициент искусственной диффузии подбирался эмпирически, исходя из требований устойчивого численного счета $C_d = 10^{-3} \frac{\text{м}^2}{\text{с}}$. Во всех расчетах интенсивность внешней силы $G = 10^{-2} \frac{\text{Н}}{\text{кг}}$. Коэффициент релаксации полимерной молекулы положим $\gamma_0 = 10^{-4} \text{ с}^{-1}$. Величина динамической вязкости $\mu = 0.5$ Па с. Частота внешней периодической силы $k = 2 \text{ м}^{-1}$.

Расчетная область представляет из себя квадрат со сторонами $L \times L = 2\pi \times 2\pi$, на границе которого поставлены периодические граничные условия и покрыта равномерной вычислительной сеткой размерностью 250×250 ячеек. Также были проведены расчеты на сектах размерностью 500×500 и 1000×1000 для проверки того, что полученные результаты не являются численным эффектом.

2. ЧИСЛЕННЫЙ МЕТОД

Для численной аппроксимации системы (1.1) применялась комбинация двух численных методов — линейризованного метода Годунова [35] и метода Курганова—Тедмора (КТ) [36]. Линеаризованным методом Годунова аппроксимировалась гидродинамическая часть модели — уравнения Навье—Стокса, уравнения, описывающие полимерную часть, аппроксимировались методом Курганова—Тедмора (КТ). Подробное описание линеаризованной схемы Годунова представлено в работе [35]. Опишем аппроксимацию уравнений на полимерную составляющую более подробно. Запишем уравнения системы (1.1) для вектора растяжения полимеров в дивергентном виде:

$$\begin{pmatrix} R^x \\ R^y \end{pmatrix}_t + \begin{pmatrix} u R^x \\ u R^y \end{pmatrix}_x - \begin{pmatrix} C_d \frac{\partial R^x}{\partial x} \\ C_d \frac{\partial R^y}{\partial x} \end{pmatrix}_x + \begin{pmatrix} v R^x \\ v R^y \end{pmatrix}_y - \begin{pmatrix} C_d \frac{\partial R^x}{\partial y} \\ C_d \frac{\partial R^y}{\partial y} \end{pmatrix}_y = \begin{pmatrix} 2R^x \frac{\partial u}{\partial x} + R^y \frac{\partial u}{\partial y} + R^x \frac{\partial v}{\partial x} - \gamma(R) R^x \\ R^x \frac{\partial v}{\partial x} + 2R^y \frac{\partial v}{\partial y} + R^y \frac{\partial u}{\partial x} - \gamma(R) R^y \end{pmatrix}.$$

Суть методики КТ состоит в алгоритме негдуновского типа для вычисления конвективных потоков $\begin{pmatrix} u R^x \\ u R^y \end{pmatrix}$, $\begin{pmatrix} v R^x \\ v R^y \end{pmatrix}$ на гранях расчетных ячеек, для аппроксимации остальных членов (диффузионных и источника) применяется обычное осреднение на гранях ячеек расчетной сетки. Обозначим конвективный поток вдоль осей x, y : $F = \begin{pmatrix} u R^x \\ u R^y \end{pmatrix}$; $G = \begin{pmatrix} v R^x \\ v R^y \end{pmatrix}$, вектор-столбец неизвестных $R = \begin{pmatrix} R^x \\ R^y \end{pmatrix}$. Конвективные потоки тогда будут вычисляться следующим образом:

$$F_{i+\frac{1}{2},j}^n = \frac{F((R_{i+\frac{1}{2},j}^+)^n) + F((R_{i-\frac{1}{2},j}^-)^n)}{2} - \frac{(a_{i+\frac{1}{2},j}^x)^n}{2} \left((R_{i+\frac{1}{2},j}^+)^n - (R_{i-\frac{1}{2},j}^-)^n \right);$$

