

Современные проблемы
вычислительной
математики
и
математического
моделирования

1

*Вычислительная
математика*

НАУКА

Современные проблемы
вычислительной
математики
и
математического
моделирования

В двух томах



МОСКВА НАУКА 2005

Современные проблемы
вычислительной
математики
и
математического
моделирования

Том 1

*Вычислительная
математика*



МОСКВА НАУКА 2005

УДК 519.6
ББК 22.19
С56

Ответственные редакторы:
академик *Н.С. БАХВАЛОВ*,
академик *В.В. ВОЕВОДИН*

Рецензенты:
академик *А.С. АЛЕКСЕЕВ*,
член-корреспондент РАН *Ю.Г. ЕВТУШЕНКО*

Современные проблемы вычислительной математики и математического моделирования : в 2 т. / Ин-т вычисл. математики. – М. : Наука, 2005. – ISBN 5-02-033436-7.

Т. 1. : Вычислительная математика / [отв. ред. Н.С. Бахвалов, В.В. Воеводин]. – 343 с. – ISBN 5-02-033716-1.

Издание настоящего двухтомника приурочено к юбилею – 80-летию академика Г.И. Марчука. В работе представлены коллективные обзорные статьи современного состояния направлений вычислительной математики и математического моделирования, которые в течение многих лет развивались в Институте вычислительной математики и в работах Г.И. Марчука.

Том 1-й посвящен современным методам вычислительной математики, высокопроизводительным вычислительным технологиям, методам Монте-Карло и статистическому моделированию, параллельным вычислениям, а также теории сопряженных уравнений и анализу сложных систем.

Для специалистов в области вычислительной математики и математического моделирования, а также для студентов и аспирантов соответствующих специальностей.

ТП 2005-I-148

Contemporary problems of numerical mathematics and mathematical modelling : in 2 vol. / Institute of Numerical Mathematics. – Moscow : Nauka, 2005. – ISBN 5-02-033436-7.

Vol. 1. : Numerical Mathematics / [ed. by N.S. Bakhvalov, V.V. Voevodin]. – 343 p. – ISBN 5-02-033716-1.

This collection of works is dedicated to the 80th birthday of Academician G.I. Marchuk. The themes of the works presented are associated with the scientific directions, in which G.I. Marchuk achieved the most outstanding results. Most of papers are of a review character demonstrating the current state of the art in numerical mathematics and mathematical modelling.

Volume 1 presents the modern methods of numerical mathematics, computational technologies, Monte Carlo methods and statistical modelling, parallel computations, adjoint equation theory and analysis of complex systems.

The book is certainly interesting to the specialists in numerical mathematics and mathematical modelling and also to students and postgraduates.

ISBN 5-02-033436-7 (общ.)
ISBN 5-02-033716-1

© Российская академия наук, 2005
© Редакционно-издательское оформление.
Издательство “Наука”, 2005



К 80-летию академика Г. И. Марчука

Действительный член Российской академии наук Г.И. Марчук — выдающийся ученый с мировым именем, крупный организатор науки. Блестящий разносторонний исследователь в области естествознания Г.И. Марчук внес огромный вклад в развитие вычислительной и прикладной математики, физики ядерных реакторов, математического моделирования, динамической метеорологии, информатики, иммунологии.

Гурий Иванович Марчук родился 8 июня 1925 года в поселке Петро-Херсонце Оренбургской области в семье сельского учителя.

После окончания математико-механического факультета Ленинградского государственного университета Г.И. Марчук поступил в аспирантуру и в 1952 г. защитил кандидатскую диссертацию "Динамика крупномасштабных полей метеорологических элементов в бароклинической атмосфере". С 1953 по 1962 гг. Г.И. Марчук работал в Физико-энергетическом институте (г. Обнинск), где заведовал лабораторией, а затем математическим

отделом. В этот период он предложил новые методы расчета ядерных реакторов, которые до настоящего времени составляют основу моделирования имитационных расчетов промышленных реакторов. Большую известность получили его работы по теории переноса излучения. Результаты этих исследований обобщены в монографии "Численные методы расчета ядерных реакторов" и в его докторской диссертации (1965 г.). В 1959–1961 гг. он принял участие в разработке требований к ядерной безопасности для заводов и других предприятий атомной промышленности, проводившейся по инициативе И.В. Курчатова. В 1961 г. за работы в области теории ядерных реакторов ему была присуждена Ленинская премия.

В 1962 г. Г.И. Марчук был избран членом-корреспондентом АН СССР (специальность "атомная энергетика"), а в 1968 г. — действительным членом АН СССР (специальность "физика атмосферы").

Г.И. Марчук — автор более 350 научных работ, в том числе 25 монографий. Научные труды Г.И. Марчука посвящены созданию и исследованию эффективных алгоритмов вычислительной математики, методов расчета ядерных реакторов, исследованию и моделированию процессов физики атмосферы и океана, математическому моделированию в проблемах охраны окружающей среды, в проблемах иммунологии и медицины, изучению актуальных задач информатики и вычислительной техники.

В области вычислительной математики Г.И. Марчуком сделан существенный вклад в развитие разностных схем. Им построены и исследованы разностные схемы для классов уравнений, возникающих в теории ядерных реакторов, предложен метод построения разностных схем на основе интегральных тождеств, который получил развитие в работах советских и зарубежных ученых. Г.И. Марчуком и его учениками решен ряд проблем в теории разностных и вариационно-разностных схем для различных задач математической физики.

Г.И. Марчук внес большой вклад в разработку методов расщепления, алгоритмов возмущений, построенных на основе использования сопряженных уравнений. Значительное внимание в работах Г.И. Марчука уделяется разработке и обоснованию новых численных методов линейной алгебры.

В методах расчета ядерных реакторов Г.И. Марчуком на основе теории сопряженных уравнений и алгоритмов возмущений разработаны принципы построения эффективных малогрупповых моделей ядерного реактора, созданы математические модели реактора в различных приближениях метода сферических гармоник и предложены численные схемы реализации возникающих уравнений. Эти модели широко использовались для расчетов критических масс промышленных реакторов.

Большой вклад Г.И. Марчук внес в решение задач численного прогноза погоды, моделирования общей циркуляции атмосферы и океана и проблему моделирования климата и его изменений. Его совместная работа с

Н.И. Булеевым, в которой была сформулирована система квазигеострофических уравнений для трехмерной атмосферы и построена функция Грина для ее решения, давно стала классической. В 60-е годы он сформулировал направление численного прогноза погоды, в основу которого были положены полные неадиабатические уравнения динамики атмосферы, а в качестве метода решения предложен метод расщепления по физическим процессам и геометрическим переменным — в то время, пожалуй, единственный метод, с помощью которого можно было решить такую сложную задачу. Важно отметить, что задача численного прогноза по полным уравнениям была доведена до оперативного использования в Западно-Сибирской Гидрометеослужбе. В 70-е годы Г.И. Марчук сформулировал новый подход к решению задачи долгосрочного прогноза погоды, основанный на так называемых сопряженных уравнениях для нелинейных уравнений термогидродинамики атмосферы и океана, дающих возможность построить функцию чувствительности для нестационарных нелинейных задач. Этот подход стал основным при выделении энергоактивных зон Мирового океана, изучению которых была посвящена программа "Разрезы", сформулированная и организованная Гурием Ивановичем. В те же 70-е годы он уделял большое внимание созданию совместной модели общей циркуляции атмосферы и океана, которая должна была стать основой для моделирования климата и его изменений. За работы по численному прогнозу погоды Г.И. Марчуку была присуждена премия им. А.А. Фридмана АН СССР, а за работы по решению задач физики атмосферы и океана — Государственная премия Российской Федерации.

Г.И. Марчуком создана теория математического моделирования оптимизационных проблем в охране окружающей среды. Им поставлены и предложены алгоритмы решения общей задачи определения допустимой области размещения промышленных предприятий, планирования строительства с учетом допустимых доз загрязнения экономически значимых зон.

Марчук Г.И. является одним из авторов нового направления прикладной математики — математического моделирования в иммунологии и медицине. Он построил систему нелинейных дифференциальных уравнений, которые описывают иммунные реакции человеческого организма, возникающие в результате вирусных и бактериальных инфекций.

Более сорока лет Гурий Иванович Марчук посвятил делу организации отечественной науки. С 1964 по 1979 год он — директор Вычислительного центра СО АН СССР; с 1969 по 1975 год — заместитель, а затем Председатель Сибирского отделения АН СССР; с 1980 по 1986 год — Председатель Государственного комитета СССР по науке и технике в ранге заместителя Председателя Совета Министров СССР; с 1975 по 1980 год — Вице-президент, а с 1986 по 1991 год Президент АН СССР. В 1980 году Г.И. Марчук создал и возглавил Отдел вычислительной математики АН

СССР. В 1991 году Отдел был преобразован в Институт вычислительной математики РАН. До 2000 года Г.И. Марчук — директор ИВМ РАН, с 2000 года — советник Президиума РАН и почетный директор ИВМ РАН.

За научные заслуги и вклад во внедрение научных достижений в народном хозяйстве Г.И. Марчук удостоен звания Героя Социалистического Труда, награжден четырьмя орденами Ленина. Ему присуждены Ленинская и Государственная премии. Он обладатель Золотой медали им. С.И. Вавилова Российской академии наук, Золотой медали им. М.В. Келдыша, а также Золотой медали им. П.Л. Чебышева. В 1998 году Г.И. Марчук награжден орденом "За заслуги перед Отечеством IV степени".

Научные успехи Г.И. Марчука высоко оценены и за рубежом. Он является почетным доктором Хьюстонского, Орегонского, Тулузского, Дрезденского, Тель-Авивского, Калькутского, Карлова и Будапештского университетов, членом Европейской академии наук, иностранным членом Академии наук Франции, Финляндии, Индии, Польши, Болгарии и др.; лауреатом Международной премии им. А.П. Карпинского; Командором Ордена "Почетного Легиона" — государственной награды Франции. Г.И. Марчук — член редколлегий пяти иностранных (США, Германия, Италия, Франция, Швеция), а также отечественных журналов, главный редактор "Российского журнала численного анализа и математического моделирования", издающегося Институтом вычислительной математики в Голландии.

На протяжении всей своей научной биографии Гурий Иванович уделял большое внимание подготовке научных кадров, возглавлял кафедры в ФЭИ (Обнинск), НГУ (Новосибирск), МФТИ (Долгопрудный). В настоящее время он работает заведующим кафедрой вычислительных технологий и моделирования сложных систем в Московском государственном университете. Под его руководством защищено свыше 35 кандидатских диссертаций. Среди учеников Г.И. Марчука 25 докторов наук.

Настоящий сборник работ включает обзоры современного состояния направлений вычислительной математики и математического моделирования, в разработке которых фундаментальную роль сыграли исследования Г.И. Марчука и его научной школы.

Академики *Н. С. Бахвалов*
В. В. Воеводин
В. П. Дымников

Численные методы решения задач математической физики

*Н.С. Бахвалов, Г.М. Кобельков, Ю.А. Кузнецов, В.И. Лебедев,
И.К. Лифанов, Ю.М. Нечепуренко, В.В. Шайдуров*

Содержание

Введение	10
1. Методы расщепления	10
1.1. Схемы расщепления по пространственным переменным . . .	11
1.2. Расщепление по физическим процессам	14
2. Многосеточный метод Федоренко	18
2.1. Простейший пример	19
2.2. Современное состояние вопроса	25
3. Жесткие системы дифференциальных уравнений, устойчи- чивые явные методы и спектральный анализ	29
3.1. Явные методы решения жестких нестационарных задач . . .	29
3.2. Спектральный анализ задач математической физики	40
4. Численные методы решения жестких эллиптических задач	45
4.1. Простейший итерационный метод на подпространстве	47
4.2. Метод потоков	49
4.3. Метод псевдопотоков	51
4.4. Вопросы численной реализации	55
5. Переобусловленные итерационные методы для линейных систем с симметричными матрицами	58
5.1. Переобусловленный метод сопряженных градиентов	60
5.2. Переобусловленный метод минимальных итераций	61
5.3. Переобуславливатели, основанные на методах релаксации .	63
6. О численном решении уравнений Навье–Стокса	68
6.1. Постановка задачи	68
6.2. Задача Стокса	71
6.3. Нелинейная стационарная задача	75

7. О применении идей метода дискретных вихрей в задаче распространения звука в мелкой воде	80
7.1. Математическая модель задачи для слоя воды на абсолютно отражающем дне на основе интегральных уравнений	82
7.2. Метод численного решения задачи для слоя воды на абсолютно отражающем дне	86
7.3. Интегрирование гиперсингулярного интеграла с весом . . .	88
7.4. Примеры численных расчетов	92
Литература	93

Введение

Статья посвящена некоторым методам решения задач математической физики; выбор материала осуществлялся по следующему критерию. В статье отражены некоторые из методов, в разработке которых фундаментальную роль сыграли исследования Г.И. Марчука, его учеников и сотрудников ИВМ РАН.

1. Методы расщепления

Рассматриваемые в этом разделе методы являются одними из наиболее используемых инструментов решения многомерных нестационарных задач математической физики. По-видимому, толчком к возникновению этих методов стал метод переменных направлений [114, 88] решения параболического уравнения, позволяющий сводить решения многомерной задачи к последовательному решению одномерных задач. При этом каждый шаг метода переменных направлений аппроксимирует исходное уравнение. Методы, возникшие в развитие этого метода, носили на разных этапах разные названия, например, методы дробных шагов, экономические методы, методы суммарной аппроксимации.

Сейчас наиболее употребительным является название "метод расщепления". Крупнейший вклад в создание этого метода и его обоснование внесли советские математики Г.И. Марчук, А.А. Самарский, Н.Н. Яненко, Е.Г. Дьяконов [50, 54, 72, 62, 27, 89].

В отличие от метода переменных направлений, в упомянутых методах удалось отказаться от требования аппроксимации, т.е. на каждом шаге аппроксимация может отсутствовать. Вместо этого требуется, чтобы имела место суммарная аппроксимация (аппроксимация на последнем шаге).

Вначале расщепление осуществлялось на основе представления исходного эллиптического оператора в виде суммы одномерных операторов. Через некоторое время появился новый подход к построению схем расщепления, основанный на представлении эллиптического оператора в виде суммы операторов, необязательно одномерных. В отличие от представления эллиптического в виде суммы одномерных операторов (*расщепление по координатам*), второй метод называют *расщеплением по физическим процессам* ввиду наличия физической интерпретации этого метода.

1.1. Схемы расщепления по пространственным переменным

Идею метода расщепления будем излагать на примере метода прямых для уравнения теплопроводности с несколькими пространственными переменными. Оставляя в стороне вопрос об аппроксимации по пространственным переменным, рассмотрим эволюционную задачу

$$\frac{du}{dt} + Au = f, \quad A = -\Delta, \quad u(x, 0) = u_0(x). \quad (1.1)$$

Для простоты полагаем, что уравнение рассматривается в параллелепипеде, а на границе задано однородное условие Дирихле, причем правая часть f равна нулю на границе.

Представим оператор A в виде

$$A = \sum_{\alpha=1}^n A_{\alpha}, \quad A_{\alpha} = -\frac{\partial^2}{\partial x_{\alpha}^2}, \quad n \geq 2; \quad (1.2)$$

здесь $x = (x_1, \dots, x_n)$. Простейший алгоритм расщепления, основанный на использовании неявных схем первого порядка точности по времени, имеет вид

$$\frac{u^{k+1/n} - u^k}{\tau} + A_1 u^{k+1/n} = 0,$$

.....

$$\frac{u^{k+1} - u^{k+(n-1)/n}}{\tau} + A_n u^{k+1} = f^k; \quad (1.3)$$

$$u^0 = u_0, \quad k = 0, 1, \dots \quad (1.4)$$

Здесь f^k — значение функции f на k -м слое. Переход от $u^{k+\alpha/n}$ к $u^{k+(\alpha+1)/n}$ называется *дробным шагом*, а переход от u^k к u^{k+1} — *шагом схемы расщепления*.

Покажем, что схема (1.3)–(1.4) является безусловно устойчивой. Действительно, умножим уравнение

$$\frac{u^{k+\alpha/n} - u^{k+(\alpha-1)/n}}{\tau} + A_\alpha u^{k+\alpha/n} = 0$$

скалярно на $2\tau u^{k+\alpha/n}$. Имеем

$$\begin{aligned} & \|u^{k+\alpha/n}\|^2 - \|u^{k+(\alpha-1)/n}\|^2 + \|u^{k+\alpha/n} - u^{k+(\alpha-1)/n}\|^2 + \\ & + 2\tau(A_\alpha u^{k+\alpha/n}, u^{k+\alpha/n}) = 0, \quad \alpha = 1, \dots, n-1. \end{aligned}$$

Поскольку операторы A_α положительны, то из последнего соотношения следует

$$\|u^{k+\alpha/n}\| \leq \|u^{k+(\alpha-1)/n}\|, \quad \alpha = 1, \dots, n-1.$$

Отсюда и из соотношения для последнего шага имеем

$$\|u^{k+1}\|^2 \leq \|u^k\|^2 + 2\tau\|f^k\|^2. \quad (1.5)$$

Это означает, что схема расщепления (1.3)–(1.4) безусловно устойчива.

Исследуем аппроксимацию. Каждое из соотношений (1.3)–(1.4) не аппроксимирует исходную задачу (1.1). Покажем, что на самом деле имеет место аппроксимация после полного шага схемы расщепления. Из (1.3)–(1.4) имеем

$$(E + \tau A_\alpha)u^{k+\alpha/n} = u^{k+(\alpha-1)/n}, \quad \alpha = 1, \dots, n-1;$$

после перехода от u^k к u^{k+1} получаем

$$\begin{aligned} (E + \tau A_1)(E + \tau A_2) \dots (E + \tau A_n)u^{k+1} &= u^k + \\ &+ \tau(E + \tau A_1) \dots (E + \tau A_{n-1})f^k. \end{aligned}$$

Раскрывая скобки, имеем

$$\left(E + \tau \sum_{j=1}^n A_\alpha + O(\tau^2)\right)u^{k+1} = u^k + (\tau E + O(\tau^2))f^k. \quad (1.6)$$

Здесь через $O(\tau^2)$ мы обозначили операторы, содержащие множители степени равной или выше τ^2 . Таким образом, из (1.6) следует, что схема формально имеет первый порядок аппроксимации по времени $O(\tau)$.

Не вдаваясь в детали, изложим метод двуциклического многокомпонентного расщепления [50], который имеет второй порядок аппроксимации

по времени. Для простоты будем рассматривать случай двух пространственных переменных для задачи (1.1) при $f = 0$. Схема расщепления в данном случае состоит из следующих частей:

$$\begin{aligned} \frac{u^{k-1/2} - u^{k-1}}{\tau} + A_1 \frac{u^{k-1/2} + u^{k-1}}{2} &= 0, \\ \frac{u^k - u^{k-1/2}}{\tau} + A_2 \frac{u^k + u^{k-1/2}}{2} &= 0; \end{aligned} \quad (1.7)$$

$$\begin{aligned} \frac{u^{k+1/2} - u^k}{\tau} + A_2 \frac{u^{k+1/2} + u^k}{2} &= 0, \\ \frac{u^{k+1} - u^{k+1/2}}{\tau} + A_1 \frac{u^{k+1} + u^{k+1/2}}{2} &= 0. \end{aligned} \quad (1.8)$$

Один шаг метода состоит в поочередном применении (1.7) и (1.8).

Исключая из соотношений (1.7), (1.8) промежуточные шаги, получим соотношение между u^{k+1} и u^k :

$$u^{k+1} = T u^{k-1}, \quad (1.9)$$

где

$$\begin{aligned} T &= \left(E + \frac{\tau}{2} A_1\right)^{-1} \left(E - \frac{\tau}{2} A_1\right) \left(E + \frac{\tau}{2} A_2\right)^{-1} \left(E - \frac{\tau}{2} A_2\right) \times \\ &\times \left(E + \frac{\tau}{2} A_2\right)^{-1} \left(E - \frac{\tau}{2} A_2\right) \left(E + \frac{\tau}{2} A_1\right)^{-1} \left(E - \frac{\tau}{2} A_1\right). \end{aligned} \quad (1.10)$$

Схема (1.9) (или, что то же самое, схема (1.7), (1.8)) аппроксимирует исходную задачу со вторым порядком по τ . Если операторы A_2 неотрицательны, то $\|T\| \leq 1$.

Устойчивость схемы следует из цепочки очевидных оценок

$$\|u^{k+1}\| \leq \|T\| \|u^{k-1}\| \leq \dots \|u^0\|.$$

Особо важную роль играют схемы расщепления для систем уравнений, не являющихся системами типа Коши-Ковалевской. Пусть

$$\begin{aligned} \frac{du}{dt} + Au + Lv = f, \quad A = \sum_{\alpha=1}^n A_{\alpha}, \\ K_1 u + K_2 v = g. \end{aligned} \quad (1.11)$$

Идея построения схемы расщепления Марчука состоит в следующем. Первое уравнение (1.11) может рассматриваться как эволюционное уравнение

с правой частью $F = f - Lv$. Поскольку для этого уравнения схемы расщепления уже разработаны, то их можно использовать в качестве схем для задачи (1.11). Более конкретно, для решения (1.11) рассмотрим следующую схему:

$$\begin{aligned}
 & \frac{u^{k+1/n} - u^k}{\tau} + \Lambda_{10}^k u^k + \Lambda_{11}^k u^{k+1/n} = B_1 F^k, \\
 & \frac{u^{k+2/n} - u^{k+1/n}}{\tau} + \Lambda_{20}^k u^k + \Lambda_{21}^k u^{k+1/n} + \Lambda_{22}^k u^{k+2/n} = B_2 F^k, \\
 & \dots\dots\dots \\
 & \frac{u^{k+1} - u^{k+(n-1)/n}}{\tau} + \Lambda_{n0}^k u^k + \Lambda_{n1}^k u^{k+1/n} + \dots + \Lambda_{nn}^k u^{k+1} = \\
 & \qquad \qquad \qquad = B_n F^{k+1}, \tag{1.12}
 \end{aligned}$$

где операторы $\Lambda_{\alpha\beta}$ и B_α такие, что $\sum_{\beta=0}^{\alpha} \Lambda_{\alpha\beta}^k$ аппроксимирует оператор A_α , а $\sum_{\alpha=1}^n B_\alpha$ аппроксимирует тождественный оператор E . При решении (1.12) для замыкания алгоритма необходимо на последнем дробном шаге использовать разностный аналог второго уравнения (1.11)

$$K_1^k u^{k+1} + K_2^k v^{k+1} = g^{k+1}; \tag{1.13}$$

здесь K_α^k — аппроксимация оператора K_α , а F^k и F^{k+1} — аппроксимации $(f - Lv)$ на k -м и $k + 1$ -м слоях.

1.2. Расщепление по физическим процессам

Идея метода расщепления по физическим процессам состоит в сведении исходной задачи к последовательности задач, описывающих более простые физические процессы. Особенно эффективен этот метод оказался в применении к решению уравнений типа Навье–Стокса. Ниже мы приведем ряд наиболее используемых схем расщепления для этой задачи.

Схема Чорина. Схема Чорина является одной из первых схем расщепления по физическим процессам для системы уравнений Навье–Стокса. Алгоритм схемы состоит в следующем.

1. Задаемся некоторым $\mathbf{u}^0 \equiv \mathbf{u}_0$.
2. При $n \geq 0$ для известного \mathbf{u}^n вычисляем $\tilde{\mathbf{u}}$ как решение линейной задачи с некоторой правой частью \mathbf{f}^{n+1} :

$$\begin{aligned}
 & \frac{\tilde{\mathbf{u}} - \mathbf{u}^n}{\tau} - \nu \Delta \tilde{\mathbf{u}} + u_k^n \tilde{\mathbf{u}}_{x_k} = \mathbf{f}^{n+1}, \\
 & \tilde{\mathbf{u}} \Big|_{\partial G} = \mathbf{0}. \tag{1.14}
 \end{aligned}$$

3. Вычисляем \mathbf{u}^{n+1} , p^{n+1} как решение следующей задачи:

$$\begin{aligned} \frac{\mathbf{u}^{n+1} - \tilde{\mathbf{u}}}{\tau} + \nabla p^{n+1} &= 0, \\ \operatorname{div} \mathbf{u}^{n+1} &= 0, \quad \mathbf{u}^{n+1} \cdot \mathbf{n} \Big|_{\partial G} = 0. \end{aligned} \quad (1.15)$$

Задача (1.15) может рассматриваться как разностная аппроксимация по времени системы уравнений типа Соболева.

Обоснование подобного рода схем расщепления для уравнений Навье–Стокса было предложено Раннахером [96]. Опишем эту идею на примере линейной задачи Стокса:

$$\begin{aligned} \mathbf{u}_t - \nu \Delta \mathbf{u} + \nabla p &= 0, \\ \operatorname{div} \mathbf{u} &= 0, \quad \mathbf{u}(x, 0) = \mathbf{u}_0(x), \quad \mathbf{u} \Big|_{\partial G} = 0. \end{aligned} \quad (1.16)$$

Один шаг дискретизации по времени схемы Чорина для задачи (1.16) заключается в следующем. Пусть найдено приближение \mathbf{u}^n , p^n для решения $\mathbf{u}(t)$, $p(t)$ в момент времени $t = n\tau$. Нахождение \mathbf{u}^{n+1} , p^{n+1} осуществляется в два шага. Вначале находится промежуточное значение $\tilde{\mathbf{u}}$ по формуле

$$\frac{\tilde{\mathbf{u}} - \mathbf{u}^n}{\tau} - \nu \Delta \tilde{\mathbf{u}} = \mathbf{f}^{n+1}. \quad (1.17)$$

После этого значения \mathbf{u}^{n+1} , p^{n+1} вычисляются как решение задачи (1.15).

Применяя к первому уравнению (1.15) оператор div и учитывая второе уравнение (1.15) и граничное условие, получим краевую задачу относительно p^{n+1} :

$$\Delta p^{n+1} = \frac{1}{\tau} \operatorname{div} \tilde{\mathbf{u}}, \quad \frac{\partial p^{n+1}}{\partial n} \Big|_{\partial G} = 0. \quad (1.18)$$

После нахождения p^{n+1} из (1.18), значение \mathbf{u}^{n+1} находится из первого уравнения (1.15) простым пересчетом:

$$\mathbf{u}^{n+1} = \tilde{\mathbf{u}} - \tau \nabla p^{n+1}. \quad (1.19)$$

Второй шаг (задача (1.15)) можно рассматривать как шаг проектирования вычисленного значения $\tilde{\mathbf{u}}$ на подпространство соленоидальных (т.е., удовлетворяющих условию $\operatorname{div} \mathbf{v} = 0$) вектор-функций.

Основная идея [96] обоснования этого метода состоит в исключении из формул (1.17), (1.15) значения \mathbf{u}^{n+1} (а не $\tilde{\mathbf{u}}$, что кажется более естественным). Действительно, из (1.15) имеем $\mathbf{u}^n = \tilde{\mathbf{u}}^{n-1} - \tau \nabla p^n$. Подставляя это выражение в (1.17) и добавляя в качестве второго уравнения (1.18), получим

$$\begin{aligned} \frac{\tilde{\mathbf{u}}^n - \tilde{\mathbf{u}}^{n-1}}{\tau} - \nu \Delta \tilde{\mathbf{u}}^n - \nabla p^n &= \mathbf{f}^{n+1}, \\ \operatorname{div} \tilde{\mathbf{u}}^n - \tau \Delta p^n &= 0, \quad \frac{\partial p^n}{\partial n} \Big|_{\partial G} = 0. \end{aligned} \quad (1.20)$$

Задача (1.20) может рассматриваться как аппроксимация по времени возмущенной задачи Стокса

$$\begin{aligned} \frac{\partial \mathbf{u}^\varepsilon}{\partial t} - \nu \Delta \mathbf{u}^\varepsilon - \nabla p^\varepsilon &= \mathbf{f}, & \mathbf{u}^\varepsilon|_{\partial G} &= \mathbf{0}, \\ \operatorname{div} \mathbf{u}^\varepsilon - \varepsilon \Delta p^\varepsilon &= 0, & \frac{\partial p^\varepsilon}{\partial n} \Big|_{\partial G} &= 0; \end{aligned} \quad (1.21)$$

здесь $\varepsilon = \tau$. Таким образом, исследование близости решений исходной задачи (1.16) и (1.17), (1.15) сводится к изучению близости решения задач (1.16) и (1.21) и исследованию аппроксимационных свойств схемы (1.20) для задачи (1.21). Вторая задача носит классический характер и может быть исследована с помощью стандартной техники метода конечных разностей или конечных элементов. Что касается оценки близости решений задач (1.16) и (1.21), то имеет место оценка

$$\max_t \{ \|\mathbf{u}^\varepsilon - \mathbf{u}\| + \sqrt{\varepsilon} \|\mathbf{u}^\varepsilon - \mathbf{u}\|_1 + \sqrt{\varepsilon} \|p^\varepsilon - p\| \} \leq c\varepsilon, \quad (1.22)$$

означающая, что схема Чорина дает решение исходной задачи с точностью $O(\tau)$ для скоростей в норме \mathbf{L}_2 и $O(\sqrt{\tau})$ для скоростей в норме \mathbf{H}^1 и давления в норме L_2 .

Схема Чорина–Узавы. Алгоритм одного шага этой схемы состоит в следующей последовательности дробных шагов.

1. Пусть $\mathbf{u}^n, p^n, \tilde{p}^n$ известны. В начальный момент времени $t = 0$ полагаем $\tilde{p}^0 = 0$. Находим вспомогательную функцию $\tilde{\mathbf{u}}^{n+1}$ как обобщенное решение задачи

$$\begin{aligned} \frac{\tilde{\mathbf{u}}^{n+1} - \mathbf{u}^n}{\tau} - \Delta \tilde{\mathbf{u}}^{n+1} + \nabla(p^n - \tilde{p}) &= \mathbf{f}^{n+1}, \\ \tilde{\mathbf{u}} \Big|_{\partial G} &= \mathbf{0}. \end{aligned} \quad (1.23)$$

2. Вычисляем проекцию $\tilde{\mathbf{u}}^{n+1}$ на подпространство соленоидальных функций в \mathbf{H}_0^1 :

$$\begin{aligned} \frac{\mathbf{u}^{n+1} - \tilde{\mathbf{u}}^{n+1}}{\tau} + \nabla \tilde{p}^{n+1} &= \mathbf{0}, \\ \operatorname{div} \mathbf{u}^{n+1} = 0, & \quad \mathbf{u}^{n+1} \cdot \mathbf{n} \Big|_{\partial G} = 0. \end{aligned} \quad (1.24)$$

3. Давление вычисляется по формуле

$$p^{n+1} = p^n - \alpha \operatorname{div} \tilde{\mathbf{u}}^{n+1}, \quad 0 < \alpha < 1. \quad (1.25)$$

В отличие от схемы Чорина, этот метод дает несколько лучшую оценку точности для вектор-функции скоростей в норме L_2 по времени.

Схема Ван Кана. Приведем схему расщепления, обладающую более высоким порядком аппроксимации.

1. Полагаем $\mathbf{u}^0 \equiv \mathbf{u}_0$ и задаемся некоторым $p^0 \approx p(0)$.

2. Для $n \geq 0$ при заданных \mathbf{u}^n , $\tilde{\mathbf{u}}^n$, p^n находим $\tilde{\mathbf{u}}^{n+1}$ как решение задачи Гельмгольца

$$\frac{\tilde{\mathbf{u}}^{n+1} - \mathbf{u}^n}{\tau} - \frac{\nu}{2} \Delta (\tilde{\mathbf{u}}^{n+1} + \tilde{\mathbf{u}}^n) + \frac{3}{2} u_k^n u_{x_k}^n = \mathbf{f}^{n+1/2} - \nabla p^n + \frac{1}{2} u_k^{n-1} u_{x_k}^{n-1},$$

$$\tilde{\mathbf{u}}^{n+1} \Big|_{\partial G} = 0. \quad (1.26)$$

3. Вычисляем \mathbf{u}^{n+1} , p^{n+1} как решение краевой задачи

$$\frac{\mathbf{u}^{n+1} - \tilde{\mathbf{u}}^{n+1}}{\tau} + \frac{1}{2} \nabla (p^{n+1} - p^n) = 0,$$

$$\operatorname{div} \mathbf{u}^{n+1} = 0, \quad \mathbf{u}^{n+1} \cdot \mathbf{n} \Big|_{\partial G} = 0. \quad (1.27)$$

На первом шаге $n = 0$ следует явную аппроксимацию нелинейных членов заменить неявной с равными весами по двум слоям.

Для задачи (1.26), (1.27) имеет место оценка [116]

$$\max_{0 \leq t = n\tau \leq T} \left(\|\mathbf{u}^n - \mathbf{u}\|_1 + \|\mathbf{u}^n - \mathbf{u}\|^{2-\delta} + \|p^n - p\| \right) \leq c\tau. \quad (1.28)$$

При обосновании этой оценки важную роль играет тот факт, что начальное условие для давления должно быть найдено с достаточно высокой точностью. А именно,

$$\|p(0) - p^0\|_1 \leq c\tau^2. \quad (1.29)$$

В последнее время достигнут значительный прогресс в численном решении задачи Стокса. Это означает, что можно рассматривать схемы расщепления, где на одном или нескольких дробных шагах необходимо решать задачу Стокса. Одна из таких схем расщепления, имеющая второй порядок точности, была предложена в [91]. Приведем эту схему для случая линейной задачи Стокса. Она состоит из трех дробных шагов. При этом первый и последний шаги совпадают с точностью до правой части:

$$\frac{\mathbf{u}^{n+\vartheta} - \mathbf{u}^n}{\vartheta\tau} - \alpha\nu\Delta\mathbf{u}^{n+\vartheta} - \beta\nu\Delta\mathbf{u}^n + \nabla p^{n+\vartheta} = \mathbf{f}^n,$$

$$\operatorname{div} \mathbf{u}^{n+\vartheta} = 0; \quad (1.30)$$

$$\frac{\mathbf{u}^{n+1-\vartheta} - \mathbf{u}^{n+\vartheta}}{\vartheta'\tau} - \alpha\nu\Delta\mathbf{u}^{n+\vartheta} - \beta\nu\Delta\mathbf{u}^{n+1-\vartheta} + \nabla p^{n+\vartheta} = \mathbf{f}^{n+1-\vartheta}; \quad (1.31)$$

$$\frac{\mathbf{u}^{n+1} - \mathbf{u}^{n+1-\vartheta}}{\vartheta\tau} - \alpha\nu\Delta\mathbf{u}^{n+1} - \beta\nu\Delta\mathbf{u}^{n+1-\vartheta} + \nabla p^{n+1} = \mathbf{f}^{n+1-\vartheta},$$

$$\operatorname{div} \mathbf{u}^{n+1} = 0. \quad (1.32)$$

Здесь $\vartheta = 1 - \sqrt{2}/2$, $\vartheta' = 1 - 2\vartheta$, а параметры процесса α и β выбираются из условий $\alpha \in (1/2, 1]$, $\beta = 1 - \alpha$.

Несомненным достоинством схемы является ее второй порядок аппроксимации. К недостаткам следует отнести следующие моменты. Во-первых, на первом и третьем шагах необходимо решать не обычную, а обобщенную задачу Стокса (задача Стокса с дополнительным членом γu в левой части уравнения движения, где $\gamma = 1/\tau$ — большой параметр). Классические методы решения задачи Стокса для этой задачи имеют скорость сходимости, зависящую от γ , что влечет за собой увеличение шагов внутреннего итерационного метода. В [101] предложен итерационный метод с переобуславливателем, являющимся аппроксимацией псевдодифференциального оператора, скорость сходимости которого не зависит от коэффициента γ . Этот метод требует на каждом шаге решения двух (в трехмерном случае — трех) задач Дирихле для уравнения Пуассона.

Во-вторых, при решении нелинейной задачи нелинейные члены входят только в правую часть второго дробного шага. Если их брать на нижнем слое, то это влечет существенное ограничение на шаг по времени. Заметим, что (1.31) может рассматриваться как разностная схема для системы нелинейных параболических уравнений

$$\frac{\partial \mathbf{v}}{\partial t} - \alpha \nu \Delta \mathbf{v} - \beta \nu \Delta \mathbf{u}^{n+1-\vartheta} + \nabla p^{n+\vartheta} = \mathbf{f}, \quad (1.33)$$

где нелинейные члены входят в правую часть. Последняя задача может быть решена прямым численным интегрированием явными методами с использованием переменного шага по времени [39]. При этом ограничения на шаг по времени становятся существенно слабее.

2. Многосеточный метод Федоренко

При численном решении задач математической физики возникает проблема выбора эффективного метода решения системы алгебраических уравнений, полученной в результате дискретизации исходной дифференциальной задачи. Для получения хорошей аппроксимации сетка должна быть мелкой и содержать большое количество узлов. В результате число неизвестных алгебраической системы может достигать сотен тысяч и даже миллионов, особенно при решении трехмерных задач.

Для решения больших систем алгебраических уравнений используют те или иные итерационные методы. Важной характеристикой итерационных методов является зависимость вычислительных затрат, требуемых для получения решения с заданной точностью, от числа неизвестных. В большинстве случаев многосеточный метод является единственным методом,

позволяющим достичь точности, обусловленной порядком сходимости решения дискретной задачи, с верхней оценкой числа арифметических операций, линейно зависящей от количества неизвестных. Другими словами, многосеточный метод требует $O(N)$ арифметических операций для решения системы уравнений с N неизвестными. Таким образом, многосеточный метод оказывается практически неулучшаемым по вычислительным затратам для достижения заданной точности. К настоящему времени многосеточный метод стал одним из самых известных и употребительных методов вычислительной математики.

Основная идея многосеточного метода была впервые предложена Р.П. Федоренко в 1961 г. Позже [68] он обосновал сходимость многосеточного метода для конечно-разностного аналога уравнения Пуассона в квадрате. В 1966 г. Н.С. Бахвалов [6] доказал оптимальность метода по числу арифметических операций для достижения точности, согласованной с порядком сходимости, и рассмотрел случай произвольного эллиптического уравнения в квадрате. В 1971 г. Г.П. Астраханцев [3] обосновал применимость многосеточного метода для случая метода конечных элементов в произвольной области. В итоге по асимптотическим оценкам эффективности метод опередил известные алгоритмы, но логическая сложность и громоздкое математическое обоснование на некоторое время завуалировали его достоинства.

Признание за рубежом многосеточный метод (multigrid method) получил во второй половине 1970-х годов после выхода работ А. Брандта и В. Хакбуша [83, 95]. С начала 1980-х годов произошел огромный всплеск интереса к многосеточному методу, и многочисленные исследователи продолжили его разработку. В это время была заложена современная теория метода. Привлечение нового математического аппарата позволило упростить обоснование многосеточного метода, сделать структуру доказательств сходимости более прозрачной.

В настоящее время многосеточный метод получил широкое развитие и распространение [95, 82, 83, 77, 80, 81, 122, 113] и стал наиболее применяемым методом решения дискретных задач, возникающих при аппроксимации задач математической физики.

2.1. Простейший пример

Основную идею многосеточного метода проще всего понять на простейшем модельном примере. Рассмотрим краевую задачу для одномерного уравнения:

$$-u''(x) = f(x) \quad \text{на } (0, 1), \quad (2.1)$$

$$u(0) = u(1) = 0. \quad (2.2)$$

На отрезке $[0, 1]$ введем сетку Ω_h с шагом $h = 1/n$, состоящую из узлов $x_j = jh$, $j = 0, \dots, n$. Во внутренних узлах сетки Ω_h заменим исходное дифференциальное уравнение (2.1) сеточным. С учетом краевых условий (2.2) получим систему линейных алгебраических уравнений:

$$\begin{aligned} -v_{h,j-1} + 2v_{h,j} - v_{h,j+1} &= h^2 f(x_j), \quad j = 1, \dots, n-1, \\ v_{h,0} &= v_{h,n} = 0, \end{aligned}$$

где $v_{h,j}$ — приближенное значение точного решения $u(x_j)$. После исключения граничных значений эту систему можно записать в матричной форме

$$L_h v_h = f_h, \quad (2.3)$$

где $v_h = (v_{h,1}, \dots, v_{h,n-1})^T$ — неизвестный вектор, $f_h = h^2(f(x_1), \dots, f(x_{n-1}))^T$ — вектор правой части и

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

— матрица размерности $(n-1) \times (n-1)$.

Обозначим через y_h^m приближенное решение системы уравнений (2.3), полученное методом простых итераций с заданным начальным приближением y_h^0 , т.е.

$$y_h^{i+1} = S_h y_h^i + \frac{1}{4} f_h, \quad i = 0, \dots, m-1,$$

где матрица итераций имеет вид

$$S_h = I - \frac{1}{4} L_h.$$

Обозначим через $e_h^i = v_h - y_h^i$ погрешность i -го приближения. Тогда

$$e_h^m = S_h^m e_h^0. \quad (2.4)$$

Разложим погрешность начального приближения в конечный ряд Фурье:

$$e_h^0 = \sum_{k=1}^{n-1} c_k \sin(jk\pi h). \quad (2.5)$$

Функции $\sin(ik\pi h)$, $k = 1, \dots, n-1$, являются собственными функциями L_h :

$$L_h \sin(jk\pi h) = \lambda_k(L_h) \sin(jk\pi h),$$

где

$$\lambda_k(L_h) = 4 \sin^2 \frac{k\pi h}{2}, \quad n = 1, \dots, n-1.$$

Из (2.4) и (2.5) следует

$$e_h^m = \sum_{k=1}^{n-1} c_k S_h^m \sin(jk\pi h). \quad (2.6)$$

Собственные функции S_h и L_h совпадают, а их собственные значения связаны соотношением

$$\lambda_k(S_h) = 1 - \frac{1}{4} \lambda_k(L_h), \quad k = 1, \dots, n-1.$$

Поэтому соотношение (2.6) можно переписать в виде

$$e_h^m = \sum_{k=1}^{n-1} c_k \lambda_k^m(S_h) \sin(jk\pi h).$$

Множитель $\lambda_k^m(S_h)$ показывает, насколько уменьшился вклад k -й составляющей погрешности начального приближения в разложении (2.5) после m итераций. При $k \geq n/2$ имеем

$$|\lambda_k(S_h)| = \left| 1 - \frac{1}{4} \lambda_k(L_h) \right| \leq \frac{1}{2}, \quad k \geq \frac{n}{2},$$

т.е. на каждой итерации коэффициенты высокочастотных составляющих погрешности уменьшаются по меньшей мере вдвое.

В результате уже после нескольких итераций коэффициенты при высокочастотных гармониках становятся близки к нулю, а основной вклад в погрешность вносят низкочастотные составляющие. Таким образом, погрешность становится плавно меняющейся функцией ("сглаживается").

Пусть y_h — приближенное решение системы уравнений (2.3). Тогда погрешность $e_h = v_h - y_h$ удовлетворяет системе уравнений

$$L_h e_h = r_h, \quad (2.7)$$

где $r_h = f_h - L_h y_h$ — невязка. Если бы мы могли найти решение e_h из системы (2.7), то мы могли бы определить и решение системы (3) из условия $v_h = y_h + e_h$. Однако решить систему (2.7) так же трудно, как и исходную.

Когда после нескольких сглаживающих итераций погрешность становится плавно меняющейся функцией, она может быть хорошо аппроксимирована на грубой сетке. Таким образом, вместо решения системы (2.7) мы будем искать приближение к e_h на более грубой сетке.

Введем сетку Ω_{2h} с шагом $2h$ и рассмотрим на ней задачу (2.7):

$$L_{2h}e_{2h} = r_{2h}. \quad (2.8)$$

Размерность системы уравнений (2.8) вдвое меньше, чем системы (2.7). Предположим, что мы можем найти ее точное решение e_{2h} . Обозначим через e_h^* интерполяцию e_{2h} с сетки Ω_{2h} на сетку Ω_h . С целью погашения ошибки интерполяции, имеющей высокочастотный характер, вновь проведем несколько сглаживающих итераций для (2.7) с начальным приближением e_h^* , получая вместо e_h поправку \bar{e}_h . Новое приближение w_h к решению v_h системы (2.3) определим следующим образом:

$$w_h = y_h + \bar{e}_h.$$

Этот шаг называется коррекцией с грубой сетки.

Таким образом, нахождение приближенного решения системы (2.3) состоит из следующих этапов.

1. Выполнить несколько сглаживающих итераций и получить приближенное решение y_h .
2. Вычислить невязку $r_h = f_h - L_h y_h$ и спроектировать ее на грубую сетку.
3. Решить систему уравнений (2.8) на грубой сетке.
4. Построить интерполяцию e_h^* решения системы (2.8) на мелкую сетку.
5. Выполнить несколько сглаживающих итераций для e_h^* и получить приближенную поправку \bar{e}_h .
6. Вычислить новое приближенное решение

$$w_h = y_h + \bar{e}_h.$$

Точное решение системы (2.8) может оказаться весьма трудоемким. Поэтому целесообразно вместо точного решения системы (2.8) найти ее приближенное решение, вновь используя двухсеточный метод с переходом к еще более грубой сетке. Многосеточный метод состоит в рекурсивном повторении этой процедуры до тех пор, пока не будет достигнута сетка, на которой система уравнений может быть эффективно решена прямым методом.

Формализуем идею многосеточного метода. Построим последовательность сеток $\Omega_0, \Omega_1, \dots, \Omega_l$. Самой грубой сетке Ω_0 соответствует шаг h_0 . Шаг h_i сетки Ω_i определяется соотношением $h_i = h_0 2^{-i}$, $i = 1, \dots, l$. Пусть задана последовательность конечномерных векторных пространств

$$M_0, M_1, \dots, M_l,$$

соответствующих этим сеткам. Предположим, что заданы операторы интерполяции

$$I_i : M_i \rightarrow M_{i+1}, \quad i = 0, \dots, l-1,$$

проектирования

$$R_i : M_i \rightarrow M_{i-1}, \quad i = 1, \dots, l,$$

а также операторы

$$L_i : M_i \rightarrow M_i, \quad i = 0, \dots, l.$$

Нашей целью является решение системы уравнений

$$L_l v_l = f_l \quad (2.9)$$

на самой мелкой сетке. Сформулируем алгоритм для любого уровня $i = 0, \dots, l$. Одна итерация многосеточного метода для решения системы уравнений

$$L_i v_i = f_i$$

определяется рекурсивной процедурой $w^{new} = MG(i, w^{old}, f_i)$. Здесь w^{old} — начальное приближение к v_i , w^{new} — полученное новое приближение. MG -алгоритм состоит в следующем: если $i = 0$, то

0) $w^{new} = L_0^{-1} f_0$ (точное решение на самой грубой сетке), далее выход из процедуры. Иначе ($i > 0$):

1) полагаем

$$u_0 = w^{old} \quad u_{k+1} = u_k - D_i(L_i u_k - f_i), \quad k = 0, \dots, m_1 - 1 \quad (2.10)$$

(m_1 предлагающих итераций);

2) $g_{i-1} = R_i(f_i - L_i u_{m_1})$ (проектирование невязки на грубую сетку);

3) $z_0 = 0$, $z_s = MG(i-1, z_{s-1}, g_{i-1})$, $s = 1, \dots, \gamma$ (γ итераций многосеточного метода на грубой сетке);

4) $y_0 = u_{m_1} + I_{i-1} z_\gamma$ (коррекция приближения u_{m_1} с грубой сетки);

5) вычисляем

$$y_{k+1} = y_k - D_i(L_i y_k - f_i), \quad k = 0, \dots, m_2 - 1 \quad (2.11)$$

(m_2 постсглаживающих итераций); полагаем

$$w^{new} = MG(i, w^{old}, f_i) = y_{m_2}.$$

Одно выполнение процедуры $w^{new} = MG(e, w^{old}, f_e)$ является одной итерацией многосеточного метода. Всего выполняют заданное количество итераций или пока норма невязки $\|f_e - L_e w^{new}\|$ не станет меньше заданного ϵ .

Примеры итераций:

1) $w^{old} = 0$;

2) пока $\|f_e - L_e w^{old}\| \geq \epsilon$, выполнять: $\{w^{new} = MG(e, w^{old}, f_e), w^{old} = w^{new}\}$.

Число предсглаживающих итераций m_1 и постсглаживающих итераций m_2 , а также число рекурсивных вызовов процедуры γ задают заранее. Обычно выбирают $\gamma = 1$ или $\gamma = 2$. При $\gamma = 1$ одна итерация метода получила название V -цикла, при $\gamma = 2$ — W -цикла. Рисунки 1 и 2, на которых показана последовательность перехода с одной сетки на другую, иллюстрируют эти названия.

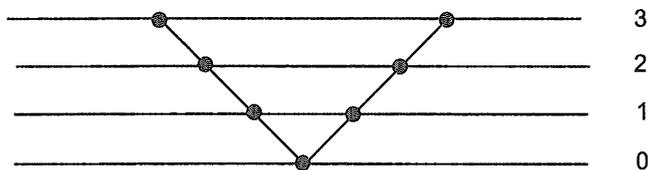


Рис. 1. V -цикл

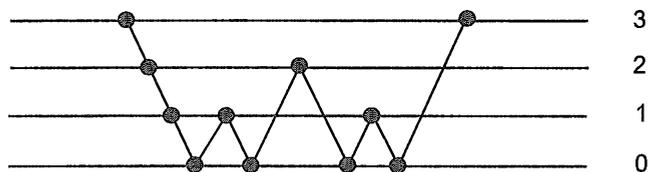


Рис. 2. W -цикл

Число предсглаживающих и постсглаживающих итераций в MG -алгоритме может быть различным. Например, можно использовать только предсглаживающие итерации ($m_2 = 0$). Такой вариант алгоритма соответствует несимметричному V - или W -циклу. В этом случае отсутствует возможность подавления погрешности, внесенной интерполяцией. Однако если все итерации выполняются на этапе предсглаживания, высокочастотная составляющая погрешности может быть уменьшена до такой степени, что вклад погрешности интерполяции не будет иметь существенного значения. Если $m_1 = m_2$ и на этапах пред- и постсглаживания используются одинаковые итерационные методы, то мы имеем симметричный V - или W -цикл.

Численные эксперименты показали, что в тех случаях, когда применимы оба алгоритма, V -цикл обычно более эффективен, чем W -цикл, причем предпочтителен симметричный вариант V -цикла. Кроме того, V -цикл допускает возможность проведения параллельных вычислений. Однако W -цикл имеет более широкую область применения, чем V -цикл. Например, он может использоваться при решении знаконеопределенных и несамосопряженных задач.

2.2. Современное состояние вопроса

Простейшей версией многосеточного метода является каскадный алгоритм, который не содержит проектирования и возврата на более грубые сетки. На последовательности сеток $\Omega_0, \dots, \Omega_l$ строится последовательность задач

$$L_i v_i = f_i, \quad i = 0, \dots, l,$$

аппроксимирующих исходную дифференциальную задачу, при этом цель состоит в решении задачи (2.9) на самой мелкой сетке.

Каскадный алгоритм начинается на самой грубой сетке, где размерность системы уравнений невелика, что позволяет эффективно решить ее прямым методом. На более мелких сетках приближенное решение получается итерационным методом, при этом в качестве начального приближения берется интерполяция приближенного решения с предыдущей более грубой сетки.

Каскадный алгоритм:

1. $u_0 = L_0^{-1} f_0$ (точное решение на самой грубой сетке).

2. Для $i = 1, \dots, l$ цикл

{ 2.1. $w_i = I_{i-1} u_{i-1}$ (интерполяция с грубой сетки);

2.2. вычисляем

$$y_0 = w_i,$$

$$y_{k+1} = y_k - D_{i,k+1}(L_i y_k - f_i), \quad k = 0, \dots, m_i - 1 \quad (2.12)$$

(m_i сглаживающих итераций);

2.3. полагаем $u_i = y_{m_i}$; }.

Итогом этого алгоритма является u_l как приближение к решению v_l задачи (2.9).

Ключевыми пунктами обоснования сходимости каскадного алгоритма являются специальный выбор итерационной процедуры сглаживания для обеспечения сходимости и определенное правило роста числа итераций m_i от верхнего уровня к нижнему при ограничении суммарного числа арифметических операций. Этот алгоритм более чувствителен к выбору числа итераций и итерационных параметров и без их оптимизации не сходится. При их оптимальном подборе он, как и V -цикл, обладает минимальным порядком $O(N)$ числа арифметических операций для решения задачи размерности N .

Логическая структура каскадного алгоритма проще, чем в V - и W -циклах. Поэтому в задачах, где он обоснован, его применение на многопроцессорных вычислительных системах более эффективно, чем у других модификаций многосеточного метода.

С 1990-х годов развитие метода идет в двух направлениях, которые более просто можно пояснить на примере решения задачи (2.9) на уровне l . В принципе, оператор многосеточного метода можно представить в виде переобуславливателя B_l^{-1} итерационного процесса

$$u_{s+1} = u_s - \tau_s B_l^{-1}(L_l u_s - f_l), \quad s = 0, 1, \dots \quad (2.13)$$

Первое направление связано с поиском матриц D_i в итерационных процессах (2.10), (2.11), возможно зависящих от номера итераций k как в (2.12), не только для обеспечения сходимости сглаживающих итераций, но и для схожести операторов B_l^{-1} и L_l^{-1} на классе допустимых сеточных решений. В понятие допустимости могут входить сеточные аналоги гладкости искомого решения. С этой целью для самосопряженных положительно определенных операторов L_l операторы $D_{i,l}$ берутся в виде скалярных множителей $\tau_{i,k}$, которые минимизируют ошибку в энергетической или среднеквадратичной норме. В случае несамосопряженного или знаконеопределенного оператора L_l операторы $D_{i,k}$ берутся в виде $\tau_{i,k} L_l^*$, где знак * указывает на сопряжение оператора. В ряде случаев, включая сеточные аппроксимации стационарных задач теплопроводности, Навье–Стокса, теории упругости, удалось подобрать эти операторы так, что B_l^{-1} оказался весьма близким к L_l^{-1} , так что τ_s полагают равным 1 и хватает одной-двух итераций вида (2.13), которые с учетом рекурсии порождают соответственно V - или W -циклы.

Второе направление нацелено на более простую структуру оператора B_l^{-1} при условии, что он "похож" на L_l^{-1} в энергетическом смысле:

$$\alpha L_l^{-1} \leq B_l^{-1} \leq \beta L_l^{-1} \quad (2.14)$$

с положительными константами α, β , не зависящими от номера уровня l и числа неизвестных N . Тогда выбором чебышевских параметров τ_s можно достигнуть подавления ошибки в конечное число раз, пропорциональное $1/s^2$. Один из приемов, называемый аддитивным многосеточным методом, состоит в проведении на каждом уровне i одной простой итерации

$$w^i = u^i - \tau_i (L_i u^i - f_i) \quad (2.15)$$

с начальными приближениями $u^l = u^s$, $u^{i-1} = R_i u^i$ и числовыми параметрами τ_i , определяемыми как обратная величина верхней оценки максимального собственного числа. После этого результат действия оператора B_l^{-1} подсчитывается путем суммирования интерполяций полученных решений на уровне l :

$$z^l = B_l^{-1} u_s = w^l + I_{l-1}(w^{l-1} + I_{l-2}(w^{l-2} + \dots + I_0 w^0) \dots).$$

Ширина интервала $[\alpha, \beta]$ в этом методе существенно больше, чем в методах первого направления, использующих большую информацию о гладко-

сти решения, зато требования к "гладкости" сеточного решения ограничиваются лишь существованием обобщенного решения, к которому сходятся приближения (2.13).

В [42] предложен метод, где выбор итерационных параметров при каждом шаге зависит от выбора итерационных параметров при более мелких шагах.

Еще один подход демонстрирует полный отказ от какой-либо гладкости сеточного решения и создает "схожесть" операторов B_l^{-1} и L_l^{-1} за счет учета структуры этих операторов. Поясним основную идею на простом примере сеточного аналога оператора Лапласа. Рассмотрим следующую типичную элементарную матрицу жесткости $A_e = (a_{ij})_{i,j=1}^6$ метода конечных элементов, соответствующую элементам, изображенным на рис. 3. Скалярное про-

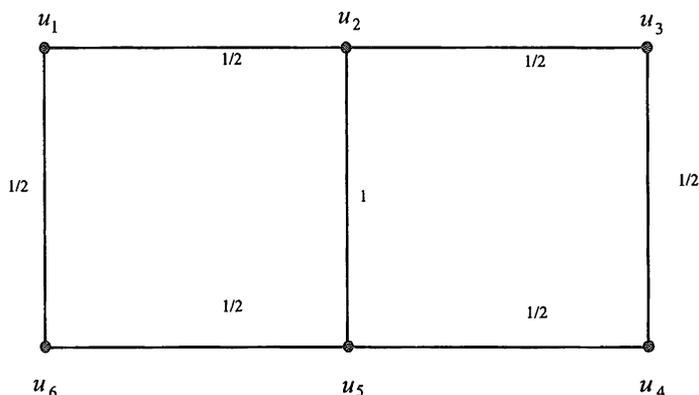


Рис. 3. Макроэлемент с числовыми связями — внедиагональными элементами

изведение $(A_e U, U)$ на векторе $U = (u_1, \dots, u_6)^T$ подсчитывается с учетом указанных связей, помеченных на ребрах графа:

$$(A_e U, U) = \frac{1}{2}(u_1 - u_2)^2 + \dots + \frac{1}{2}(u_6 - u_1)^2 + (u_2 - u_5)^2.$$

Отбросим связь между u_2 и u_5 , что эквивалентно обнулению двух внедиагональных элементов $a_{2,5}$ и $a_{5,2}$, а также уменьшению на единицу диагональных элементов $a_{2,2}$, $a_{5,5}$. Обозначим полученную матрицу жесткости через B_e и заметим, что

$$(B_e U, U) = (A_e U, U) - (u_2 - u_5)^2.$$

Простыми выкладками можно показать, что

$$\alpha A_e \leq B_e \leq \beta A_e.$$

Суммирование этих неравенств по всей области определения дает оценку

$$\alpha L_l \leq B_l \leq \beta L_l$$

для глобального оператора жесткости $B_l = \sum B_e$, эквивалентную (2.14). Отметим, что в системе с оператором B_l неизвестные типа u_2 и u_5 могут быть исключены, что уменьшает число неизвестных примерно вдвое и дает структуру элементарных матриц жесткости, аналогичную исходной. Поэтому возможно дальнейшее повторение такой редукции числа неизвестных. Этот простой пример демонстрирует лишь идею редукции и не является оптимальным. Более оптимальным является шахматное исключение неизвестных или редукция неизвестных в четыре раза за счет использования макроэлементов из четырех треугольников или прямоугольников.

Оба направления становятся более эффективными в трехмерном случае, когда число неизвестных уменьшается примерно в 8 раз (а не в 4 раза, как в двумерном случае) при увеличении шага сетки вдвое. Эта редукция допускает большую вариацию параметров и расширяет область применимости многосеточного метода.

Расширяется также применение многосеточного метода для решения сеточных систем уравнений на каждом шаге по времени у нестационарных задач. Дело в том, что в этом случае метод имеет тот же линейный порядок числа арифметических операций cN , что и явные методы, правда, с большей константой c . Но неявные аппроксимации в некоторых случаях имеют неоспоримые преимущества либо из-за отсутствия существенных ограничений на шаги по времени, либо из-за наличия оценок устойчивости только для неявных схем. В этом случае многосеточные методы применяются как переобуславливатели в (2.13) с начальными приближениями, экстраполируемыми с предыдущих временных шагов. Последнее в значительной мере уменьшает объем вычислений, необходимых для достижения требуемой точности.

В итоге четырех десятилетий развития многосеточный метод стал эффективным средством решения сеточных задач с оптимальной линейной оценкой числа арифметических операций cN , пропорциональной числу неизвестных N для достижения заданной точности. Его различные модификации покрывают довольно широкий спектр сеточных задач и в зависимости от используемых свойств задачи могут быть настроены с различной степенью эффективности (с точки зрения уменьшения константы c , ибо линейный порядок неулучшаем). В последние годы на основе многосеточного метода был создан многоуровневый метод (multilevel method), имеющий более простую структуру и особенно удобный в случае параллельных вычислений [113, 81, 82].

3. Жесткие системы дифференциальных уравнений, устойчивые явные методы и спектральный анализ

3.1. Явные методы решения жестких нестационарных задач

Если аппроксимировать в линейных нестационарных уравнениях математической физики дифференциальные по пространственным переменным операторы конечномерными, то мы получим систему линейных обыкновенных дифференциальных уравнений (нестационарные уравнения метода прямых). Таким образом получаем на отрезке $0 \leq t \leq T$ задачу Коши для системы дифференциальных уравнений n -го порядка вида

$$\frac{du}{dt} = -Au + f, \quad u|_{t=0} = u_0, \quad (3.1)$$

где $u = (u_1, \dots, u_n)$, $u_0 = (u_{01}, \dots, u_{0n})$, $f = (f_1, \dots, f_n)$, A — $n \times n$ -матрица с постоянными коэффициентами. Пусть (λ_i, φ_i) — собственные пары матрицы A , $Sp(A) = \{\lambda_i\}$, $\{\varphi_i\}$ образуют базис в пространстве R^n , а лежащая в правой полуплоскости комплексного переменного $z = x + iy$ часть спектра $\{\lambda_i\}$ принадлежит замкнутой области B , содержащейся в круге K_r минимального радиуса $r > 0$ вида

$$K_r = \{z: |z - r| \leq r\}. \quad (3.2)$$

Область может состоять из нескольких изолированных подобластей. Величина r для задач математической физики есть $O(h^{-p})$, где $p \geq 1$, а h — характерный размер сетки. Она может быть достаточно точно оценена сверху по теореме Гершгорина. Пусть ∂B — граница области B . Назовем задачу (3.1) жесткой, если $rT \gg 1$.

Пусть $t_0 = 0 < t_1 < t_2 < \dots < t_k < t_N \leq T$; $\tau_{k+1} = t_{k+1} - t_k$, $\tau_{k+1} \leq \tau$ — временные шаги, $u_k = u(t_k)$. Для простоты за приближенными решениями мы оставим те же обозначения. Пусть $t_k \leq t \leq t_{k+1}$ и для заданного $\varepsilon > 0$ (точности локальной аппроксимации) мы знаем $\tau > 0$ ($\tau < T - t_0$) такое, что при $|\Delta t| \leq \tau$

$$\frac{du}{dt} = \frac{u(t + \Delta t) - u(t)}{\Delta t} + \eta(t), \quad \|\eta(t)\| \leq \varepsilon \text{ при } t_0 \leq t, \quad t + \Delta t \leq t_N. \quad (3.3)$$

Итак, пусть для выбранной нами точности локальной аппроксимации $\varepsilon > 0$ условиями (3.3) определен размер максимального временного шага τ (вопрос о том, как зависит от ε ошибка, т.е. разность между точным и разностным решениями, — более сложный и здесь мы его не рассматриваем).

Как известно, существуют явные и неявные методы решения нестационарных задач математической физики. В неявном методе для нахождения u_{k+1} приходится решать уравнение. Явный метод свободен от этого недостатка, однако для его устойчивой реализации могут возникнуть сильные ограничения на временные шаги.

Обоснованные неявные методы обладают таким ценным качеством, как устойчивость того или иного типа, которая предполагает выполнение для оператора перехода метода определенных условий. В этих методах временной шаг лимитируется условиями аппроксимации. В явных методах с постоянным по времени шагом интегрирования помимо условий аппроксимации возникают обременительные ограничения на временной шаг, связанные с выполнением условий устойчивости счета. Для явных методов Эйлера с постоянным шагом τ по времени для устойчивого счета должно быть выполнено неравенство $\tau \leq \text{cou} = r^{-1}$, которое накладывает сильное ограничение на размер временного шага.

Рассматриваемые в едином комплексе три проблемы: методы решения жестких систем обыкновенных дифференциальных уравнений большого порядка, разностные или вариационные методы решения задач математической физики и методы распараллеливания для многопроцессорных ЭВМ побуждают еще раз вернуться к исследованию эффективности явных разностных схем, которые в разбираемой ситуации допускают почти идеальное распараллеливание вычислений. Исследуем эффективность и возможности явных устойчивых алгоритмов с переменными шагами τ_{k+1} вида [39, 40, 41, 103, 105]:

$$u_{k+1} = u_k - \tau_{k+1}(Au_k - f_k), \quad k = 0, 1, \dots, N - 1, \quad (3.4)$$

интегрирующих жесткую задачу Коши при затрате существенно меньшего количества шагов $N = N(g)$, где $g = \tau/\text{cou}$. После N шагов сумма их равна

$$l_N = \sum_{i=1}^N \tau_i. \quad (3.5)$$

Естественно потребовать выполнения следующих ограничений, обеспечивающих малую величину локальной аппроксимации в (3.1) и (3.3) для цикла из N шагов:

$$\max_{1 \leq k \leq N} \tau_k = \tau, \quad \tau \leq l_N \leq B_0 \tau, \quad (3.6)$$

где $B_0 > 1$ — независящая от N величина. Сделав N шагов, в новой точке определяем новые значения τ, N и продолжаем процесс интегрирования.

Положим для простоты в (3.1) $f = 0$. Тогда решение задачи Коши равно

$$u(t) = \exp(-At)u_0. \quad (3.7)$$

После N шагов по методу (3.4) имеем

$$u_N = P_N(A)u_0, \quad (3.8)$$

где многочлен $P_N(z)$ имеет вид

$$P_N(z) = \prod_{i=1}^N (1 - \tau_i z) \quad (3.9)$$

и обладает следующими свойствами:

$$P_N(0) = 1, \quad P'_N(0) = -l_N.$$

Поэтому

$$|\exp(-l_N z) - P_N(z)| \leq O((l_N z)^d) \quad (3.10)$$

при $d = 2$, т.е. для малых $l_N z$ аппроксимация оператора $\exp(-At)$ оператором $P_N(A)$ выполнена. Таким образом мы имеем разностную схему 1-го порядка точности.

Пусть при выполнении условий аппроксимации (3.6) многочлен $P_N(z)$ удовлетворяет условию устойчивости

$$\max_{z \in B} |P_N(z)| \leq q_N \leq 1 \quad (3.11)$$

и выполнены условия устойчивой реализации вычислений внутри цикла из N шагов [41, 45, 46].

Для наибольшего продвижения по времени за N шагов требуется найти решение обобщенной задачи Маркова:

$$d_N = \sup_{R_N} (-R'_N(0)), \quad P_N(z) = \arg \sup_{R_N} (-R'_N(0)), \quad (3.12)$$

где \sup взят по всем многочленам вида (3.9), удовлетворяющих условиям (3.6), (3.11). Если обозначить через z_i , $i = 1, \dots, N$, корни экстремального многочлена (3.9), то в оптимальном методе (3.4) на классе задач со спектром из B временные шаги определяются из множества чисел: z_i^{-1} , $i = 1, \dots, N$. При этом, как это будет показано далее, пара комплексно сопряженных шагов может быть реализована в действительной арифметике, а в формулах (3.5), (3.6) τ_i следует заменить на $Re \tau_i$.

Поскольку любой многочлен от z неограничен при $z \rightarrow \infty$, заслуживает особого рассмотрения вопрос о разумном приближении многочленами убывающих операторных экспонент для жестких сеточных задач, ибо их спектральные компоненты, соответствующие большим по модулю собственным значениям, плохо аппроксимируют аналогичные компоненты исходной дифференциальной задачи, которые с увеличением времени быстро стремятся к нулю (скорость этого убывания, как будет показано далее, можно

оценить методами, основанными на интегральных критериях качества дихотомии). Для простоты рассмотрим случай, когда $S\rho(A) \in [0, 2r]$. Выполнение условий аппроксимации следует требовать лишь на отрезке $[0, \tau^{-1}]$ или на части его, а для остальных $z \in [\tau^{-1}, 2r]$ потребовать выполнения неравенства (3.11). Расположенные на отрезке $[\tau^{-1}, 2r]$ нули многочлена определяют некоторый алгоритм фильтрации решения. Известны аналитические формулы для таких многочленов: при $d = 2$ в (3.10) через многочлены Чебышева 1-го рода [39, 40, 41, 103, 105], при $d = 3$ через многочлены Золотарева 2-го типа 1-го рода [44]–[104], при $d = 4$ через функции Шоттки [69].

Сформулируем три теоремы теории аналитических функций.

Явное решение задачи (3.12) в общем случае неизвестно, однако теорема Сеге позволяет дать оценки сверху и снизу для $|P'_N(0)|$. Пусть B — односвязная ограниченная область с границей L , составленной из конечного числа аналитических дуг, и пусть точка $z = z_0 = 0 \in L$. Внешним углом $\alpha\pi$, $0 \leq \alpha \leq 2$, для L в точке z_0 назовем точную верхнюю границу углов, имеющих вершину в точке z_0 и таких, что достаточно близкие к z_0 внутренние точки этих углов лежат вне L . Пусть \mathcal{P}_N — класс многочленов $R_N(z)$ степени не выше N , удовлетворяющих условию $\sup_{z \in B} |R_N(z)| \leq C$, $C > 0$.

Положим $\mu = \mu(L, z_0) = \sup_{R_N \in \mathcal{P}_N} |R'_N(z_0)|$.

Теорема 3.1. (Сеге [66].) *Существуют два числа $A' > A > 0$, зависящие от свойств кривой L (но не зависящие от N), такие, что при всех достаточно больших N будет*

$$AN^\alpha \leq \mu_N(L, z_0) \leq A'N^\alpha. \quad (3.13)$$

Таким образом, наилучшие результаты будут при $\alpha = 2$. Тогда точка $z = 0$ является вершиной острого с нулевым углом клюва при $z = 0$, внутри которого или отсутствуют, или присутствуют почти действительные собственные значения $S\rho(A)$. И чем больше будет расположено корней z_i^{-1} , $i = 1, \dots, N$ многочлена $P_N(z)$ внутри этого клюва и близких к точке $z = 0$, тем больше будет величина d_N . Однако в этой теореме отсутствует информация о $|P_N(z_0)|$. Точных решений задачи (3.12) известно мало. Упомянем два случая. Наименьшее убывание величины d_N наблюдается, когда $B = K_r$, тогда в (3.9) $P_N(z) = (1 - z/r)^N$ и $-P'_N(0) = N\text{сои}$, $\tau_k = \text{сои}$. Для сравнения с этим методом рассмотрим величину

$$V_N = \frac{l_N}{N \text{сои}}, \quad (3.14)$$

которая равна отношению суммы длин отрезков интегрирования за N шагов метода (3.4) с переменными шагами и при шаге, равном сои . При заданном N степень эффективности метода (3.4) зависит от B . Наибольшее убывание величины d_N (3.12) будет согласно теореме Маркова, когда

$B = [m, M], 0 \leq m \leq M \leq 2r$. Так, при $m = 0, M = 2r$ имеем $l_N = N^2$ соц, то есть $V_N = N$ (см. ниже). В связи с тем, что алгоритмы решения задачи (3.12) сложны, рассмотрим возможности другого алгоритма для определения параметров многочлена $P_N(z)$, для которого $V_N = O(N)$.

Выясним свойства многочленов, лемнискаты которых приближают границу области B . Сначала оценим величину q_N в (3.11) для многочленов специального вида (3.9).

Пусть $Q_l(z)$ — многочлен степени l , z_1, \dots, z_l — его корни, $\rho > 0$, а ∂B_ρ — линия уровня (лемниската) для $Q_l(z)$:

$$\partial B_\rho = \{z: |Q_l(z)| = \rho\}, \quad (3.15)$$

ограничивающая замкнутую область $B_\rho = \{z: |Q_l(z)| \leq \rho\}$. При $l > 1$ и достаточно малом ρ область B_ρ будет состоять в окрестности точек z_1, \dots, z_l из нескольких изолированных частей, а при достаточно большом — односвязной. Лемнискату можно определить как геометрическое место точек плоскости, для которых произведение расстояний до l данных точек плоскости z_1, \dots, z_l , называемых фокусами (и являющихся корнями $Q_l(z)$), есть величина постоянная.

Пусть $z_0 \notin B_\rho$ и $\overline{\prod}_N(z_0)$ — класс многочленов $p_N(z)$ степени N , удовлетворяющих условию $p_N(z_0) = 1$. Справедлива [66]

Теорема 3.2. Для любого $m > 0$

$$P_{ml}(z) = \arg \inf_{p_{ml} \in \overline{\prod}_{ml}(z_0)} \max_{z \in B_\rho} |p_{ml}(z)| = \left(\frac{Q_l(z)}{Q_l(z_0)} \right)^m \quad (3.16)$$

и

$$\bar{E}_{ml} = \inf_{p_{ml}(z) \in \overline{\prod}_N(z_0)} \max_{z \in B_\rho} |p_{ml}(z)| = \left| \frac{\rho}{Q_l(z_0)} \right|^m < 1. \quad (3.17)$$

Пусть

$$|P_N(z)| = \rho, \quad 0 \leq \rho \leq 1, \quad (3.18)$$

— лемниската для $P_N(z)$, B_N — замкнутая область, лежащая внутри лемнискаты с $\rho = 1$. Из (3.16) следует, что $B \in B_N$. Исследуем поведение величины V_N тех многочленов $P_N(z)$, для которых достаточно близки границы областей B и B_N . О существовании таких лемнискат утверждает следующая теорема Гильберта [48].

Теорема 3.3. Конечное число произвольных взаимно-внешних, замкнутых, изолированных кривых Жордана без самопересечения, расположенных в ограниченной части комплексной плоскости, можно равномерно сколь угодно точно аппроксимировать одной и той же лемнискатой.

Определим один из классов множеств B , для которых возможно конструктивно построить алгоритмы с $V_N = O(N)$.

Пусть, далее, $0 \leq m \leq M \leq 2r$ — вспомогательный параметр, и

$$\begin{aligned} s &= s(z) = (M + m - 2z)/(M - m), \\ \theta &= (M + m)/(M - m) \geq 1, \\ \overline{\text{cou}} &= 2/M. \end{aligned} \quad (3.19)$$

Тогда $s(0) = \theta$, $s(m) = 1$, $s(M) = -1$. Пусть $w = Q_l(s)$ — многочлен степени l с вещественными коэффициентами, $Q_l(1) = 1$ и линия уровня $\rho = Q_l(1)$ лежит в полуплоскости $\text{Res} \leq 1$, тогда $Q'_l(s) > 0$ при $\text{Res} \geq 1$. Положим $N = nl$. Пусть $[m_1, 1]$ — максимальный по длине отрезок действительной оси, целиком принадлежащий внутренности и границе единичной лемнискаты. Не уменьшая общности можно считать, что $m_1 = -1$.

Замечая, что $T'_n(1) = n^2$, рассмотрим множество областей типа $B = B_N$, для которых при $V_N = O(N)$ устойчивые многочлены имеют вид $T_n(Q_l(s(z)))/T_n(Q_l(\theta))$ и $Q_l(T_n(s(z)))/Q_l(T_n(\theta))$. Далее рассмотрим лишь первый случай:

$$P_N(z) = \frac{T_n(Q_l(s(z)))}{T_n(Q_l(\theta))}, \quad (3.20)$$

а

$$\max_{m \leq z \leq M} |P_N(z)| = \frac{1}{|T_n(Q_l(\theta))|} = q_N \leq 1. \quad (3.21)$$

В этом случае множество B_N содержит точки, окружающие l отрезков кривых $S_l = \{z : -1 \leq Q_l(s(z)) \leq 1\}$, которые назовем "скелетом". Он расположен симметрично относительно действительной оси и ему принадлежит отрезок действительной оси $[m, M]$, принадлежащей B_N . Пусть l_N^0, V_N^0 — значения l_N, V_N при $m = 0$. Если $h_0 = 2/(M + m)$, то

$$\tau_i \in \{h_0/(1 - \theta^{-1}s_{jk}) : k = 1, \dots, l, j = 1, \dots, n\}, \quad (3.22)$$

где $s_{jk}, j = 1, \dots, l$, — корни уравнений

$$Q_l(s) = \cos \frac{2k-1}{2n} \pi, \quad k = 1, \dots, n. \quad (3.23)$$

Тогда

$$l_N = \frac{nU_{n-1}(Q_l(\theta))Q'_l(\theta)}{T_n(Q_l(\theta))} \cdot \frac{2}{M-m}, \quad l_N^0 = n^2Q'_l(1)\overline{\text{cou}}, \quad (3.24)$$

$$V_N = \frac{2U_{n-1}(Q_l(\theta))Q'_l(\theta)}{(M-m)l\text{cou}}, \quad V_N^0 = \frac{nQ'_l(1)\overline{\text{cou}}}{l\text{cou}}, \quad (3.25)$$

и граница области B_N есть линия уровня $|P_N(z)| = 1$. Эта область устойчивости ($|P_N(z)| \leq 1$) содержит множество S_l .

Пусть s_0 — ближайший к единице (при $k = 1$) корень уравнения (3.23). Для него справедливо приближенное равенство:

$$1 + Q'_l(1)(s_0 - 1) = \cos \frac{\pi}{2n},$$

из которого получаем

$$1 - s_0 = \frac{\sin^2 \frac{\pi}{4n}}{Q'_l(1)}.$$

Пусть

$$\tau_0 = \max_k \tau_k = \frac{h_0}{(1 - \theta^{-1}s_0)}. \quad (3.26)$$

Теперь определим n и m . Сначала определим то минимальное n , при котором в предельном случае $\theta = 1$ (т.е. $m = 0$) максимальный шаг τ_0 (3.6) не меньше τ . Полагая в (3.26) $m = 0$, $\tau_0 = \tau$, получаем

$$\tau = \frac{\overline{\cos}}{(1 - s_0)} = Q'_l(1)\overline{\cos} \left[2 \sin^2 \frac{\pi}{4n} \right]^{-1}.$$

Отсюда при больших τM получаем

$$n = \left[\frac{\pi}{4} \sqrt{\frac{\tau M}{Q'_l(1)}} \right] + 1. \quad (3.27)$$

Затем определяем m из равенства

$$\tau = h_0 / (1 - \theta^{-1}s_0). \quad (3.28)$$

Тогда

$$l_N = \pi^2 \tau M / (8(M - m)), \quad V_N = \frac{2nQ'_k(\theta)}{(M - m)l \cos}. \quad (3.29)$$

Таким образом, оба условия (3.6) выполнены. Областями устойчивости в этом случае являются зависящие от n , M , m и $Q_l(s)$ области B_N , ограниченные лемнискатами, окружающими l отрезков кривых $z = (M + m - (M - m)s)/2$, $-1 \leq Q_l(s) \leq 1$. Можно дать многопараметрическое, легко анализируемое и алгоритмически реализуемое представление многочлена $Q_l(s)$ следующего частного вида. Пусть

$$t_k = t_k(s) = \frac{2t_{k-1}^{2l_k}(s) - 1 + p_k}{1 + p_k}, \quad k = 1, \dots, K, \quad (3.30)$$

$t_0 = s$, $t_k(\pm 1) = 1$, где $l_k > 0$ — целые, а $p_k > -1$. Тогда полагаем [40]

$$w = Q_l(s) = t_K(s) \quad \text{при} \quad l = 2 \sum_{k=1}^K l_k, \quad (3.31)$$

а

$$l_N = \frac{nU_{n-1}(Q_l(\theta))}{T_n(Q_l(\theta))(M-m)} 2^{K+1} \prod_{k=1}^K l_k t_k^{2l_k-1}(\theta)(1+p_k)^{-1} \quad (3.32)$$

и

$$l_N^0 = n^2 \operatorname{coi} 2^K \prod_{k=1}^K l_k (1+p_k)^{-1}. \quad (3.33)$$

Рассмотрим несколько случаев.

1. Предположим, что $Sp(A)$ действителен и принадлежит отрезку $[0, M]$, тогда $w = Q_1(z) = s$.

Мы видим, что многочлен $P_N(z)$ на отрезке $[0, m]$ убывает. Согласно теореме Маркова все другие многочлены N -ой степени, удовлетворяющие условию (3.11) при $z \in [m, M]$, имеют меньшую по модулю производную при $z = 0$ и

$$l_N^0 = N^2 \operatorname{coi}, \quad V_N^0 = N. \quad (3.34)$$

Таким образом, употребляя многочлен Чебышева, мы можем за N шагов уйти по времени в N раз дальше по сравнению с явным методом Эйлера (3.4) с постоянным шагом, равным coi . Из условий (3.6) получаем, что $N = O(\sqrt{g})$. Для разностных схем повышенной аппроксимации имеем $l_N^0 = 0.4109N^2 \operatorname{coi}$ при $d = 3$ и $l_N^0 = 0.25N^2 \operatorname{coi}$ при $d = 4$.

2. Пусть в (3.30), (3.31) $K = 1, l_1 = 1$,

$$w = Q_2(s) = \frac{2s^2 - 1 + p_1}{1 + p_1}. \quad (3.35)$$

Множество S_l — это крест или два отрезка, положение и размеры которых определяются параметром p_1 . В [41] изображены линии уровня этой функции.

3. Рассмотрим преобразования (3.30), (3.31) при $K = 2, l_1 = l_2 = 1$:

$$w = \frac{2v^2 - 1 + p_2}{1 + p_2}, \quad v = \frac{2s^2 - 1 + p_1}{1 + p_1}, \quad (3.36)$$

где p_1, p_2 — параметры. Множеством S_l могут быть: два креста, крест и ветви гиперболы, два отрезка и ветви гиперболы, два отрезка и крест, четыре отрезка [40]. Положение и размеры этих множеств определяются параметрами p_1, p_2 .

4. Если спектр задачи содержит чисто мнимые или тесно лежащие в окрестности мнимой оси собственные значения, то в этом неподходящем под наши предположения случае нельзя за счет увеличения N получить существенное увеличение величины V_N . Поэтому достаточно ограничиться $N = 4$ и в качестве $P_4(z)$ взять аппроксимирующий экспоненту $e^{-l_4 t}$ с 5-й степенью точности многочлен

$$P_4(z) = Q_4(l_4 z), \quad Q_4(t) = 1 - t + t^2/2 - t^3/6 + t^4/24, \quad l_4 = 2\sqrt{2}. \quad (3.37)$$

Многочлен $Q_4(t)$ имеет наибольший отрезок устойчивости вдоль мнимой оси ($[-l_4i, l_4i]$) среди многочленов $Q_4(t) = 1 - t + at^2 + bt^3 + ct^4$. Вдоль действительной оси область устойчивости простирается от 0 до 2.7853. Изображение области устойчивости в этом случае содержится в [69].