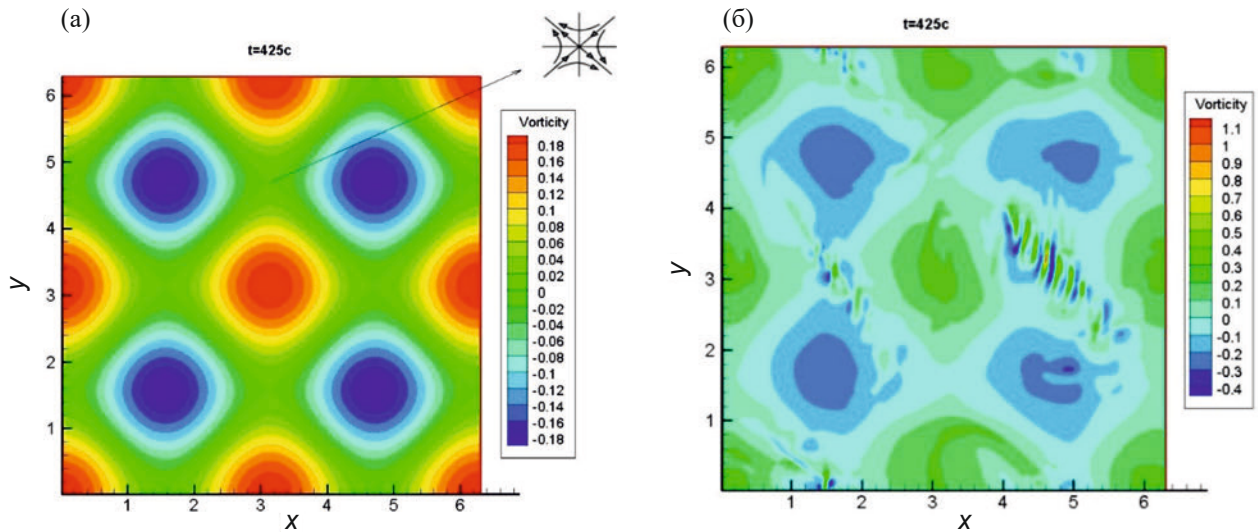
$$G_{i+\frac{1}{2},j}^n = \frac{G((R_{i+\frac{1}{2},j}^+)^n) + G((R_{i-\frac{1}{2},j}^-)^n)}{2} - \frac{(a_{i+\frac{1}{2},j}^y)^n}{2} \left((R_{i+\frac{1}{2},j}^+)^n - (R_{i-\frac{1}{2},j}^-)^n \right);$$

$$(R_{i+\frac{1}{2},j}^\pm)^n = R_{i+1,j}^n \mp \frac{\Delta x}{2} (R_x)_{i+\frac{1}{2} \pm \frac{1}{2},j}^n; \quad (R_{i,j+\frac{1}{2}}^\pm)^n = R_{i,j+1}^n \mp \frac{\Delta y}{2} (R_y)_{i,j+\frac{1}{2} \pm \frac{1}{2}}^n;$$

$$(a_{i+\frac{1}{2},j}^x)^n = \max_{\pm} \left(\frac{\partial F}{\partial R} ((R_{i+\frac{1}{2},j}^\pm)^n) \right) = |u_{i+\frac{1}{2},j}|; \quad (a_{i+\frac{1}{2},j}^y)^n = \max_{\pm} \left(\frac{\partial G}{\partial R} ((R_{i,j+\frac{1}{2}}^\pm)^n) \right) = |v_{i+\frac{1}{2},j}|;$$

$$u_{i+\frac{1}{2},j} = \frac{1}{2} (u_{i,j} + u_{i+1,j}); \quad v_{i+\frac{1}{2},j} = \frac{1}{2} (v_{i,j} + v_{i+1,j});$$

$$(R_x)_{i,j}^n = \min \text{ mod} \left(\theta \frac{(R_x)_{i,j}^n - (R_x)_{i-1,j}^n}{\Delta x}, \frac{(R_x)_{i+1,j}^n - (R_x)_{i-1,j}^n}{2\Delta x}, \theta \frac{(R_x)_{i+1,j}^n - (R_x)_{i,j}^n}{\Delta x} \right);$$



Фиг. 1. Завихренность течения, характеризующегося числом Рейнольдса $Re \sim 10$, в отсутствии полимерной примеси (а) и при ее наличии (б).

$$(R_y)_{i,j}^n = \min \operatorname{mod} \left(\theta \frac{(R_y)_{i,j}^n - (R_y)_{i-1,j}^n}{\Delta y}, \frac{(R_y)_{i+1,j}^n - (R_y)_{i-1,j}^n}{2\Delta y}, \theta \frac{(R_y)_{i+1,j}^n - (R_y)_{i,j}^n}{\Delta y} \right);$$

$\theta = 1.5$ — весовой коэффициент ограничителя.

Здесь полуцелые индексы обозначают потоки на гранях ячеек, целые — значения величин в центре ячейки, u, v суть x, y -компоненты скорости течения, a — локальная скорость распространения возмущений. Для ограничения наклонов реконструированного решения в ячейке использовался $\min \operatorname{mod}$ ограничитель.