5. Рассмотрим разностные схемы для решения жестких задач с отделимым от нуля спектром. Пусть $0 < \rho < 1$, а ∂B_ρ — линия уровня (3.18) для $Q_l(z)$, внутри которой лежит спектр. Свойства решений таких задач исследованы в [69]. Пусть многочлен (3.9) имеет вид

$$P_N(z) = R_k(z)Q_l^n(z), \quad R_k(z) = \prod_{j=1}^k (1 - \tau_j z), \quad (3.38)$$

где $N = k + nl$, $k \ll nl$, а $\tau_j > 0$ — заданные шаги, обеспечивающие необходимую степень аппроксимации ($\max \tau_j = \tau$) и спектральную устойчивость (3.11) при $B = B_\rho$. Тогда $l_N = \sum_{j=1}^k \tau_j - n(Q_l'(0))$. Используя теорему 3.2, мы устанавливаем, что неравенство (3.38) будет выполнено, если минимальную степень n определить из неравенства

$$\max_{z \in \partial B_\rho} |R_k(z)|\rho^n = q_N \leq 1. \quad (3.39)$$

Изложенный метод предполагает, что корни многочлена (3.9) действительны. Для общего случая получены другие формулы явного метода. Для простоты пусть $N = 2n$. Тогда многочлен (3.9) можно факторизовать на квадратичные множители, пользуясь тем, что

$$(1 - \tau_{2i-1}z)(1 - \tau_{2i}z) = (1 - h_i z)^2 - \gamma_i h_i^2 z^2, \quad (3.40)$$

где

$$h_i = \frac{1}{2}(\tau_{2i-1} + \tau_{2i}), \quad \gamma_i = \left(\frac{\tau_{2i} - \tau_{2i-1}}{\tau_{2i} + \tau_{2i-1}} \right)^2.$$

Подставляя (3.40) в (3.9) получаем

$$P_N(z) = \prod_{i=1}^n [(1 - h_i z)^2 - \gamma_i h_i^2 z^2]. \quad (3.41)$$

При этом будем считать, что если τ_{2i-1} — комплексно, то $\tau_{2i} = \bar{\tau}_{2i-1}$, а если τ_{2i-1}, τ_{2i} — действительны, то в эти пары объединяем последовательно максимальные и минимальные значения. Тогда два шага метода (3.4) можно реализовать по формулам

$$\begin{aligned} y_{k+1/2} &= u_k - h_{k+1} A u_k, & y_{k+1} &= y_{k+1/2} - h_{k+1} A y_{k+1/2}, \\ u_{k+2} &= y_{k+1} + \gamma_{k+1} (A y_{k+1/2} - A u_k). \end{aligned} \quad (3.42)$$

В этом методе при $g \gg 1$ максимальный шаг h_i приблизительно в два раза меньше, а минимальный — в два раза больше по сравнению с предыдущим методом.

Пусть область $\Omega \in R^m$ содержится в прямоугольном параллелепипеде P^m . Построим в P^m ортогональную, равномерную с шагом h сетку P_h . Пусть $x_i = (x_{1i}, \dots, x_{mi}), i = 1, \dots, n$ — узлы сетки. Из них Ω_{h1} — узлы, лежащие внутри Ω в количестве n штук, а Ω_{h2} — остальные узлы. Пусть (3.1) является системой уравнений метода прямых для решения нестационарной задачи математической физики в области Ω с однородными условиями Дирихле, в которой пространственные производные, образующие эллиптический оператор, заменены разностными отношениями. Продолжим непрерывным образом на Ω_{h2} коэффициенты оператора A с выполнением условия эллиптичности и условия Дирихле на границе P_h . Пусть на Ω_{h1} $q(t) = Au - f$, а на Ω_{h2} $u_0 = 0$ и $q(t) = Au + bI$ [37], где $b = O(r) > 0$ — числовой параметр. Для снижения жесткости системы (3.1) предлагается следующий способ ее регуляризации [106]:

$$\frac{du}{dt} = -R(t, \tau(t))(Au - f), \quad u|_{t=0} = u_0, \quad (3.43)$$

где $R(t, \tau(t))$ — линейный псевдоразностный на P_h оператор (ПРО), зависящий от параметра $\tau(t)$ и являющийся дискретным аналогом псевдодифференциального эллиптического дифференциального оператора. Пусть $w = w(x, t), d(x), (d(x) \neq 0)$ — функции, определенные на P_h , частный вид ПРО — $v(s) = F(w(x))$ — m -мерное дискретное преобразование Фурье на P_h , а $w(x) = F^{-1}(v(s))$ — обратное к нему преобразование. Пусть $S(s) = S(s, t, \tau(t))$ — символ ПРО такой, что $S(s) \rightarrow 1$ при $\tau \rightarrow 0$, например, $S = (1 + a\tau|s|^2)^{-1}$. Определим семейство ПРО формулой:

$$R(t, \tau(t))w = d^{-1}(x)F^{-1}(S(s, t, \tau(t))F(d(x)w(x))). \quad (3.44)$$

Вместо преобразования Фурье могут быть использованы взвешенные чебышевские фильтры [42] из произведения одномерных по x_i фильтров, или фильтры, если в них за оператор L в [42] взять, например, определенный на P_h с точностью до нормировочного множителя разностный оператор Лапласа с условием Дирихле. Анализ эффективности такого способа регуляризации упрощается, если операторы R и A коммутируют. В этом случае известные неявные схемы и некоторые методы расщепления можно реализовать как явные методы для (3.43).

Задачу Коши для n нелинейных дифференциальных уравнений

$$\frac{du}{dt} = f(u, t), \quad u|_{t=0} = u_0 \quad (3.45)$$

решаем явным методом

$$u_{k+1} = u_k + \tau_{k+1}f(u_k, t_k), \quad k = 0, 1, \dots, N - 1, \quad (3.46)$$

где временные шаги определяются выполнением условий аппроксимации и спектром оператора Фреше (Якобиана) $J(f_k)$. Для учета комплексного спектра систему (3.46) заменим следующей:

$$\begin{aligned} y_{k+1/2} &= u_k + h_{k+1} f(u_k, t_k), & t_{k+1/2} &= t_k + h_{k+1}, \\ y_{k+1} &= y_{k+1/2} + h_{k+1} f(y_{k+1/2}, t_{k+1/2}), & t_{k+1} &= t_{k+1/2} + h_{k+1}, \\ u_{k+1} &= y_{k+1} + \gamma_{k+1} h_{k+1} (f(u_k, t_k) - f(y_{k+1/2}, t_{k+1/2})), \\ k &= 0, 1, \dots, n-1. \end{aligned} \tag{3.47}$$

Этот метод состоит из двух эйлеровых шагов и поправки. В нем параметры h_i, γ_i легко определяются по корням $P_N(z)$.

Метод (3.47), учитывающий и комплексный спектр, оказался достаточно гибким, алгоритмы его основываются на свойствах T-последовательности корней многочленов Чебышева. Он был запрограммирован в программе DUMKA; в ней в зависимости от условий аппроксимации $2 \leq N \leq 2^{16} \cdot 3^6 = 47\,775\,744$. С помощью этой программы был успешно решен ряд линейных и нелинейных многомерных задач математической физики с действительным и комплексным спектром.

Более жестким требованием, чем устойчивость, к качеству численного метода решения системы дифференциальных уравнений вида (3.1), или более общего вида (3.45), является контрактивность, что означает невозрастание расстояния между решениями. Например, для контрактивности метода решения задачи (3.1) с функцией устойчивости $R(z)$ необходимо и достаточно, чтобы $\|R(-A)\| \leq 1$. Если матрица A нормальная, то понятия контрактивности и устойчивости совпадают. Для матрицы A , отличной от нормальной, из контрактивности следует устойчивость, но не наоборот.

Следующая теорема является конечномерным аналогом теоремы, установленной фон Нейманом для операторов в гильбертовом пространстве [111], и позволяет установить контрактивность многих неявных методов для обладающих контрактивностью систем дифференциальных уравнений вида (3.1) и (3.45).

Теорема 3.4. ([22]) Пусть $R(z)$ ограничена при $\operatorname{Re} z \leq 0$ и хаусдорфово множество

$$\mathcal{H} = \{(Bv, v) : v \in \mathbb{C}^n\}$$

$n \times n$ матрицы B лежит в левой полуплоскости: $\mathcal{H} \subset \mathbb{C}^-$. Тогда в норме, связанной со скалярным произведением, имеем

$$\|R(B)\| \leq \sup_{\operatorname{Re} z \leq 0} |R(z)|.$$

Напомним, что метод решения дифференциальных уравнений называется A -устойчивым [87], если область его устойчивости S содержит левую полуплоскость:

$$S \supset \mathbb{C}^- = \{z : \operatorname{Re} z \leq 0\}.$$

Например, неявный метод Эйлера с функцией устойчивости $R(z) = 1/(1 - z)$ и областью устойчивости $\{z : |1 - z| \geq 1\}$ является, в силу этого определения, A -устойчивым.

Если функция устойчивости $R(z)$ такова, что $|R(iy)| \leq 1$ при любых вещественных y и $R(z)$ — аналитична при $\operatorname{Re} z \leq 0$, то соответствующий метод A -устойчив. Функцию устойчивости, обладающую указанными выше свойствами, обеспечивающими A -устойчивость, называют A -устойчивой.

Если функция устойчивости $R(z)$ — A -устойчива, то метод $u_{j+1} = R(\tau B)u_j$ в силу теоремы 3.4 контрактивен для матрицы B , хаусдорфово множество которой лежит в левой полуплоскости.

Вопрос о контрактивности явных методов, функции устойчивости которых являются многочленами и очевидно не могут быть A -устойчивыми, в настоящее время остается открытым.

3.2. Спектральный анализ задач математической физики

Численное интегрирование жестких систем обыкновенных дифференциальных уравнений, возникающих в результате аппроксимации уравнений в частных производных, не позволяет полностью предсказать их возможное поведение. Более полную информацию дает спектральный анализ, сводящий исходную задачу к проблемам собственных значений для соответствующих матриц или матричных пучков. Он состоит в решении полной проблемы собственных значений, отвечающей системе, полученной в результате аппроксимации на самой грубой сетке, подробном исследовании свойств такой системы с помощью найденных спектральных характеристик и в последующем уточнении отдельных результатов посредством перехода к более мелким сеткам и решению соответствующих частичных проблем собственных значений.

В условиях приближенных вычислений полную проблему собственных значений можно решить достаточно точно лишь в случае, когда все собственные значения различны и хорошо отделены друг от друга, а матрица, составленная из отвечающих им собственных векторов, хорошо обусловлена. Если это не так, то разумно ограничиться вычислением инвариантных подпространств, отвечающих хорошо отделенным друг от друга частям спектра, и сужений исходной матрицы на эти подпространства. Иначе говоря, вычислительно корректная полная проблема собственных значений, в общем случае, состоит в нахождении достаточно хорошо обусловленных

матриц преобразования подобия, приводящих исходную матрицу к блочно-диагональному виду, и исследовании полученных разложений.

Эффективные методы решения полных неэрмитовых проблем собственных значений, разработанные в последние десятилетия на основе разложения Шура [21], и реализация этих методов в широко распространенных пакетах программ таких, как LAPACK [73], сделали широко доступным спектральный анализ жестких систем с матрицами не слишком большого порядка. Однако полный спектральный анализ таких систем оставался очень вычислительноемкой задачей. В работе [109] была предложена новая технология численного спектрального анализа на основе разложения Шура. В отличие от стандартной технологии, требующей $O(n^4)$ операций для детального спектрального анализа системы (3.1), новая технология позволяет выполнять его за $O(n^3)$ операций. Ускорение достигается за счет замены многократных решений уравнений Сильвестра однократным вычислением разложения Шура матрицы, сопряженной исходной.

Известно большое число методов вычисления отдельных собственных значений неэрмитовых матриц и отвечающих им инвариантных подпространств на основе аппроксимаций в подпространствах Крылова. Это методы одновременных итераций, Арнольди, Ланцоша, Дэвидсона и различные их модификации, например [117, 118]. Все они не требуют преобразования исходной матрицы и используют в своей схеме лишь операцию ее умножения на вектор. Этим же свойством обладают и градиентные методы такие, как метод сингулярной функции [108], различные варианты метода Ньютона и метод Ньютона-Канторовича, полученный введением в схему метода Ньютона [20] вспомогательного оператора [60].

Одной из наиболее важных задач, возникающих при анализе жестких систем обыкновенных дифференциальных или разностных уравнений, является задача оценки для заданной матрицы A и заданного контура γ , не проходящего через ее собственные значения, норм матриц вида

$$P_\gamma(f, A) = \frac{1}{2\pi i} \oint_\gamma f(z)(zI - A)^{-1} dz,$$

где f — функция комплексного переменного, непрерывная на контуре γ , и задача оценки устойчивости таких матриц к возмущениям элементов матрицы A . Требуемые оценки можно получать [18, 59] на основе интегральных критериев качества дихотомии вида

$$\kappa_{\gamma,p}(A; B, C) = \nu_{B,C} \max_{\|v\|_2=1} \left(\int_\gamma \|C(zI - A)^{-1} Bv\|_2^p |dz| \right)^{1/p}, \quad v \in \mathbb{C}^n,$$

где $\nu_{B,C} = \|B^{-1}\|_2 \|C^{-1}\|_2 / (2\pi)$, $p \geq 1$, а B и C — произвольные квадратные невырожденные матрицы порядка n .

Теорема 3.5. $\|P_\gamma(f, A)\|_2 \leq \|f\|_{\gamma, q} \kappa_{\gamma, p}(A; B, C)$.

Теорема 3.6. Для любой матрицы $\Delta \in \mathbf{C}^{n \times n}$ такой, что

$$\delta = \|B^{-1}\Delta C^{-1}\|_2 \kappa_{\gamma, \infty}(A; B, C) / \nu_{B, C} < 1, \quad (3.48)$$

выполняются следующие неравенства:

$$\begin{aligned} \|P_\gamma(f, A) - P_\gamma(f, A + \Delta)\|_2 &\leq \|f\|_{\gamma, q} \kappa_{\gamma, p}(A; B, C) \frac{\delta}{1 - \delta}, \\ |\kappa_{\gamma, p}(A; B, C) - \kappa_{\gamma, p}(A + \Delta; B, C)| &\leq \kappa_{\gamma, p}(A; B, C) \frac{\delta}{1 - \delta}. \end{aligned}$$

Здесь $q = p/(p - 1)$, а

$$\|f\|_{\gamma, q} = \left(\int_\gamma |f(z)|^q |dz| \right)^{1/q}.$$

Пусть $f(z) = \exp(tz)$. Тогда $P_\gamma(f, A) = \exp(tA)P$, где P — спектральный проектор (проектор Рисса) матрицы A , отвечающий ее собственным значениям, лежащим внутри контура γ , и, в силу теоремы 3.5, $\|\exp(tA)P\|_2 \leq \varphi(t)$, где

$$\varphi(t) = \left(\int_\gamma \exp(q \operatorname{Re} z t) |dz| \right)^{1/q} \kappa_{\gamma, p}(A; B, C).$$

Опираясь на это неравенство, можно показать, что для всех $u^0 \in \mathbf{C}^n$ и всех вещественных t для сужения Pu решения задачи Коши $u(0) = u^0$, $du/dt = Au$ на инвариантное подпространство матрицы A , отвечающее ее собственным значениям, лежащим внутри контура γ , справедливы следующие двусторонние оценки:

$$\|Pu^0\|_2 / \varphi(-t) \leq \|Pu(t)\|_2 \leq \|Pu^0\|_2 \varphi(t).$$

Для решения вопроса о том, насколько могут отличаться собственные значения возмущенной матрицы от собственных значений исходной, широко используют псевдоспектр [18, 120]. Он дает точные границы всех возможных смещений собственных значений при любых возмущениях, по норме не превосходящих заданную величину. Однако для возмущений, имеющих специальный вид, эти границы часто оказываются слишком завышенными. Проблему иногда удается решить используя различные обобщения псевдоспектра.

Структурированным псевдоспектром матрицы $A \in \mathbf{C}^{n \times n}$ называют множество вида

$$\Lambda_\varepsilon(A; B, C) = \bigcup_{\|S\|_2 \leq \varepsilon} Sp(A + BSC), \quad S \in \mathbf{C}^{n \times n},$$

где B и $C \in \mathbb{C}^{n \times n}$ — фиксированные матрицы. В частности, множество $\Lambda_\varepsilon(A) = \Lambda_\varepsilon(A; I, I)$, где I — единичная матрица порядка n , — это обычный псевдоспектр матрицы A . Справедливо следующее утверждение [97]:

$$\Lambda_\varepsilon(A; B, C) = Sp(A) \bigcup \{z : \|C(zI - A)^{-1}B\|_2 \geq 1/\varepsilon\}.$$

Структурированный псевдоспектр с невырожденными матрицами B и C называют регулярно структурированным [110]. Для него справедливы следующие равенства:

$$\Lambda_\varepsilon(A; B, C) = \bigcup_{\|B^{-1}EC^{-1}\|_2 \leq \varepsilon} Sp(A + E), \quad E \in \mathbb{C}^{n \times n},$$

и

$$\Lambda_\varepsilon(A; B, C) = \{z : \|C(zI - A)^{-1}B\|_2 \geq 1/\varepsilon\}.$$

В силу этих равенств, условие (3.48) означает, что контур γ не имеет общих точек с $\Lambda_\varepsilon(A; B, C)$, где $\varepsilon = \|B^{-1}\Delta C^{-1}\|_2$, т.е. γ осуществляет дихотомию этого регулярно структурированного псевдоспектра.

Представляет также интерес задача о дихотомии спектра мнимой осью. Соответствующие оценки можно получать [18, 19, 57, 58, 107] на основе интегрального критерия $\|H(A)\|_2^{1/2}$ качества дихотомии спектра матрицы A мнимой осью, где

$$H(A) = \frac{1}{\pi} \int_{-\infty}^{\infty} (i\omega I - A)^{-*} (i\omega I - A)^{-1} d\omega.$$

Если A не имеет собственных значений на мнимой оси, то матричный интеграл $H(A)$ — единственное эрмитово положительно определенное решение обобщенного уравнения Ляпунова

$$A^*H + HA = (I - P)^*(I - P) - P^*P, \quad HP = P^*H,$$

где P означает спектральный проектор, отвечающий части спектра матрицы A , лежащей в левой полуплоскости.

Пусть A — квадратная матрица порядка n такая, что мнимая ось не содержит точек ее спектра и разбивает его на два непустых подмножества, лежащих соответственно в левой и правой полуплоскостях. Обозначим через $h_{\min} = h_{\min}(A) < 0$ и $h_{\max} = h_{\max}(A) > 0$ соответственно левую и правую границы хаусдорфова множества

$$W(A) = \{(Av, v)/(v, v) : v \in \mathbb{C}^n, v \neq 0\}$$

матрицы A :

$$h_{\min}(A) = \min\{Re\lambda : \lambda \in W(A)\}, \quad h_{\max}(A) = \max\{Re\lambda : \lambda \in W(A)\}.$$

Теорема 3.7. ([57]) *Справедливы неравенства*

$$\begin{aligned}\|\exp(tA)(I - P)\|_2 &\leq \sqrt{c_1} \exp(t/\kappa), \quad t \leq 0, \\ \|\exp(tA)P\|_2 &\leq \sqrt{c_2} \exp(-t/\kappa), \quad t \geq 0,\end{aligned}\tag{3.49}$$

где $\kappa = \|H(A)\|_2$,

$$c_1 = \frac{1}{\gamma} \left(\frac{\alpha + \gamma}{\alpha + \beta} \right)^{1+\alpha}, \quad c_2 = \frac{1}{\gamma} \left(\frac{\beta + \gamma}{\alpha + \beta} \right)^{1+\beta},$$

а $\alpha = 1/(|h_{\min}|\kappa)$, $\beta = 1/(h_{\max}\kappa)$, $\gamma = \max\{\alpha, \beta\}$.

Неравенства (3.49) являются верхними оценками нормы матрицы Грина

$$G(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (i\omega I - A)^{-1} \exp(i\omega t) d\omega = \exp(tA) \times \begin{cases} P - I, & t < 0, \\ P, & t > 0, \end{cases}$$

системы обыкновенных дифференциальных уравнений $du/dt = Au + f$, $-\infty < t < \infty$ и имеют много других полезных приложений. В частности, они позволяют оценить норму спектрального проектора:

$$\|P\|_2 = \|I - P\|_2 \leq \sqrt{\min\{c_1, c_2\}}$$

и для решения $u(t) = \exp(tA)u^0$ задачи Коши $u(0) = u^0$, $du/dt = Au$ оценить скорость убывания жесткой компоненты по сравнению с главной:

$$\frac{\|Pu(t)\|_2}{\|(I - P)u(t)\|_2} \leq \sqrt{c_1 c_2} \exp(-2t/\kappa) \frac{\|Pu^0\|_2}{\|(I - P)u^0\|_2}, \quad t \geq 0.$$

Можно показать, что константы из теоремы 3.7 удовлетворяют следующим неравенствам:

$$c_1 \leq \min\{h_{\max}, 4|h_{\min}|\}\kappa, \quad c_2 \leq \min\{4h_{\max}, |h_{\min}|\}\kappa.$$

Отсюда для нормы матрицы Грина имеем

$$\|G(t)\|_2 \leq \sqrt{c} \exp(-|t|/\kappa), \quad -\infty < t < \infty,$$

где $c = 4 \min\{h_{\max}, |h_{\min}|\}\kappa$.

Если матрица A является конечномерным аналогом линейного неограниченного замкнутого оператора L , действующего в сепарабельном гильбертовом пространстве \mathcal{H} , с областью определения $D(L) \subset \mathcal{H}$, то чтобы оценки (3.49) оставались нетривиальными при $n \rightarrow \infty$, необходимо и достаточно, чтобы был ограничен интегральный критерий качества дихотомии спектра оператора L мнимой ось, т.е.

$$\sup_{x \in \mathcal{H}} \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\|(i\omega - L)^{-1}x\|_{\mathcal{H}}^2}{\|x\|_{\mathcal{H}}^2} d\omega < \infty,$$

и хаусдорфово множество оператора L было ограничено слева или справа, т.е. выполнялось хотя бы одно из двух следующих условий:

$$h_{\min}(L) = \inf \{ \operatorname{Re}(Lx, x)_{\mathcal{H}} / (x, x)_{\mathcal{H}} : x \in D(L), x \neq 0 \} > -\infty,$$

$$h_{\max}(L) = \sup \{ \operatorname{Re}(Lx, x)_{\mathcal{H}} / (x, x)_{\mathcal{H}} : x \in D(L), x \neq 0 \} < \infty.$$

Во многих задачах математической физики, таких как уравнения конвекции-диффузии, переноса нейтронов и т.п., второе из этих условий выполнено.

Для теоремы 3.7 существенно предположение о том, что мнимая ось разбивает спектр матрицы A на два непустых подмножества. Поэтому установленные в этой теореме оценки нельзя использовать в случаях, когда $P = 0$ или I . Для этих случаев, требующих отдельного рассмотрения, оценки, аналогичные оценкам из теоремы 3.7, получены в работе [58]. Для произвольной квадратной матрицы A эти оценки позволяют, в свою очередь, получать оценки нормы матричной экспоненты вида

$$\| \exp(tA) \|_2 \leq \nu(\theta) \exp(\theta t), \quad t \geq 0, \quad (3.50)$$

с любым θ из интервала $r_{\max} < \theta < h_{\max}$, где r_{\max} означает максимальную вещественную часть собственных значений матрицы A . Учитывая, что

$$\exp(r_{\max} t) \leq \| \exp(tA) \|_2 \leq \exp(h_{\max} t), \quad t \geq 0,$$

нетрудно видеть, что экспоненциальная часть полученных оценок вида (3.50) может принимать все представляющие интерес значения. В работе [58] ставится и решается задача оптимизации константы $\nu(\theta)$.

Описанные выше методы применялись, в частности, для анализа переходных процессов в ядерных реакторах [23, 102]. В результате впервые удалось полностью описать структуру подпространства главных мод реакторных моделей, учитывающих запаздывающие нейтроны, и объяснить спектральный состав решения.

4. Численные методы решения жестких эллиптических задач

Под жесткими эллиптическими задачами (ж.э.з.) понимают краевые задачи для эллиптических уравнений и систем с большим разбросом коэффициентов. Такие задачи возникают, например, в следующих ситуациях.

1. Находятся эффективные характеристики композиционных материалов периодической структуры, состоящие из материалов с существенно различающимися свойствами. Для отыскания этих характеристик приходится решать задачу на ячейке периодичности для жестких эллиптических уравнений и систем [8].

2. Для определенности рассматриваем далее двумерный случай. В области сложной формы решается краевая задача для эллиптического уравнения или системы. Например, в ограниченной липшицевой области G_1 с границей ∂G_1 решается краевая задача для уравнения Пуассона

$$Lu = \frac{\partial}{\partial x_j} \left(k \frac{\partial u}{\partial x_j} \right) = f, \quad k = 1, \quad u|_{\partial G_1} = 0; \quad (4.1)$$

по повторяющимся индексам производится суммирование в пределах [1, 2].

При численном решении задачи в случае сложной формы области G_1 может оказаться целесообразным использование метода фиктивных областей (м.ф.о.) [65], [37]. Выбирается область простой формы G , $G_1 \subset G$, и находится решение вспомогательной краевой задачи

$$L_\omega u_\omega = \frac{\partial}{\partial x_j} \left(K_\omega \frac{\partial u_\omega}{\partial x_j} \right) = f, \quad u_\omega \in H_0^1(G); \quad (4.2)$$

здесь $K_\omega = 1$ в области G_1 , и $K_\omega = \omega$ в области $G_2 = G \setminus G_1$, $\omega \gg 1$, правая часть f как-то продолжена в G_2 , например, нулем. При $f \in L_2(G)$ можно показать, что $\|u - u_\omega\|_{H^1(G_1)} = O(\|f\|_{L_2(G_1)} / \omega)$ и, таким образом, вместо исходной краевой задачи можно решать задачу (4.2) с большим ω .

Под простотой формы области понимается возможность решения некоторых вспомогательных краевых задач с малыми вычислительными затратами. Обычно в качестве G берется некоторый параллелепипед, круг или цилиндр. Поскольку коэффициент K_ω разрывен, то решение (4.2) понимается в смысле удовлетворения интегральному тождеству С.Л. Соболева. А именно, как функция $u_\omega \in H_0^1(G)$, удовлетворяющая интегральному тождеству

$$\Lambda_\omega(u_\omega, \varphi) = \int_G (f, \varphi) dx \quad \forall \varphi \in H_0^1(G), \quad \text{где} \quad \Lambda_\omega(u, \varphi) = \int_G \left(K_\omega \frac{\partial u}{\partial x_j} \frac{\partial \varphi}{\partial x_j} \right) dx.$$

Наиболее распространенными численными итерационными методами решения краевых задач являются итерационные методы с переобуславлителем, предложенные в [26, 29]:

$$\bar{L} \frac{u^{n+1} - u^n}{\theta} = L_\omega u^n - f, \quad u \in H_0^1(G). \quad (4.3)$$

В случае дискретной аппроксимации в прямоугольнике в качестве \bar{L} целесообразно взять сеточный оператор Лапласа. Тогда u^{n+1} может быть найдено с малыми вычислительными затратами, например, с помощью быстрого преобразования Фурье или устойчивой реализацией марш-алгоритма. Описанный вариант применения м.ф.о. обладает следующим недостатком. Для повышения точности требуется увеличивать ω ; использование

классического итерационного метода (4.3), вообще говоря, требует числа итераций, растущего с ростом ω . На алгебраическом уровне аналогом м.ф.о. выступает метод фиктивных компонент (неизвестных), предложенный в [52, 4].

Далее мы рассматриваем методы, скорость сходимости которых оценивается геометрической прогрессией со знаменателем q меньшим 1 и не зависящим от величины параметра ω . Существенным моментом, позволяющим построить такой итерационный процесс, является использование идеи итераций на подпространстве, имеющей более чем полувековую историю. Методы ускорения итераций в подпространствах были предложены в [38], [53] при создании KP -метода для решения кинетического уравнения. Современное состояние теории таких процессов в алгебраическом варианте подробно обсуждается в [34].

Однако применение этой идеи к случаю ж.э.з. и их дискретных аппроксимаций требует привлечения дополнительных нестандартных соображений. Описание методов решения ж.э.з. проводится далее в континуальном варианте. Переход к случаю дискретных аппроксимаций обсуждается позднее.

Для случая ж.э.з. впервые метод со скоростью сходимости, не зависящей от разброса коэффициентов, был предложен и обоснован в [98]. Подробный обзор работ по таким методам приведен в [7] (см. также [10]).

Рассматриваемые далее методы решения ж.э.з. существенно используют теоремы продолжения функций [61]. Имеется в виду справедливость следующего утверждения. Если функция f определена в некоторой липшицевой области G_1 и $G_1 \subset G$, то существует функция \tilde{f} , определенная в G , удовлетворяющая некоторым условиям сопряжения на $\partial G_2 \cap \partial G_1$ и в некоторой норме $\|\cdot\|$ неравенству $\|\tilde{f}\|_{G_2} \leq C\|f\|_{G_1}$. Постоянная C не зависит от конкретной функции f .

4.1. Простейший итерационный метод на подпространстве

Проиллюстрируем применение идеи итераций на подпространстве на простейшем примере. Далее $(\mathbf{f}, \mathbf{g})_k = \int_{G_k} (\mathbf{f}, \mathbf{g}) dx$ при $k = 1, 2$,

$$(\mathbf{f}, \mathbf{g}) = \int_G (\mathbf{f}, \mathbf{g}) dx,$$

$$S = \text{grad}, \quad S^* = \text{div}, \quad \|Su\|^2 = \frac{\partial u}{\partial x_j} \frac{\partial u}{\partial x_j}.$$

Рассмотрим итерационный процесс

$$\Delta \frac{u^{n+1} - u^n}{\theta} + \frac{\partial}{\partial x_j} \left(K_\omega \frac{\partial u^n}{\partial x_j} \right) = f, \quad u^n \in H_0^1(G). \quad (4.4)$$

В случае симметричных операторов A и B скорость сходимости итерационного процесса с переобуславливателем

$$B \frac{u^{n+1} - u^n}{\theta} + Au^n - f = 0 \quad (4.5)$$

оценивается через границы отношения $\Lambda(r) = (Ar, r)/(Br, r)$. Для функций $r \in H_0^1(G)$ имеем

$$\Lambda(r) = \int_G (K_\omega \|Sr\|^2) dx / \int_G \|Sr\|^2 dx.$$

Из определения K_ω следует оценка $m = 1 \leq \Lambda(r) \leq M = \omega$ и эта оценка не может быть улучшена. Таким образом, на классе произвольных значений начальных ошибок при оптимальном значении параметра $\theta = \frac{2}{M+m}$ мы имеем оценку скорости сходимости $O\left(\left(\frac{M-m}{M+m}\right)^n\right)$ метода (4.4), в рассматриваемом случае — $O\left(\left(\frac{\omega-1}{\omega+1}\right)^n\right)$. В соответствии с этой оценкой для уменьшения погрешности в $1/\varepsilon$ раз потребуется $\sim 2\omega \ln(1/\varepsilon)$ итераций.

Если сочетать применение переобуславливателя с идеями трехслойных итерационных процессов, методом сопряженных градиентов [51] или Чебышевскими ускорениями [45], то число итераций уменьшится до $O(\sqrt{\omega} \log(1/\varepsilon))$. В случае дополнительного использования в качестве переобуславливателя оператора попеременно-треугольного метода [63, 64] число итераций уменьшится до $O(\sqrt[3]{\omega} \log(1/\varepsilon))$.

Ошибка $r^n = u^n - u$ приближения u^n удовлетворяет соотношению

$$\Delta \frac{r^{n+1} - r^n}{\theta} + \frac{\partial}{\partial x_j} \left(K_\omega \frac{\partial r^n}{\partial x_j} \right) = 0, \quad r^n \in H_0^1(G). \quad (4.6)$$

Поэтому выполняется равенство

$$\Delta r^{n+1} = (1 - \theta) \Delta r^n \quad \text{при } x \in G_1. \quad (4.7)$$

Возьмем в качестве начального приближения u^0 решение краевой задачи $\Delta u^0 = f$; тогда $\Delta r^0 = 0$ в G_1 . По индукции получаем, что $\Delta r^n = 0 \forall n$ в G_1 , иначе говоря, $r^n \in H_0^1(G)$ принадлежит подпространству Q функций, гармонических в области G_1 .

Произведем более точную оценку снизу $\Lambda(r)$ на подпространстве функций Q . Справедливо неравенство $\Lambda(r) \geq \omega \int_{G_2} (\|Sr\|^2) dx / \int_G (\|Sr\|^2) dx$.

Согласно теореме о продолжении существует функция $\tilde{r} \in H_0^1(G)$, совпадающая с r в G_2 такая, что $\int_{G_2} (\|Sr\|^2) dx \geq \kappa \int_G (\|S\tilde{r}\|^2) dx$, где постоянная

$\kappa > 0$ зависит только от совокупности областей G_1, G_2 . Функция r удовлетворяет уравнению Лапласа в G_1 и поэтому она минимизирует интеграл Дирихле $\int_{G_1} (\|Sr\|^2) dx$ среди всех функций, совпадающих с r на ∂G_1 ; поэтому $\int_G (\|S\tilde{r}\|^2) dx \geq \int_{G_1} (\|Sr\|^2) dx$. Из приведенных соотношений получаем, что $\Lambda(r) \geq \kappa\omega$, то есть $m \geq \kappa\omega$.

Поэтому при $\theta = \frac{2}{(1+\kappa)\omega}$ рассматриваемый итерационный процесс сходится как $O\left(\left(\frac{1-\kappa}{1+\kappa}\right)^n\right)$.

Разделим соотношение (4.4) на ω и перейдем к пределу при $\omega \rightarrow \infty$. Получим итерационный процесс

$$\Delta \frac{u^{n+1} - u^n}{\theta_0} + \frac{\partial}{\partial x_j} \left(K \frac{\partial u^n}{\partial x_j} \right) = 0, \quad \theta_0 = \frac{2}{1+\kappa}, \quad K(x) = \begin{cases} 0 & \text{в } G_1, \\ 1 & \text{в } G_2, \end{cases}$$

сходящийся непосредственно к решению исходной задачи (4.1) со скоростью геометрической прогрессии.

Существенным моментом применения описанного алгоритма является возможность эффективного решения краевой задачи для уравнения с коэффициентами равными, значениям коэффициентов в области, где их значения малы.

Если бы решалась краевая задача для уравнения, в котором коэффициент теплопроводности в области G_2 был бы равен $\omega k(x)$, $0 < \underline{k} \leq k(x) \leq \bar{k} < \infty$, где $k(x)$ не зависит от ω , то, используя (4.7), аналогично получаем, что рассматриваемый итерационный процесс имел бы также скорость сходимости, не зависящую от ω при $\omega \geq 1$.

4.2. Метод потоков

В случае, когда коэффициенты переменные в области, где они малы, можно применить описываемый далее метод потоков [10]. Пусть исходно решается краевая задача

$$L_\omega u = \frac{\partial}{\partial x_j} \left(K_\omega \frac{\partial u_\omega}{\partial x_j} \right) = f, \quad u_\omega|_{\partial G_1} = 0, \quad 0 < \underline{k} \leq k(x) \leq \bar{k} < \infty, \quad (4.8)$$

а продолжение в область G_2 осуществляется, как и выше, с $k = \omega \geq \bar{k}$. Запишем оператор L_ω в виде $L_\omega = S^* K_\omega S$. Поскольку $S^* S = \Delta$, то краевая задача

$$S^* S w = \mathbf{h}, \quad w \in H_0^1(G)$$

может быть решена с малыми вычислительными затратами.

Обозначим $K_\omega S u_\omega$ через \mathbf{p} . Уравнение (4.2) равносильно системе уравнений

$$S u_\omega = K_\omega^{-1} \mathbf{p}, \quad S^* \mathbf{p} = f; \quad (4.9)$$

далее обозначаем u_ω как u .

Обобщенная постановка краевой задачи для системы уравнений (4.9) формулируется следующим образом. Находятся $u \in H_0^1(G)$, $\mathbf{p} \in L^2(G)$ такие, что

$$S u = K_\omega^{-1} \mathbf{p}, \quad (\mathbf{p}, S \varphi) = (f, \varphi) \quad \forall \varphi \in H_0^1(G).$$

Рассмотрим оператор проектирования С.Л. Соболева $P_1 = S(S^* S)^{-1} S^*$ над подпространством $L^2(G)$; под $z = (S^* S)^{-1} S^* \mathbf{p}$ понимается решение задачи:

$$z \in H_0^1(G), \quad (S z, S \varphi) = (\mathbf{p}, S \varphi) \quad \forall \varphi \in H_0^1(G).$$

Оператор P_1 является тождественным над подпространством R_1 функций вида $\mathbf{p} = S u$, $u \in H_0^1(G)$. Оператор $P_2 = E - P_1$ является на этом подпространстве нулевым. Применяя оператор P_2 к первому из соотношений (4.9), получаем равенство

$$P_2(K_\omega^{-1} \mathbf{p}) = \mathbf{0}. \quad (4.10)$$

Пусть H^0 подпространство вектор-функций из $L_2(G)$, удовлетворяющих условию $(\mathbf{h}, S \varphi) = 0 \quad \forall \varphi \in H_0^1(G)$. Оператор P_1 является нулевым, а оператор P_2 — тождественным на этом подпространстве. Операторы P_k являются проекторами: $P_k^2 = P_k$ при $k = 1, 2$.

Рассмотрим итерационный процесс

$$\frac{\mathbf{p}^{n+1} - \mathbf{p}^n}{\theta} + P_2 K_\omega^{-1} \mathbf{p}^n = \mathbf{0} \quad (4.11)$$

решения уравнения (4.10) при начальном приближении

$$\mathbf{p}^0 = P_1 \mathbf{p} = S(S^* S)^{-1} S^* \mathbf{p} = S(S^* S)^{-1} f.$$

Ошибки итераций $\mathbf{r}^n = \mathbf{p}^n - \mathbf{p}$ удовлетворяют равенству

$$\frac{\mathbf{r}^{n+1} - \mathbf{r}^n}{\theta} + P_2 K_\omega^{-1} \mathbf{r}^n = \mathbf{0}. \quad (4.12)$$

Поскольку $\mathbf{r}^0 = P_1 \mathbf{p} - \mathbf{p} = P_2(-\mathbf{p})$, $P_2^2 = P_2$, то $\mathbf{r}^0 \in H_0$.

Если $\mathbf{r}^n = P_2 \mathbf{r}^n$, то применив к (4.12) оператор P_2 , получим

$$\frac{P_2 \mathbf{r}^{n+1} - \mathbf{r}^n}{\theta} + P_2 K_\omega^{-1} \mathbf{r}^n = \mathbf{0}.$$

Отсюда и из (4.12) получаем, что $\mathbf{r}^{n+1} = P_2 \mathbf{r}^{n+1}$. Следовательно, при любом n выполнено равенство $\mathbf{r}^n = P_2 \mathbf{r}^n$, то есть все $\mathbf{r}^n \in H_0$. Поэтому соотношение (4.12) может быть записано в виде

$$\frac{\mathbf{r}^{n+1} - \mathbf{r}^n}{\theta} + P_2 K_\omega^{-1} P_2 \mathbf{r}^n = \mathbf{0}.$$

Если существует $\psi^n \in H_0^1(G)$ такая, что при $x \in G_2$ имеет место равенство $\mathbf{r}^n = S\psi^n$, то при $x \in G_2$ вследствие (4.12) также имеем $\mathbf{r}^{n+1} = S\psi^{n+1}$, где

$$\psi^{n+1} = \psi^n - \theta \left(\frac{\psi^n}{\omega} - (S^* S)^{-1} S^* \frac{\mathbf{r}^n}{\omega} \right).$$

Поскольку $\mathbf{r}^0 = S\psi^0$, $\psi^0 = (S^* S)^{-1} S^* \mathbf{p} - \omega u$, при $x \in G_2$, то индукцией по n получаем, что при любом n и $x \in G_2$ выполнено равенство $\mathbf{r}^n = S\psi^n$, где $\psi^n \in H_0^1(G)$. Таким образом, ошибки \mathbf{r}^n всех приближений принадлежат подпространству R вектор-функций \mathbf{r} , удовлетворяющих условиям:

- 1) $r \in H^0(G)$;
- 2) при $x \in G_2$ выполнено равенство $\mathbf{r} = G\psi$, где $\psi \in H_0^1(G)$.

Следовательно, скорость сходимости итерационного процесса (4.11) определяется границами спектра самосопряженного оператора $P_2 K_\omega^{-1} P_2$ на этом подпространстве.