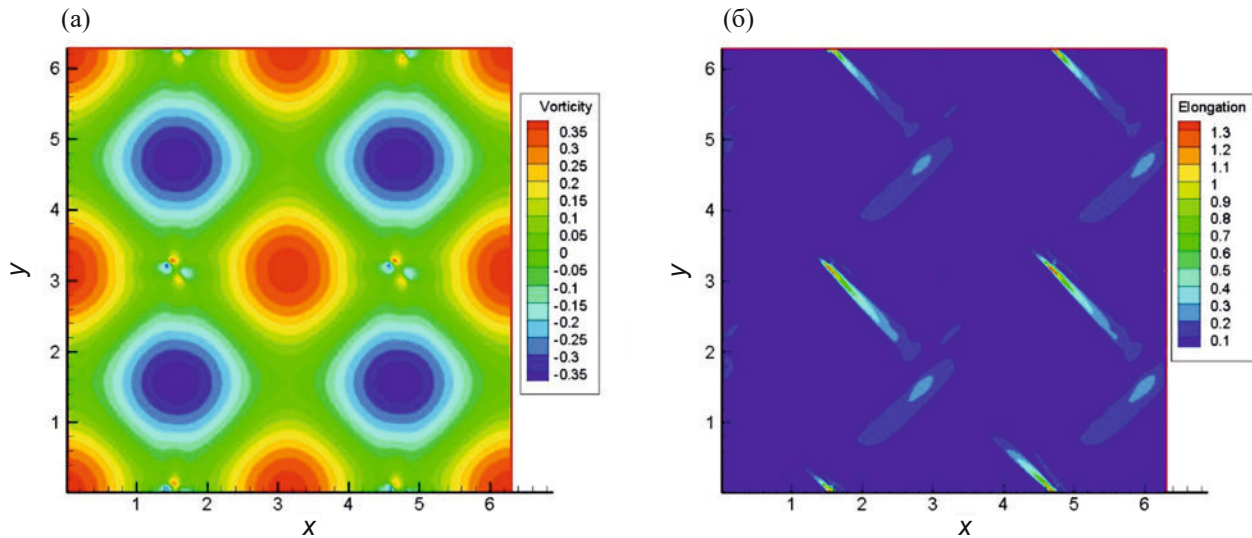
3. РЕЗУЛЬТАТЫ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ

Основным результатом численного моделирования следует выделить разрушение ламинарного режима течения и переход во вторичный колебательный режим, показанный на фиг. 1. Для данного режима характерно колебательное движение потока около равновесного положения, обусловленного действием внешней силы. В данном расчете были исследованы следующие значения полимерных параметров: $R_m = 0.1$ м, $A = 1000 \frac{\text{кг}}{\text{м}^3 \text{с}}$. Значение для коэффициента обратного влияния полимеров на поток A было подобрано эмпирически. Данный коэффициент увеличивался до тех пор, пока течение оставалось устойчивым. На фиг. 1 для сравнения показана завихренность ламинарного течения в отсутствии полимерной примеси и вторичного с полимерной примесью для числа $Re \sim 10$ и числа $Wi \sim 100$. Если говорить о переходном режиме — когда происходит переход от ламинарного течения ко вторичному, то он характеризуется тем, что неустойчивость возникает в областях, где течение имеет гиперболический характер (фиг.1). Это объясняется максимальным растяжением полимеров в этих областях и, соответственно, возникновением обратного влияния полимеров на гидродинамический поток.

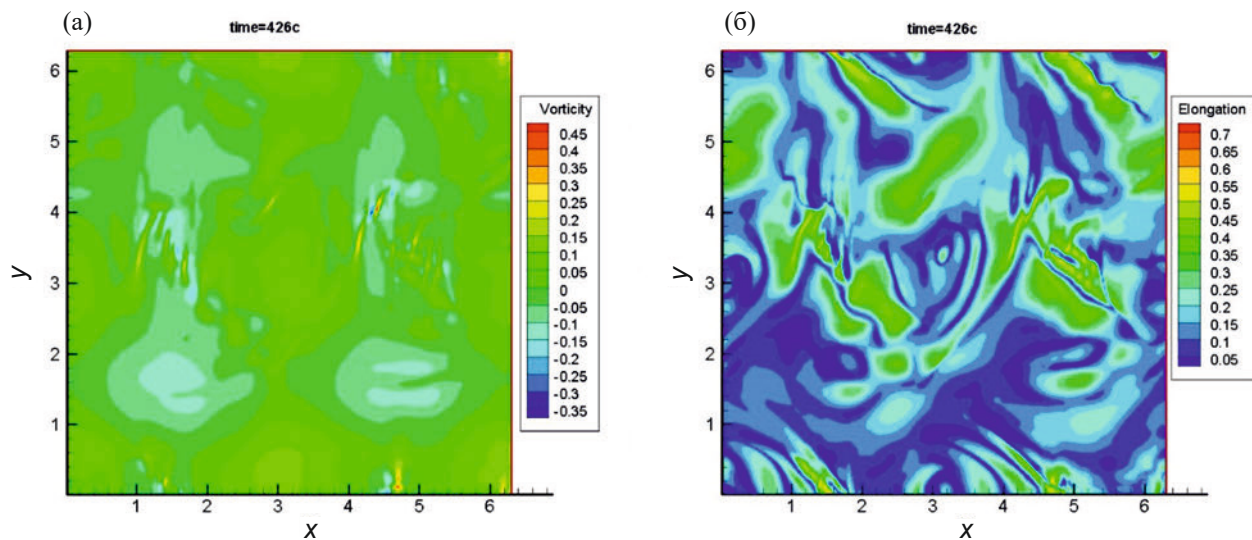
На фиг. 2 показан начальный момент смены режима течения. Видно, что в областях гиперболичности течения растяжение полимеров принимает наибольшее значение и именно здесь начинает быть заметно обратное влияние полимеров на поток.