Поскольку $P_2 \mathbf{r} = \mathbf{r}$ при $\mathbf{r} \in R$, то воспользовавшись (4.8) получаем цепочку соотношений

$$\lambda(\mathbf{r}) = \frac{(P_2 K_\omega^{-1} P_2 \mathbf{r}, \mathbf{r})}{(\mathbf{r}, \mathbf{r})} = \frac{(K_\omega^{-1} \mathbf{r}, \mathbf{r})}{(\mathbf{r}, \mathbf{r})} \leq (\underline{k})^{-1}.$$

С другой стороны,

$$\lambda(\mathbf{r}) = \frac{(K_\omega^{-1} \mathbf{r}, \mathbf{r})}{(\mathbf{r}, \mathbf{r})} \geq (\bar{k})^{-1} \frac{(\mathbf{r}, \mathbf{r})_1}{(\mathbf{r}, \mathbf{r})}.$$

Далее доказательство равномерной по ω оценки $m \leq \lambda(\mathbf{r})$ проводится аналогично соответствующему доказательству для рассматриваемого далее метода псевдопоток, ориентированного на более общий случай. При соответствующем выборе θ итерационный процесс (4.11) сходится со скоростью $O(q^n)$, $q < 1$, равномерно по ω . После нахождения \mathbf{p} значения u_ω могут быть найдены по формуле $u_\omega = \bar{L}^{-1} S^* K_\omega^{-1} \mathbf{p}$.

4.3. Метод псевдопоток

Рассмотрим общий случай, когда коэффициент теплопроводности $k(x)$ переменный на подмножествах G_1 и G_2 и имеет вид [99]

$$k(x) = \begin{cases} a(x) & \text{в } G_1, \\ \omega a(x) & \text{в } G_2. \end{cases}$$

За счет умножения уравнения на некоторый постоянный множитель можно добиться равенства $a \geq 2$. Далее предполагаем $2 \leq a(x) \leq \bar{a} < \infty$, $1 \leq \omega$. Положим

$$\chi_\omega(x) = \begin{cases} 1 & \text{в } G_1, \\ \omega & \text{в } G_2 \end{cases}$$

и запишем исходную систему уравнений в виде

$$S^*Su + L'_\omega u = f, \quad (4.13)$$

где $L'_\omega = S^*M\chi_\omega S$, $M = a(x) - 1/\chi_\omega$. Справедливы неравенства

$$1 \leq M(x) \leq \bar{a}, \quad M_\omega^{-1} \leq M^{-1} \leq 1, \quad \bar{a}^{-1} \leq M^{-1} \quad \forall x \in G. \quad (4.14)$$

Обозначим $M\chi_\omega = M_\omega$, $M_\omega Su = \mathbf{p}$. Уравнение (4.13) можно записать в виде системы уравнений метода, называемого нами методом псевдопотокков [99], [7]:

$$S^*Su + S^*\mathbf{p} = f, \quad Su = (M_\omega)^{-1}\mathbf{p}.$$

Рассматриваемый итерационный процесс имеет вид :

$$\begin{aligned} B \frac{u^{n+1} - u^n}{\theta} + S^*Su^{n+1} + S^*\mathbf{p}^{n+1} &= f, \\ T \frac{\mathbf{p}^{n+1} - \mathbf{p}^n}{\theta} + (M_\omega)^{-1}\mathbf{p}^{n+1} - Su^{n+1} &= 0; \end{aligned} \quad (4.15)$$

здесь T — скаляр, а B — некоторый оператор.

Если выразить \mathbf{p}^{n+1} из второго уравнения через остальные величины и подставить в первое, то получим уравнение относительно u^{n+1} вида

$$(B + D)u^{n+1} = G(u^n, \mathbf{p}^n, f),$$

где $D = \theta S^*(E + (T + \theta(M_\omega)^{-1})^{-1})S$. Далее оператор B выбирается так, чтобы краевая задача для уравнения $(B + D)w = g$ решалась с малыми вычислительными затратами, например, можно взять $B = S^*S - D$. Определенную проблему при использовании этого метода представляет выбор разумного значения параметра θ , поскольку для этого требуется информация о значениях постоянных в теоремах о продолжении функций.

Для рассматриваемого далее симметричного варианта (4.16) метода (4.15) обоснование сходимости со скоростью, не зависящей от ω , производится проще и выбор итерационных параметров может быть осуществлен с помощью вариационного принципа. В то же время метод (4.15) обладает [79] преимуществами в случае дискретной реализации.

Выразив u из первого уравнения и подставив во второе, получим систему уравнений

$$A\mathbf{p} = M_{\omega}^{-1}\mathbf{p} + A_0\mathbf{p} = \mathbf{F},$$

где $A_0\mathbf{p} = S(S^*S)^{-1}S^*\mathbf{p}$, $\mathbf{F} = S(S^*S)^{-1}f$, при этом операторы A и A_0 симметричные.

Обозначим через R_M подпространство вектор-функций из $L_2(G)$, представимых в G_2 в виде $\mathbf{p} = MS\psi_2$, где $\psi_2 \in H_0^1(G)$.

Рассмотрим итерационный процесс

$$M^{-1}\frac{\mathbf{p}^{n+1} - \mathbf{p}^n}{\theta} + A\mathbf{p}^n = \mathbf{F} \quad (4.16)$$

при произвольном начальном приближении $\mathbf{p}^0 \in R_M$. В качестве такого начального приближения можно взять, например, $\mathbf{p}^0 = 0$. Имеем

$$(A_0\mathbf{p}, \mathbf{p}) \leq (\mathbf{p}, \mathbf{p}) \quad \forall \mathbf{p} \in L^2(G).$$

Учитывая, в частности, (4.14), имеем

$$((M_{\omega})^{-1}\mathbf{p}, \mathbf{p}) \leq (M^{-1}\mathbf{p}, \mathbf{p}), \quad (A_0\mathbf{p}, \mathbf{p}) \leq (\mathbf{p}, \mathbf{p}) \leq \bar{a}(M^{-1}\mathbf{p}, \mathbf{p}).$$

Поэтому

$$(A\mathbf{p}, \mathbf{p}) \leq (1 + \bar{a})(M^{-1}\mathbf{p}, \mathbf{p}) \quad \forall \mathbf{p} \in L^2(G).$$

Если $\mathbf{p}^n = MS\psi^n$, $\psi^n \in H_0^1(G)$ в G_2 , то в G_2 вследствие (4.16) также имеем равенство

$$\mathbf{p}^{n+1} = MS\psi^{n+1},$$

где $\psi^{n+1} = \left(1 - \frac{\theta}{\omega}\right)\psi^n - \theta((S^*S)^{-1}S^*\mathbf{p}^n - (S^*S)^{-1}f)$.

В точках области G_2 точное решение задачи $\mathbf{p} = MS\psi$, $\psi = \omega u$, и поэтому $\mathbf{r}^0 = \mathbf{p}^0 - \mathbf{p} \in R_M$. Индукцией по n убеждаемся, что при начальном приближении $\mathbf{p}^0 \in R_M$ все приближения $\mathbf{p}^n \in R_M$.

Оценим снизу $(A\mathbf{r}, \mathbf{r})$ при $\mathbf{r} \in R_M$. Согласно определению A имеем

$$(A\mathbf{r}, \mathbf{r}) = (M^{-1}\mathbf{r}, \mathbf{r}) + (A_0\mathbf{r}, \mathbf{r}) \geq (M^{-1}\mathbf{r}, \mathbf{r})_1 + (A_0\mathbf{r}, \mathbf{r}).$$

Обозначим как Q_M^0 подпространство вектор-функций $\mathbf{q} \in L_2(G)$, представимых в виде $\mathbf{q} = MS\psi$, $\psi \in H_0^1(G)$. Вследствие ортогональности подпространств H_0 и Q_M^0 в скалярном произведении $[\mathbf{f}, \mathbf{g}] = (M^{-1}\mathbf{f}, \mathbf{g})$ всякая функция, в частности функция $\mathbf{r} \in R_M$, представима в виде

$$\mathbf{r} = \mathbf{h} + \mathbf{q}, \quad \mathbf{h} \in H^0, \quad \mathbf{q} = MS\psi \in Q_M^0, \quad (4.17)$$

причем $(M^{-1}\mathbf{r}, \mathbf{r}) = [\mathbf{r}, \mathbf{r}] = [\mathbf{h}, \mathbf{h}] + [\mathbf{q}, \mathbf{q}] = (M^{-1}\mathbf{h}, \mathbf{h}) + (MS\psi, S\psi)$.

Из определения подпространства H^0 следует равенство $A_0\mathbf{h} = 0$; поэтому с учетом симметричности оператора A_0 имеем

$$(A_0\mathbf{r}, \mathbf{r}) = (A_0MS\psi, MS\psi).$$

Можно последовательно доказать справедливость неравенств

$$(A_0\mathbf{r}, \mathbf{r}) \geq \sup_{\varphi \in H_0^1(G)} \frac{(\mathbf{r}, S\varphi)^2}{(S\varphi, S\varphi)},$$

$$(A_0MS\psi, MS\psi) \geq (MS\psi, S\psi) \quad \text{при } M \geq 1, \psi \in H_0^1(G).$$

Покажем, что при $\mathbf{r} \in R_M$ верна оценка $(A\mathbf{r}, \mathbf{r}) \geq \kappa(M^{-1}\mathbf{r}, \mathbf{r}) > 0$. Если \mathbf{r} представлена в виде (4.17), то

$$(A\mathbf{r}, \mathbf{r}) \geq (M^{-1}\mathbf{r}, \mathbf{r})_1 + (A_0\mathbf{r}, \mathbf{r}) = (M^{-1}\mathbf{r}, \mathbf{r})_1 + (MS\psi, S\psi). \quad (4.18)$$

При любых \mathbf{h} и \mathbf{q} и $\eta < 1$ и любом скалярном произведении $[\mathbf{h}, \mathbf{q}]$ имеет место оценка

$$[\mathbf{h} + \mathbf{q}, \mathbf{h} + \mathbf{q}] \geq \eta[\mathbf{h}, \mathbf{h}] - \frac{\eta}{1 - \eta}[\mathbf{q}, \mathbf{q}], \quad (4.19)$$

получаемая в результате вычитания неравенств

$$\begin{aligned} [\mathbf{h} + \mathbf{q}, \mathbf{h} + \mathbf{q}] &\geq [\mathbf{h}, \mathbf{h}] + [\mathbf{q}, \mathbf{q}] - 2|[\mathbf{h}, \mathbf{q}]|, \\ 2|[\mathbf{h}, \mathbf{q}]| &\leq (1 - \eta)[\mathbf{h}, \mathbf{h}] + \frac{1}{1 - \eta}[\mathbf{q}, \mathbf{q}]. \end{aligned}$$

Из (4.19) имеем

$$(M^{-1}\mathbf{r}, \mathbf{r})_1 \geq \eta(M^{-1}\mathbf{h}, \mathbf{h})_1 - \frac{\eta}{1 - \eta}(MS\psi, S\psi)_1.$$

Отсюда и из (4.18) получаем

$$(A\mathbf{r}, \mathbf{r}) \geq \eta(M^{-1}\mathbf{h}, \mathbf{h})_1 - \frac{\eta}{1 - \eta}(MS\psi, S\psi)_1 + (MS\psi, S\psi). \quad (4.20)$$

Покажем, что $(M^{-1}\mathbf{h}, \mathbf{h})_1 \geq \text{const}(M^{-1}\mathbf{h}, \mathbf{h})$. Из определения \mathbf{h} следует

$$(\mathbf{h}, S\varphi)_2 = -(\mathbf{h}, S\varphi)_1 \quad \forall \varphi \in H_0^1(G); \quad (4.21)$$

при этом $\mathbf{h} = MS\psi$ в G_2 . Пусть $\tilde{\psi} \in H_0^1(G)$ продолжение ψ в G такое, что $(S\tilde{\psi}, S\tilde{\psi})_1 \leq C(S\psi, S\psi)_2$.

Положим $\varphi = \tilde{\psi}$, тогда

$$\begin{aligned} |(\mathbf{h}, S\varphi)_2| &= |(M^{-1}\mathbf{h}, \mathbf{h})_2| = |(\mathbf{h}, S\varphi)_1| \leq \\ &\leq \sqrt{(\mathbf{h}, \mathbf{h})_1 (MS\tilde{\psi}, S\tilde{\psi})_1} \leq \sqrt{(\mathbf{h}, \mathbf{h})_1 C(S\psi, S\psi)_2} \leq \sqrt{(\mathbf{h}, \mathbf{h})_1 C(M^{-1}\mathbf{h}, \mathbf{h})_2}. \end{aligned}$$

После возведения в квадрат неравенства

$$|(\mathbf{h}, S\varphi)_2| \leq \sqrt{(\mathbf{h}, \mathbf{h})_1} C(M^{-1}\mathbf{h}, \mathbf{h})_2$$

и сокращения на $(\mathbf{h}, S\varphi)_2 = (M^{-1}\mathbf{h}, \mathbf{h})_2$ получаем

$$(M^{-1}\mathbf{h}, \mathbf{h})_2 \leq C(\mathbf{h}, \mathbf{h})_1 \leq C\bar{a}(M^{-1}\mathbf{h}, \mathbf{h})_1.$$

Отсюда следует оценка $C_0(M^{-1}\mathbf{h}, \mathbf{h}) \leq (M^{-1}\mathbf{h}, \mathbf{h})_1$, где $C_0 = 1/(C\bar{a} + 1)$.

Воспользовавшись (4.20), получаем

$$(\mathbf{A}\mathbf{r}, \mathbf{r}) \geq C_0\eta(M^{-1}\mathbf{h}, \mathbf{h}) - \frac{\eta}{1-\eta}(MS\psi, S\psi)_1 + (MS\psi, MS\psi).$$

Отсюда следует, что

$$(\mathbf{A}\mathbf{r}, \mathbf{r}) \geq C_0\eta(M^{-1}\mathbf{h}, \mathbf{h}) + \left(1 - \frac{\eta}{1-\eta}\right)(MS\psi, S\psi) \geq \kappa(M^{-1}\mathbf{r}, \mathbf{r}),$$

где $\kappa = \min\left(C_0\eta, 1 - \frac{\eta}{1-\eta}\right)$. Наилучшая оценка получится, если величина η будет выбрана из условия

$$C_0\eta = 1 - \frac{\eta}{1-\eta}.$$

Из сказанного выше следует, что при подходящем выборе θ итерационный процесс сходится со скоростью $O(q^n)$ равномерно при $\omega \geq 1$.

Выше нигде не использовалось то обстоятельство, что $\mathbf{r} = MS\psi$ при $x \in G_1$. Поэтому все проведенные построения подвергнутся лишь несущественным изменениям, если вместо переобуславливателя M^{-1} взять переобуславливатель M_0^{-1} , где M_0 совпадает с M в G_2 .

4.4. Вопросы численной реализации

При решении нестационарных задач с помощью неявных разностных схем возникают уравнения относительно значений решения на верхнем слое, содержащие слагаемое с большим множителем $1/\tau^m$, где τ шаг по времени. Оказывается, что для решения таких систем также можно построить [7] модификации описанных методов фиктивных областей с равномерной скоростью сходимости по параметрам τ и параметру продолжения ω .

Описанные методы применимы и в случае более сложных задач, например, в случае эллиптических систем. При этом в [7] рассматриваются и другие многопараметрические задачи. Сюда относятся, в частности, задачи слабосжимаемой упругости, включая случай системы уравнений упругости

для несжимаемых сред, совпадающей с системой уравнений Стокса [7, 100].
Случай квазилинейных задач рассмотрен в [13].

Для всех этих методов получена оценка скорости сходимости, не зависящая от параметра ω в области больших или, соответственно, малых значений этого параметра. Таким образом, в случае метода фиктивных областей, соответствующего граничным условиям Дирихле (ω в G_2 велико), целесообразно применение метода (4.4). В случае метода фиктивных областей, соответствующего граничным условиям Неймана (ω в G_2 мало), целесообразно применение метода (4.11).

В случае смешанных граничных условий через часть границы с граничным условием Дирихле можно производить продолжение с большим значением ω_1 коэффициента k , а через часть границы с граничным условием Неймана производить продолжение с малым значением ω_2 коэффициента k . В [79] для задач с коэффициентом k , принимающим три различных по порядку значения, предложен итерационный процесс с равномерной оценкой скорости сходимости, в последнем случае по обоим этим параметрам ω_1, ω_2 .

Заметим, что ранее в [5, 55, 30] были предложены итерационные методы решения краевых задач, близкие к получающимся при предельном переходе $\omega \rightarrow \infty$ или $\omega \rightarrow 0$ из описанных выше (или совпадающие с ними). Преимущество рассматриваемых нами методов заключается, в частности, в возможности применения их к допредельному случаю $0 < \omega < \infty$, существенному при рассмотрении задач механики композиционных материалов.

Описанные методы в дискретном варианте непосредственно применимы для решения задач с неоднородными граничными условиями: после дискретизации неоднородность граничного условия преобразуется в неоднородность правой части переносом в правую часть уравнений слагаемых, соответствующих граничным условиям.

Существенным моментом возможности применения описанных методов является возможность эффективного решения некоторой вспомогательной сеточной задачи, обычно с сеточными операторами Лапласа или Гельмгольца. Эти задачи можно решать с помощью быстрого преобразования Фурье по одной из переменных (в трехмерном случае по двум переменным) и последующего решения возникающих одномерных дискретных задач методом прогонки. Однако быстрое решение краевых задач с сеточным оператором Лапласа не снимает всех проблем.

Рассмотрим случай, когда краевая задача решается в прямоугольнике G методом конечных элементов с разбиением его на равные прямоугольники G и дальнейшим разбиением их диагональю. При непосредственном применении метода конечных элементов не возникает никаких дополнительных вопросов. Однако если граница между областями с различным порядком значений коэффициентов не проходит по границам треуголь-

ников триангуляции, мы имеем плохую аппроксимацию. Для получения лучшей аппроксимации имеет смысл применить так называемые *локально модифицированные* сетки.

В случае, когда исходная область не прямоугольник, имеет смысл предварительно отобразить исходную область на прямоугольник в непрерывном или дискретном варианте. При этом полезно записать рассматриваемые уравнения в самосопряженной форме. Хотя непосредственное применение многих алгоритмов описанного вида оказывается невозможным, в [9, 79] путем численных экспериментов показывается высокая эффективность предлагаемых алгоритмов, если их преобразовать в некоторые двухступенчатые.

Практически может оказаться наиболее перспективным [79] применение метода (4.15), не требующее двухступенчатости. Дело заключается в следующем. Применение остальных методов требует обязательного обращения какого-то конкретного оператора, определяемого самой задачей; в методе (4.15) можно взять $B = A^{-1} - D$, где A — оператор "спектрально эквивалентный" S^*S и такой, что значения Az вычисляются с малыми вычислительными затратами. В случае иерархически организованных сеток в качестве такого оператора может быть взят оператор, построенный на основе одного шага многосеточного метода.

Основную сложность при теоретическом обосновании применимости описанных методов представляет получение новых теорем продолжения. Для многих практически важных случаев не верна или не доказана необходимая теорема о продолжении. Однако численные эксперименты показали эффективность рассматриваемых методов и ряде таких случаев. Дело в том, что в дискретном случае при шаге дискретизации h часто оказывается, что теорема о продолжении выполняется с константой порядка $O(\ln(1/h))$ и поэтому имеет место сходимость со скоростью $O(q^n)$, где $0 < q \leq 1 - c/\ln(1/h)$.

В случае задач с большим разбросом коэффициентов рассмотренные методы в определенном смысле конкурируют с методом разделения областей [43]. Оказывается [76], что при специальном выборе начального приближения получаемые некоторыми из рассматриваемых методов приближения совпадают с приближениями этого метода. Если G_1 и G_2 не связны и состоят из большого числа параллелепипедов, то, по-видимому, более целесообразно применение описанных методов. Если число таких параллелепипедов мало, то, по-видимому, целесообразнее применение метода разделения областей.

Отметим также, что в случае иерархически организованных сеток для некоторых вариантов многосеточного метода объем вычислительной работы при решении задач с большим разбросом коэффициентов сравним с числом арифметических действий для описанных выше методов.

5. Переобусловленные итерационные методы для линейных систем с симметричными матрицами

Решение систем линейных алгебраических уравнений итерационными методами является одним из важнейших разделов вычислительной математики. В данном разделе мы попытаемся кратко описать только два подхода к эффективному решению систем линейных алгебраических уравнений с симметричными матрицами. А именно, мы рассмотрим переобусловленный метод сопряженных градиентов для систем с положительно определенными (полуопределенными) матрицами и переобусловленный метод минимальных итераций для систем с симметричными матрицами.

Термин "переобуславливатель" появился в математической литературе сравнительно недавно и является прямым переводом англоязычного термина "preconditioner". Во второй части раздела мы дадим описание двух подходов к построению симметричных положительно определенных переобуславливателей. Первый из них основан на методах последовательной релаксации, которые широко используются в приложениях как для непосредственного решения алгебраических систем, так и для построения эффективных переобуславливателей. Второй подход был основан на методах одновременной релаксации. Оба подхода имеют тесную связь с мультипликативным и аддитивным методами Шварца, разработанными для переобуславливания симметричных положительно определенных и полуопределенных матриц, возникающих при дискретизации эллиптических дифференциальных операторов различными вариантами методов конечных элементов. На протяжении этого раздела неявно предполагается, что рассматриваемые алгебраические системы возникают при аппроксимации уравнений в частных производных сеточными методами и имеют достаточно высокую размерность. В то же время описание и анализ рассматриваемых методов и переобуславливателей в явной форме этого не предполагают.

Имеется обширная литература по итерационным методам решения линейных алгебраических систем. В первую очередь это классическая монография [67], в которой с достаточной полнотой изложена теория классических итерационных методов сопряженных направлений (без переобуславливателей) и методов релаксации. Более подробное описание рассматриваемых в разделе методов (см. [74, 17, 28, 49, 53, 56]).

Нашим объектом является система линейных алгебраических уравнений

$$Au = f \tag{5.1}$$

с симметричной вещественной матрицей $A \in \mathbb{R}^{n \times n}$ и вещественным вектором правой части $f \in \mathbb{R}^n$. Если матрица A особенная, то система предполагается совместной, т. е. $f \in im A$ или, что эквивалентно, $f \perp ker A$,

где imA и $ker A$ обозначают образ и ноль-пространство матрицы A , соответственно. Все итерационные методы и способы построения переобуславливателей, которые здесь рассматриваются, применимы как к особенным, так и к неособенным матрицам. Поэтому в тех случаях, где это специально не оговорено, без потери общности мы будем предполагать, что матрица A в системе (5.1) неособенная. По поводу обоснования методов для случая особенных матриц (см. [17, 51]).

Процедура переобусловливания заключается в умножении системы (5.1) на некоторую неособенную матрицу H . В результате мы получаем переобусловленную систему

$$HAu = Hf, \quad (5.2)$$

которая эквивалентна исходной системе (5.1). Основной целью переобусловливания является уменьшение числа обусловленности матрицы системы по отношению к заданной норме $\|\cdot\|_*$ в пространстве \mathbb{R}^n . А именно, нужна такая матрица H , чтобы число обусловленности

$$\text{cond}_*(HA) = \|HA\|_* \|(HA)^{-1}\|_* \quad (5.3)$$

матрицы HA было бы значительно меньше числа обусловленности

$$\text{cond}_*(A) = \|A\|_* \|(A)^{-1}\|_*. \quad (5.4)$$

Если $\|\cdot\|_*$ является эллиптической нормой, генерируемой симметричной положительно определенной матрицей D , т.е. $\|v\|_* = \|v\|_D = (Dv, v)^{1/2}$, $v \in \mathbb{R}^n$, и матрица DHA симметрична и положительно определена, то

$$\text{cond}_D(HA) = \lambda_{\max}/\lambda_{\min},$$

где $\lambda_{\max} = \lambda_{\max}(HA)$ и $\lambda_{\min} = \lambda_{\min}(HA)$ являются максимальным и минимальным собственными числами HA . В этом случае $\text{cond}_D(HA)$ иногда называют спектральным числом обусловленности матрицы HA . При аппроксимации уравнений с частными производными сеточными методами мы имеем дело с последовательностью матриц на последовательности измельчающихся сеток. Для такой последовательности матриц A в ряде случаев удастся построить последовательности матриц H и D таких, что DHA — симметричные и положительно определенные матрицы и последовательность $\text{cond}_D(HA)$ ограничена константой, не зависящей от сеток. Тогда матрицы H^{-1} и A называются спектрально эквивалентными [28].

При реализации итерационных методов для переобусловленных систем умножение векторов на матрицу H является одной из основных численных операций. Очевидно, что она не должна быть арифметически намного "дороже", чем операция умножения на матрицу A . В большинстве практических ситуаций матрица H явно неизвестна. Она задается либо проце-

дурой умножения (например, в многосеточном переобуславливателе), либо матрицей $B = H^{-1}$. В последнем случае на каждом шаге итерационного метода требуется эффективно решать систему с матрицей B .

5.1. Переобусловленный метод сопряженных градиентов

Предположим, что матрицы A и H в (5.2) симметричны и положительно определены. Тогда переобусловленный метод сопряженных градиентов является одним из наиболее эффективных итерационных методов решения системы (5.1). Трехточечные формулы этого метода имеют следующий вид:

$$u^{k+1} = \frac{1}{q_k} [H(Au^k - f) - e_{k-1}(u^k - u^{k-1})], \quad k = 1, 2, \dots \quad (5.5)$$

Здесь u^0 — начальное приближение, а коэффициенты вычисляются по формулам:

$$\begin{aligned} e_{-1} &= 0, \\ q_k &= \frac{\|H\xi^k\|_A^2}{\|\xi^k\|_H^2} - e_{k-1}, \\ e_k &= q_k \frac{\|\xi^{k+1}\|_H^2}{\|\xi^k\|_H^2} - e_{k-1}, \quad k = 0, 1, 2, \dots, \end{aligned} \quad (5.6)$$

где $\xi^k = Au^k - f$ обозначает вектор невязки.

Для метода (5.5)–(5.6) справедлива оценка

$$\|u^k - u^*\|_A \leq 2 \left(\frac{\nu - 1}{\nu + 1} \right)^k \|u^0 - u^*\|_A,$$

где $u^* = A^{-1}f$, а

$$\nu = \text{cond}_A(HA) \quad (5.7)$$

— спектральное число обусловленности матрицы HA . Для больших значений числа обусловленности метод сходится очень медленно, что отражает важность выбора переобуславливателя H с целью минимизации величины ν .

Если матрица A не симметрична, то умножив систему (5.1) на матрицу $A^T H_1$ с некоторой симметричной положительно определенной матрицей H_1 , мы получаем эквивалентную систему

$$Au = f \quad (5.8)$$

с симметричной положительно определенной матрицей

$$A = A^T H_1 A. \quad (5.9)$$

Здесь верхний индекс "T" обозначает транспонирование. Теперь для решения системы (5.8) может быть применен метод (5.5)–(5.6) с некоторым симметричным положительно определенным переобуславливателем H_2 .

Изложенному подходу можно дать другую интерпретацию. Мы можем сразу умножить систему (5.1) на матрицу $H = H_2 A^T H_1$, рассматривая ее как переобуславливатель для системы с несимметричной матрицей. При $D = H_2^{-1}$ матрица $DHA = A^T H_1 A$ симметрична и положительно определена. Тогда, как показано в [51], для ее решения может быть применен специальный вариант метода (5.5), в котором коэффициенты вычисляются по формулам

$$\begin{aligned} e_{-1} &= 0, \\ q_k &= \frac{\|A^T H_1 \xi^k\|_{H_2}^2}{\|\xi^k\|_{H_1}^2} - e_{k-1}, \\ e_k &= q_k \frac{\|\xi^{k+1}\|_{H_1}^2}{\|\xi^k\|_{H_1}^2}, \quad k = 0, 1, \dots, \end{aligned} \quad (5.10)$$

где $\xi^k = Au^k - f$.

В частном случае $H_1 = H_2 = I$, где I — единичная матрица, первый шаг метода (5.5), (5.10) совпадает с одним шагом метода минимальных ошибок [67].

5.2. Переобусловленный метод минимальных итераций

Пусть матрица A системы (5.1) симметрична и H — симметричная положительно определенная матрица. Тогда k -тый шаг переобусловленного варианта классического метода минимальных итераций реализуется следующим образом:

$$g_k = \begin{cases} H\xi^0, & k = 1 \\ HAg_1 - \alpha_2 g_1, & k = 2 \\ HAg_{k-1} - \alpha_k g_{k-1} - \gamma_k g_{k-2}, & k \geq 3, \end{cases} \quad (5.11)$$

где коэффициенты вычисляются по формулам

$$\begin{aligned} \alpha_k &= \frac{(AHAg_{k-1}, HAg_{k-1})}{\|Ag_{k-1}\|_H^2}, \\ \gamma_k &= \frac{\|Ag_{k-1}\|_H^2}{\|Ag_{k-2}\|_H^2}, \\ \beta_k &= \frac{(\xi^{k-1}, Ag_k)}{\|Ag_k\|_H^2}, \quad k = 1, 2, \dots \end{aligned} \quad (5.12)$$

Наиболее важным объектом для применения переобусловленного метода минимальных итераций являются системы (5.1) с блочными матрицами

$$A = \begin{pmatrix} M & C^T \\ C & -\Sigma \end{pmatrix}, \quad (5.13)$$

где матрица M симметрична и положительно определена, а матрица Σ либо положительно определена, либо положительно полуопределена.

Системы с матрицами (5.13) возникают при сеточных аппроксимациях задачи Стокса в механике жидкостей, в смешанных и гибридных методах конечных элементов для эллиптических уравнений, в методах конечных элементов на разрывных сетках (так называемые мортарные методы конечных элементов), в задачах минимизации квадратичных функционалов с ограничениями в виде линейных уравнений.

Эффективным способом переобусловливания системы (5.1) с матрицей (5.13) является выбор H в виде блочно-диагональной матрицы:

$$H = \begin{pmatrix} H_1 & 0 \\ 0 & H_2 \end{pmatrix}, \quad (5.14)$$

где H_1 и H_2 — симметричные положительно определенные матрицы тех же порядков, что и матрицы M и Σ , соответственно.

Как выбирать матрицы H_1 и H_2 ? Хорошую основу для построения этих матриц дает следующий простой факт. Предположим, что $\Sigma = 0$. Выберем $H_1^{-1} = M$, $H_2^{-1} = CM^{-1}C^T$ и рассмотрим задачу на собственные значения

$$Aw = \lambda H^{-1}. \quad (5.15)$$

Прямые вычисления показывают, что все собственные числа $\lambda \neq 0$ являются корнями уравнения

$$\lambda^2 - \lambda - 1 = 0.$$

Отсюда следует, что матрица HA имеет не более трех различных собственных значений: отрицательного $\lambda_1 = (1 - \sqrt{5})/2$ и двух положительных $\lambda_2 = 1$ и $\lambda_3 = (1 + \sqrt{5})/2$. Это означает, что метод (5.11)–(5.13) сходится для любого начального приближения u^0 не более, чем за три итерации.

Рассмотренный выбор H является абсолютно теоретическим и не может быть использован на практике.

С другой стороны, если выбрать H_1^{-1} и H_2^{-1} достаточно "хорошими" приближениями для матриц M и $CM^{-1}C^T$, соответственно, то метод (5.11)–(5.13) имеет очень хорошую скорость сходимости. В частности, если для матриц A , возникающих в сеточных аппроксимациях эллиптических задач, выбрать H_1^{-1} спектрально эквивалентной M , а матрицу H_2^{-1}

спектрально эквивалентной $CM^{-1}C^T$, то матрица H из (5.14) будет спектрально эквивалентной матрице A из (5.13) в том смысле, что собственные числа матрицы HA будут принадлежать объединению отрезков $[a; b]$ и $[c; d]$ с границами $a \leq b < 0 < c \leq d$, не зависящими от размерности n .

Для практики интересной является ситуация, когда $\Sigma = 0$ и матрица A особенная. В этом случае ноль-пространство матрицы A состоит из векторов $w = \begin{pmatrix} 0 \\ v \end{pmatrix}$, где подвектор v принадлежит ноль-пространству матрицы C^T . В этом случае матрица H_2 может быть положительно полуопределенной с $\ker H_2 = \ker C^T$.

5.3. Переобуславливатели, основанные на методах релаксации

Методы релаксации — одни из наиболее развитых в теории итерационных методов. Наиболее известными представителями этой группы методов являются блочные методы последовательной и одновременной релаксации, которые также хорошо известны как блочные методы Гаусса–Зейделя и Якоби, соответственно. Мы изложим общую конструкцию этих методов в стиле, удобном для построения симметричных положительно определенных переобуславливателей для методов, рассмотренных в двух предыдущих разделах.

Предположим, что матрица A симметрична и положительно определена. Разобьем множество целых чисел $N = \{1, 2, \dots, n\}$ на m , вообще говоря, непересекающихся групп N_1, N_2, \dots, N_m . Пусть множество N_s состоит из целых чисел $l_1^{(s)} \leq l_2^{(s)} < \dots < l_{n_s}^{(s)}$, $1 \leq s \leq m$. Определим матрицы

$$Q_s = (q_{ij}^{(s)}) \in R^{n \times n_s}, \quad s = 1, \dots, m, \quad (5.16)$$

где

$$q_{ij}^{(s)} = \begin{cases} 1, & \text{если } j = l_i^{(s)}, \\ 0, & \text{в противном случае,} \end{cases}$$

$i, j = 1, \dots, n_s$, $s = 1, \dots, m$. Например, в случае, когда $N_1 = \{1, 2, \dots, n_1\}$, получаем блочную матрицу

$$Q_1 = \begin{bmatrix} I_1 \\ 0 \end{bmatrix}, \quad (5.17)$$

где I_1 — единичная $n_1 \times n_1$ матрица.

Матрицы

$$A_s = Q_s^T A Q_s \quad (5.18)$$

являются диагональными блоками матрицы A , составленными из ее элементов, расположенных на пересечении строк и столбцов с номерами $l_1^s, \dots, l_{n_s}^s$, $s = \overline{1, m}$. В силу положительной определенности A ее диагональные подматрицы A_s неособенные.

Зададим матрицы

$$H_s = Q_s A_s^{-1} Q_s^T \quad (5.19)$$

и рассмотрим m -шаговый итерационный процесс:

$$\begin{aligned} v^0 &\in \mathbb{R}^n, \\ v^s &= v^{s-1} - H_s(Av^{s-1} - f), \quad s = \overline{1, m}. \end{aligned} \quad (5.20)$$

Этот процесс можно записать в матричном виде

$$v^m = v^0 - H(Av^0 - f), \quad (5.21)$$

где

$$\begin{aligned} H &= I - TA^{-1}, \\ T &= (I - H_m A) \cdots (I - H_1 A). \end{aligned} \quad (5.22)$$

Известно, что матрицы $H_s A$ являются A -ортогональными проекторами. Поэтому, используя предположение $N = \bigcup_{s=1}^m N_s$, можно доказать, что

$$\|T\|_A < 1. \quad (5.23)$$

Рассмотрим другой, так называемый симметричный, вариант релаксационного метода:

$$\begin{aligned} v^0 &\in \mathbb{R}^n, \\ v^s &= v^{s-1} - H_s(Av^{s-1} - f), \quad s = 1, 2, \dots, m, \\ v^s &= v^{s-1} - H_{m-s+1}(Av^{s-1} - f), \quad s = m+1, \dots, 2m. \end{aligned} \quad (5.24)$$

В матричном виде этот процесс может быть записан формулой

$$v^{2m} = v^0 - H(Av^0 - f), \quad (5.25)$$

где

$$H = I - TA^{-1}, \quad (5.26)$$

и

$$T = (I - H_1 A) \cdots (I - H_m A)(I - H_m A) \cdots (I - H_1 A). \quad (5.27)$$

В силу того что $T_m = I - H_m A$ является проектором, $T_m^2 = T_m$, формула (5.26) может быть упрощена. Например, в случае $m = 2$ мы имеем

$$T = (I - H_1 A)(I - H_2 A)(I - H_1 A). \quad (5.28)$$

Нетрудно видеть, что матрица H из (5.26) симметрична и положительно определена. Таким образом, она может быть использована как переобуславливатель H в методах сопряженных градиентов и минимальных итераций. Нужно заметить, что для построения H не обязательно использовать матрицу исходной системы. Например, H может быть построена на основе некоторой матрицы $A = A^T > 0$, которая спектрально эквивалентна матрице A исходной системы.

Рассмотренный вариант блочного релаксационного метода допускает пересечение различных множеств N_s из разбиения $N = \bigcup_{s=1}^m N_s$. Применительно к системам уравнений, возникающих из аппроксимации уравнений эллиптического типа, метод (5.20) известен как мультипликативный метод Шварца, а метод (5.24) — как симметричный мультипликативный метод Шварца. Таким образом, имеется весьма тесная связь между методами, основанными на разбиении сеточных областей на подобласти (domain decomposition method), и релаксационными методами.

В случае, когда множества $\{N_s\}$ взаимно не пересекаются, мы имеем

$$\sum_{s=1}^m Q_s^T Q_s = I, \quad (5.29)$$

где I — единичная $n \times n$ матрица. Предполагая без потери общности, что каждое из N_s содержит только последовательный набор чисел, т. е. $l_j^{(s)} = l_{j-1}^{(s)} + 1$, мы можем представить матрицу A в блочном виде как

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & \cdots & \vdots \\ A_{m1} & \cdots & A_{mm} \end{pmatrix}, \quad (5.30)$$

где $A_{st} \in R^{n_s \times n_s}$ и

$$A_{ss} = Q_s^T A Q_s, \quad s = 1, \dots, m. \quad (5.31)$$

В этом случае матрица $B = H^{-1}$ из (5.22) может быть записана в блочном виде как

$$B = \begin{pmatrix} B_{11} & \cdots & B_{1m} \\ \vdots & & \vdots \\ B_{m1} & \cdots & B_{mm} \end{pmatrix}, \quad (5.32)$$

где

$$B_{st} = \begin{cases} A_{st}, & \text{если } s \geq t; \\ 0, & \text{в противном случае.} \end{cases} \quad (5.33)$$

Таким образом, процедура (5.20) задает один шаг классического блочного метода Гаусса–Зейделя.

Блочная матрица A из (5.30) может быть представлена в виде

$$A = \Lambda + L + L^T, \quad (5.34)$$

где

$$\Lambda = \text{diag}\{A_{11}, \dots, A_{mm}\} \quad (5.35)$$

является блочно-диагональной матрицей и $L = (L_{st})$ является строго нижней треугольной частью A с блоками

$$L_{st} = \begin{cases} A_{st}, & \text{если } s > t; \\ 0, & \text{если } s \leq t. \end{cases} \quad (5.36)$$

Матрица B из (5.32) может быть представлена в виде

$$B = \Lambda + L. \quad (5.37)$$

Определим матрицу

$$B_\omega = \frac{1}{\omega} \Lambda + L, \quad (5.38)$$

где ω — вещественный релаксационный параметр. Тогда итерационный метод

$$B_\omega(u^k - u^{k-1}) = -(Au^{k-1} - f), \quad (5.39)$$

$k = 1, 2, \dots$, называется методом последовательной верхней релаксации (SOR в англоязычной литературе). Хорошо известно [17, 67], что этот метод сходится в R^n для симметричных положительно определенных и полуопределенных матриц в том и только в том случае, когда параметр $\omega \in (0; 2)$.

Симметричный вариант метода имеет вид

$$\begin{aligned} B_\omega(u^{k-\frac{1}{2}} - u^{k-1}) &= -(Au^{k-1} - f), \\ B_\omega^T(u^k - u^{k-\frac{1}{2}}) &= -(Au^{k-\frac{1}{2}} - f), \end{aligned} \quad (5.40)$$

$k = 1, 2, \dots$, и тоже сходится для любого $\omega \in (0; 2)$.

Несложные вычисления показывают, что метод (5.40) может быть также представлен в виде

$$\begin{aligned} (\Lambda + \tau A_1)(u^{k-\frac{1}{2}} - u^{k-1}) &= -\tau(Au^{k-1} - f), \\ (\Lambda + \tau A_2)(u^k - u^{k-\frac{1}{2}}) &= -\tau(Au^{k-\frac{1}{2}} - f), \end{aligned}$$

где

$$A_1 = \frac{1}{2} \Lambda + L, \quad A_2 = \frac{1}{2} \Lambda + L^T. \quad (5.41)$$

Очевидно, что

$$A = A_1 + A_2. \quad (5.42)$$

Продолжая преобразования, получаем

$$(\Lambda + \tau A_1)\Lambda^{-1}(\Lambda + \tau A)(u^k - u^{k-1}) = -2\tau(Au^{k-1} - f), \quad (5.43)$$

где

$$\tau = \frac{\omega}{2 - \omega}, \quad (5.44)$$

или, эквивалентно,

$$\left(\frac{1}{\omega}\Lambda + L\right)\Lambda^{-1}\left(\frac{1}{\omega}\Lambda - L^T\right)(u^k - u^{k-1}) = -(Au^{k-1} - f). \quad (5.45)$$

Таким образом, симметричный метод последовательной верхней релаксации (5.40) может быть записан в форме (5.45) и этот метод также сходится для любого $\omega \in (0; 2)$. Согласно формуле (5.44) метод (5.43) сходится для любого $\tau \in (0; 2)$.

Симметричный метод последовательной верхней релаксации в форме (5.45) предоставляет формулу для симметричного положительно определенного переобуславливателя

$$H_\omega = \left(\frac{1}{\omega}\Lambda - L^T\right)^{-1} \Lambda \left(\frac{1}{\omega}\Lambda + L\right), \quad (5.46)$$

который может быть использован в методах предыдущих двух разделов с любым положительным значением параметра ω .