На фиг. 3 показаны картины установившегося вторичного периодического течения; а — завихренность, б — растяжение полимерных молекул.

На фиг. 4 показана зависимость кинетической энергии течения от времени для различных режимов, отличающихся числами Re и Wi . Отчетливо виден момент, когда течение перестает быть ламинарным — зависимость энергии становится более хаотической. Как и следовало ожидать, при увеличении чисел Re и Wi , течение становится менее устойчивым, что не связано с влиянием только числа Re . Влияние числа Рейнольдса обуславливает перенос энергии из гидродинамического потока в полимерную часть, вследствие чего полимеры сильнее растягиваются и начинают оказывать обратное влияние на течение. В результате чего ламинарный режим теряет свою устойчивость и развивается вторичный колебательный режим течения.



Фиг. 2. Завихренность течения (а) и растяжение полимеров (б) в начальные моменты смены режима течения.



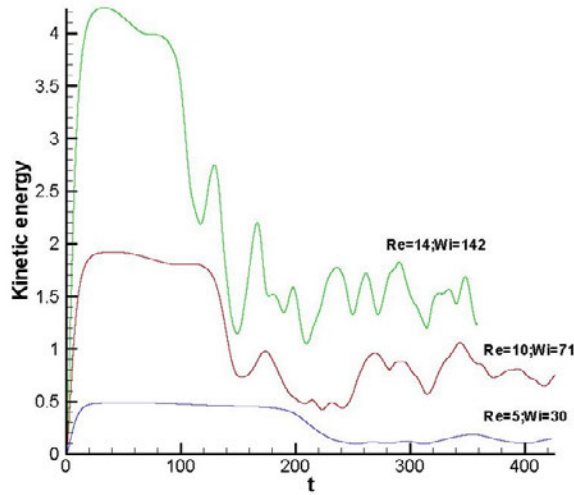
Фиг. 3. Установившееся турбулентное течение полимерного раствора, (а) - распределение в расчетной области завихренности, (б) – растяжения полимеров.

Фиг. 5 иллюстрирует энергетический спектр (см. фиг. 5б) сформировавшегося вторичного течения (см. фиг. 5а), который вычислялся следующим образом:

$$\varepsilon(k_x, k_y) = \frac{1}{2} \left[\sum_{i=0}^1 \sum_{j=1}^4 u_{i(j)}^2(k_x, k_y) \right];$$

$$u_{i(1)}(k_x, k_y) = \frac{1}{\pi} \int_0^{2\pi} u_i(x, y) \cos(k_x x) \cos(k_y y) dx dy,$$

$$u_{i(2)}(k_x, k_y) = \frac{1}{\pi} \int_0^{2\pi} u_i(x, y) \cos(k_x x) \sin(k_y y) dx dy,$$



Фиг. 4. Кинетическая энергия течения в зависимости от времени для различных течений, характеризующихся различными Re и Wi .

$$u_{i(3)}(k_x, k_y) = \frac{1}{\pi} \int_0^{2\pi} u_i(x, y) \sin(k_x x) \cos(k_y y) dx dy,$$

$$u_{i(1)}(k_x, k_y) = \frac{1}{\pi} \int_0^{2\pi} u_i(x, y) \sin(k_x x) \sin(k_y y) dx dy.$$

Осреднение по направлениям вектора $\mathbf{k}$ осуществлялось по кольцу шириной $\delta = 0.5$:

$$E_k(k) = \sum_{k_x, k_y: |k - \sqrt{k_x^2 + k_y^2}| < \delta} \varepsilon(k_x, k_y).$$

На фиг. 5 изображен спектр растяжений полимерных молекул (см. фиг. 5г):

$$R(k_x, k_y) = \sqrt{r_{i(1)}^2(k_x, k_y) + r_{i(2)}^2(k_x, k_y) + r_{i(3)}^2(k_x, k_y) + r_{i(4)}^2(k_x, k_y)};$$