Блочное разбиение (5.30) матрицы A предоставляет еще одну полезную возможность для выбора симметричной положительно определенной матрицы H на основе метода одновременной блочной релаксации (блочный метод Якоби). А именно, в качестве H может быть выбрана матрица, обратная к матрице Λ из (5.35), т.е.

$$H = \Lambda^{-1}. \quad (5.47)$$

В случае, когда множества N_s пересекаются, матрица H задается формулой

$$H = \sum_{s=1}^m Q_s A_s^{-1} Q_s^T. \quad (5.48)$$

Матрицы A_s^{-1} в (5.48) могут быть заданы любыми симметричными положительно определенными матрицами H_s , которые достаточно близки к ним по спектральным свойствам, например, спектрально эквивалентны. Тогда результирующий переобуславливатель задается формулой

$$H = \sum_{s=1}^m Q_s H_s S_s^T. \quad (5.49)$$

Для систем, возникающих при аппроксимации эллиптических задач сеточными методами, итерационные методы с переобуславливателями (5.49), называемые аддитивными методами Шварца, предложены и исследованы в [56]. Эти методы значительно более эффективны при реализации на параллельных компьютерах, чем рассмотренные ранее методы, которые используют последовательную блочную релаксацию.

6. О численном решении уравнений Навье–Стокса

Уравнения Навье–Стокса, описывающие течение вязкой несжимаемой жидкости, в течение многих десятилетий привлекают внимание ученых, занимающихся проблемами разрешимости уравнений в частных производных, и специалистов в области численного анализа из-за многочисленных приложений. Несмотря на такой интерес, до сих пор остается открытым вопрос о существовании и единственности ”в целом” решения нестационарных уравнений Навье–Стокса в случае трех пространственных переменных.

Ситуация с численным решением этих уравнений носит более сложный характер. Дело в том, что численные методы, хорошо зарекомендовавшие себя при решении одного класса задач, являются неэффективными при решении другого класса. Наличие малого параметра влечет за собой проблемы, связанные с измельчением шага сетки, а решение задач геофизики (динамики атмосферы и океана) требует настолько большого числа узлов сетки, что решение получающихся систем уравнений не под силу современным суперкомпьютерам. С точки зрения обоснования численных методов отсутствует возможность использования ряда результатов теории дифференциальных уравнений, поскольку, как уже упоминалось выше, они являются открытыми для системы уравнений Навье–Стокса.

В настоящей статье приводятся некоторые результаты исследований в области численных методов решения уравнений Навье–Стокса, проводимых в Институте вычислительной математики РАН. При этом основное внимание уделяется не проблемам аппроксимации, а численным алгоритмам решения получающихся в результате аппроксимации систем уравнений.

6.1. Постановка задачи

Итак, пусть $\mathbf{u} = (u_1, u_2, u_3)$ — вектор скоростей, p — функция гидростатического давления, $\mathbf{a} = (f_1, f_2, f_3)$ — вектор-функция источников. Система уравнений Навье–Стокса, описывающая движение вязкой несжимаемой

жидкости, имеет вид [35]

$$\begin{aligned} \mathbf{u}_t - \nu \Delta \mathbf{u} + \nabla p + u_k \mathbf{u}_{x_k} &= \mathbf{f}, \\ \operatorname{div} \mathbf{u} &= 0, \quad \mathbf{u}(x, 0) = \mathbf{u}^0(x), \quad \mathbf{u}|_{\partial G} = \mathbf{0}; \end{aligned} \quad (6.1)$$

здесь $\operatorname{div} \mathbf{u}^0 = 0$, ν — коэффициент вязкости (величина, обратно пропорциональная числу Рейнольдса), решение ищется в ограниченной области G с липшицевой границей ∂G трехмерного пространства R^3 . По одинаковым индексам в произведении здесь и далее подразумевается суммирование в пределах от 1 до 3. С физической точки зрения первое уравнение (6.1) (*уравнение движения*) является законом сохранения моментов движения, в то время как второе уравнение (*уравнение неразрывности*) означает, что жидкость несжимаема.

Следует отдельно рассматривать стационарную и нестационарную задачи. Как правило, решение нестационарной задачи является более сложным, поскольку на каждом шаге по времени необходимо удовлетворять уравнению неразрывности с достаточно высокой степенью точности.

Рассмотрим стационарную задачу. В этом случае (6.1) имеет вид

$$\begin{aligned} -\nu \Delta \mathbf{u} + \nabla p + u_k \mathbf{u}_{x_k} &= \mathbf{f}, \\ \operatorname{div} \mathbf{u} &= 0, \quad \mathbf{u}|_{\partial G} = \mathbf{0}. \end{aligned} \quad (6.2)$$

Заменой

$$\mathbf{u} = \nu \mathbf{v}, \quad p = \nu^2 q, \quad \mathbf{f} = \nu^2 \mathbf{g}$$

задача (6.2) сводится к следующей задаче без коэффициента ν (который присутствует в неявном виде в правой части):

$$\begin{aligned} -\Delta \mathbf{v} + \nabla q + v_k \mathbf{v}_{x_k} &= \mathbf{g}, \\ \operatorname{div} \mathbf{v} &= 0, \quad \mathbf{v}|_{\partial G} = \mathbf{0}. \end{aligned} \quad (6.3)$$

Один из первых подходов к решению задачи (6.3) относится к случаю двух пространственных переменных. Тогда возможно ввести новую неизвестную скалярную функцию (*функцию тока*) ψ следующим образом:

$$\psi_{x_2} = v_1, \quad \psi_{x_1} = -v_2. \quad (6.4)$$

Задача (6.3) в этом случае приобретает вид

$$\Delta^2 \psi + \psi_{x_2} \partial_1 \Delta \psi - \psi_{x_1} \partial_2 \Delta \psi = h \equiv g_{1x_2} - g_{2x_1}. \quad (6.5)$$

Обычно для удобства записи вводят вихрь $\omega = \Delta \psi$. В этом случае уравнение (6.5) преобразуется к системе уравнений второго порядка

$$\begin{aligned} \Delta \omega + \psi_{x_2} \omega_{x_1} - \psi_{x_1} \omega_{x_2} &= h, \\ \Delta \psi - \omega &= 0. \end{aligned} \quad (6.6)$$

Перейдем теперь к постановке граничных условий для системы (6.6). Из определения функции тока ψ и граничных условий (6.1) следует

$$\frac{\partial\psi}{\partial n} = 0, \quad \frac{\partial\psi}{\partial\tau} = 0, \quad (6.7)$$

где n и τ — векторы внешней нормали и касательной к границе. Поскольку из уравнений (6.6) и граничных условий (6.7) функция ψ находится с точностью до постоянного слагаемого, то в случае односвязной области мы можем зафиксировать значение ψ равным нулю в какой-либо точке границы и условия (6.7) в этом случае переходят в

$$\psi = 0, \quad \frac{\partial\psi}{\partial n} = 0. \quad (6.8)$$

Таким образом, в случае односвязной области система уравнений Навье–Стокса сводится к нелинейному бигармоническому уравнению (6.5) и классическим краевым условиям (6.8). Хотя эта задача хорошо исследована с теоретической точки зрения, на практике ее решение вызывает трудности из-за высокого (четвертого) порядка уравнения.

Рассмотрим случай неодносвязной области. Для простоты предположим, что область является двусвязной, т.е. граница ∂G состоит из двух связанных контуров ∂G_1 и ∂G_2 . В этом случае краевые условия (6.7) имеют вид

$$\psi = 0 \text{ на } \partial G_1, \quad \psi = c_2 \text{ на } \partial G_2, \quad \frac{\partial\psi}{\partial n} = 0 \text{ на } \partial G, \quad (6.9)$$

где c_2 — неизвестная постоянная. Таким образом, в случае неодносвязной области итерационный метод решения задачи должен обеспечивать также нахождение значений решения на каждом контуре (за исключением контура, где значение полагается равным нулю).

В случае нестационарной задачи ситуация ухудшается, поскольку в первое уравнение (6.6) добавляется производная по времени от ψ . Кроме этого, для функции тока имеются два граничных условия, в то время как для вихря граничное условие отсутствует. Все это требует разработки специальных численных методов. Как правило, эти методы основаны на введении искусственного граничного условия для вихря (условия Тома, Вудса и т.д.). При этом близость решений приближенной и исходной задач имеет место на достаточно мелкой сетке, поскольку законы сохранения при этом нарушаются. С математической точки зрения эти методы для большинства задач плохо обоснованы — для них отсутствуют теоремы сходимости по сетке.

В связи с высказанными выше замечаниями по поводу использования записи уравнений Навье–Стокса в форме ”функция тока — вихрь” с семидесятих годов прошлого века началось интенсивное исследование системы

уравнений Навье–Стокса и численных методов их решения в исходных переменных ”скорость — давление”. На этом пути удалось продвинуться гораздо дальше и достичь результатов, которые образовали целые направления в области численного анализа. Как правило, полученные на этом пути результаты являются обоснованными с математической точки зрения, и они имеют место как в двумерном так и в трехмерном случаях.

6.2. Задача Стокса

Рассмотрим линейризованную задачу (6.2) — задачу Стокса (для простоты полагаем $\nu = 1$):

$$\begin{aligned} -\Delta \mathbf{u} + \nabla p &= \mathbf{f}, \\ \operatorname{div} \mathbf{u} &= 0, \quad \mathbf{u}|_{\partial G} = \mathbf{0}. \end{aligned} \quad (6.10)$$

Поскольку p определяется из (6.10) с точностью до постоянного слагаемого, будем в дальнейшем считать, что p удовлетворяет условию $\int G p(x) dx = 0$. Известно [35], что обобщенное решение этой задачи $\mathbf{u} \in \mathbf{H}_0^1(G)$, $p \in L_2/R$ существует и единственно; здесь L_2/R — подпространство функций из L_2 , ортогональных единице. (Далее вместо $\mathbf{H}_0^1(G)$ пишем просто \mathbf{H}_0^1 .)

В теории уравнений Навье–Стокса важную роль играет следующее неравенство, позволяющее получать априорные оценки для функции гидростатического давления.

Лемма. Пусть L_2/R — подпространство функций из $L_2(G)$, ортогональных единице. Тогда для любой функции $p \in L_2/R$ справедливо неравенство

$$\|p\| \leq c_0 \sup_{\mathbf{v} \in \mathbf{H}_0^1} \frac{(p, \operatorname{div} \mathbf{v})}{\|\mathbf{v}\|_1}. \quad (6.11)$$

Неравенство (6.11) носит название ”inf-sup”, или LBB-неравенства.

Построим итерационные методы для задачи Стокса. Выпишем двухслойный итерационный процесс с переобуславливателем для задачи (6.10). Вначале итерационный метод запишем в полностью неявном виде:

$$B \frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\tau} - \Delta \mathbf{u}^{n+1} + \nabla p^{n+1} = \mathbf{f},$$

$$\beta(p^{n+1} - p^n) + \operatorname{div} \mathbf{u}^{n+1} = 0,$$

$$\mathbf{u}^{n+1}|_{\partial G} = \mathbf{0}, \quad \mathbf{u}^0 = \mathbf{u}_0 \in \mathbf{H}_0^1, \quad p^0 = p_0 \in L_2/R. \quad (6.12)$$

Здесь B — симметричный положительно определенный оператор спектрально эквивалентный оператору $-\Delta$:

$$-\kappa_1 \Delta \leq B \leq -\kappa_2 \Delta, \quad (6.13)$$

а β и τ — итерационные параметры.

Исследуем алгоритм реализации процесса (6.12). Заметим, что второе уравнение (6.12) явно разрешимо относительно p^{n+1} :

$$p^{n+1} = p^n - \frac{1}{\beta} \operatorname{div} \mathbf{u}^{n+1}. \quad (6.14)$$

Подставляя это выражение в первое уравнение (6.12) и учитывая тот факт, что $\mathbf{u}^{n+1} = \mathbf{u}^n + \tau \frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\tau}$, получим

$$\left(B - \tau \Delta - \frac{\tau}{\beta} \nabla \operatorname{div} \right) \frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\tau} - \Delta \mathbf{u}^n - \frac{1}{\beta} \nabla \operatorname{div} \mathbf{u}^n + \nabla p^n = \mathbf{f}. \quad (6.15)$$

Для реализации (6.15) необходимо, чтобы оператор

$$C = B - \tau \Delta - \frac{\tau}{\beta} \nabla \operatorname{div}$$

был легко обратим. В дальнейшем, для определенности, будем считать, что в области G мы умеем обращаться оператор Лапласа. Тогда естественно положить

$$C = -\Delta. \quad (6.16)$$

В этом случае имеем

$$B = -\Delta + \tau \Delta + \frac{\tau}{\beta} \nabla \operatorname{div}.$$

Оператор B симметричен, а при выполнении неравенства

$$\tau < (1 + 1/\beta)^{-1} \quad (6.17)$$

— положительно определен. В дальнейшем предполагаем, что оператор B положительно определен. Алгоритм реализации итерационного процесса в этом случае имеет вид

$$\begin{aligned} -\Delta \frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\tau} - \Delta \mathbf{u}^n - \frac{1}{\beta} \nabla \operatorname{div} \mathbf{u}^n + \nabla p^n &= \mathbf{f}, \\ p^{n+1} &= p^n - \frac{1}{\beta} \operatorname{div} \mathbf{u}^{n+1}; \quad \mathbf{u}^0 \in \mathbf{H}_0^1, p^0 \in L_2/R. \end{aligned} \quad (6.18)$$

Если \mathbf{u}^n , p^n известны, то \mathbf{u}^{n+1} находится из первого уравнения (6.18) решением задачи Дирихле для уравнения Пуассона. После этого p^{n+1} находится из второго уравнения (6.18) по явной формуле.

Исследуем сходимость итерационного процесса (6.18). Введем обозначения

$$\hat{q} = q^{n+1}, \quad q = q^n, \quad q_t = \frac{\hat{q} - q}{\tau}.$$

Из предыдущих преобразований следует, что метод (6.18) может быть записан в виде полностью неявного итерационного процесса следующим образом:

$$\begin{aligned} B\mathbf{u}_t - \Delta \hat{\mathbf{u}} + \nabla \hat{p} &= \mathbf{f}, \\ \beta\tau p_t + \operatorname{div} \hat{\mathbf{u}} &= 0, \\ \hat{\mathbf{u}}|_{\partial G} &= \mathbf{0}. \end{aligned}$$

Тогда погрешность $\mathbf{v} = \mathbf{u}^n - \mathbf{u}$, $q = p^n - p$ удовлетворяет системе уравнений

$$\begin{aligned} B\mathbf{v}_t - \Delta \hat{\mathbf{v}} + \nabla \hat{q} &= \mathbf{0}, \\ \beta\tau q_t + \operatorname{div} \hat{\mathbf{v}} &= 0, \\ \hat{\mathbf{v}}|_{\partial G} = \mathbf{0}, \quad \mathbf{v}^0 = \mathbf{v}_0 \in \mathbf{H}_0^1, \quad q^0 = q_0 \in L_2/R. \end{aligned} \quad (6.19)$$

Умножим первое уравнение (6.19) скалярно в \mathbf{L}_2 на $2\tau\hat{\mathbf{v}}$, а второе — скалярно в L_2 на $2\tau\hat{q}$. Складывая полученные результаты и учитывая тот факт, что $(B\mathbf{v}_t, \mathbf{v}) = \|\hat{\mathbf{v}}\|_B^2 - \|\mathbf{v}\|_B^2 + \tau^2\|\mathbf{v}_t\|_B^2$, получим

$$\|\hat{\mathbf{v}}\|_B^2 - \|\mathbf{v}\|_B^2 + \tau^2\|\mathbf{v}_t\|_B^2 + 2\tau\|\hat{\mathbf{v}}\|_1^2 + \beta\tau\|\hat{q}\|^2 - \beta\tau\|q\|^2 + \beta\tau^3\|q_t\|^2 = 0. \quad (6.20)$$

Оценим норму $\|\hat{q}\|$ из первого уравнения (6.19). Умножим обе части первого уравнения (6.19) скалярно в \mathbf{L}_2 на $\mathbf{w}/\|\mathbf{w}\|_1$, где $\mathbf{w} \in \overset{\circ}{\mathbf{W}}_2^1$. После интегрирования по частям получим

$$\frac{(\hat{q}, \operatorname{div} \mathbf{w})}{\|\mathbf{w}\|_1} = \frac{(B\mathbf{v}_t, \mathbf{w})}{\|\mathbf{w}\|_1} + \frac{(\nabla \hat{\mathbf{v}}, \nabla \mathbf{w})}{\|\mathbf{w}\|_1}.$$

Оценивая правую часть соотношения по неравенству Коши–Буняковского, имеем

$$\frac{(\hat{q}, \operatorname{div} \mathbf{w})}{\|\mathbf{w}\|_1} \leq \|\mathbf{v}_t\|_B + \|\hat{\mathbf{v}}\|_1.$$

Поскольку правая часть полученного соотношения не зависит от \mathbf{w} , в левой части можно взять \sup по \mathbf{w} . Из полученного соотношения и ”inf-sup”-неравенства следует оценка для $\|\hat{q}\|$:

$$\|\hat{q}\| \leq c_0 (\|\mathbf{v}_t\|_B + \|\hat{\mathbf{v}}\|_1).$$

Возведем обе части неравенства в квадрат и оценим правую часть:

$$\|\hat{q}\|^2 \leq 2c_0^2 (\|\mathbf{v}_t\|_B^2 + \|\hat{\mathbf{v}}\|_1^2).$$

Умножим обе части (6.2.) на $\beta\gamma\tau^2$, где γ — произвольное положительное число. Складывая полученное неравенство с (6.20), имеем

$$\begin{aligned} \|\hat{\mathbf{v}}\|_B^2 - \|\mathbf{v}\|_B^2 + \tau^2 (1 - 2c_0^2\beta\gamma) \|\mathbf{v}_t\|_B^2 + 2\tau(1 - c_0^2\beta\gamma\tau) \|\hat{\mathbf{v}}\|_1^2 + \\ + \beta\tau(1 + \gamma\tau) \|\hat{q}\|^2 - \beta\tau\|q\|^2 \leq 0. \end{aligned} \quad (6.21)$$

Зафиксируем $\beta > 0$. Из (6.17) следует, что существует $\tau_1 > 0$ такое, что при всех $\tau \in (0, \tau_1)$ оператор B будет положительно определенным. Далее, выберем γ , чтобы выполнялось неравенство

$$1 - c_0^2 \beta \gamma \geq 0.$$

Теперь выберем $\tau_2 \leq \tau_1$ таким образом, чтобы выполнялось неравенство

$$1 - c_0^2 \beta \gamma \tau_2 \geq 0.5.$$

Тогда для всех $\tau \in (0, \tau_2)$ из (6.21) и очевидного соотношения $B \leq I$ следует окончательное неравенство

$$(1 + c\tau) \left(\|\hat{\mathbf{v}}\|_B^2 + \beta\tau \|\hat{q}\|^2 \right) \leq \|\mathbf{v}\|_B^2 + \beta\tau \|q\|^2, \quad (6.22)$$

где постоянная $c = \min \{1, \gamma\}$ зависит от c_0 . Таким образом, доказана

Теорема 6.1. *Для любого $\beta > 0$ существует такое $\tau_0 > 0$, что для всех $\tau \in (0, \tau_0)$ оператор B будет положительно определен, итерационный процесс (6.18) будет сходиться со скоростью геометрической прогрессии, а для погрешности будет справедлива оценка (6.22).*

На самом деле можно показать, что для сходимости итерационного процесса (6.18), (6.14) необходимо и достаточно, чтобы оператор B был неотрицателен.

Итерационный метод (6.18) отличается от метода Эрроу-Гурвица наличием на нижнем слое дополнительного члена $\beta^{-1} \nabla \operatorname{div} \mathbf{u}^n$, который можно рассматривать как некоторый "штрафной функционал", который не позволяет итерациям сильно отклоняться от подпространства соленоидальных (удовлетворяющих условию несжимаемости) вектор-функций. Введение этого дополнительного члена в итерационный метод впервые было предложено в [32]. Численные эксперименты показали, что введение этого члена существенным образом ускоряет скорость сходимости метода в нелинейных задачах. Более детальное исследование добавления такого слагаемого было проведено в [70], а в [112] было также установлено, что наличие этого члена улучшает свойства численных методов при решении нестационарных задач.

Другой итерационный метод решения задачи Стокса основан на использовании дополнения по Шуру. А именно, выразим \mathbf{u} из первого уравнения (6.10):

$$\mathbf{u} = \Delta^{-1} (\nabla p - \mathbf{f}).$$

После подстановки этого выражения в уравнение неразрывности получим уравнение относительно функции гидростатического давления

$$Ap \equiv \operatorname{div} \Delta^{-1} \nabla p = \operatorname{div} \Delta^{-1} \mathbf{f}. \quad (6.23)$$

Используя "inf-sup"-неравенство, можно показать (см., например, [33]), что оператор A спектрально эквивалентен единичному оператору I на пространстве L_2/R . А именно, существует постоянная c такая, что имеет место неравенство

$$cI \leq A \leq I. \quad (6.24)$$

При этом постоянная в (6.24) зависит от постоянной c_0 в "inf-sup"-неравенстве.

Оценка (6.24) означает, что для решения уравнения (6.23) могут быть применены классические итерационные методы, например, метод сопряженных градиентов, метод Ричардсона и т.п. При этом на каждом шаге итерационного процесса необходимо вычислять значение оператора A от функции из L_2/R , а именно, значение $q = \operatorname{div} \Delta^{-1} \nabla p^n$. Это можно осуществить следующим образом. Обозначим $\mathbf{v} = \Delta^{-1} \nabla p^n$. Последнее соотношение эквивалентно нахождению обобщенного решения следующей краевой задачи:

$$\Delta \mathbf{v} = \nabla p^n, \quad \mathbf{v}|_{\partial G} = \mathbf{0}.$$

После этого значение $q = \operatorname{div} \mathbf{v}$ находится прямым дифференцированием.

Таким образом, один шаг итерационного метода требует решения двух (трех, в случае трех пространственных переменных) задач Дирихле для уравнения Пуассона.

Скорость сходимости данного итерационного метода будет по порядку такой же, что и в предыдущем случае, однако показатель скорости сходимости будет лучше. К недостаткам метода следует отнести необходимость точного обращения оператора Лапласа при вычислении значения оператора A . Как показали численные эксперименты, это требование является достаточно жестким и оно трудно выполнимо в областях сложной формы.

6.3. Нелинейная стационарная задача

Перейдем к построению и исследованию итерационных методов для решения нелинейной системы уравнений (6.3). Прежде всего следует отметить, что нелинейные члены в (6.3) обладают свойством *кососимметричности* (иногда это свойство называют свойством *нейтральности*). Умножая нелинейный член из (6.3) скалярно в L_2 на \mathbf{v} , получим

$$(v_k \mathbf{v}_{x_k}, \mathbf{v}) = \frac{1}{2} \left(v_k, \frac{\partial \mathbf{v}^2}{\partial x_k} \right) = -\frac{1}{2} \left(\operatorname{div} \mathbf{v}, \mathbf{v}^2 \right) = 0.$$

Более того, из полученного выше соотношения следует, что нелинейность вида $v_k \mathbf{v}_{x_k} + \frac{1}{2} \operatorname{div} \mathbf{v} \cdot \mathbf{v}$ обладает свойством кососимметричности на вектор-функциях из \mathbf{H}_0^1 .

Умножая обе части первого уравнения (6.3) скалярно на \mathbf{v} и учитывая кососимметричность нелинейных членов, получим

$$\|\mathbf{v}\|_1^2 = (\mathbf{f}, \mathbf{v}).$$

Отсюда следует неравенство

$$\|\mathbf{v}\|_1 \leq \|\mathbf{g}\|_{-1}, \quad \|\mathbf{g}\|_{-1} = \sup_{\mathbf{h} \in \mathbf{H}_0^1} \frac{(\mathbf{g}, \mathbf{h})}{\|\mathbf{h}\|_1}. \quad (6.25)$$

Выясним условия [35], которые обеспечивают единственность обобщенного решения задачи (6.3). Предположим, что существует пара \mathbf{u} , p , также являющаяся обобщенным решением задачи (6.3). Тогда разность $\mathbf{w} = \mathbf{u} - \mathbf{v}$, $r = p - q$ является решением краевой задачи

$$\begin{aligned} -\Delta \mathbf{w} + \nabla r + u_k \mathbf{w}_{x_k} + w_k \mathbf{v}_{x_k} &= \mathbf{0}, \\ \operatorname{div} \mathbf{w} &= 0, \quad \mathbf{w}|_{\partial G} = \mathbf{0}. \end{aligned} \quad (6.26)$$

Умножая первое уравнение (6.26) скалярно на \mathbf{w} , получим

$$\|\mathbf{w}\|_1^2 + (w_k \mathbf{v}_{x_k}, \mathbf{w}) = 0.$$

Оценим скалярное произведение в последнем соотношении:

$$|(w_k \mathbf{v}_{x_k}, \mathbf{w})| = |(\mathbf{v}_{x_k}, \mathbf{w} w_k)| \leq c \|\mathbf{v}\|_1 \|\mathbf{w}^2\| = c \|\mathbf{v}\|_1 \|\mathbf{w}\|_{L^4}^2 \leq c_1 \|\mathbf{g}\|_{-1} \|\mathbf{w}\|_1^2,$$

где постоянная c_1 зависит от области G .

Таким образом, решение \mathbf{w} , r задачи (6.26) удовлетворяет неравенству

$$\|\mathbf{w}\|_1^2 \leq c_1 \|\mathbf{g}\|_{-1} \|\mathbf{w}\|_1^2. \quad (6.27)$$

Поэтому, если правая часть \mathbf{f} удовлетворяет неравенству

$$c_1 \|\mathbf{g}\|_{-1} = \frac{c_1 \|\mathbf{f}\|_{-1}}{\nu^2} < 1, \quad (6.28)$$

то (6.27) имеет место лишь при $\|\mathbf{w}\|_1 = 0$, т.е. в этом случае решение единственно.

По аналогии с линейным случаем рассмотрим итерационный метод для решения задачи (6.3). При этом для упрощения изложения нелинейные члены линеаризуем по слоям:

$$\begin{aligned} C \frac{\mathbf{v}^{n+1} - \mathbf{v}^n}{\tau} - \Delta \mathbf{v}^n - \frac{1}{\beta} \nabla \operatorname{div} \mathbf{v}^n + \nabla q^n + N(\mathbf{v}^n, \mathbf{v}^{n+1}) + \\ + K(\mathbf{v}^n, \mathbf{v}^{n+1}) = \mathbf{g}, \quad q^{n+1} = q^n - \frac{1}{\beta} \operatorname{div} \mathbf{v}^{n+1}, \\ \mathbf{v}^{n+1}|_{\partial G} = \mathbf{0}, \quad \mathbf{v}^0 = \mathbf{u}^0 \in \mathbf{H}_0^1, \quad q^0 = p^0 \in L_2/R; \end{aligned} \quad (6.29)$$

здесь $N(\mathbf{v}, \mathbf{w}) = v_k \mathbf{w}_{x_k}$, $K(\mathbf{v}, \mathbf{w}) = \frac{1}{2} \operatorname{div} \mathbf{v} \cdot \mathbf{w}$.

Так же как и при исследовании сходимости итерационного метода в линейном случае, перейдем к записи (6.29) в форме полностью неявного метода. Используя введенные ранее обозначения, имеем

$$\begin{aligned} B\mathbf{v}_t - \Delta \hat{\mathbf{v}} + \nabla \hat{q} + N(\mathbf{v}, \hat{\mathbf{v}}) + K(\mathbf{v}, \hat{\mathbf{v}}) &= \mathbf{g}, \\ \beta\tau q_t + \operatorname{div} \hat{\mathbf{v}} &= 0, \\ \hat{\mathbf{v}}|_{\partial G} = 0, \quad \mathbf{v}^0 = \mathbf{u}^0 \in \mathbf{H}_0^1, \quad q^0 = p^0 \in L_2/R; \end{aligned} \quad (6.30)$$

здесь $B = C + \tau\Delta + \frac{\tau}{\beta}\nabla\operatorname{div}$. В качестве оператора C положим, как и в линейном случае, $C = -\Delta$. Тогда оператор B является симметричным, а при параметрах итерационного процесса β и τ , удовлетворяющих соотношению

$$\tau \left(1 + \frac{1}{\beta}\right) \leq 0.5, \quad (6.31)$$

оператор B будет положительно определенным и спектрально эквивалентным оператору $-\Delta$ так, что

$$-0.5\Delta \leq B \leq -\Delta. \quad (6.32)$$

Зафиксируем $\beta > 0$. Найдем τ_1 из соотношения $\tau_1(1 + 1/\beta) = 0.5$. Тогда для всех $\tau \in (0, \tau_1)$ выполняется неравенство (6.32). Поскольку итерационный процесс нелинеен, докажем вначале равномерную ограниченность итераций (6.30). Умножим первое уравнение (6.30) скалярно в L_2 на $2\tau\hat{\mathbf{v}}$, а второе — скалярно в L_2 на $2\tau\hat{q}$. Складывая результаты, получим

$$\begin{aligned} \|\hat{\mathbf{v}}\|_B^2 - \|\mathbf{v}\|_B^2 + \tau^2\|\mathbf{v}_t\|_B^2 + 2\tau\|\hat{\mathbf{v}}\|_1^2 + \\ + \beta\tau\|\hat{q}\|^2 - \beta\tau\|q\|^2 + \beta\tau^3\|q_t\|^2 = 2\tau(\mathbf{g}, \hat{\mathbf{v}}). \end{aligned} \quad (6.33)$$

Оценим норму $\|\hat{q}\|$ из первого уравнения (6.30). Имеем

$$\|\hat{q}\| \leq c_0 \left(\|\mathbf{v}_t\|_B + \|\hat{\mathbf{v}}\|_1 + c_1\|\mathbf{v}\|_B\|\hat{\mathbf{v}}\|_1 + \|\mathbf{g}\|_{-1} \right).$$

Возводя обе части в квадрат и оценивая правую часть, получим

$$\|\hat{q}\|^2 \leq 4c_0^2 \left(\|\mathbf{v}_t\|_B^2 + \|\hat{\mathbf{v}}\|_1^2 + c_1^2\|\mathbf{v}\|_B^2\|\hat{\mathbf{v}}\|_1^2 + \|\mathbf{g}\|_{-1}^2 \right).$$

Умножим обе части полученного неравенства на $\gamma\beta\tau^2$, где γ — положительное число, которое будет определено далее. Складывая результат с (6.34), имеем

$$\begin{aligned} \|\hat{\mathbf{v}}\|_B^2 - \|\mathbf{v}\|_B^2 + \tau^2(1 - 4c_0^2\gamma\beta)\|\mathbf{v}_t\|_B^2 + 2\tau(1 - 2c_0^2\gamma\beta\tau(1 + c_1^2\|\mathbf{v}\|_B^2))\|\hat{\mathbf{v}}\|_1^2 + \\ + \beta\tau(1 + \gamma\tau)\|\hat{q}\|^2 - \beta\tau\|q\|^2 - 4c_0^2\gamma\beta\tau^2\|\mathbf{g}\|_{-1}^2 \leq \tau\|\mathbf{g}\|_{-1}^2. \end{aligned} \quad (6.34)$$

Пусть $\|\mathbf{v}\|_B^2 + \beta\tau\|q\|^2 \leq M^2$. Найдем условия, которым должны удовлетворять параметры процесса и постоянная M , чтобы выполнялось неравенство $\|\hat{\mathbf{v}}\|_B^2 + \beta\tau\|\hat{q}\|^2 \leq M^2$. Из (6.34) следует оценка

$$\begin{aligned} & \|\hat{\mathbf{v}}\|_B^2 + \tau^2(1 - 4c_0^2\gamma\beta)\|\mathbf{v}_t\|_B^2 + 2\tau(1 - 2c_0^2\gamma\beta\tau(1 + c_1^2M^2))\|\hat{\mathbf{v}}\|_1^2 + \\ & + \beta\tau(1 + \gamma\tau)\|\hat{q}\|^2 \leq \|\mathbf{v}\|_B^2 + \beta\tau\|q\|^2 + \tau(1 + 4c_0^2\gamma\beta\tau)\|\mathbf{g}\|_{-1}^2. \end{aligned} \quad (6.35)$$

Пусть M фиксировано; выберем γ , зависящее только от β , и $\tau_2 \leq \tau_1$ такие, чтобы при всех $\tau \leq \tau_2$ выполнялись неравенства

$$1 - 4c_0^2\gamma\beta \geq 0, \quad 1 - 2c_0^2\gamma\beta\tau(1 + c_1^2M^2) \geq 0.5.$$

Тогда из (6.35) имеем

$$(1 + \tau/2)\|\hat{\mathbf{v}}\|_B^2 + \beta\tau(1 + \gamma\tau)\|\hat{q}\|^2 \leq \|\mathbf{v}\|_B^2 + \beta\tau\|q\|^2 + 2\tau\|\mathbf{g}\|_{-1}^2. \quad (6.36)$$

Из (6.36) следует, что для выполнения условия $\|\hat{\mathbf{v}}\|_B^2 + \beta\tau\|\hat{q}\|^2 \leq M^2$ достаточно, чтобы имела место оценка

$$(1 + c_2\tau)^{-1} \left(M^2 + 2\tau\|\mathbf{g}\|_{-1}^2 \right) \leq M^2,$$

где $c_2 = \min\{0.5, \gamma\}$. Из последнего неравенства получаем оценку для M :

$$M^2 \geq \max \left\{ \|\mathbf{v}^0\|_B^2 + \beta\tau\|q^0\|^2, \frac{2}{c_2} \|\mathbf{g}\|_{-1}^2 \right\}. \quad (6.37)$$

Таким образом, мы доказали, что для любого $\beta > 0$ можно найти τ_2 такое, что при всех $\tau \leq \tau_2$ оператор B будет спектрально эквивалентен оператору $-\Delta$ и все итерации процесса (6.29) будут в совокупности ограничены постоянной, удовлетворяющей (6.37). Отметим, что постоянная M по порядку совпадает с постоянной априорной оценки (6.25).

Исследуем теперь сходимость итерационного процесса. Обозначим через \mathbf{u} , p решение задачи (6.3). Погрешность $\mathbf{w} = \mathbf{v} - \mathbf{u}$, $r = q - p$ удовлетворяет системе уравнений

$$\begin{aligned} B\mathbf{w}_t - \Delta\hat{\mathbf{w}} + \nabla\hat{r} + N(\mathbf{w}, \hat{\mathbf{v}}) + N(\mathbf{u}, \hat{\mathbf{w}}) + K(\mathbf{w}, \hat{\mathbf{v}}) &= \mathbf{0}, \\ \beta\tau r_t + \operatorname{div}\hat{\mathbf{w}} &= 0, \\ \hat{\mathbf{w}}|_{\partial G} &= \mathbf{0}, \quad \mathbf{w}^0 = \mathbf{v}^0 - \mathbf{u} \in \mathbf{H}_0^1, \quad r^0 = q^0 - p \in L_2/R. \end{aligned} \quad (6.38)$$

Умножим первое уравнение (6.38) скалярно в L_2 на $2\tau\hat{\mathbf{w}}$, а второе — скалярно в L_2 на $2\tau\hat{r}$. Складывая результаты, получим

$$\begin{aligned} & \|\hat{\mathbf{w}}\|_B^2 - \|\mathbf{w}\|_B^2 + \tau^2\|\mathbf{w}_t\|_B^2 + 2\tau\|\hat{\mathbf{w}}\|_1^2 + \beta\tau\|\hat{r}\|^2 - \beta\tau\|r\|^2 + \\ & + \beta\tau^3\|r_t\|^2 + 2\tau(N(\mathbf{w}, \hat{\mathbf{v}}), \hat{\mathbf{w}}) + 2\tau(K(\mathbf{w}, \hat{\mathbf{v}}), \hat{\mathbf{w}}) = 0. \end{aligned} \quad (6.39)$$

Оценим скалярные произведения в (6.39). Имеем

$$\begin{aligned} 2\tau |(N(\mathbf{w}, \hat{\mathbf{v}}), \hat{\mathbf{w}})| &= 2\tau |(w_k \hat{\mathbf{v}}_{x_k}, \hat{\mathbf{w}})| \leq c\tau \|\hat{\mathbf{v}}\|_1 \|\mathbf{w}\|_{L_4} \|\hat{\mathbf{w}}\|_{L_4} \leq \\ &\leq c_3\tau M (\|\mathbf{w}\|_B^2 + \|\hat{\mathbf{w}}\|_B^2). \end{aligned}$$

Оценка второго скалярного произведения имеет аналогичный вид.

Таким образом, из (6.39) получаем

$$\begin{aligned} \|\hat{\mathbf{w}}\|_B^2 + \tau^2 \|\mathbf{w}_t\|_B^2 + \tau(2 - c_3M) \|\hat{\mathbf{w}}\|_1^2 + \beta\tau \|\hat{r}\|^2 &\leq \\ &\leq (1 + c_3M\tau) \|\mathbf{w}\|_B^2 + \beta\tau \|r\|^2. \end{aligned} \quad (6.40)$$

Оценим норму $\|\hat{r}\|$ из первого уравнения (6.38) как это делалось ранее. Имеем

$$\|\hat{r}\| \leq c_0 (\|\mathbf{w}_t\|_B + \|\hat{\mathbf{w}}\|_1 + cM \|\mathbf{w}\|_B + cM \|\hat{\mathbf{w}}\|_B).$$

Возведем обе части неравенства в квадрат и оценим правую часть:

$$\|\hat{r}\|^2 \leq 4c_0^2 \left(\|\mathbf{w}_t\|_B^2 + \|\hat{\mathbf{w}}\|_1^2 + c_4M^2 \|\mathbf{w}\|_B^2 + c_4M^2 \|\hat{\mathbf{w}}\|_B^2 \right).$$

Умножим обе части на $\beta\gamma\tau^2$, где $\gamma > 0$ будет выбрано далее, и сложим с (6.40). В результате получим

$$\begin{aligned} \|\hat{\mathbf{w}}\|_B^2 + \tau^2(1 - 4c_0^2\beta\gamma) \|\mathbf{w}_t\|_B^2 + \tau(2 - c_3\tau M^2 - 4c_0^2c_4\beta\gamma\tau M^2) \|\hat{\mathbf{w}}\|_1^2 + \\ + \beta\tau(1 + \gamma\tau) \|\hat{r}\|^2 \leq (1 + c_3M^2\tau + 4c_0^2c_4\beta\gamma\tau M^2) \|\mathbf{w}\|_B^2 + \beta\tau \|r\|^2. \end{aligned} \quad (6.41)$$

Из (6.41) следует, что при выполнении условия $M < c_5$ или, что то же самое,

$$\frac{\|\mathbf{f}\|_{-1}}{\nu^2} < c_5, \quad (6.42)$$

и начальном условии, удовлетворяющем неравенству

$$\|\mathbf{v}^0\|_B^2 + \beta\tau \|q^0\|^2 \leq \|\mathbf{g}\|_{-1}^2$$

(последнего легко добиться полагая, например, $\mathbf{v}^0 = \mathbf{0}$, $q^0 = 0$), при достаточно малом τ итерационный процесс будет сходиться со скоростью геометрической прогрессии.

В исследуемом итерационном процессе на каждом шаге необходимо решать систему линейных алгебраических уравнений с несимметричной матрицей. Если нелинейные члены брать с нижнего слоя, то матрица будет симметричной, однако для равномерной ограниченности итераций необходимо брать достаточно малое β . В остальном техника исследования остается прежней. Стоит отметить тот факт, что условие на норму правой ча-

сти, обеспечивающее сходимость итераций, совпадает по порядку с условием [35], гарантирующим существование и единственность решения исходной задачи. Отметим также, что все рассуждения остаются в силе, если оператор Лапласа в переобуславливателе заменить другим спектрально-эквивалентным ему оператором (аналогично тому, как это предлагается в п. 4.4).

Сделаем несколько замечаний относительно численного решения нестационарной задачи. В простых ситуациях возможен стандартный путь. А именно, аппроксимируем задачу по времени какой-либо разностной схемой, беря нелинейные члены с предыдущего шага по времени. Тогда на каждом шаге по времени необходимо будет решать "обобщенную задачу Стокса" (задачу Стокса с дополнительным членом вида $\alpha \mathbf{u}$, где α — большой параметр). Тем не менее, при решении более сложных задач прямое решение уравнений Навье–Стокса вызывает существенные трудности. В связи с этим на практике используют методы расщепления, которые позволяют сводить задачу к последовательности более простых задач.

Мы не касались в статье методов аппроксимации уравнений Навье–Стокса. Эти методы довольно детально исследованы в ряде монографий, а некоторые из них легли в основу пакетов программ для моделирования задач динамики вязкой несжимаемой жидкости. Тем не менее, несмотря на наличие таких пакетов, существует большое число задач, которые не могут быть успешно решены таким образом и требуют разработки специальных методов решения.