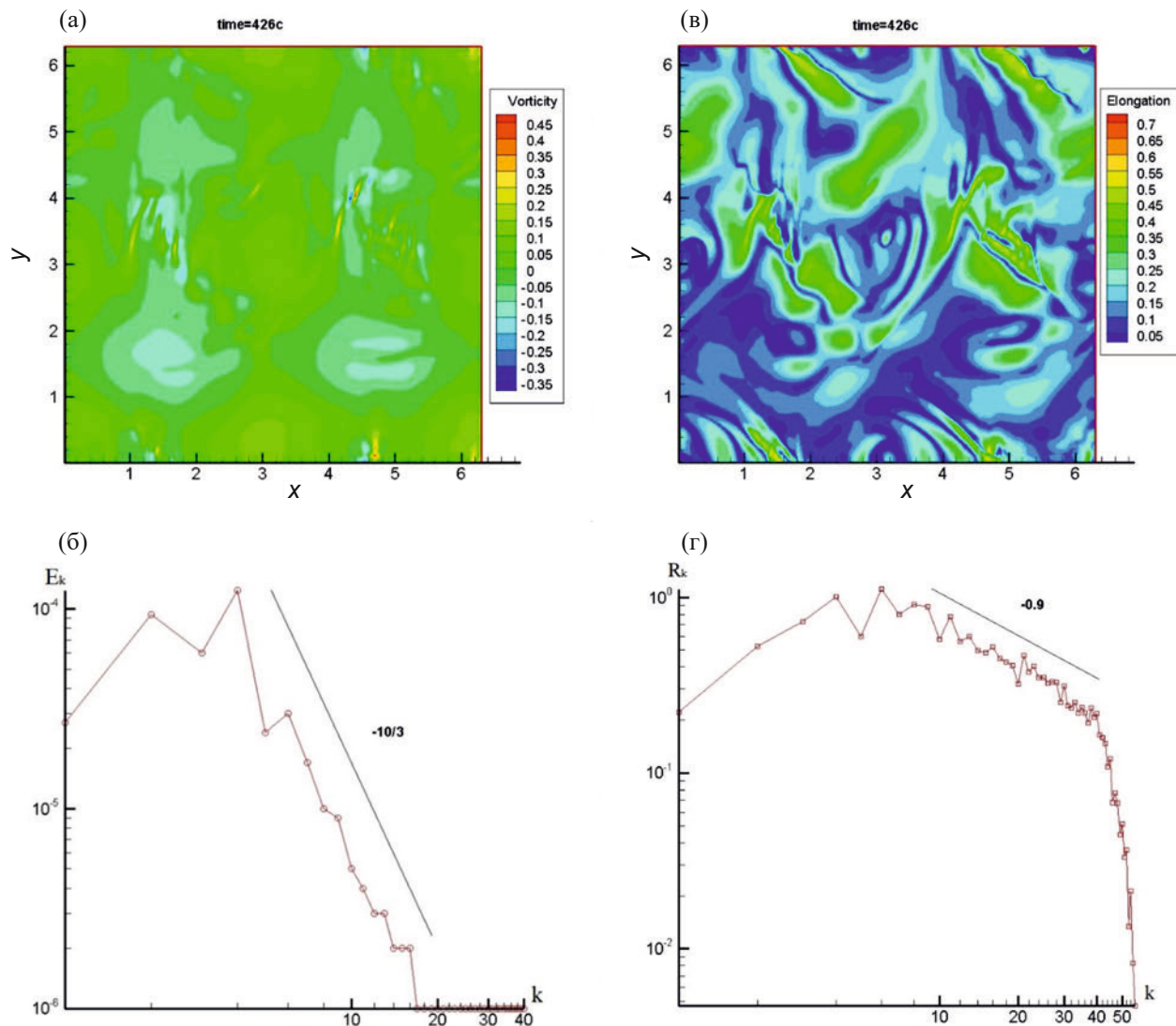
$$r_{i(1)}(k_x, k_y) = \frac{1}{\pi} \int_0^{2\pi} R^i(x, y) \cos(k_x x) \cos(k_y y) dx dy,$$

$$r_{i(2)}(k_x, k_y) = \frac{1}{\pi} \int_0^{2\pi} R^i(x, y) \cos(k_x x) \sin(k_y y) dx dy,$$

$$r_{i(3)}(k_x, k_y) = \frac{1}{\pi} \int_0^{2\pi} R^i(x, y) \sin(k_x x) \cos(k_y y) dx dy,$$

$$r_{i(4)}(k_x, k_y) = \frac{1}{\pi} \int_0^{2\pi} R^i(x, y) \sin(k_x x) \sin(k_y y) dx dy.$$

Осреднение по направлениям вектора $\mathbf{k}$ осуществлялось аналогично случаю с энергией. На фиг. 6 изображены течения, посчитанные на более мелких сетках для параметров задачи $R_m = 0.1$ м, $A = 1000 \frac{\text{кг}}{\text{м}^3 \text{с}}$. Видно, что также нарушается ламинарность течения, что говорит о том, что рассмотренный нами вторичный режим не является численным эффектом.



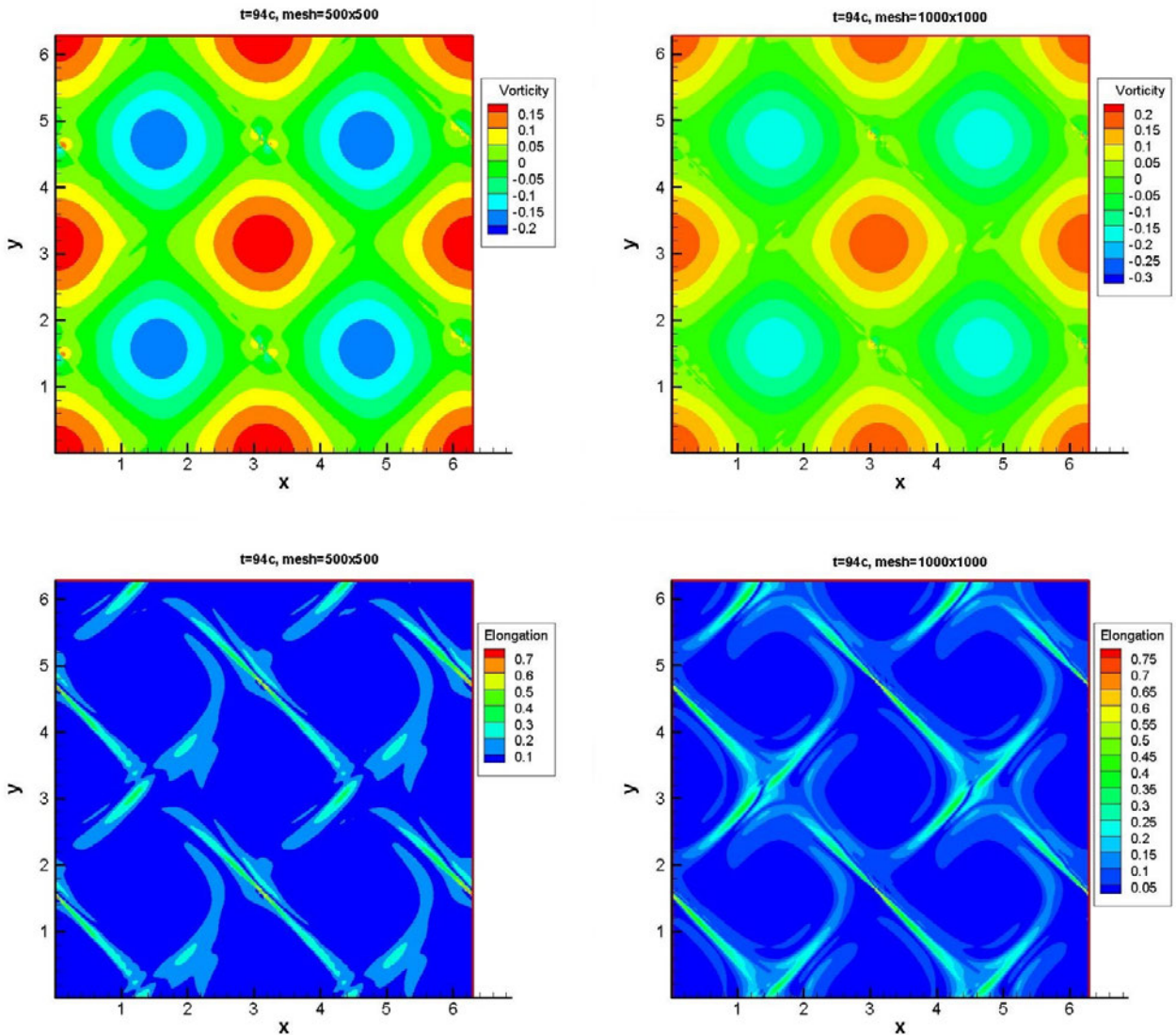
Фиг. 5. Энергетический спектр (б) $E_k(k)$ и спектр растяжений полимерных молекул (г) $R_k(k)$, изображенный в системе координат двойного логарифмического масштаба, вторичного течения (в) полимерного раствора для чисел $Re \sim 10$ и $Wi \sim 100$.

Наличие степенной зависимости в спектре энергии говорит о том, что наблюдаемый нами вторичный режим течения показывает наличие влияния полимеров на поток и требует дальнейшего тщательного исследования в части изучения влияния различных параметров на характеристики потока.

4. ЗАКЛЮЧЕНИЕ

В работе предложена гибридная численная схема второго порядка точности, аппроксимирующая уравнения динамики полимерного раствора (1.1). Гидродинамическая часть системы (1.1) – уравнения Навье–Стокса аппроксимировались линейризованным методом Годунова [35], полимерная часть – методикой Курганова–Тедмора [36].

Основным результатом численного моделирования колмогоровского течения при наличии полимеров является наличие обратного влияния полимеров на гидродинамический поток и формирование вторичного колебательного режима течения. Неустойчивость рождается из областей т.н. гиперболичности течения, где растяжение полимеров имеет максимальную величину. Следует отметить, что параметром, характеризующим устой-



Фиг. 6. Возникновение вторичного режима в расчетах на сетках размерностью 500×500 и 1000×1000 . Вверху показаны завихренности, ниже — растяжения полимерных молекул.

чивость течения здесь не является по отдельности ни число Re , ни число Wi , а их комбинация, коэффициент эластичности $El = \frac{Wi}{Re}$ (см. [1]).