7. О применении идей метода дискретных вихрей в задаче распространения звука в мелкой воде

Звуковые волны — единственные из волн известной физической природы способны распространяться под водой на расстояние многих сотен километров. Эта уникальная способность акустических волн и обуславливает тот практический интерес, который проявляют к особенностям распространения звука в океане ученые и инженеры самых разных специальностей.

Распространение звуковых волн под водой — крайне сложный процесс, трудность описания которого обусловлена большим многообразием явлений и свойств, присущих различным районам Мирового океана. В первую очередь это относится к шельфовым зонам океана, где особенности распространения звуковых волн связаны с их взаимодействием с морским дном, а также со специфическим характером гидродинамических возмущений. Вместе с тем океанский шельф наиболее важен для жизнедеятельности человека и, следовательно, больше всего интересен для исследова-

дователей. Континентальные шельфы являются первоочередными объектами для постановки геолого-разведочных работ на нефть и газ. К настоящему времени усилиями разных групп получено большое количество экспериментальных и теоретических результатов [2, 14, 78, 85], позволяющих выделить акустику мелкого моря в самостоятельный раздел акустики океана.

Большой вклад в развитие теории распространения звука в слое воды внес академик Л.М. Бреховских [15], который изучал распространение как электромагнитных, так и звуковых волн в слоистых средах. Он ввел лучевые представления для электромагнитных и звуковых волн в случае, когда поля создаются монохроматическим источником. В этом случае задача сводится к решению краевой задачи для скалярного уравнения Гельмгольца относительно звукового давления.

Лучевые представления позволяют решать задачи в тех случаях, когда поверхность воды и дно являются плоскостями или дно может иметь небольшой наклон.

Метод интегральных уравнений налагает меньше ограничений на геометрию границы воды с воздухом и на дно. Однако специалисты в области распространения звука в воде не применяли метод интегральных уравнений, вероятно, по следующим обстоятельствам. Если применить теорию потенциала, то поскольку на границе воды и воздуха давление должно быть непрерывно, на этой границе необходимо расположить потенциал простого слоя. Так как на дне непрерывна нормальная производная от давления звука, на дне надо расположить потенциал двойного слоя. Выполняя теперь граничные условия, мы получим систему двух интегральных уравнений с двумя неизвестными функциями. При этом одно из интегральных уравнений будет гиперсингулярным. Однако численных методов решения гиперсингулярных интегральных уравнений до 80-х годов прошлого столетия не было и поэтому ученые, занимающиеся теорией распространения звука в слое воды, уходили от решения таких интегральных уравнений. Теперь, когда создана технология численного решения гиперсингулярных интегральных уравнений в задачах аэродинамики [11, 47, 16], такие краевые задачи можно решать численно, в том числе и в тех случаях, когда дно может иметь ребра и конические вершины, т.е. когда дно может иметь разломы и горы.

В данной работе дано сведение различных краевых задач для уравнения Гельмгольца, к которым сводятся некоторые задачи распространения звука в мелком море, к системе интегральных уравнений и предложен метод их численного решения, использующий идеи метода замкнутых дискретных вихревых рамок (МЗДВР). В некоторых случаях даны аналитические решения интегральных уравнений, в других — численное решение.

7.1. Математическая модель задачи для слоя воды на абсолютно отражающем дне на основе интегральных уравнений

Пусть имеется слой воды, занимающий область D (см. рис. 4).

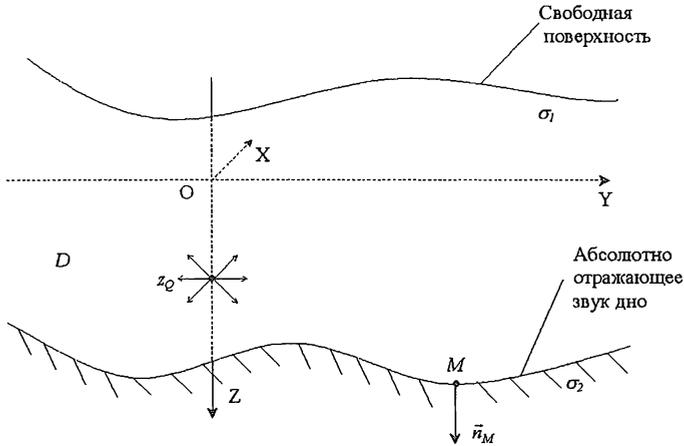


Рис. 4. Область D

Границы этой области — поверхность σ_1 , являющаяся границей между водой и воздухом, и поверхность σ_2 , являющаяся абсолютно отражающим дном. Пусть в точке $M_Q(0, 0, z_Q) \in D$ расположен источник звука, создающий звуковое давление $p(M)$ по формуле

$$p_Q(M) = \frac{Q}{4\pi} \frac{e^{ikr_{MM_Q}}}{r_{MM_Q}}, \quad (7.1)$$

где k — волновое число, Q — мощность источника.

Требуется найти звуковое давление $p(M)$, возникающее в области D , удовлетворяющее следующим условиям:

$$\Delta p(M) + k^2 p(M) = 0, \quad M \in D \setminus M_Q; \quad (7.2)$$

$$p(M) = 0, \quad M \in \sigma_1; \quad (7.3)$$

$$\frac{\partial p(M)}{\partial n_M} = 0, \quad M \in \sigma_2; \quad (7.4)$$

и условию излучения на бесконечности:

$$\left(\frac{\vec{r}_M}{r_M}, \text{grad } p_Q(M) \right) - ikp_Q(M) = O\left(\frac{1}{r_{MM_Q}} \right), \quad (7.5)$$

при $r_{MM_Q} \rightarrow \infty$, где $r_M = |\vec{r}_M| = |x\vec{i} + y\vec{j} + z\vec{k}| = \sqrt{x^2 + y^2 + z^2}$.

Функцию $p(M)$ будем искать в виде

$$p(M) = p_Q(M) + p_1(M) + p_2(M), \quad (7.6)$$

где

$$p_1(M) = \frac{1}{4\pi} \int_{\sigma_1} \frac{e^{ikr_{MM_0}}}{r_{MM_0}} g_1(M_0) d\sigma_{1,M_0}, \quad M \notin \sigma_1,$$

$$p_2(M) = \frac{1}{4\pi} \int_{\sigma_2} \frac{\partial}{\partial n_{M_0}} \left(\frac{e^{ikr_{MM_0}}}{r_{MM_0}} \right) g_2(M_0) d\sigma_{2,M_0}, \quad M \notin \sigma_2. \quad (7.7)$$

Тогда функция $p(M)$, задаваемая формулой (7.6), удовлетворяет уравнению (7.2) и будет удовлетворять условиям (7.3) и (7.4), если функции $g_1(M)$, $M \in \sigma_1$ и $g_2(M)$, $M \in \sigma_2$ будут решениями следующей системы интегральных уравнений:

$$\begin{aligned} & \frac{1}{4\pi} \int_{\sigma_1} \frac{e^{ikr_{MM_0}}}{r_{MM_0}} g_1(M_0) d\sigma_{1,M_0} + \\ & + \frac{1}{4\pi} \int_{\sigma_2} \frac{\partial}{\partial n_{M_0}} \left(\frac{e^{ikr_{MM_0}}}{r_{MM_0}} \right) g_2(M_0) d\sigma_{2,M_0} = \\ & = -\frac{Q}{4\pi} \frac{e^{ikr_{MM_Q}}}{r_{MM_Q}}, \quad M \in \sigma_1; \end{aligned} \quad (7.8)$$

$$\begin{aligned} & \frac{1}{4\pi} \int_{\sigma_1} \frac{\partial}{\partial n_M} \left(\frac{e^{ikr_{MM_0}}}{r_{MM_0}} \right) g_1(M_0) d\sigma_{1,M_0} + \\ & + \frac{1}{4\pi} \int_{\sigma_2} \frac{\partial}{\partial n_M} \frac{\partial}{\partial n_{M_0}} \left(\frac{e^{ikr_{MM_0}}}{r_{MM_0}} \right) g_2(M_0) d\sigma_{2,M_0} = \\ & = -\frac{Q}{4\pi} \frac{\partial}{\partial n_M} \left(\frac{e^{ikr_{MM_Q}}}{r_{MM_Q}} \right), \quad M \in \sigma_2. \end{aligned} \quad (7.9)$$

Уравнение (7.8) является слабосингулярным интегральным уравнением первого рода относительно функции $g_1(M)$ на поверхности σ_1 , а уравнение (7.9) — гиперсингулярным интегральным уравнением первого рода относительно функции $g_2(M)$ на поверхности σ_2 .

Так как в достаточно большом числе задач распространения звука в слое воды можно считать свободную поверхность воды σ_1 плоскостью, то решение задачи (7.2)–(7.4) можно свести не к решению системы интегральных уравнений (7.8), (7.9), а к решению одного гиперсингулярного интегрального уравнения на поверхности σ_2 . Покажем как это сделать.

Будем считать, что σ_1 — плоскость $z = 0$. В точке $M_Q(0, 0, z_Q)$, $z_Q > 0$ и $M_Q \in D$, расположен источник звука, который порождает звуковое давление $p_Q(M)$ как указано выше. Возьмем теперь источник звука кососимметричный относительно плоскости $z = 0$ исходному, т.е. источник звука, помещенный в точке $\tilde{M}_Q(0, 0, -z_Q)$ и дающий звуковое давление $\tilde{p}_Q(M)$ по формуле

$$\tilde{p}_Q(M) = -\frac{Q}{4\pi} \frac{e^{ikr_{M\tilde{M}_Q}}}{r_{M\tilde{M}_Q}}. \quad (7.10)$$

Далее, на поверхности σ_2 расположим потенциал двойного слоя с плотностью $g_2(M)$, дающий звуковое давление $p_2(M)$ по формуле (7.7). Возьмем поверхность σ_2^* , симметричную относительно плоскости $z = 0$ поверхности σ_2 . Например, если уравнение поверхности σ_2 имеет вид

$$z = h(x, y),$$

где $h(x, y)$ достаточно гладкая на пространстве R^2 (плоскости OXY) функция и $h(x, y) \geq a > 0$ для любой точки $M(x, y) \in R^2$. Тогда уравнение поверхности σ_2^* будет иметь вид

$$z = -h(x, y), (x, y) \in R^2.$$

На поверхности σ_2^* расположим потенциал двойного слоя кососимметричный исходному, т.е. имеющий в точке $M(x, y, -h(x, y))$ плотность $g_2^*(M) = -g_2(x, y, h(x, y))$, который создает звуковое давление $p_2^*(M)$ по формуле

$$p_2^*(M) = \frac{1}{4\pi} \int_{\sigma_2^*} \frac{\partial}{\partial n_{M_0}} \left(\frac{e^{ikr_{MM_0}}}{r_{MM_0}} \right) g_2^*(M_0) d\sigma_{2, M_0}^*,$$

$$g_2^*(x, y, -h(x, y)) = -g_2(x, y, h(x, y)).$$

Тогда функция $p(M)$, определяемая формулой

$$p(M) = p_Q(M) + \tilde{p}_Q(M) + p_2(M) + p_2^*(M), \quad M \in D,$$

удовлетворяет условию (7.3) на границе σ_1 , т.е. на плоскости $z = 0$. Эта функция $p(M)$ будет удовлетворять условию (7.4) на поверхности σ_2 , если функция $g_2(x, y, h(x, y))$ будет решением интегрального уравнения:

$$\begin{aligned} & \frac{1}{4\pi} \int_{\sigma_2} \frac{\partial}{\partial n_M} \frac{\partial}{\partial n_{M_0}} \left(\frac{e^{ikr_{MM_0}}}{r_{MM_0}} \right) g_2(M_0) d\sigma_2 + \\ & + \frac{1}{4\pi} \int_{\sigma_2^*} \frac{\partial}{\partial n_M} \frac{\partial}{\partial n_{M_0}} \left(\frac{e^{ikr_{MM_0}}}{r_{MM_0}} \right) g_2^*(M_0) d\sigma_2^* = \\ & = -\frac{Q}{4\pi} \frac{\partial}{\partial n_M} \left(\frac{e^{ikr_{MM_Q}}}{r_{MM_Q}} \right) + \frac{Q}{4\pi} \frac{\partial}{\partial n_M} \left(\frac{e^{ikr_{M\tilde{M}_Q}}}{r_{M\tilde{M}_Q}} \right), \quad M \in \sigma_2. \end{aligned} \quad (7.11)$$

Уравнение (7.11) можно записать как уравнение на области определения функции $z = h(x, y)$, т.е. на поверхности $\sigma_1 = R^2$. Для этого заметим, что

$$d\sigma_2^* = d\sigma_2 = \sqrt{1 + h'_{x_0}{}^2(x_0, y_0) + h'_{y_0}{}^2(x_0, y_0)} dx_0 dy_0,$$

где $M_0(x_0, y_0, h(x_0, y_0))$ и $M_0^*(x_0, y_0, -h(x_0, y_0))$.

Далее имеем

$$g_2^*(M_0^*) = g_2^*(x_0, y_0, -h(x_0, y_0)) = -g_2(x_0, y_0, h(x_0, y_0)) = -g_2(M_0),$$

где $M_0 \in \sigma_2$, $M_0^* \in \sigma_2^*$. Если точка $M_0 \in \sigma_2$, то

$$\vec{n}_{M_0} = \frac{-h'_{x_0}(x_0, y_0)\vec{i} - h'_{y_0}(x_0, y_0)\vec{j} + \vec{k}}{\sqrt{1 + h'_{x_0}{}^2(x_0, y_0) + h'_{y_0}{}^2(x_0, y_0)}},$$

а если $M_0^* \in \sigma_0^*$, то

$$\vec{n}_{M_0^*} = \frac{-h'_{x_0}(x_0, y_0)\vec{i} - h'_{y_0}(x_0, y_0)\vec{j} - \vec{k}}{\sqrt{1 + h'_{x_0}{}^2(x_0, y_0) + h'_{y_0}{}^2(x_0, y_0)}},$$

где через $\vec{n}_{M_0^*}$ обозначен орт нормали к поверхности σ_2^* в точке M_0^* , симметричный вектору \vec{n}_{M_0} относительно плоскости $z = 0$. Положим

$$\begin{aligned} K_{N,H}(M, M_0) &= \frac{\partial}{\partial n_M} \frac{\partial}{\partial n_{M_0}} \left(\frac{e^{ikr_{MM_0}}}{r_{MM_0}} \right) = \\ &= e^{ikr_{MM_0}} \left[\frac{(1 - ikr_{MM_0})(\vec{n}_M, \vec{n}_{M_0})}{r_{MM_0}^3} + \right. \\ &\left. + \left(\frac{k^2}{r_{MM_0}^3} - \frac{3(1 - ikr_{MM_0})}{r_{MM_0}^5} \right) (\vec{r}_{MM_0}, \vec{n}_M) (\vec{r}_{MM_0}, \vec{n}_{M_0}) \right], \end{aligned} \quad (7.12)$$

$$\begin{aligned} f(M, M_0) &= \frac{\partial}{\partial n_M} \left(\frac{e^{ikr_{MM_Q}}}{r_{MM_Q}} \right) = \\ &= e^{ikr_{MM_Q}} \frac{(1 - ikr_{MM_Q})(\vec{r}_{MM_Q}, \vec{n}_M)}{r_{MM_Q}^3}. \end{aligned} \quad (7.13)$$

Тогда уравнение (7.11) получит вид

$$\begin{aligned} \frac{1}{4\pi} \int_{R^2} (K_{N,H}(M, M_0) - K_{N,H}(M, M_0^*)) g_2(M_0) \times \\ \times \sqrt{1 + h'_{x_0}{}^2(x_0, y_0) + h'_{y_0}{}^2(x_0, y_0)} dx_0 dy_0 = \\ = \frac{Q}{4\pi} (-f(M, M_Q) + f(M, \tilde{M}_Q)), \quad M \in \sigma_2, \end{aligned} \quad (7.14)$$

где $M = M(x, y, h(x, y))$, $M_0 = M_0(x_0, y_0, h(x_0, y_0))$, $M_Q = M(0, 0, z_Q)$, $M_0^* = M_0(x_0, y_0, -h(x_0, y_0))$, $M_Q = M(0, 0, -z_Q)$. Из формулы (7.12) следует, что уравнение (7.14) является гиперсингулярным интегральным уравнением на плоскости R^2 .

В случае прибрежной полосы задача сводится опять к гиперсингулярному интегральному уравнению вида (7.14), только теперь интеграл будет браться не по плоскости R^2 , а по полупространству R_L^2 , ограниченному кривой L , являющейся береговой линией, то есть $L = \sigma_1 \cap \sigma_2$, и параметрическая точка (x, y) также принадлежит этому полупространству R_L^2 .

7.2. Метод численного решения задачи для слоя воды на абсолютно отражающем дне

Как было показано в предыдущем подразделе, решение задачи (7.2)–(7.4) сводится к решению гиперсингулярного интегрального уравнения (7.11) для случая, когда поверхность σ_1 — плоскость $z = 0$. Очевидно, что для произвольной гладкой поверхности σ_2 уравнение (7.11) необходимо решать численно. В данной работе предлагается использовать метод замкнутых дискретных вихревых рамок, широко применяемый в аэродинамике.

Для простоты рассуждений рассмотрим решение гиперсингулярного интегрального уравнения на части плоскости σ , являющемся гладким образом квадрата, т.е. уравнения

$$\frac{1}{4\pi} \int_{\sigma} \frac{\partial}{\partial n_M} \frac{\partial}{\partial n_{M_0}} \left(\frac{e^{ikr_{MM_0}}}{r_{MM_0}} \right) g(M_0) d\sigma_2 = -\frac{\partial p_Q(M)}{\partial n_M}. \quad (7.15)$$

Для численного решения зададим поверхность σ параметрически $\sigma = \sigma(\xi, \zeta)$, где ξ и ζ — параметры, заданные на множестве $\Pi = \{0 \leq \xi < 1, 0 \leq \zeta < 1\}$ — единичном квадрате. В соответствии с МЗДВР [47, 25] разобьем поверхность равномерно по параметрам на $s_1 \cdot s_2$ частей. При этом $s_1 + 1$ — число координатных линий $\xi_i = const$, $i = 1, 2, \dots, s_1$, а $s_2 + 1$ — число линий $\zeta_i = const$, $i = 1, 2, \dots, s_2$. Точки пересечения координатных линий обозначим через $M_{ij}^0 (x_{ij}^0, y_{ij}^0, z_{ij}^0)$. Таким образом, поверхность σ состоит из $s_1 \cdot s_2$ ячеек, которые имеют вершины $M_{i-1j-1}^0, M_{ij-1}^0, M_{ij}^0, M_{i-1j}^0$. Заменяем эти ячейки четырехугольниками с теми же вершинами. Расчетную точку M_{ij} следует помещать на пересечении отрезков, соединяющих середины противоположных боковых сторон соответствующей ячейки.

Согласно классическому подходу [25] для численного решения гиперсингулярного интегрального уравнения (7.15) полагается, что величины,

$g_{ij} = g(M_{ij})$ постоянны для каждой ячейки и уравнение (7.15) записывается в расчетных точках M_{ij} :

$$\sum_{i=1}^{s_1} \sum_{j=1}^{s_2} g_{ij} \Phi_{ij}(M_{lm}) = -f(M_{lm}), \quad l = 1, 2, \dots, s_1, \quad l = 1, 2, \dots, s_2, \quad (7.16)$$

где $f(M)$ — правая часть уравнения (7.15), а $\Phi_{ij}(M)$ вычисляется по формуле [25]

$$\Phi_{ij}(M) = - \int_{L\sigma_{ij}} \vec{n}_M \cdot (\vec{dl} \times \nabla_{M_0} F(M, M_0)) + k^2 \int_{\sigma_{ij}} F(M, N) \vec{n}_M \cdot \vec{n}_N d\sigma \quad (7.17)$$

при

$$F(M, M_0) = \frac{1}{4\pi} \frac{e^{ikr_{MM_0}}}{r_{MM_0}}. \quad (7.18)$$

Однако если мы вычислим функцию $g(N)$ для следующих параметров задачи: $\nu = 4$ Гц, $z_Q = 50$ м, $c = 1450$ м/с, то учитывая осесимметричность задачи (то есть $g(M) = g(\sqrt{x_M^2 + y_M^2}) = g(r_M)$) получим графики зависимости действительной и мнимой частей $g(r_M)$, изображенные на рисунке 5. Из рисунка видно, что искомая функция является осциллирующей функцией. Для аппроксимации осциллирующей функции кусочно-постоянными

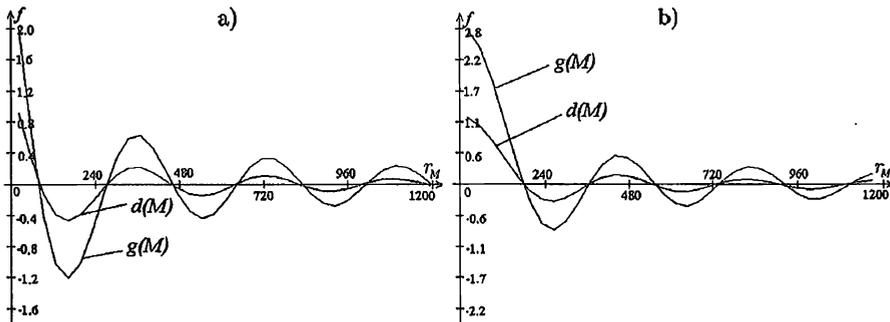


Рис. 5. Графики зависимости действительной и мнимой частей:
а) — действительная часть; б) — мнимая часть.

функциями требуется минимум 10 ячеек на период. То есть при данных параметрах задачи для получения $g(M)$ необходимо решить систему порядка 10^3 уравнений. Однако в случае увеличения частоты до $\nu = 30$ Гц возникает необходимость в решении 10^5 уравнений.

На рисунке 5 изобразим функцию $d(M) = 1000 \cdot p_Q(M)$. Как видно из рисунка, решение интегрального уравнения целесообразно искать в виде

$$g(M) = p_Q(M)h(M), \quad (7.19)$$

где $h(M)$ — новая неизвестная функция. В результате рассуждений, аналогичных классическому МЗДВР, полагая постоянной в каждой ячейке функцию $h_{ij}(M)$, приходим к системе алгебраических уравнений

$$\sum_{i=1}^{s_1} \sum_{j=1}^{s_2} h_{ij} \Upsilon_{ij}(M_{lm}) = -f(M_{lm}), \quad l = 1, 2, \dots, s_1, \quad l = 1, 2, \dots, s_2, \quad (7.20)$$

где интеграл

$$\Upsilon_{ij}(M) = \int_{\sigma} p_Q(N) \frac{\partial^2 F(M, M_0)}{\partial n_M \partial n_{M_0}} d\sigma_{M_0} \quad (7.21)$$

понимается в смысле главного значения по Адамару [1].

В результате решения системы (7.20) получается гладкая функция $h(M)$. Однако для получения решения необходимо брать интегралы вида (7.21).

7.3. Интегрирование гиперсингулярного интеграла с весом

Рассмотрим вычисление интеграла вида

$$I(M) = \int_{\sigma} p_Q(N) \frac{\partial^2 F(M, N)}{\partial n_M \partial n_N} d\sigma_N. \quad (7.22)$$

Представим интеграл в виде $I(M) = \vec{n}_M \cdot \vec{\Phi}(M)$, где $\vec{\Phi}(M) = \Phi_x(M)\vec{i} + \Phi_y(M)\vec{j} + \Phi_z(M)\vec{k}$, $\vec{n}_M = \cos \alpha_M \vec{i} + \cos \beta_M \vec{j} + \cos \gamma_M \vec{k}$. Здесь $\cos \alpha_M$, $\cos \beta_M$, $\cos \gamma_M$ — направляющие косинусы вектора \vec{n}_M , а также введены обозначения

$$\begin{aligned} \Phi_x(M) &= \int_{\sigma} p_Q(N) \frac{\partial^2 F(M, N)}{\partial x_M \partial n_N} d\sigma_N; \\ \Phi_y(M) &= \int_{\sigma} p_Q(N) \frac{\partial^2 F(M, N)}{\partial y_M \partial n_N} d\sigma_N; \\ \Phi_z(M) &= \int_{\sigma} p_Q(N) \frac{\partial^2 F(M, N)}{\partial z_M \partial n_N} d\sigma_N. \end{aligned} \quad (7.23)$$

Пусть $\vec{n}_N = \cos \alpha \vec{i} + \cos \beta \vec{j} + \cos \gamma \vec{k}$, где $\cos \alpha$, $\cos \beta$, $\cos \gamma$ — направляющие косинусы вектора \vec{n}_N . Принимая во внимание тот факт, что $\nabla_M F(M, N) = -\nabla_N F(M, N)$, получаем, что

$$\begin{aligned} \Phi_x(M) &= - \int_{\sigma} p_Q(N) \frac{\partial^2 F(M, N)}{\partial x_N^2} \cos \alpha d\sigma_N - \\ &- \int_{\sigma} p_Q(N) \frac{\partial^2 F(M, N)}{\partial x_N \partial y_N} \cos \beta d\sigma_N - \int_{\sigma} p_Q(N) \frac{\partial^2 F(M, N)}{\partial x_N \partial z_N} \cos \gamma d\sigma_N. \end{aligned}$$

Представим $\Phi_x(M)$ в следующем виде:

$$\Phi_x(M) = -I_1(M) - I_2(M) - I_3(M), \quad (7.24)$$

где через $I_1(M)$, $I_2(M)$, $I_3(M)$ обозначены следующие интегралы:

$$\begin{aligned} I_1(M) &= \int_{\sigma} p_Q(N) \frac{\partial^2 F(M, N)}{\partial x_N^2} \cos \alpha d\sigma_N; \\ I_2(M) &= \int_{\sigma} p_Q(N) \frac{\partial^2 F(M, N)}{\partial x_N \partial y_N} \cos \beta d\sigma_N; \\ I_3(M) &= \int_{\sigma} p_Q(N) \frac{\partial^2 F(M, N)}{\partial x_N \partial z_N} \cos \gamma d\sigma_N. \end{aligned} \quad (7.25)$$

Так как $F(M, N)$ — фундаментальное решение уравнения Гельмгольца (7.2), то справедливо следующее соотношение:

$$\frac{\partial^2 F(M, N)}{\partial x_N^2} = -\frac{\partial^2 F(M, N)}{\partial y_N^2} - \frac{\partial^2 F(M, N)}{\partial z_N^2} - k^2 F(M, N),$$

или после умножения правой и левой частей на $p_Q(N)$ получаем

$$\begin{aligned} p_Q(N) \frac{\partial^2 F(M, N)}{\partial x_N^2} &= -p_Q(N) \frac{\partial^2 F(M, N)}{\partial y_N^2} - \\ &- p_Q(N) \frac{\partial^2 F(M, N)}{\partial z_N^2} - k^2 p_Q(N) F(M, N). \end{aligned} \quad (7.26)$$

Для дважды дифференцируемых функций $A(x, y)$ и $B(x, y)$ справедливы соотношения

$$A(x, y) \frac{\partial^2 B(x, y)}{\partial x^2} = \frac{\partial}{\partial x} \left(A(x, y) \frac{\partial B(x, y)}{\partial x} \right) - \frac{\partial A(x, y)}{\partial x} \frac{\partial B(x, y)}{\partial x}, \quad (7.27)$$

$$A(x, y) \frac{\partial^2 B(x, y)}{\partial x \partial y} = \frac{\partial}{\partial x} \left(A(x, y) \frac{\partial B(x, y)}{\partial y} \right) - \frac{\partial A(x, y)}{\partial x} \frac{\partial B(x, y)}{\partial y}. \quad (7.28)$$

Применяя (7.27) для $A(x, y) = p_Q(N)$, $B(x, y) = F(M, N)$ и заменяя $\frac{\partial}{\partial x}$ на $\frac{\partial}{\partial y_N}$, получаем

$$p_Q(N) \frac{\partial^2 F(M, N)}{\partial y_N^2} = \frac{\partial}{\partial y_N} \left(p_Q(N) \frac{\partial F(M, N)}{\partial y_N} \right) - \frac{\partial p_Q(N)}{\partial y_N} \frac{\partial F(M, N)}{\partial y_N}, \quad (7.29)$$

а если заменить $\frac{d}{dx}$ на $\frac{\partial}{\partial z_N}$, то имеем

$$p_Q(N) \frac{\partial^2 F(M, N)}{\partial z_N^2} = \frac{\partial}{\partial z_N} \left(p_Q(N) \frac{\partial F(M, N)}{\partial z_N} \right) - \frac{\partial p_Q(N)}{\partial z_N} \frac{\partial F(M, N)}{\partial z_N}. \quad (7.30)$$

Подставим (7.29) и (7.30) в (7.26); тогда для интеграла $I_1(M)$ из (7.25) будет справедливо следующее соотношение:

$$I_1(M) = \int_{\sigma} \left[\frac{\partial p_Q(N)}{\partial y_N} \frac{\partial F(M, N)}{\partial y_N} + \frac{\partial p_Q(N)}{\partial z_N} \frac{\partial F(M, N)}{\partial z_N} - \right. \\ \left. - k^2 p_Q(N) F(M, N) \right] \cos \alpha d\sigma + \int_{\sigma} \cos \alpha \left[\frac{\partial}{\partial y_N} \left(-p_Q(N) \frac{\partial F(M, N)}{\partial y_N} \right) - \right. \\ \left. - \frac{\partial}{\partial z_N} \left(p_Q(N) \frac{\partial F(M, N)}{\partial z_N} \right) \right] d\sigma.$$

Применим формулу (7.28) к $p_Q(N) \frac{\partial^2 F(M, N)}{\partial x_N \partial y_N}$; полагая $A(x, y) = p_Q(N)$ и $B(x, y) = F(M, N)$ и заменяя $\frac{\partial}{\partial x}$ на $\frac{\partial}{\partial x_N}$, а $\frac{\partial}{\partial y}$ на $\frac{\partial}{\partial y_N}$, получаем

$$p_Q(N) \frac{\partial^2 F(M, N)}{\partial x_N \partial y_N} = \frac{\partial}{\partial x_N} \left(p_Q(N) \frac{\partial F(M, N)}{\partial y_N} \right) - \frac{\partial p_Q(N)}{\partial x_N} \frac{\partial F(M, N)}{\partial y_N}. \quad (7.31)$$

Рассуждая аналогичным образом и заменяя в формуле (7.28) $\frac{\partial}{\partial x}$ на $\frac{\partial}{\partial x_N}$, а $\frac{\partial}{\partial z}$ на $\frac{\partial}{\partial z_N}$, приходим к

$$p_Q(N) \frac{\partial^2 F(M, N)}{\partial x_N \partial z_N} = \frac{\partial}{\partial x_N} \left(p_Q(N) \frac{\partial F(M, N)}{\partial z_N} \right) - \frac{\partial p_Q(N)}{\partial x_N} \frac{\partial F(M, N)}{\partial z_N}. \quad (7.32)$$

Используя формулы (7.31) и (7.32), перепишем интегралы $I_2(M)$ и $I_3(M)$ соответственно в виде:

$$I_2(M) = - \int_{\sigma} \cos \beta \frac{\partial p_Q(N)}{\partial x_N} \frac{\partial F(M, N)}{\partial y_N} d\sigma_N + \\ + \int_{\sigma} \cos \beta \left[- \frac{\partial}{\partial x_N} \left(-p_Q(N) \frac{\partial F(M, N)}{\partial y_N} \right) \right] d\sigma_N$$

и

$$I_3(M) = \int_{\sigma} \cos \gamma \left[- \frac{\partial}{\partial x_N} \left(p_Q(N) \frac{\partial F(M, N)}{\partial z_N} \right) \right] d\sigma_N - \\ - \int_{\sigma} \cos \gamma \frac{\partial p_Q(N)}{\partial x_N} \frac{\partial F(M, N)}{\partial z_N} d\sigma_N.$$

Для дальнейших рассуждений воспользуемся формулой Стокса

$$\int_{\sigma} \left[\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) \cos \alpha + \left(\frac{\partial P}{\partial y} - \frac{\partial R}{\partial x} \right) \cos \beta + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \cos \gamma \right] d\sigma = \\ = \int_{L_{\sigma}} P dx + Q dy + R dz, \quad (7.33)$$

где L_{σ} — граница поверхности σ . Полагая в формуле (7.33)

$$P = 0, \quad Q = p_Q(N) \frac{\partial F(M, N)}{\partial z_N} \quad \text{и} \quad R = -p_Q(N) \frac{\partial F(M, N)}{\partial y_N},$$

получаем с учетом выражений I_1, I_2, I_3 следующее соотношение для $\Phi_x(M)$:

$$\Phi_x(M) = - \int_{L_{\sigma}} p_Q(N) \left(\frac{\partial F(M, N)}{\partial z_N} dy_N - \frac{\partial F(M, N)}{\partial y_N} dz_N \right) - \\ - \int_{\sigma} \left\{ \left[\frac{\partial p_Q(N)}{\partial y_N} \frac{\partial F(M, N)}{\partial y_N} + \frac{\partial p_Q(N)}{\partial z_N} \frac{\partial F(M, N)}{\partial z_N} - k^2 p_Q(N) F(M, N) \right] \cos \alpha - \right. \\ \left. - \cos \beta \frac{\partial p_Q(N)}{\partial x_N} \frac{\partial F(M, N)}{\partial y_N} - \cos \gamma \frac{\partial p_Q(N)}{\partial x_N} \frac{\partial F(M, N)}{\partial z_N} \right\} d\sigma_N,$$

или в компактной форме:

$$\Phi_x(M) = - \int_{L_{\sigma}} p_Q(N) \left(\vec{dl} \times \nabla_N F(M, N) \right)_x + \\ + k^2 \int_{\sigma} p_Q(N) F(M, N) \cos \alpha d\sigma_N - \\ - \int_{\sigma} \cos \alpha \left[\nabla_N F(M, N) \times (\vec{n}_N \times \nabla_N p_Q(N)) \right]_x d\sigma_N, \quad (7.34)$$

где $\vec{dl} = dx\vec{i} + dy\vec{j} + dz\vec{k}$, а под $(\cdot)_x$ понимается координата x соответствующего вектора.

Проводя аналогичные рассуждения для проекций $\Phi_y(M)$ и $\Phi_z(M)$ из (7.23), получаем, что интеграл (7.22) вычисляется по формуле

$$I(M) = - \int_{L_{\sigma}} p_Q(N) \vec{n}_M \cdot \left(\vec{dl} \times \nabla_N F(M, N) \right) + \\ + k^2 \int_{\sigma} p_Q(N) F(M, N) (\vec{n}_M \cdot \vec{n}_N) d\sigma_N - \\ - \int_{\sigma} \vec{n}_M \cdot \left[\nabla_N F(M, N) \times (\vec{n}_N \times \nabla_N p_Q(N)) \right] d\sigma_N. \quad (7.35)$$

7.4. Примеры численных расчетов

В качестве приложения рассмотрим поле давления в волноводе конечной глубины. На рисунках 6а и 6б изображены величины давления $p(M)$ для источника, расположенного на глубине $z_Q = 30$ м, в случае, когда поверхность σ_2 (дно моря) является плоской (рис. 6а) и изогнутой (рис. 6б). В расчетах частота источника полагалась 30 Гц.

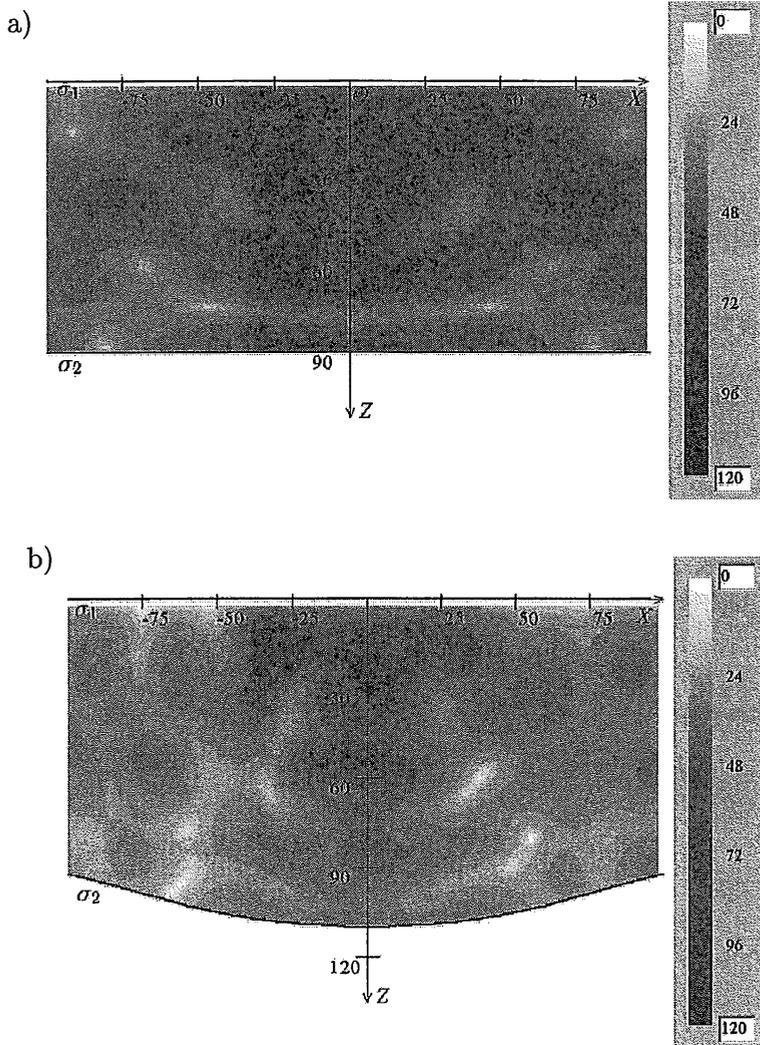


Рис. 6. Поле давления в волноводе

Работа выполнена при финансовой поддержке грантов РФФИ №№ 02-01-00490, 02-01-00263, 02-01-00651, 02-01-00045, 04-01-00045, 02-01-00523.

Литература

- [1] *Адамар Ж.* Задача Коши для линейных уравнений с частными производными гиперболического типа. М.: Наука, 1978.
- [2] *Андреев М.Ю., Кацнельсон Б.Г., Кулапин Л.Г., Петников.* Влияние гидродинамической изменчивости на фазу звукового поля в мелком море // Акуст. журнал. 1996. Т.42, № 4. С. 459–464.
- [3] *Астраханцев Г.П.* Об одном итерационном методе решения сеточных эллиптических задач // ЖВМ и МФ. 1971. Т. 11, № 2. С. 439–448.
- [4] *Астраханцев Г.П.* Итерационные методы решения вариационно-разностных схем для двумерных эллиптических уравнений второго порядка. Л.: ЛОМКИ АН СССР, 1972.
- [5] *Астраханцев Г.П.* Метод фиктивных областей для эллиптического уравнения второго порядка с естественными граничными условиями // ЖВМ и МФ. 1978. Т. 18, № 1. С. 118–125.
- [6] *Бахвалов Н.С.* О сходимости одного релаксационного метода для эллиптического оператора с естественными ограничениями // ЖВМ и МФ. 1966. № 6. С. 101–135.
- [7] *Бахвалов Н.С.* Эффективные методы решения жестких многомерных многопараметрических задач // ЖВМ и МФ. 1999. Т. 39, № 12. С. 2019–2049.
- [8] *Бахвалов Н.С., Панасенко Г.П.* Осреднение процессов в периодических средах. Математические задачи механики композиционных материалов. М.: Наука, 1984.
- [9] *Бахвалов Н.С., Богачев К.Ю., Мэтр Ж.Ф.* Эффективный алгоритм решения жестких эллиптических задач с приложениями к методу фиктивных областей // ЖВМ и МФ. 1999. Т. 39, № 6. С. 919–931.
- [10] *Бахвалов Н.С., Князев А.В.* Методы эффективного вычисления осредненных характеристик композитов периодической структуры из существенно разнородных материалов // Вычислительные процессы и системы. Т. 8. М.: Наука, 1991. С. 52–94.
- [11] *Белоцерковский С.М., Лифанов И.К.* Численные методы в сингулярных интегральных уравнениях. М.: Наука, 1985.
- [12] *Богатырев А.Б.* Вычисление наилучших многочленов устойчивости // Мат. сборник. 2004 (в печати).
- [13] *Богачев К.Ю.* Итерационный метод решения смешанной задачи для квазилинейных эллиптических уравнений в областях сложной формы // Доклады РАН. 1995. Т. 340. № 6. С. 727–730.
- [14] *Бондарь Л.Ф., Борисов С.В., Гриценко А.В.* Результаты исследований флуктуаций интенсивности и фазы акустических сигналов на стационарных трассах и шельфовой зоне Японского моря // Акуст. журнал. 1994. Т. 40, № 4. С. 561–570.
- [15] *Бреховских Л.М.* Волны в слоистых средах. М.: Изд-во АН СССР, 1957.
- [16] *Вайникко Г.М., Лифанов И.К., Полтавский Л.Н.* Численные методы в гиперсингулярных интегральных уравнениях и их приложения. М.: ТОО Янус, 2001.
- [17] *Воеводин В.В., Кузнецов Ю.А.* Матрицы и вычисления. М.: Наука, 1984.
- [18] *Годунов С.К.* Лекции по современным аспектам линейной алгебры. Новосибирск: Научная Книга, 2002.