Анализ спектральных характеристик энергии полученного течения показывает, что режим данного течения является неустойчивым. Отметим, что энергетический спектр течения вязкого полимерного раствора характеризуется более крутым углом наклона с коэффициентом ~ -3 , чем известный спектр турбулентного колмогоровского течения вязкой жидкости $\sim -5/3$, что согласуется с экспериментальными данными [37, 38]. Спектр растяжений полимерных молекул имеет зависимость ~ -1 .

СПИСОК ЛИТЕРАТУРЫ

1. Victor Steinberg. Elastic Turbulence: An Experimental View on Inertialess Random // Flow. Annu. Rev. Fluid Mech. 2021. V. 53. P. 27–58.
2. Armin Shahmardi, Sagar Zade, Mehdi N. Ardekani, Rob J. Poole, Fredrik Lundell, Marco E. Rosti, and Luca Brandt. Turbulent duct flow with polymers // J. Fluid Mech. 2019. V. 859. P. 1057–1083.
3. Belan S., Chernykh A., Lebedev V. Boundary layer of elastic turbulence // J. Fluid Mech. 2018. V. 855. P. 910–921.

4. *Zhang Hong-Na, Li Feng-Chen, Cao Yang, Kunugi Tomoaki, Yu Bo.* Direct numerical simulation of elastic turbulence and its mixing-enhancement effect in a straight channel flow // *Chin. Phys. B* 2013. V. 22. No 2. 024703.
5. *Larson R.G, Shaqfeh E.S.G., Muller S.J.* A purely elastic instability in Taylor–Couette flow // *J. Fluid Mech.* 1990. V. 218. P. 573–600.
6. *Groisman A., Steinberg V.* Elastic versus inertial instability in a polymer solution flow // 1998a. *Europhys. Lett.* 43:165–70.
7. *Groisman A., Steinberg V.* Mechanism of elastic instability in Couette flow of polymer solutions: experiment // 1998b. *Phys. Fluids* 10:2451–63.
8. *Steinberg V, Groisman A.* Elastic versus inertial instability in Couette–Taylor flow of a polymer solution: review // *Philos. Mag. B* 1998. V. 78. P. 253–63.
9. *McKinley G.H, Byars J.A, Brown R.A, Armstrong R.C.* Observations on the elastic instability in cone-and-plate and parallel-plate flows of a polyisobutylene Boger fluid // *J. Non-Newton. Fluid Mech.* 1991. V. 40:201–29.
10. *Oeztekin A., Brown R.A.* Instability of a viscoelastic fluid between rotating parallel disks: analysis for the Oldroyd-B fluid // *J. Fluid Mech.* 1993. 255:473–502.
11. *Byars J.A, Oeztekin A., Brown R.A., McKinley G.H.* Spiral instabilities in the flow of highly elastic fluids between rotating parallel disks // 1994. *J. Fluid. Mech.* 271:173–218.
12. *Oeztekin A., Brown R.A., McKinley G.H.* Quantitative prediction of the viscoelastic instability in coneand plate flow of a Boger fluid using a multi-mode Giesekus model // *J. Non-Newton. Fluid Mech.* 1994. 54:351–77.
13. *Joo Y.L, Shaqfeh E.S.G.* A purely elastic instability in Dean and Taylor–Dean flow // 1992. *Phys. Fluids. A.* 1992. V. 4:524–42.
14. *Burghilea T., Segre E., Bar-Joseph I., Groisman A., Steinberg V.* Chaotic flow and efficient mixing in a microchannel with a polymer solution // *Phys. Rev.* 2004. E 69:066305.
15. *Burghilea T., Segre E., Steinberg V.* Mixing by polymers: experimental test of decay regime of mixing // 2004b. *Phys. Rev. Lett.* 2004. V. 92:164501.
16. *Groisman A., Steinberg V.* Elastic turbulence in curvilinear flows of polymer solutions // *New J. Phys.* 2004. V. 6:29–48.
17. *Jun Y., Steinberg V.* Elastic turbulence in a curvilinear channel flow // *Phys. Rev. E* 2011. V. 84:056325.
18. *Soulies A., Aubril J., Castelain C., Burghilea T.* Characterization of elastic turbulence in a serpentine micro-channel // *Phys. Fluids.* 2017. V. 29:083102.
19. *Arratia P.E., Thomas C.C., Diorio J., Gollub J.P.* Elastic instabilities of polymer solutions in cross-channel flow // *Phys. Rev. Lett.* 2006. V. 96:144502.
20. *Poole R.J., Alves M.A., Oliveira P.J.* Purely elastic flow asymmetries // *Phys. Rev. Lett.* 2007. V. 99:164503.
21. *Haward S.J., McKinley G.H.* Instabilities in stagnation point flows of polymer solutions. // *Phys. Fluids.* 2013. V. 25:083104.
22. *Varshney A., Afik E., Kaplan Y., Steinberg V.* Oscillatory elastic instabilities in an extensional viscoelastic flow // *Soft Matter.* 2016. 12:2186–91.
23. *McKinley G.H., Armstrong R.C., Brown R.A.* The wake instability in viscoelastic flow past confined circular cylinders // *Philos. Trans. R. Soc. A.* 1993. 344:265–304.
24. *Vazquez-Quesada A., Ellero M.* SPH simulations of a viscoelastic flow around a periodic array of cylinders confined in a channel // *J. Non-Newton. Fluid Mech.* 2012. 167–68:1–8.
25. *Grilli M., Vazquez-Quesada A., Ellero M.* Transition to turbulence and mixing in a viscoelastic fluid flowing inside a channel with a periodic array of cylindrical obstacles. // *Phys. Rev. Lett.* 2013. V. 110:174501.
26. *Varshney A., Steinberg V.* Elastic wake instabilities in a creeping flow between two obstacles // *Phys. Rev. Fluids.* 2017. 2:051301(R).

27. *Larson R.G.* Instabilities in viscoelastic flows // *Rheol. Acta*. 1992. 31:213–63.
28. *Shaqfeh E.S.G.* Purely elastic instabilities in viscometric flows // *Annu. Rev. Fluid Mech.* 1996. V. 28:129–85.
29. *Berti S., Bistagnino A., Boffetta G., Celani A., and Musacchio S.* Two-dimensional elastic turbulence // *Physical Review*. E2008. 77, 055306 R.
30. *Anupam Gupta, Dario Vincenzi.* Effect of polymer-stress diffusion in the numerical simulation of elastic turbulence // 7 Mar 2019. arXiv:1809.09648v2 [physics.flu-dyn].
31. *Berti S., Boffetta G.* Elastic waves and transition to elastic turbulence in a two-dimensional viscoelastic Kolmogorov flow // *Physical Review*. 2010 E 82, 036314.
32. *Dzanic V., From C.S., and Saurt E.* Conserving elastic turbulence numerically using artificial diffusivity // 20 Jun 2022. arXiv: 2203.12962v2 [physics.flu-dyn].
33. *Денисенко В.В., Фортובה С.В.* Численное моделирование эластической турбулентности в ограниченной двумерной ячейке // *Сиб. журнал индустриальной матем.* 2023. Т. 26. № 1. С. 55–64.
34. *Zhang H. N., Li F. C., Cao Y., Yang J. C., Li X. B., Cai W. H.* Proceedings of the 6th International Conference on Fluid Mechanics, 2011, June 30–July 3, Guangzhou, China, p. 107.
35. *Годунов С.К., Денисенко В.В., Ключинский Д.В., Фортובה С.В., Шепелев В.В.* Исследование энтропийных свойств линеаризованной редакции метода Годунова // *Ж. вычисл. матем. и матем. физ.* 2020. Т. 60. № 4. С. 639–651.
36. *Alexander Kurganov, Eitan Tadmor.* New High-Resolution Central Schemes for Nonlinear Conservation Laws and Convection–Diffusion Equations // *J. of Comput. Phys.* 2000. V. 160. P. 241–282.
37. *Groisman A., Steinberg V.* *Nature*. 2000. 405, 53.
38. *Burgheslea T., Segre E., Steinberg V.* Elastic turbulence in von Karman swirling flow between two disks // *Phys. Fluids*. 2007. 19, 053104.

Numerical simulation of polymer solution flow for Kolmogorov flow

V. V. Denisenko^{a,*}, S. V. Fortova^a, V. V. Lebedev^b, I. V. Kolokolov^b

^a 123056 Moscow, 2nd Brestskaya str., 19/18, Institute of Design Automation of the Russian Academy of Sciences, Russia

^b 142432 Chernogolovka, Moscow region, Ave. Akad Semenova, 1A, L.D. Landau Institute of Theoretical Physics of the Russian Academy of Sciences, Russia

*e-mail: ned13@rambler.ru

Received: 22.02.2024

Revised: 22.02.2024

Accepted: 31.05.2024

Abstract. A numerical method approximating the equations of the dynamics of the polymer solution flow is proposed. The developed methodology is based on a hybrid approach. The hydrodynamic component of the flow is described by a system of Navier–Stokes equations and is numerically approximated by the linearized Godunov method. The polymer component of the flow is described by a system of equations for the polymer stretching vector $\mathbf{R}$ and is numerically approximated by the Kurganov–Tadmor method. Using the constructed scheme, the problem of the stability of the polymer solution flow at low Reynolds numbers $Re \sim 10$ in a square periodic cell under the action of an external periodic force is investigated. The instability of this type of flow, characterized by a violation of its laminarity, has been studied by numerical experiment. Spectral characteristics of the polymer solution at low Reynolds numbers are constructed.

Keywords: numerical modeling, elastic turbulence, hydrodynamic instability, Kolmogorov flow.