- [19] Годунов С.К., Нечепуренко Ю.М. Оценки для главной и жесткой компонент на основе интегрального критерия качества дихотомии // ЖВМ и МФ. 2000. Т.40, № 1. С. 35–42.
- [20] Годунов С.К., Нечепуренко Ю.М. Оценки скорости сходимости метода Ньютона для вычисления инвариантных подпространств // ЖВМ и МФ. 2002. Т. 42, № 6. С. 771–779.
- [21] Голуб Дэс., Ван Лоун Ч. Матричные вычисления. М.: Мир, 1999.
- [22] Деккер К., Вервер Я. Устойчивость методов Рунге-Кутты для жестких нелинейных дифференциальных уравнений. М.: Мир, 1988.
- [23] Деметьев В.Г., Лебедев В.И., Нечепуренко Ю.М. Спектральный анализ модели ядерного реактора с запаздывающими нейтронами / Алгоритмы и программы для нейтронно-физических расчетов ядерных реакторов. Обнинск: ФЭИ, 1999. С. 143–150.
- [24] Де Санто Дэс. Акустика океана. М.: Мир, 1982.
- [25] Довгий С.А., Лифанов И.К. Методы решения интегральных уравнений. Киев: Наукова думка, 2002.
- [26] Дьяконов Е.Г. О одном итерационном методе решения систем конечно-разностных уравнений эквивалентных по спектру // ЖВМ и МФ. 1961. Т. 138, № 3. С. 522–525.
- [27] Дьяконов Е.Г. Разностные методы решения краевых задач. Вып.2. М.: Изд-во МГУ, 1972.
- [28] Дьяконов Е.Г. Минимизация вычислительной работы / Асимптотически оптимальные алгоритмы для эллиптических задач. М.: Наука, 1989.
- [29] Канторович Л.В. Функциональный анализ и прикладная математика // УМН. 1948. Т. XVI, № 7. С. 1211–1225.
- [30] Капорин И.Е., Николаев Е.С. Метод фиктивных неизвестных для решения разностных эллиптических краевых задач в нерегулярных областях // Дифф. уравнения. 1980. Т. 16, № 7. С. 1211–1225.
- [31] Кобельков Г.М. Об эквивалентных нормировках подпространств L_2 // Analysis Mathematica. 1977. Т. 3, № 3. С. 177–186.
- [32] Кобельков Г.М. О методах решения уравнений Навье–Стокса // ДАН СССР. 1978. Т. 243, № 4. С. 843–846.
- [33] Кобельков Г.М. О численных методах решения уравнений Навье–Стокса в переменных скорость–давление / Вычислительные процессы и системы. Вып. 8. М.: Наука, 1991. С. 204–256.
- [34] Кузнецов Ю.А. Вычислительные методы в подпространствах / Вычислительные процессы и системы. Вып. 2. М.: Наука, 1985. С. 265–350.
- [35] Ладыженская О.А. Математические проблемы динамики вязкой несжимаемой жидкости. М.: Наука, 1970.
- [36] Ладыженская О.А., Солонников В.А. О некоторых задачах векторного анализа и обобщенных постановках краевых задач для уравнений Навье–Стокса // Зап. научн. семин. ЛОМИ. 1976. Т. 59. С. 81–116.
- [37] Лебедев В.И. Разностные аналоги ортогональных разложений, основных дифференциальных операторов и некоторых краевых задач математической физики // ЖВМ и МФ. 1964. Т. 4, № 3, 4. С. 449–465, 649–659.
- [38] Лебедев В.И. О КР-методе ускорении сходимости итераций при решении кинетического уравнения / Численные методы решения задач математической физики. М.: Наука, 1966. С. 154–170.

- [39] *Лебедев В.И.* Как решать явными методами жесткие системы дифференциальных уравнений / Вычислительные процессы и системы. М.: Наука, 1991. С. 237–291.
- [40] *Лебедев В.И.* Явные разностные схемы для решения жестких задач с комплексным или разделимым спектром // ЖВМ и МФ. 2000. Т. 40, № 12. С. 1801–1812.
- [41] *Лебедев В.И.* Функциональный анализ и вычислительная математика. М.: Физматлит, 2000.
- [42] *Лебедев В.И.* Экстремальные многочлены и методы оптимизации вычислительных алгоритмов // Матем. сборник. 2004. Т. 195, № 10. С. 21–66.
- [43] *Лебедев В.И., Агошков В.И.* Операторы Пуанкаре-Стеклова и их приложения в анализе. М.: ОВМ АН СССР, 1983.
- [44] *Лебедев В.И., Медовиков А.А.* Методы второго порядка точности с переменными шагами по времени // Известия ВУЗов: Математика. 1998, № 9. С. 1–10.
- [45] *Лебедев В.И., Финогенов С.А.* О порядке выбора итерационных параметров в чебышевском циклическом итерационном методе // ЖВМ и МФ. 1971. Т. 11, № 2. С. 425–439.
- [46] *Лебедев В.И., Финогенов С.А.* Об использовании упорядоченных чебышевских параметров в итерационных методах // ЖВМ и МФ. 1976. Т. 16, № 4. С. 895–912.
- [47] *Лифанов И.К.* Метод сингулярных интегральных уравнений и численный эксперимент. М.: ТОО Янус, 1995.
- [48] *Маркушевич А. И.* Теория аналитических функций. М.;Л.: ГИТТЛ, 1950.
- [49] *Марчук Г.И.* Методы вычислительной математики. М.: Наука, 1980.
- [50] *Марчук Г.И.* Методы расщепления. М.: Наука, 1988.
- [51] *Марчук Г.И., Кузнецов Ю.А.* Итерационные методы и квадратичные функционалы. Новосибирск: Наука, 1972.
- [52] *Марчук Г.И., Кузнецов Ю.А.* Некоторые вопросы итерационных методов / Вычислительные методы линейной алгебры. Новосибирск: Наука, 1972. С. 4–20.
- [53] *Марчук Г.И., Лебедев В.И.* Численные методы переноса нейтронов. М.: Атомиздат, 1981.
- [54] *Марчук Г.И., Яненко Н.Н.* Применение метода расщепления (дробных шагов) для решения задач математической физики / Некоторые вопросы вычислительной и прикладной математики. Новосибирск: Наука, 1966.
- [55] *Мацокин А.М.* Метод фиктивных компонент и модифицированный разностный аналог метода Шварца // Вычислительные методы линейной алгебры. Новосибирск: ВЦ СОАН, 1980. С. 66–77.
- [56] *Мацокин А.М., Непомнящих С.В.* Метод альтернирования Шварца в подпространстве // Изв. ВУЗов: Математика. 1985. № 10. С. 61–66.
- [57] *Нечепуренко Ю.М.* Оценка нормы матрицы Грина через интегральный критерий качества дихотомии и границы хаусдорфова множества // Мат. заметки. 2002. Т. 71, вып. 2. С. 232–238.
- [58] *Нечепуренко Ю.М.* Оценка нормы матричной экспоненты через норму решения уравнения Ляпунова и границы хаусдорфова множества // ЖВМ и МФ. 2002. Т. 42, № 2. С. 131–141.

- [59] *Нечепуренко Ю.М.* О дихотомии спектра матрицы замкнутым контуром // Доклады РАН. 2004. Т. 397, № 4. С. 459–460.
- [60] *Нечепуренко Ю.М., Саджане М.* О сходимости метода Ньютона-Канторовича для вычисления инвариантных подпространств // Мат. заметки. 2004. Т. 73, вып. 1. С. 109–114.
- [61] *Никольский С.М.* О продолжении функций многих переменных с сохранением дифференциальных свойств // Мат. сб. 1956. Т. 40(82), № 2. С. 244–268.
- [62] *Самарский А.А.* Теория разностных схем. М.: Наука. 1982.
- [63] *Самарский А.А., Николаев Е.С.* Методы решения сеточных уравнений. М.: Наука, 1978.
- [64] *Саульев В.К.* Интегрирование уравнений параболического типа методом сеток. М.: Физматгиз. 1960.
- [65] *Саульев В.К.* О решении некоторых краевых задач на быстродействующих вычислительных машинах методом фиктивных областей // Сиб. матем. ж. 1963. Т. 4, № 4. С. 912–925.
- [66] *Смирнов В.И., Лебедев Н.А.* Конструктивная теория функций комплексного переменного. М.: Наука, 1964.
- [67] *Фадеев Д.К., Фадеева В.И.* Вычислительные методы линейной алгебры. М.;Л.: Физматгиз, 1963.
- [68] *Федоренко Р.П.* О скорости сходимости одного итерационного процесса // ЖВМ и МФ. 1964. № 4. С. 227–235.
- [69] *Хайер Э., Ваннер Г.* Решение обыкновенных дифференциальных уравнений. Жесткие и дифференциально-алгебраические задачи. М.: Мир, 1999.
- [70] *Чижонков Е.В.* Релаксационные методы решения седловых задач. М.: ИВМ РАН, 2002.
- [71] *Шайдулов В.В.* Многосеточные методы конечных элементов. М.: Наука, 1989.
- [72] *Яценко Н.Н.* Метод дробных шагов решения многомерных задач математической физики. Новосибирск: Наука, 1967.
- [73] *Anderson E., Bai Z., Bischof C., Demmel J., Dongarra J., Du Croz J., Greenbaum A., Hammarling S., McKenney A., Ostrouchov S., Sorensen D.* LAPACK users guide. Philadelphia: SIAM, 1992.
- [74] *Axelsson O.* Iterative Solution Methods. Cambridge: University Press, 1994.
- [75] *Aziz K., Babuška I.* The Mathematical Foundations of the Finite Element Method with Application to Partial Differential Equations. N.Y.: Academic Press Inc., 1972.
- [76] *Bakhvalov N.S., Knyazev A.V., Kobelkov G.M.* Iterative methods for solving equations with highly varying coefficients // Domain decomposition methods for partial differential equations. Philadelphia: SIAM, 1991. P. 197–205.
- [77] *Bastian P., Hackbush W., Wittum G.* Additive and multiplicative multigrid — a comparison // Computing. 1998. V. 62. P. 345–364.
- [78] *Belov A.I., Serebryany A.N., Zhuravlev V.A.* Observation of internal wave effects on acoustic and oceanographic field study in the Middle Atlantic Bight // Proceed. of the International conference on shallow water acoustics. Beijing HEP, 1997.
- [79] *Bogachev K.Yu.* Efficient algorithms for stiff elliptic problems with large parameters // Russian J. Numer. Anal. and Math. Modelling. 2002. V. 17, № 4. P. 347–366.

- [80] *Bramble J.H., Pasciak J.E.* A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems // *Math. Comp.* 1988. V. 50. P. 1–17.
- [81] *Bramble J.H., Pasciak J.E., Xu J.* Parallel multilevel preconditioners // *Math. Comp.* 1990. V. 55. P. 1–22.
- [82] *Brandt A.* Multi-level adaptive technique (MLAT) for fast numerical solution to boundary value problems // *Proc. 3rd Int. Conf. of Numerical Methods in Fluid Mechanics.* V. 1. Berlin: Springer, 1973. P. 82–89.
- [83] *Brandt A.* Multi-level adaptive solutions to boundary value problems // *Math. Comput.* 1977. V. 31. P. 333–390.
- [84] *Brezzi F.* On the existence, uniqueness and approximation of saddle point problems arising from Lagrangian multipliers // *RAIRO Anal. Numer.* 1974. № 8, P. 129–151.
- [85] *Caille G.W., Dahl P.H., Zijun Gan et al.* Overview of the joint China-U.S. "Yellow Sea'96" experiment // *Proceed. of the International conference on shallow water acoustics.* Beijing: HEP, 1997.
- [86] *Chorin A. J.* Numerical solution of the Navier–Stokes equations // *Math. Comp.* 1968. № 22. P. 745–762.
- [87] *Dahlquist G.* A special stability problem for linear multistep methods // *BIT.* 1963. V. 3. P. 27–43.
- [88] *Douglas J., Rachford H.* On the numerical solution of heat conduction problems in two and three space variables // *Trans. Amer. Math. Soc.* 1956. V. 82, № 2. P. 421.
- [89] *Dyakonov E.G.* Optimization in Solving Elliptic Problems. N.Y.: CRC Press, 1996.
- [90] *Feistauer M.* Mathematical Methods in Fluid Dynamics. London: Longman Scientific&Technical, 1993.
- [91] *Glowinski R.* Le ϑ -scheme // *Numerical methods for the Navier–Stokes equations* // *Comp. Phys. report.* 1987. № 6. P. 73–87.
- [92] *Glowinski R.* Finite element methods for incompressible viscous flows // *Handbook of Numerical Analysis.* V. 9. Amsterdam: North-Holland, 2003.
- [93] *Gunzburger M.* Finite Element Methods for Viscous Incompressible Flows: a Guide to Theory, Practice and Algorithms. Boston: Academic, 1989.
- [94] *Gunzburger M.* Incompressible Computational Fluid Dynamics: Trends and Advances. Cambridge: University Press, 1993.
- [95] *Hackbusch W.* Multi-grid Methods and Applications. Berlin: Springer, 1985.
- [96] *Heywood J.G., Rannacher R.* Finite element approximation of the nonstationary Navier–Stokes problem / II. Stability of solutions and error estimates uniform in time // *SIAM J. Numer. Anal.* 1986. V. 25. P. 750–777.
- [97] *Hinrichsen D., Kelb B.* Spectral value sets: a graphical tool for robustness analysis // *Systems Control Lett.* 1993. V. 21. P. 127–136.
- [98] *Kobelkov G.M.* Fictitious domain method and the solution of elliptic equations with highly varying coefficients // *Sov. J. Numer. Anal. Math. Modelling.* 1987. V. 2, № 6. P. 407–418.
- [99] *Kobelkov G.M.* On the solution of the boundary value problem for the diffusion equation with highly varying coefficient // *Russ. J. Numer. Anal. Math. Modelling.* 1996. V. 11. № 6. P. 487–495.

- [100] *Kobelkov G.M., Arbash J.* Numerical methods for theory elasticity equations with highly varying coefficients // Russ. J. of Numer. Anal. and Mathem. Modelling. 1993. V. 8, № 5. P. 371–384.
- [101] *Kobelkov G.M., Olshanskii M.A.* Effective preconditioning of Uzawa type schemes for a generalized Stokes problem // Numerische Mathematik. 2000. V. 86, № 3. P. 443–470.
- [102] *Kotsarev A., Lebedev V., Shishkov L., Nechepurenko Yu., Dementiev V.* Spectral analysis of VVER-1000 reactor model at high negative reactivities // Proceedings of the ninth Symposium of AER, 4–8 October, Slovakia, 1999. Budapest: Kiadja a KFKI Atomenergia Kutato Intezet, 1999. P. 453–468.
- [103] *Lebedev V.I.* How to solve stiff systems of differential equations by explicit method // Numerical method and applications. Boca Raton: CRC Press, 1994. P. 45–80.
- [104] *Lebedev V.I.* Zolotarev polynomials and extremum problem // Russ. J. of Numer. Anal. and Mathem. Modelling. 1994. V. 9, № 3. P. 191–214.
- [105] *Lebedev V.I.* Explicit difference schemes with variable time steps for solving stiff systems of equation / Numerical Analysis and its Applications // Proc. Lecture Notes in Computer Science 1196. Berlin: Springer, 1997. P. 274–283.
- [106] *Lebedev V.I.* Extremal polynomials, pseudo-difference operators, and stable explicit schemes for the solution of stiff nonstationary problems of mathematical physics // Proc. of the International Conference on Computational Mathematics. Part I. Novosibirsk: Computing Center, 2002. P. 60–63.
- [107] *Nechepurenko Yu.M.* An upper bound for stiff parts of solutions to a linear system of ODEs with a given set of initial values // Russ. J. Numer. Anal. Math. Modelling. 1998. V. 13, № 5. P. 405–409.
- [108] *Nechepurenko Yu.M.* On the singular-function approach to eigenproblems // Russ. J. Numer. Anal. Math. Modelling. 1998. V. 13, № 3. 219–233.
- [109] *Nechepurenko Yu.M.* New spectral analysis technology based on the Schur decomposition // Russ. J. Numer. Anal. Math. Modelling. 1999. V. 14, № 3. P. 265–274.
- [110] *Nechepurenko Yu.M.* The regularly structured pseudospectrum // Russ. J. Numer. Anal. Math. Modelling. 2004. V. 19, № 3. P. 265–268.
- [111] *Neumann J.* Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes // Math. Nachrichten. 1951. V. 4. P. 258–281.
- [112] *Olshanskii M.* A low order Galerkin finite element method for the Navier–Stokes equations of steady incompressible flows: a stabilization issue and iterative methods // Comput. Meth. Appl. Mech. Engineering. 2002. V. 191. P. 5515–5536.
- [113] *Oswald P.* Multilevel Finite Element Approximation. Stuttgart: Teubner, 1994.
- [114] *Peaceman O., Rachford H.* The numerical solution of parabolic and elliptic differential equations // J. Soc. and Appl. Math. 1955. V. 3, № 1. P. 28.
- [115] *Pekeris C.L.* Theory of propagation of explosive sound in shallow water // Geol. Soc. Am. Mem. 1948. V. 27. P. 1–117.
- [116] *Prohl A.* Projection and quasi-compressibility methods for solving the incompressible Navier–Stokes equations. Stuttgart: Teubner, 1997.
- [117] *Sadkane M.* Block–Arnoldi and Davidson methods for unsymmetric large eigenvalue problems // Numer. Math. 1993. V. 64, № 2. P. 195–211.

- [118] *Sleijpen G.L.G., Van der Vorst.* A Jacobi–Davidson iteration method for linear eigenvalue problems //SIAM J. Matrix Anal. Appl. 1996. V. 17. P. 410–425.
- [119] *Temam R.* Navier–Stokes equaions. Amsterdam: Horth–Holland, 1977.
- [120] *Trefethen L.N.* Pseudospectra of linear operators // SIAM Review. 1997. V. 39. P. 383–406.
- [121] *Turek S.* Efficient Solvers for Incompressible Flow Problems: An Algorithmic Approach in View of Computational Aspects. Berlin: Springer, 1998.
- [122] *Verürth R.* A multilevel algorithm for mixed problems // SIAM J. Numer. Anal. 1984. V. 21. P. 264–271.

Вычислительные технологии

Ю.В. Василевский, В.П. Ильин, Е.Е. Тьртышников

Содержание

Введение	100
1. Методы сжатия и структуризации данных	104
1.1. Структуры и методы для матриц специального вида	104
1.2. Сепарабельные аппроксимации в теории матриц	109
1.3. Разделение переменных и многоуровневые матрицы	111
1.4. Дискретные вейвлет-преобразования	114
1.5. Тензорные аппроксимации с использованием вейвлет-преобразований	115
2. Методы решения сеточных систем	117
2.1. Методы фиктивных областей	118
2.2. Многосеточные методы	121
2.3. Методы декомпозиции области	126
2.4. Технологии "черного ящика"	130
3. Примеры организации технологических процессов	132
3.1. Цели и назначение ППП БАЗИС-А	133
3.2. Математические постановки задач моделирования процессов в алюминиевом электролизере	134
3.3. Краткое описание алгоритмов	138
3.4. Вычислительно-информационные технологии	139
3.5. Общая структура пакета программ БАЗИС-А	141
Литература	141

Введение

Вычислительные технологии — это совокупность алгоритмов, структур данных, расчетных методик и программных реализаций для решения вычислительных задач на вычислительных системах. Элементы вычислительных технологий обсуждаются, конечно, в различных разделах вычислительной математики. Однако с точки зрения развития науки и адекватной

системы образования целесообразно выделить их в особый раздел. Главных причин для этого, по меньшей мере, две.

Во-первых, вычислительные технологии оцениваются с точки зрения совершенно особой целевой функции. Это не оценки скорости сходимости или точности, не количество операций, не ускорение или эффективность при распараллеливании. Несмотря на то, что перечисленные характеристики исключительно важны, при разработке вычислительных технологий они имеют лишь вспомогательное значение. Главной целью является получение ответа в нужной форме, к заданному сроку и при известных ограничениях на доступ к машинным и человеческим ресурсам.

Во-вторых, базовые вычислительные технологии имеют "универсальный" характер и поэтому полезны при численном решении самых различных математических и прикладных задач. Можно выделить следующие базовые разделы:

- методы представления, сжатия и структуризации данных;
- алгоритмические аспекты: методы геометрического моделирования, построения расчетных сеток, дискретизации и сопряженные с ними методы решения систем сеточных уравнений;
- инструментальные средства вычислительного эксперимента (библиотеки алгоритмов и пакеты прикладных программ, электронные системы и базы знаний и т.д.).

Предмет вычислительных технологий является синтетическим по существу. Выход на принципиально новые возможности в реализации вычислений связан, бесспорно, с внедрением мощной вычислительной техники и средств программирования. Но в настоящее время в еще большей степени он определяется развитием новых *математических идей*. Например, совсем недавно (20–30 лет назад) решение систем линейных уравнений с заполненной матрицей коэффициентов, не имеющих какой-либо специфики, считалось трудной задачей, если число неизвестных превышало несколько тысяч. Теперь можно успешно решать системы с числом неизвестных порядка нескольких сотен тысяч. Нетрудно понять, что один лишь рост производительности компьютеров, в том числе и за счет параллельных вычислений, до сих пор не может сделать решение задач подобного размера обыденным. В данном случае прорыв в решении больших задач обеспечивается новыми математическими технологиями сжатия и структуризации данных. Другим фактором прорыва является *технологичность* и *безотказность* технологических компонент вычислительного процесса. Решение практических междисциплинарных задач невозможно без использования методов с экономичной программной реализуемостью, обеспечивающих надежность исполнения своих функций (от английского *robust* в русской литературе даже появился не вполне формализованный термин "робастный алгоритм"). На этапе дискретизации здесь

ярким примером могут служить методы конечных элементов, завоевавших широкое признание не только в силу фундаментального обоснования алгоритмов практически произвольного порядка точности для самого широкого класса задач математической физики, но и в силу своей уникальной поэлементной технологии, основанной на построении локальных матриц и обеспечившей им глобальную универсальность, простую реализуемость и естественную распараллеливаемость. В последнее десятилетие такой же толчок получили методы конечных объемов, или балансные (консервативные) конечно-разностные схемы, благодаря определению локальных матриц баланса и сборки глобальных матриц на сетках разного типа [17]⁺, [77]⁺.¹

Другой важной особенностью многих современных вычислительных технологий является их относительная универсальность и тенденция работать с каждой частной задачей как с "черным ящиком". При решении системы линейных алгебраических уравнений, например, последнее означает, что задан способ получения матрицы коэффициентов и правой части, но нет полной информации о происхождении системы. Несмотря на это, метод решения системы строится на основе определенных предположений. Если они не выполнены, то метод "имеет право на ошибку" или может оказаться не очень эффективным. В любом случае от "хорошей" технологии следует требовать информации о достоверности полученного ответа. Заметим также, что "универсальность" технологии обычно означает возможность ее применения в ситуациях, когда формальные предположения стоящей за ней теории не проверены или нарушены — реальная область применения метода часто оказывается шире той области, для которой имеются теоремы и доказательства. Определенной альтернативой универсальным технологиям типа "черного ящика" являются подходы, использующие дополнительную, но не обременительную для пользователя, информацию о задаче. Подобные технологии (типа "серого ящика") имеют более широкую область применения, а также более эффективны и просты в программной реализации. В случае решения сеточных систем примером дополнительной информации могут быть данные о расчетной сетке и геометрической модели.

В прикладной задаче цель вычислений чаще всего достигается в ходе решения последовательности частных задач, следуя так называемому модульному принципу расщепления сложной проблемы на достаточно самостоятельные фрагменты. Для реализации такой цели требуется объединять (комбинировать) различные вычислительные технологии для отдельных этапов в единый технологический процесс. Среди главных компонент вычислительных технологий, ориентированных на "большие" задачи ма-

¹Здесь и в дальнейшем символ "+" при ссылках на литературу будет означать "смотри также представленную в цитируемой работе библиографию по соответствующей теме".

тематического моделирования (в том числе междисциплинарные), можно выделить следующие, см. [15]–[18]:

- средства геометрического и функционального моделирования, обеспечивающие описание и модификацию сложных расчетных областей с криволинейными многосвязными границами и подобластями, в которых могут решаться различные типы систем дифференциальных уравнений с кусочно-гладкими коэффициентами, определяющими материальные свойства разных сред;
- генераторы адаптивных сеток, учитывающие априорные и/или апостериорные свойства искомых решений, в том числе неструктурированные, вложенные и состоящие из различных типов объектов;
- средства аппроксимации исходных дифференциальных постановок методами конечных объемов или конечных элементов (МКО, МКЭ);
- алгебраические решатели для систем линейных или нелинейных уравнений, в том числе с разреженными матрицами высокого порядка, представленных с помощью различных экономичных форматов;
- средства управления вычислительным процессом, включающие предварительную обработку и контроль исходных данных (препроцессинг), обработку и визуализацию результатов (постпроцессинг и графический пользовательский интерфейс).

Эти компоненты функционируют на базе согласованных структур данных: геометрических, сеточных, алгебраических и графических, — которые обеспечивают внутренние интерфейсы различных расчетных модулей.

Отдельная важная тема — это использование вычислительных технологий на многопроцессорных вычислительных системах, включающая всегда актуальную проблему, в определении Г.И. Марчука, "отображения алгоритмов на архитектуру ЭВМ". Хотя мы неизбежно упоминаем о распараллеливании алгоритмов, в данной области переплетается слишком много фундаментальных математических и технологических проблем, требующих специального исследования. Мы также не будем касаться обширной области информационных технологий (хотя термины "структуры данных" и "базы" знаний — это их прерогатива), сосредоточиваясь только на аспектах, связанных с реализацией задач большой вычислительной сложности. Рассматриваемые в представленной работе вопросы исследовались в огромном количестве монографий и статей, среди которых следует отметить пионерские работы Г.И. Марчука [98], [97], а также А.П. Ершова [12], А.А. Самарского [38]⁺ и Н.Н. Яненко [4], [47], [48], сформировавших вместе с руководимыми ими коллективами в 1960–1970-е годы основы нового направления вычислительных наук (за рубежом одновременно появился термин Computer Science), связанного с понятиями "крупномасштабный вычислительный эксперимент", "методология па-

кетов прикладных программ”, ”технологическая цепочка математического моделирования” и т.д. В приводимом списке упоминаемых работ мы особо выделяем публикации, выполненные в ИВМ РАН и ВЦ СО АН СССР (ныне — ИВМиМГ СО РАН), отнюдь не претендуя на исчерпывающий обзор исследований, выполненный отечественными и зарубежными авторами.

В данной работе мы обсудим подробнее два фундаментальных раздела современных вычислительных технологий: методы сжатия и структуризации данных в алгебраических задачах с плотными матрицами (раздел 1) и методы решения сеточных систем (раздел 2). В заключительной части рассматриваются принципы организации вычислительных процессов на примере решения междисциплинарной задачи комплексного моделирования технологических процессов в алюминиевом электролизере (раздел 3). Заметим, что каждый из трех разделов имеет свой стиль изложения, отражающий особенности представляемого материала. Наиболее законченная форма второго раздела (от определений основных объектов до таблиц с данными численных экспериментов) объясняется тем, что технологии сжатия сводятся практически к двум идеям: аппроксимации малого ранга для блоков и отбрасыванию относительно малых элементов после преобразований вейвлетовского типа. Многообразие идей, лежащих в основе различных технологий решения сеточных систем, предопределяет лишь обзорный характер их представления в разделе 2, с упором на классификацию по целям и используемым средствам. Представление в разделе 3 технологической цепочки решения сложной практической задачи также является описательным в силу своей специфики.

1. Методы сжатия и структуризации данных

1.1. Структуры и методы для матриц специального вида

Матричные вычисления — ключевая часть многих приложений. Однако несмотря на успехи в построении общих алгоритмов для классических задач, имеется много задач, которые не могут быть решены с помощью существующих алгоритмов из-за большого размера или неустойчивости (чаще всего требующей пересмотра самой постановки задачи). Можно утверждать, что трудная прикладная задача может быть решена только на основе выявления и использования структурных свойств, отличающих ее от задач общего вида.

Структура в данных связана с особым способом их представления с помощью относительно малого числа параметров. Последнее свойство означает в широком смысле *разреженность данных*. Важный пример разре-

женности данных дают так называемые *разреженные* матрицы, в которых ненулевых элементов относительно мало, а в качестве определяющих параметров можно взять позиции и значения ненулевых элементов. Иной характер имеет разреженность данных в плотных матрицах. Рассмотрим типичные примеры.

Дискретное преобразование Фурье. Матрица Фурье

$$F_n = \left[\exp \left(i \frac{2\pi}{n} kl \right) \right], \quad 0 \leq k, l \leq n-1,$$

— плотная матрица, определяемая всего одним параметром n . Но она может быть умножена на вектор всего лишь за $O(n \log n)$ операций (вместо n^2 для матриц общего вида) — при помощи *быстрого преобразования Фурье* (БПФ). Для этого используются специальные разреженные факторизации матрицы F_n . Например, если $n = 2m$, то

$$F_n = P_n \begin{bmatrix} F_m & 0 \\ 0 & F_m \end{bmatrix} \begin{bmatrix} I_m & 0 \\ 0 & W_m \end{bmatrix} \begin{bmatrix} I_m & I_m \\ I_m & -I_m \end{bmatrix},$$

$$W_m = \text{diag} \left\{ \exp \left(i \frac{\pi}{m} k \right) \right\},$$

матрица P_n осуществляет четно-нечетную перестановку:

$$[v_0, v_1, \dots, v_{m-1}, v_m, v_{m+1}, \dots, v_{2m-1}] P_n =$$

$$= [v_0, v_2, \dots, v_{2(m-1)}, v_1, v_3, \dots, v_{2m-1}].$$

Если $n = 2^L$, то аналогичным образом можно факторизовать матрицы $F_m, F_{m/2}, \dots, F_4$. В итоге мы получаем рекурсивный алгоритм, в котором $(1/2)n \log_2 n$ умножений и $(3/2)n \log_2 n$ сложений комплексных чисел. Идеи и алгоритмы типа БПФ появлялись еще в первой половине XX в., но их бурное внедрение в практику вычислений началось с работы Кули и Тьюки 1965 года (см. [5]).

Теплицевы матрицы и их обобщения. Теплицева матрица

$$A = [a_{k-l}], \quad 0 \leq k, l \leq n-1,$$

полностью определяется своим первым столбцом и первой строкой. Если $a_k = a_l$ при $k = l \pmod{n}$, то A называется также *циркулянтной* матрицей. Теплицевы матрицы связаны с различными отделами классического анализа (ортогональные многочлены, ряды Фурье, проблема моментов и др.) и многими приложениями из статистики и вычислительной математики (например, интегральные уравнения с ядром типа свертки) (см. [5, 6, 39, 84]).

Умножение теплицевой матрицы A на вектор выполняется за $O(n \log_2 n)$ операций в два приема:

- 1) A достраивается (как подматрица) до циркулянтной матрицы C порядка $N \geq n$;
- 2) C диагонализуется с помощью БПФ: $(1/n) F_N^* C F_N = \text{diag}(F_N c)$, где c — первый столбец матрицы C .

К матрицам *типа теплицевых* обычно относят $n \times n$ -матрицы вида

$$A = U_1 V_1 + \dots + U_r V_r, \quad (1.1)$$

где $r \ll n$ и все сомножители U_k, V_k — теплицевы матрицы. Оказывается, для теплицевой матрицы обратная матрица имеет вид (1.1) с $r \leq 2$. Этот факт был открыт в работе [11] и впоследствии привел к важному понятию *теплицевых рангов*, или *рангов смещения* (Th. Kailath и др.; см. [39, 84]). В общем случае под (U, V) -рангом смещения для произвольной матрицы A понимается ранг $r(U, V)$ матрицы

$$\Delta(U, V)(A) = AU - VA, \quad (1.2)$$

где U, V — некоторые фиксированные матрицы. При специальном выборе U и V возникают интересные классы матриц A , для которых $\text{rank}(\Delta(U, V)(A)) \ll n$. Преобразования вида (1.2) удобны при построении "малопараметрических" обобщений различных матриц специального вида: ганкелевых, Коши, Вандермонда и др. При этом матрица A^{-1} имеет схожее "малопараметрическое" описание — в силу очевидного равенства

$$\text{rank}(\Delta(U, V)(A)) = \text{rank}(\Delta(V, U)(A^{-1})).$$

Общий характер преобразования (1.2) позволяет обнаружить полезные связи между формально разными классами матриц. Например, пусть $Z = [\delta_{ij-1}]$ — матрица сдвига и $P = \text{diag}(p_1, \dots, p_n)$, $Q = \text{diag}(q_1, \dots, q_n)$ — диагональные матрицы такие, что $p_i \neq q_i$. Тогда:

- $r(Z, Z)(A) = \text{rank}(\Delta(Z, Z)(A)) \leq 2$ для любой теплицевой матрицы A (в случае $r \ll n$ это обобщенные теплицевы матрицы);
- $r(P, Z)(A) = \text{rank}(\Delta(P, Z)(A)) \leq 1$ для матрицы Вандермонда $A = [p_i^j]$, $0 \leq i, j \leq n - 1$ (в случае $r \ll n$ это обобщенные матрицы Вандермонда);
- $r(P, Q)(A) = \text{rank}(\Delta(P, Q)(A)) \leq 2$ для матрицы Коши $A = [1/(p_i - q_j)]$ (в случае $r \ll n$ это обобщенные матрицы Коши).

В совершенно общем случае легко заметить, что

$$r(U, Q)(ABC) \leq r(U, V)(A) + r(V, Z)(B) + r(Z, Q)(C).$$

Отсюда, в частности, вытекает, что обобщенную теплицевую матрицу B можно преобразовать в обобщенную матрицу Коши путем двустороннего умножения на матрицы Вандермонда A и C с увеличением обобщенного ранга максимум на 2 (поскольку $r(P, Q) \leq r(Z, Z) + 2$).

Для решения линейных систем со строго регулярной теплицевой матрицей коэффициентов и вычисления "малопараметрического" разложения соответствующей обратной матрицы известны алгоритмы сложности $O(n^2)$ — *быстрые* (Левинсон, Шур и др.) и $O(n \log^2 n)$ — *супербыстрые* (Морф, Юн и др.). До сих пор остается актуальным изучение быстрых и супербыстрых прямых методов для разных типов матриц с малым рангом смещения.

Матрицы с блоками малого ранга. Малоранговые матрицы

$$A = \sum_{k=1}^r u_k v_k^T, \quad r \ll n,$$

определяются элементами векторов u_k, v_k (так называемых *скелетонов*) и играют важнейшую роль при аппроксимации матриц и в матричных вычислениях (см. [3, 71]). При $r < n$ такие матрицы являются вырожденными и поэтому не могут использоваться непосредственно для аппроксимации обратимых матриц. Однако они могут хорошо приближать достаточно большие области в обратимых матрицах, возникающих при рассмотрении нелокальных операторов (см. [123, 81, 107]). Часто такие области являются подматрицами (блоками), но это могут быть подмножества и более сложного вида. Например, в *семисепарабельной* матрице это верхний и нижний треугольники — оба вкладываются в некоторые (разные) матрицы малого ранга. В общем случае семисепарабельными оказываются матрицы, обратные к ленточным матрицам (см. [33]).

Матрицы малого тензорного ранга. Пусть $n = p^2$, U_k, V_k — матрицы порядка p , а матрица A имеет вид

$$A = \sum_{k=1}^r U_k \otimes V_k, \quad r \ll n.$$

Такая матрица определяется элементами кронекеровых сомножителей U_k, V_k ; их число равно $rp^2 = O(n)$ при ограниченном r . Матрицы, представленные суммой кронекеровых произведений с числом факторов больше двух, играют важную роль в многомерном анализе [53].

Несмотря на успехи в изучении вышеперечисленных классов матриц (в частности, алгоритмы быстрого умножения, прямые методы решения систем типа теплицевых с асимптотикой $O(n^2)$ и даже $O(n \log^2 n)$), имеется разрыв между нуждами приложений и нынешним состоянием дел в теории и алгоритмах. Наиболее трудные прикладные задачи — многомерные, они приводят к многоуровневым структурированным матрицам, а наши знания именно в этой части наименее развиты.

Пример 1. Рассмотрим обратимую дважды теплицеву матрицу порядка $n = p^2$:

$$A = [a_{k_1-l_1}], \quad 0 \leq k_1, l_1 \leq p-1,$$

$$a_{k_1-l_1} = [a_{k_1-l_1, k_2-l_2}], \quad 0 \leq k_2, l_2 \leq p-1,$$

и поставим вопрос: существует ли пригодная для вычислений специфика в A^{-1} ? Конечно, A можно рассматривать как блочно-теплицеву матрицу с блоками $p \times p$, тогда обратная матрица описывается формулами Гохберга-Хайнига. Но число параметров в этом описании составляет $O(n^{3/2})$ — такой подход малоинтересен: он не выходит за рамки техники работы с одноуровневыми матрицами и игнорирует теплицеву структуру блоков, тогда как дважды теплицева матрица существенным образом двухуровневая. (Недавно в работе по авторегрессивным фильтрам по двум переменным J.S. Gerónimo и H.J. Woerdeman рассматривали сходный вопрос и получили некоторое описание, но оно не привело к быстрым вычислительным процедурам.)

Пример 2. Рассмотрим обратимую дважды трехдиагональную матрицу порядка $n = p^2$:

$$A = \begin{bmatrix} a_1^0 & a_1^1 & & & & \\ a_2^{-1} & a_2^0 & a_2^1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & a_{n-1}^{-1} & a_{n-1}^0 & a_{n-1}^1 & \\ & & & a_n^{-1} & a_n^0 & \end{bmatrix},$$

$$a_i^j = \begin{bmatrix} a_{i,1}^{j,0} & a_{i,1}^{j,1} & & & & \\ a_{i,2}^{j,-1} & a_{i,2}^{j,0} & a_{i,2}^{j,1} & & & \\ & \ddots & \ddots & \ddots & & \\ & & a_{i,n-1}^{j,-1} & a_{i,n-1}^{j,0} & a_{i,n-1}^{j,1} & \\ & & & a_{i,n}^{j,-1} & a_{i,n}^{j,0} & \end{bmatrix}$$

и снова поставим вопрос о существовании пригодной для вычислений структуры в A^{-1} . Если A рассматривается как блочно-трехдиагональная матрица, то обратную дает блочное обобщение [33] известной формулы обращения трехдиагональных матриц

$$A^{-1} = L \circ E + U \circ E^T - \frac{1}{2} \text{diag}(L + U),$$

где \circ означает адамарово (поэлементное) произведение,

$$E = \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ \dots & \dots & \dots & & \\ 1 & \dots & 1 & 1 & \end{bmatrix}, \quad \text{rank } L = \text{rank } U = 1, \quad \text{diag}(L) = \text{diag}(U).$$

Прямое блочное обобщение вновь содержит $O(n^{3/2})$ параметров и поэтому неудовлетворительно. Блочный подход не выходит за рамки методов для одноуровневых матриц, а двухуровневый случай требует, очевидно, новых идей. Вопрос о том, каких именно, остается открытым.

Заметим также, что область применения прямых методов ограничена даже для большинства матриц специального вида, в том числе для дважды теплицевых. Более перспективно развитие итерационных методов, в особенности методов типа метода минимальных невязок, подробно изложенного в работе Г.И. Марчука и Ю.А. Кузнецова [30] — одной из самых ранних в данном направлении, а также, например, в [109]. Вычислительная работа в итерационном методе определяется двумя особенностями: а) сложностью одной итерации; б) числом итераций. Первое определяется сложностью матрично-векторного умножения, второе — спектральными свойствами матрицы коэффициентов и согласованностью матрицы и правой части. Для рассмотренных выше матриц специального вида умножение на вектор осуществляется с помощью известных быстрых алгоритмов (часто связанных с БПФ). Поэтому основное внимание в исследованиях направляется на способы улучшения спектральных свойств с помощью *переобуславливания*. Многие работы, выполненные в ИВМ РАН, позволили сформировать очень эффективные методы переобуславливания для конкретных задач и общие инструменты теоретического анализа, связанные с такими понятиями как *равнораспределенность* и *спектральные кластеры* [122].

Современные вычислительные технологии решения матричных задач основаны на изучении матричных (обычно нелинейных) аппроксимаций. Построение переобуславливателей рассматривается как задача аппроксимации матриц с помощью тех или иных матриц специального вида. В случае больших плотных матриц аппроксимация исходной матрицы подходящей "малопараметрической" матрицей дает возможность снизить сложность матрично-векторного умножения.

1.2. Сепарабельные аппроксимации в теории матриц

Малоранговые аппроксимации — это частный пример дискретных *сепарабельных* аппроксимаций (аппроксимаций функции $a_{ij\dots k}$ путем разделения переменных i, j, \dots, k). Теория таких аппроксимаций в случае двух индексов обладает особой элегантностью и завершенностью (Шмидт, Алахвердиев, Мирский [8]). Она связана с *сингулярным разложением матрицы* [40, 8], которое в общем случае эффективно вычисляется по алгоритму Голуба-Кахана [71], а в малоранговом случае — по методу bidi-агонализации Ланцоша [69]. Оба подхода теряют эффективность, конечно, в практически важном случае очень больших матриц, когда слож-

ность $O(n^3)$ и даже $O(n^2)$ нельзя считать приемлемой. Практически полезные алгоритмы для больших задач должны иметь сложность $O(n)$ или $O(n \log^\alpha n)$.

Таким образом, аппроксимации следует искать используя лишь малую часть всех элементов аппроксимируемой матрицы. То, что это возможно в принципе, впервые доказано в [10, 70]. При построении аппроксимаций такого типа фундаментальную роль играет следующий принцип максимального объема.

Теорема [68]. Пусть M имеет порядок n и сингулярные числа $\sigma_1 \geq \dots \geq \sigma_n$. Предположим, что

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix},$$

где блок M_{11} невырожден, $r \times r$, и имеет максимальный объем (т.е. модуль детерминанта) среди всех подматриц $r \times r$. Тогда

$$|\{M_{22} - M_{21}M_{11}^{-1}M_{12}\}_{ij}| \leq (r+1)\sigma_{r+1}, \quad 1 \leq i, j \leq n.$$

Можно заметить, что в обычном алгоритме LU -разложения выбор ведущего элемента ведет к увеличению объема базовой ведущей подматрицы. Если ранг матрицы равен r , то на r -м шаге получается аппроксимация ранга r — при этом достаточно вычислить только строки и столбцы, отвечающие базовой ведущей подматрице, и отказаться от вычисления всех элементов очередной активной подматрицы. Область матрицы, занятая заданным числом строк и столбцов, имеет форму креста — поэтому соответствующие алгоритмы были названы алгоритмами *неполной крестовой аппроксимации* [126], [64]. Важно, что матрица задается виртуально — процедурой $\mathcal{M}(i, j)$, вычисляющей по требованию элементы в позиции (i, j) .

Алгоритм [64]. Для заданной функции $\mathcal{M}(i, j)$, $1 \leq i, j \leq n$, и заданной точности ε найти вектор-столбцы $u_1, v_1, \dots, u_r, v_r$ размера n такие, что

$$\left\| M - \sum_{k=1}^r u_k v_k^T \right\|_F \leq \tilde{\varepsilon} \|M\|_F,$$

где $M = [m_{ij}]$, $m_{ij} = \mathcal{M}(i, j)$, $1 \leq i, j \leq n$, $\tilde{\varepsilon} \approx \varepsilon$.

- 1) Положить $\delta = \varepsilon$, $k = 1$ и $\mathcal{I}[1 : n] = \mathcal{J}[1 : n] = [1, 2, \dots, n]$.
- 2) Вычислить $m_{ij} = \mathcal{M}(i, j)$ для

$$(i, j) \in \mathcal{P}_k = \{(\mathcal{I}(k), \mathcal{J}(k)), (\mathcal{I}(k+1), \mathcal{J}(k+1)), \dots, (\mathcal{I}(n), \mathcal{J}(n))\}$$

и положить $m'_{ij} = m_{ij} - \sum_{l=1}^{k-1} u_l(i)v_l(j)$.

- 3) Найти пару индексов (i'_k, j_k) такую, что $|m'_{i'_k j_k}| = \max_{(i,j) \in \mathcal{P}_k} |m'_{i,j}|$.
- 4) Вычислить $m_{i,j_k} = \mathcal{M}(i, j_k)$ для всех $i \in \mathcal{I}_k = \{\mathcal{I}(k), \mathcal{I}(k+1), \dots, \mathcal{I}(n)\}$ и положить $m'_{i j_k} = m_{i j_k} - \sum_{l=1}^{k-1} u_l(i) v_l(j_k)$.
- 5) Найти индекс i_k такой, что $d_k \equiv |m'_{i_k j_k}| = \max_{i \in \mathcal{I}_k} |m'_{i j_k}|$. (Теперь k -й ведущий элемент находится в позиции (i_k, j_k) матрицы M .)
- 6) Остановиться, если d_k меньше машинной точности. В противном случае вычислить $\tilde{\varepsilon} = d_k(n-k) / \left\| \sum_{l=1}^{k-1} u_l v_l^T \right\|_F$ и, если $\tilde{\varepsilon} \leq \varepsilon$, то положить $r = k - 1$ и остановиться.
- 7) Вычислить $m_{i_k j} = \mathcal{M}(i_k, j)$ для $j = 1, \dots, n$, $\alpha = m_{i_k j_k} / \sqrt{|m_{i_k j_k}|}$, $\beta = \sqrt{|m_{i_k j_k}|}$, $u_k(i) = m_{i j_k} / \alpha$, $i = 1, \dots, n$, $v_k(j) = m_{i_k j} / \beta$, $j = 1, \dots, n$.
- 8) В системе \mathcal{I} переставить индексы в позициях k и l , где $\mathcal{I}(l) = i_k$. В системе \mathcal{J} переставить индексы в позициях k и l , где $\mathcal{J}(l) = j_k$.
- 9) Если $k < n$, то положить $k \leftarrow k + 1$ и перейти к шагу 2.

Приведенный алгоритм, конечно, не может гарантировать получение правильной аппроксимации ранга r для произвольной матрицы (пусть и допускающей такую аппроксимацию). Но он прекрасно работает во многих практически важных случаях. Более того, при определенных предположениях относительно функции $\mathcal{M}(i, j)$ корректность данного алгоритма (и его простых модификаций, ориентированных на дальнейшее увеличение объема выбираемой базовой подматрицы) можно доказать [42, 126].

1.3. Разделение переменных и многоуровневые матрицы

Пусть A — матрица размеров $n \times n$ и $\{A_k\}$ — конечная система ее блоков такая, что каждый элемент входит в один и только один блок. Будем называть такую систему блоков *мозаичным биением*. Если A_k имеет размеры $m_k \times n_k$ и ранг r_k , то при хранении только скелетонов сепарабельных разложений ранга r_k получаем, что общее число параметров (память) для представления матрицы A имеет вид

$$\text{mem}(A) = \sum_k \min\{m_k n_k, (m_k + n_k)\}.$$

Величину

$$\text{mr}(A) = \text{mem}(A) / (2n)$$

назовем *мозаичным рангом* (системы блоков) матрицы A . С точки зрения экономии памяти и арифметической работы роль мозаичного ранга аналогична роли классического ранга матрицы: число операций для умножения

матрицы A ранга r на вектор (и память для хранения скелетонов) имеет вид $r(m+n)$; то же самое сохраняется, если r — мозаичный ранг матрицы A .

Мозаичное биение строится с помощью *мозаичного дерева*, или *дерева кластеров*. Можно считать, что элемент a_{ij} в матрице характеризует "взаимодействие" элементов сетки (базисных функций, конечных элементов) с номерами i и j . Поставим в соответствие элементу i точку x_i в пространстве (например, геометрический центр носителя базисной функции). Под *кластером* понимается совокупность каких-то точек x_i . Кластер нулевого уровня (корень дерева) — это все множество точек. Далее, каждый кластер разбивается на непересекающиеся подкластеры в соответствии с тем или иным методом сепарации. После какого-то разрешенного числа сепараций появляется дерево кластеров. Любые две вершины дерева кластеров определяют блок в матрице. Любое мозаичное биение определяется некоторым набором пар вершин. Таким образом, построение мозаичного биения можно рассматривать как оптимизационную комбинаторную задачу на мозаичном дереве (цель — минимизация мозаичного ранга). Однако на практике применяется простой эвристический прием: дерево просматривается начиная с корня, для каждой пары вершин l -ого уровня проверяется некоторый *признак*, делающий блок *кандидатом* на включение в список блоков, допускающих малоранговую аппроксимацию; блок-кандидат включается в систему блоков искомого мозаичного биения. Данный подход нацелен на получение малого ранга для как можно более крупных блоков.

Пример. Рассмотрим интегральное уравнение с логарифмическим ядром

$$-\frac{1}{2\pi} \int_{\partial\Omega} \log|x-y|U(y)ds(y) = F(x), \quad x \in \partial\Omega.$$

Пусть $\partial\Omega$ — эллипс с полуосями $a = 1$ и $b = 0.5$. При использовании кусочно-постоянных базисных функций получаются матрицы с малым мозаичным ϵ -рангом ($\epsilon = 10^{-4}$):

Порядок матрицы	512	1024	2048	4096	8192	16384	32768
Мозаичный ранг	63.46	71.48	78.44	86.84	93.80	100.76	106.50
Фактор сжатия	24.79%	13.96%	7.66%	4.24%	2.29%	1.23%	0.65%

Фактор сжатия — это отношение памяти для хранения скелетонов к полной памяти. При увеличении порядка в 2 раза фактор сжатия уменьшается примерно вдвое.

Для получения строгих утверждений об аппроксимациях малого мозаичного ранга рассматриваются последовательности матриц A_n и B_n порядка n для некоторой возрастающей последовательности значений n и изучаются общие предположения, при которых справедливы следующие оценки:

$$\text{mr} B_n = O(\log^p n \log^q \varepsilon^{-1}), \quad p, q \geq 0, \quad (1.3)$$

$$\|A_n - B_n\| \leq \varepsilon n^\sigma, \quad (1.4)$$

где σ зависит от вида матричной нормы в левой части неравенства. Последовательность матриц A_n может ассоциироваться с функцией $f(x, y)$ от точек в ν -мерном евклидовом пространстве и последовательностью сеток $x_i = x_i^{(n)}$, $y_i = y_i^{(n)}$ следующим образом:

$$a_{ij} = f(x_i, y_j), \quad i, j = 1, \dots, n. \quad (1.5)$$

В более общем случае рассматривается последовательность функций

$$U_i(x) = U_i^{(n)}(x), \quad V_i = V_i^{(n)}(x)$$

и при этом

$$a_{ij} = (U_i, V_j)_f, \quad i, j = 1, \dots, n,$$

$$(U, V)_f \equiv \int_{\mathcal{D}} \int_{\mathcal{D}} U(x) f(x, y) V(y) dx dy. \quad (1.6)$$

Для построения аппроксимаций $B_n \approx A_n$ и получения оценок (1.3), (1.4) нужны определенные предположения относительно функции $f(x, y)$ и распределения точек x_i, y_i . Это предположения разных типов, связанные с двумя вполне независимыми этапами вывода оценок. На первом этапе изучаются специфические (мозаично-скелетонные) функции $g(x, y)$, аппроксимирующие исходную функцию $f(x, y)$ в равномерной метрике. На втором этапе матрицы малого мозаичного ранга B_n определяются как матрицы с элементами $g(x_i, y_j)$.

Первые быстрые алгоритмы приближенного матрично-векторного умножения получены в рамках *мультипольного метода* [74, 107, 59, 117] (см. также [81]). Эти алгоритмы строятся для конкретных функций $f(x, y)$ и в действительности дают особый *метод дискретизации* интегральных уравнений. В ИВМ РАН аналогичные задачи изучались как задачи аппроксимации матриц [10, 70, 123, 126, 68]. Оценки мозаичных рангов получены для широкого класса функций, включающего все практически интересные случаи. Это класс *асимптотически сепарабельных* функций.

Пусть $V(a, h)$ обозначает куб с центром в точке a и стороной $2h$, а $W(a, h)$ — его дополнение в "большом кубе" Π . Функция $f(x, y)$ называется *асимптотически сепарабельной*, если существуют $\rho > 1$, $0 < q < 1$ и $\alpha > 0$ такие, что для любого целого $m > 0$ существует целое $n > 0$ и сепарабельная функция $p_n(x, y)$ с n членами с разделенными переменными x и y , обеспечивающая оценки

$$|f(x, y) - p_n(x, y)| \leq c_1 q^m, \quad n \leq c_2 m^\alpha \quad \forall x \in V(a, h), y \in W(a, \rho h).$$

По сути дела в этом определении существование n -членных сепарабельных аппроксимаций предполагается для всех точек x и y , принадлежащих "достаточно отдаленным" областям точек.

1.4. Дискретные вейвлет-преобразования

Использование вейвлет-преобразований вызвано следующими причинами. В матрицах, связанных с интегральными уравнениями или со свойством сдвиговой инвариантности, обычно нельзя пренебречь какой-либо существенной частью элементов, так как они примерно одинаковы по порядку величины. Тем не менее разделенные разности элементов (рассматриваемых как значения функции на целочисленной сетке) могут отличаться по величине существенным образом. В таких случаях дискретное вейвлет-преобразование (DWT) дает способ разделения информации, содержащейся в "гладкой" матрице, на блоки элементов, соответствующих взвешенным средним и взвешенным разделенным разностям элементов исходной матрицы. Первые велики по сравнению со вторыми, что и позволяет получить разреженную аппроксимацию трансформированной матрицы путем отбрасывания малых элементов. Классические вейвлеты в своей основе связаны с равномерными сетками и преобразованием Фурье. Однако на практике часто требуются неравномерные сетки. В случае нерегулярных сеток построение функций и преобразований, имеющих аналогичные классическим вейвлетам свойства, требует совершенно иного подхода.

Для заданной сетки x_i , $i = 1, \dots, n + k + 1$, на отрезке и подсетке \tilde{x}_i , $i = 1, \dots, N + k + 1$ с $N < n$, введем B -сплайны порядка k :

$$\begin{aligned} B_i(x) &= [x_i; \dots; x_{i+k+1}]_+(y - x)^k, \quad i = 1, \dots, n, \\ \tilde{B}_i(x) &= [\tilde{x}_i; \dots; \tilde{x}_{i+k+1}]_+(y - x)^k, \quad i = 1, \dots, N. \end{aligned}$$

Квадратные скобки означают разделенные разности по переменному y , а $(z)_+ = z$ при $z \geq 0$ и 0 в противном случае.

Определим V и \tilde{V} как подпространства, натянутые соответственно на B_i , $i = 1, \dots, n$ и \tilde{B}_i , $i = 1, \dots, N$. Заметим, что V и \tilde{V} — пространства сплайн-функций на сетках x_i и \tilde{x}_i . Поскольку \tilde{x}_i — подсетка x_i , заключаем, что \tilde{V} — подпространство для V . Следовательно,

$$\tilde{B}_i = \sum_s r_{is} B_s r.$$

Вейвлет-пространство W определяется как дополнение \tilde{V} в V . Рассмотрим любое удобное пространство W с известным базисом $\tilde{\psi}_i = \sum_s \beta_{is} B_s$, $i = 1, \dots, n - N$, и "улучшим" его при помощи *лифтинговой схемы* (Sweldens):

$$\tilde{\psi}_i = \sum_s \beta_{is} B_s - \sum_{j=j_{\min}}^{j=j_{\max}} \alpha_{ij} \tilde{B}_j, \quad i = 1, \dots, n - N.$$

Чтобы определить неизвестные лифтинговые коэффициенты α_{ij} , потребуем, чтобы функции $\tilde{\psi}_i$ имели предписанное число (скажем, m) исчезающих моментов:

$$\int \tilde{\psi}_i x^p dx = 0, \quad p = 0, \dots, m.$$

(Чтобы уравнивать число неизвестных и число уравнений, положим $j_{\max} = j_{\min} + m$.) Требование m исчезающих моментов дает линейную систему для лифтинговых коэффициентов; эффективный метод явного решения этой системы (а значит, и построения нестандартных вейвлетов) получен недавно в работе [35] и успешно применен при численном решении интегральных уравнений в [36, 65].

1.5. Тензорные аппроксимации с использованием вейвлет-преобразований

Пусть A — двухуровневая матрица с размерами уровней n_1 и n_2 . Попытаемся приблизить ее суммой кронекеровых произведений вида $A_r = \sum_{k=1}^r A_k^1 \otimes A_k^2$, где размеры A_k^1 и A_k^2 соответственно $n_1 \times n_1$ и $n_2 \times n_2$. Пусть $A = A_r$, и r — наименьшее возможное число кронекеровых произведений, сумма которых есть A . Тогда r называется тензорным рангом A . Наилучшие приближения, минимизирующие $\|A - A_r\|_F$, могут быть получены алгоритмом SVD, примененным к матрицам, полученным специальным "перемешиванием" элементов.

Логика подхода, предложенного в [64], очень проста. Для заданной линейной системы $Ax = b$ (имеется в виду, что задан b и процедура вычисления произвольного элемента A) и величины ϵ допустимой относи-

тельной погрешности (во фробениусовой норме) мы выполняем следующие шаги:

(А) Аппроксимировать A суммой кронекеровых произведений

$$A \approx B = \sum_{k=1}^r U_k \otimes V_k, \quad \|B - A\|_F \leq \varepsilon \|A\|_F,$$

где U_k и V_k , соответственно, $p \times p$ и $q \times q$ и $n = pq$ есть порядок матрицы A . Для простоты предполагаем ниже $p = q = n^{1/2}$.

(В) Применить дискретное вейвлет-преобразование (DWT) с предписанным числом исчезающих моментов m к каждому кронекерову фактору:

$$P_k = WU_kW^T, \quad Q_k = WV_kW^T, \quad 1 \leq k \leq r.$$

Здесь W — матрица DWT степени m . P_k и Q_k — псевдоразреженные матрицы. Выбирая подходящий порог $\tau = \tau(\varepsilon, P_k, Q_k)$ и зануляя все элементы P_k и Q_k , меньшие τ по модулю, сформировать окончательное приближение B матрицей

$$C = W^{-T} \otimes W^{-T} DW \otimes W \approx B, \quad D = \sum_{k=1}^r P_k^\tau \otimes Q_k^\tau,$$

приближающей B с заданной погрешностью ε : $\|B - C\|_F \leq \varepsilon \leq \|B\|_F$.

(С) Построить переобуславливатель M^{-1} для матрицы A . Мы сосредоточимся на следующих опциях:

- 1) многоуровневые циркулянтные переобуславливатели с масштабированием;
- 2) приближенные обратные, полученные модифицированными итерациями Ньютона.

(D) Применить GMRES или PCG для решения $CM^{-1}y = b$ и определить $M^{-1}y$ как приближение к точному решению x .

Ожидается, что $r \ll n$. Таким образом, формат кронекеровых произведений требует хранения всего лишь $2rn \ll n^2$ чисел. Шаг (А) позволяет хранить приближение для A в оперативной памяти.

Чтобы дать представление об эффективности вычислительных технологий, реализующих рассмотренный выше круг идей, приведем численные результаты для матриц, возникающих при решении уравнения Прандтля

$$-\int_0^1 \int_0^1 \frac{U(x, y)}{[(x - x_0)^2 + (y - y_0)^2]^{3/2}} dx dy = F(x_0, y_0)$$

методом дискретных вихрей [36] в случае равномерных меток. В табл. 1 отражено построение обратных матриц, в табл. 2 — времена работы нашего

тензор-вейвлетовского метода решения соответствующих линейных алгебраических систем (чтобы оценивать погрешность наиболее консервативным образом, правая часть была выбрана равной сумме столбцов матрицы с номерами 1, 5, 10).

Таблица 1. Приближенное обращение матриц

$n = p^2$	4096	16384	65536	262144
ϵ	10^{-5}	10^{-5}	10^{-5}	10^{-5}
Тензорный ранг A	8	8	9	10
Время на обращение	2.68 sec	15.39 sec	1.47 min	7.29 min
Число итераций	7	8	9	10
Тензорный ранг A^{-1}	14	15	15	15
Невязка	$4 \cdot 10^{-6}$	$3 \cdot 10^{-6}$	$1.5 \cdot 10^{-4}$	$1.1 \cdot 10^{-4}$

Таблица 2. Решение систем с использованием масштабированных блочно-циркулянтных переобуславливателей

n	16129	65025	261121	1046529
Тензорный ранг A	11	12	12	13
Погрешность тензорной аппроксимации	10^{-7}	10^{-7}	10^{-7}	10^{-7}
Время на матрично-векторное умножение	0.05 sec	0.28 sec	1.43 sec	7.6 sec
Число итераций	17	21	22	22
Построение переобуславливателя	1.8 sec	7.2 sec	29.1 sec	1.9 min
Время на решение	2.3 sec	7.9 sec	41.6 sec	3.3 min
Относительная погрешность	$3.1 \cdot 10^{-6}$	$5.8 \cdot 10^{-6}$	$1.93 \cdot 10^{-5}$	$1.1 \cdot 10^{-5}$

2. Методы решения сеточных систем

В данном разделе мы рассмотрим один из основных этапов технологического процесса получения численного решения — решение систем сеточных уравнений. Основное внимание будет уделено методам решения систем, которые используют дополнительную информацию о сетке и геометрической модели, а также ограничениям на сетку, накладываемым каждым из этих методов. Отметим, что в статье [1] настоящего сборника эти методы представлены с *алгоритмической* точки зрения, в то время как здесь будут обсуждаться их *технологические* аспекты. Кроме того, будут рассмотрены технологии "черного ящика", использующие минимум данных о сеточной системе.

2.1. Методы фиктивных областей

Метод фиктивных областей (Саульев, Лебедев, Руховец) сводит решение краевых задач в областях сложной формы к решению краевых задач в областях простой формы (например, в прямоугольнике или параллелепипеде). Сведение обусловлено тем фактом, что в областях простой формы краевые задачи и/или их сеточные аналоги могут быть решены эффективными и экономичными методами. Существует несколько принципиально разных подходов для подобного сведения. Прежде чем приступить к их обсуждению, перечислим общие технологические компоненты, необходимые при реализации этих подходов.

1. Сетки, используемые при аппроксимации краевых задач в областях сложной формы, должны быть прямоугольными локально модифицированными, т.е. аппроксимировать границу со вторым порядком точности и являться логическим подмножеством некоей прямоугольной сетки, целиком покрывающей расчетную область. При этом отличаться от исходных прямоугольных ячеек могут только те ячейки, которые пересекаются границей расчетной области, или их соседи (рис. 1). Отличие может быть двояким: узлы модифицированных ячеек могут быть сдвинуты, а также некоторые части ячеек могут быть отсечены, образуя треугольники (2D) или тетраэдры (3D). Существуют алгоритмы построения подобных сеток как на плоскости, так и в пространстве [31], [52].

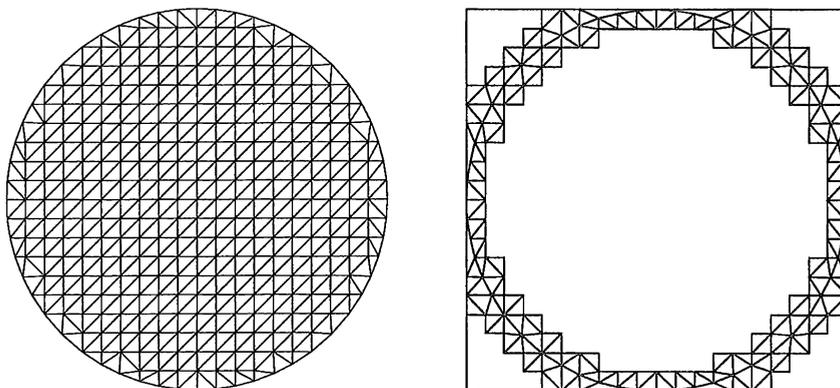


Рис. 1. Локально модифицированная сетка

2. Сеточные задачи, возникающие в расширенных областях простой формы, должны допускать эффективное численное решение, арифметическая сложность которого пропорциональна количеству узлов в прямоуголь-

ной сетке, с (поли)логарифмическим множителем от числа узлов. Быстрые прямые решатели рассматривались в работах [22], [108]⁺.

Необходимость использования локально модифицированных сеток обусловлена тем, что подмножество ячеек прямоугольной сетки не может аппроксимировать произвольную область с кусочно-гладкой границей со вторым порядком точности, который достигается во внутренних узлах области во многих методах аппроксимации краевых задач. Треугольные или тетраэдральные покрытия приграничной зоны обеспечивают второй порядок точности. Другим важным обстоятельством оказывается то, что сеточные уравнения, порожденные в исходной области, отличаются от систем уравнений, эффективно разрешимых быстрыми прямыми методами, только в узкой приграничной полосе.

Один из подходов метода фиктивных областей основан на известном факте, что в среде с относительно большими значениями коэффициента диффузии изменение плотности диффундируемого вещества достаточно мало. В связи с этим можно заменить первую краевую задачу (с однородным краевым условием и оператором диффузии) на задачу в расширенной области, где коэффициенты диффузии доопределены достаточно большими значениями. При этом решение расширенной задачи стремится к решению исходной при увеличении коэффициента диффузии в расширенной области. Аналогично можно осуществить замену для третьей краевой задачи (с однородным краевым условием), однако в расширенной области коэффициенты диффузии должны быть доопределены достаточно малыми значениями. В обоих случаях краевая задача в области сложной формы заменяется расширенной краевой задачей в области простой формы, но с сильным перепадом коэффициента диффузии вдоль границы исходной области.

Для решения подобных задач предложен целый ряд быстро сходящихся итерационных процессов [2]⁺, где на каждом шаге применяется быстрый прямой метод. Отметим, что в некоторых методах итерации осуществляются в подпространстве дискретно-гармонических векторов [91], что приводит к так называемым частичным задачам [22], [51], [99]⁺: найти решение сеточного аналога уравнения Пуассона в прямоугольнике (параллелепипеде) в узлах около границы исходной области, когда правая часть отлична от нуля также только в приграничной полосе. Решение частичных задач можно осуществить еще более эффективно, храня в памяти компьютера только степени свободы из подпространства дискретно-гармонических векторов.

Альтернативным подходом является метод фиктивных компонент (Астраханцев, Марчук, Кузнецов), который осуществляет расширение на матричном уровне. Для аппроксимации кусочно-линейными конечными элементами оператора Лапласа с краевыми условиями Неймана на локально-модифицированной сетке, и ее аналога в дополнение до объем-

лющего прямоугольника (параллелепипеда), верно блочное представление

$$\begin{pmatrix} A_{11} & A_{1\Gamma} & 0 \\ A_{\Gamma 1} & A_{\Gamma}^{(1)} & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & A_{\Gamma}^{(2)} & A_{\Gamma 2} \\ 0 & A_{2\Gamma} & A_{22} \end{pmatrix} = \begin{pmatrix} B_{11} & B_{1\Gamma} & 0 \\ B_{\Gamma 1} & B_{\Gamma} & B_{\Gamma 2} \\ 0 & B_{2\Gamma} & B_{22} \end{pmatrix},$$

где матрица B — конечно-элементное представление оператора Лапласа в расширенной области на локально модифицированной сетке. Здесь индекс Γ обозначает степени свободы, лежащие в модифицированных ячейках, а индексы 1 и 2 — оставшиеся степени свободы в расчетной области и в ее дополнении, соответственно. Отметим, что отличие матрицы B от матрицы \tilde{B} , эффективно обрабатываемой быстрыми прямыми методами, ограничено блоками $B_{1\Gamma}$, $B_{\Gamma 1}$, B_{Γ} , $B_{\Gamma 2}$, $B_{2\Gamma}$. Метод фиктивных компонент может быть определен как некий итерационный процесс [99]⁺ с исходной матрицей A_1 и переобуславливателем \tilde{B}_1 , где

$$A_1 = \begin{pmatrix} A_{11} & A_{1\Gamma} \\ A_{\Gamma 1} & A_{\Gamma}^{(1)} \end{pmatrix}, \quad \tilde{B}_1 = \begin{pmatrix} \tilde{B}_{11} & \tilde{B}_{1\Gamma} \\ \tilde{B}_{\Gamma 1} & \tilde{B}_{\Gamma} - \tilde{B}_{\Gamma 2} \tilde{B}_{22}^{-1} \tilde{B}_{2\Gamma} \end{pmatrix},$$

или итерации с расширенной нулями матрицей A_1 и переобуславливателем \tilde{B} . В обоих вариантах итераций векторы ошибки являются дискретно-гармоническими и невязки отличны от нуля лишь в приграничной полосе, что позволяет использовать технологию решателей частичных задач на всех итерациях, кроме начальной. Существуют варианты метода фиктивных компонент на случай задачи Дирихле, когда исходная матрица A_1 расширяется несимметрично некоторыми блоками матрицы \tilde{B} [99]⁺, [63]. Сохранение ошибок в подпространстве дискретно-гармонических функций позволяет использовать метод сопряженных градиентов как в случае несимметричного расширения, так и в случае вырожденного расширения.

Критерием независимости скорости сходимости итерационных процессов методов фиктивных областей от параметров дискретизации и/или геометрических параметров является независимость от этих параметров энергетической нормы оператора продолжения сеточной функции из расчетной области в окаймляющий ее прямоугольник (параллелепипед). Поэтому анализ сеточных продолжений представляется необходимым этапом предварительной оценки целесообразности применимости метода. Определенную трудность в методах фиктивных областей представляют точки смены типов краевых условий: имеется потребность в быстросходящихся и экономичных итерационных процессах, удобных в вычислительной практике.

Отметим, что метод фиктивного пространства [102] может быть успешно применен к постановкам метода фиктивных областей. Один из

приемов основан на использовании переобуславливателя для конечно-элементного оператора Лапласа в расширенной области и операторов продолжения сеточных конечно-элементных функций [102]. Другое применение метода фиктивного пространства для построения переобуславливателя заключается в переинтерполяции сеточных значений на иерархическую сетку в объемлющей области и использовании многоуровневого переобуславливателя с выборочным суммированием, необходимым для обработки различных типов краевых условий и минимизации вычислительной работы [104], [67]. Оба приложения метода фиктивного пространства не используют свойство локальной модифицированности сетки.

2.2. Многосеточные методы

Многосеточный метод (Федоренко, Бахвалов) является одним из наиболее популярных методов решения сеточных задач [46], [79]. Главной предпосылкой этому можно назвать его численную эффективность, выражаемую через асимптотическую оптимальность совершаемой арифметической работы. Несмотря на многообразие многосеточных методов, предложенных за последние десятилетия, основными составляющими многосеточной технологии являются несколько компонент: последовательность иерархически вложенных сеток, операторы перехода с грубой сетки на мелкую и обратно, выбор оператора на грубой сетке, сглаживатель, и процедура решения, объединяющая вышеприведенные компоненты. Для простоты приведем простейший пример организации $V(1, 1)$ -цикла для K -уровневого метода:

- 1) Set $u_K = v$;
- 2) For $l = K, K - 1, \dots, 1$: $x_l = R_l u_l$, $u_{l-1} = P_l(u_l - M_l x_l)$;
- 3) $w_0 = M_0^{-1} u_0$;
- 4) For $l = 1, 2, \dots, K$: $y_l = x_l + I_l w_{l-1}$, $w_l = y_l + R_l(u_l - M_l y_l)$;
- 5) Set $B_K v = w_K$.

Здесь I_l — оператор продолжения с грубой сетки $l - 1$ на мелкую l , P_l — оператор интерполяции с мелкой сетки l на грубую $l - 1$, R_l — оператор сглаживания на сетке l , M_l — оператор на сетке уровня l . Отметим, что в некоторых процедурах решения могут отсутствовать те или иные компоненты. Технологически для каждой из базовых процедур существует несколько реализаций. Рассмотрим некоторые из них.

Последовательность иерархически вложенных сеток может быть получена разными способами, которые зависят от топологии сеток. Так, в сетках с четырехугольными (2D) или шестигранными (3D) ячейками иерархия может быть получена равномерным измельчением каждой ячейки

на 4 (2D) или 8 (3D) ячеек. Произвольное локальное измельчение невозможно в рамках конформных сеток, однако допущение подчиненных узлов (рис. 2) делает процедуру измельчения очень простой и обеспечивает последовательность вложенных сеток эффективной структурой данных типа четверо-дерево или восьми-дерево.

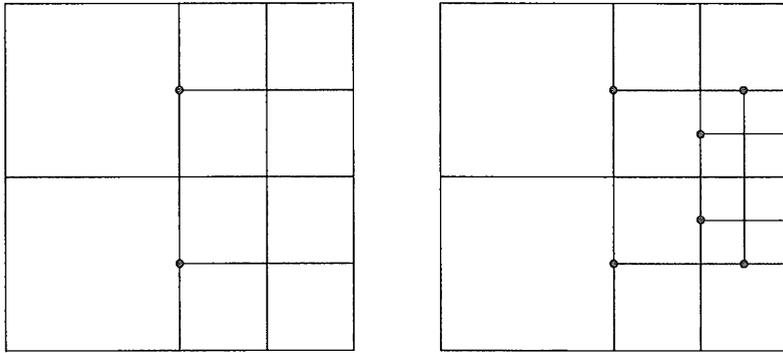


Рис. 2. Локально измельченная сетка с подчиненными узлами

Для конформных сеток с треугольными (2D) или тетраэдральными (3D) ячейками возможно как равномерное, так и произвольное локальное измельчения. Отметим два наиболее распространенных способа локального измельчения треугольных и тетраэдральных сеток: метод бисекции [134] и метод красно-зеленого разбиения [141]. Суть метода бисекции заключается в разбиении назначенных к разбиению ячеек и еще некоторых их соседей пополам за счет проведения медианы или медианной плоскости. Разбиение дополнительных ячеек необходимо для восстановления конформности сетки (рис. 3). Метод красно-зеленого разбиения разбивает назначенные к измельчению ячейки на четыре (2D) или восемь (3D) ячеек. Восстановление конформности сетки достигается доразбиением конфликтных (соседних к измельченному) ячеек.

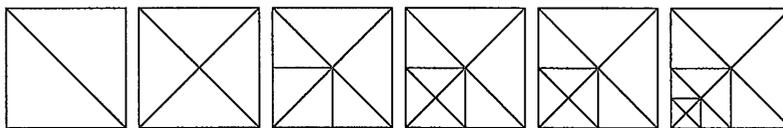


Рис. 3. Локальное измельчение методом бисекций треугольников

Построение операторов межсеточных переходов базируется на сеточной иерархии, особенностях сеточных пространств и свойствах исходного дифференциального оператора. Особенности сеточных пространств проявляются в носителе степени свободы и шаблоне оператора продолжения сеточной функции с грубой сетки на мелкую. Носителями степеней свободы могут быть узлы, ребра, грани, ячейки сетки, а также ансамбли узлов, ребер, граней и ячеек. В случае конформных иерархических треугольных сеток треугольники мелкой сетки являются частью треугольника или целым треугольником более грубой сетки; узлы мелкой сетки являются либо узлами грубой сетки, либо новыми (дополнительными) узлами; ребра мелкой сетки — ребрами грубой сетки или их частями, либо новыми, соединяющими новые узлы. В случае областей с криволинейными границами строгая геометрическая вложенность треугольников и ребер может нарушаться, однако логическая иерархия сохраняется, что обусловлено способом построения сетки. В силу этого многосеточные методы могут быть обобщены на случай подобных квази-иерархических сеток. Непосредственный выбор носителя степени свободы зависит от способа аппроксимации дифференциального оператора, а также от способа построения операторов сеточных уровней. Шаблон оператора продолжения I_l с грубой сетки на мелкую выбирается исходя из структуры связей носителей степеней свободы на сетке, а также соображений устойчивости проектирования с грубой сетки на мелкую. Так, для конечно-разностной аппроксимации оператора Лапласа на квадратной сетке простейший шаблон схематически представим как

$$\begin{array}{ccccccc}
 \mathbf{1} & \rightarrow & 1/2 & \leftarrow & \mathbf{1} & & \\
 \downarrow & \searrow & & \swarrow & \downarrow & & \\
 1/2 & & 1/4 & & 1/2 & , & \\
 \uparrow & \nearrow & & \nwarrow & \uparrow & & \\
 \mathbf{1} & \rightarrow & 1/2 & \leftarrow & \mathbf{1} & &
 \end{array}$$

где жирным шрифтом выделены веса узлов грубой сетки. Отметим, что этот шаблон может изменяться при переходе к оператору диффузии со скачками изотропного коэффициента диффузии или анизотропным тензором диффузии, а также при переходе к оператору конвекции-диффузии [73].

Выбор оператора M_l на грубой сетке может осуществляться несколькими способами. Первый, являющийся таковым и исторически, использует заданный способ аппроксимации дифференциального оператора, примененный к задаче на более грубой сетке. Например, в случае конечно-разностной аппроксимации двумерного оператора Лапласа на квадратной

сетке шаблоны сеточных операторов на мелкой (с шагом h) и грубой сетках (с шагом $H = 2h$) представимы в виде:

$$\begin{array}{ccc}
 -1/h^2 & & -1/H^2 \\
 | & & | \\
 -1/h^2 & - & 4/h^2 & - & -1/h^2 & & -1/H^2 & - & 4/H^2 & - & -1/H^2 \\
 | & & | \\
 -1/h^2 & & -1/H^2
 \end{array}$$

Другим примером может служить конечно-элементная SUPG аппроксимация оператора конвекции-диффузии, которая приводит к немасштабируемому (т.е. не сводящимся друг к другу умножением на некое число) шаблонам на мелкой и грубой сетках [34]. Второй способ построения оператора на грубой сетке — проектирование по Галеркину оператора M_l с мелкой сетки l на грубую $l - 1$. При этом используются пространства сеточных функций на мелкой и грубой сетках V_l, V_{l-1} , а также оператор продолжения сеточной функции $I_l: V_{l-1} \rightarrow V_l$ с грубой сетки на мелкую. Результирующий оператор на грубой сетке представляется как $M_{l-1} = I_l^T M_l I_l$ и вычисляется явно. Ключевым условием выбора оператора продолжения является его устойчивость в соответствующих пространствах с нормой, порождаемой исходным дифференциальным оператором: $\|I_l u\|_* \leq C \|u\|_*$. Отметим, что проектирование по Галеркину не гарантирует сохранение шаблона M_l у сеточного оператора M_{l-1} . В зависимости от сеточных пространств и оператора продолжения, шаблон может расширяться [57] или оставаться неизменным. Поэтому предварительный анализ арифметической сложности алгоритма должен учитывать возможную динамику шаблона сеточных операторов.

Третий способ построения оператора на грубой сетке — алгебраический. Некоторые многосеточные методы ($[92]^+$, $[82]^+$, [23], [25], [26], [27]), выстраивают иерархическую последовательность операторов на основе алгебраического исключения некоторых степеней свободы. Явное исключение возможно для матриц, представимых в ассемблированном по суперэлементам (понимаемым здесь как ячейка грубого уровня, разбитая на ячейки мелкого уровня) виде. Матрица на грубой сетке ассемблируется по суперэлементам, причем локальная матрица вычисляется явно. Определенным преимуществом работы с ассемблируемыми матрицами является точная оценка скорости сходимости итерационного многосеточного метода, не зависящая от скачков изотропного коэффициента диффузии.

Процедура сглаживания R_l — определенный итерационный процесс, который может рассматриваться как метод простой итерации с одним из об-

щепринятых переобуславливателей: диагональным (релаксированный метод Якоби), верхне-треугольным (метод Зейделя), блочно-диагональным, неполного LU -разложения. Отметим, что реализации каждого из этих методов широко известны, поэтому дальнейшая детализация, по-видимому, не имеет смысла.

Эффективность (малая арифметическая цена) многосеточных методов основана на быстром уменьшении числа степеней свободы при переходе от мелкого уровня на более грубый. Так, например, если число степеней свободы уменьшается в четыре раза при переходе на более грубую сетку (равномерное 2D измельчение), то асимптотическую оптимальность арифметической работы будут иметь как V -, так W -циклы. На локально сгущающихся сетках коэффициент уменьшения может быть существенно ниже, что может привести к потере асимптотической оптимальности даже V -циклом. Сохранение эффективности возможно двумя способами. Во-первых, можно переопределить понятие сеточного уровня, который больше не будет связан с процессом построения иерархической сетки, а будет абстрактным набором степеней свободы, соответствующим равномерному укрупнению мелкой сетки. При этом необходимо переопределить все компоненты многосеточного метода. Во-вторых, можно использовать многоуровневые методы [54], реализация которых возможна как в рамках вычисления двойной суммы (по уровням сетки и по степеням свободы на каждом уровне), так и в рамках упрощенного V -цикла.

Иногда особенности задачи не позволяют иметь многоуровневую иерархию вложенных сеток, допуская не более двух или трех уровней равномерного измельчения. Поскольку исходная (самая грубая) сетка является весьма мелкой и неструктурированной, подобные двух- или трехуровневые сетки удобно называть почти неструктурированными [86], [19]. Сложность в использовании почти неструктурированных сеток заключается в арифметической цене решения задачи на самом грубом уровне, размерность которой достаточно велика, чтобы ее можно было решить прямым методом. В некоторых случаях ([82], [26], [27]), будучи примененным к неволновому оператору Гельмгольца $-\Delta + C$ на мелкой сетке K , многосеточный метод порождает оператор $-\Delta + q^K C$ на грубой сетке, где $q \simeq 2$ (2D) или $q \simeq 4$ (3D). Таким образом, при $K = 2$ и $C = 10$ число обусловленности матрицы на грубой 3D сетке ($l = 0$) уменьшается в $4^2 \cdot 16$ вместо 4^2 раз при сохранении оператора на грубом сеточном уровне. Наличие нежесткого оператора на грубом сеточном уровне позволяет обращаться его итерационно с диагональным переобуславливателем. Помимо применимости к почти неструктурированным сеткам, метод обладает хорошим внутренним параллелизмом: распределив по процессорам степени свободы грубого уровня, можно добиться эффективного ускорения за счет равномерной загрузки всех процессоров [26]⁺.

Многосеточные технологии применимы и к полностью неструктурированным сеткам. Например, используя технику метода фиктивного пространства [102], можно построить иерархическую локально измельченную сетку, аппроксимирующую исходную неструктурированную (рис. 4), и сформировать переобуславливатель на основе многосеточного метода и оператора интерполяции с иерархической сетки на неструктурирован-

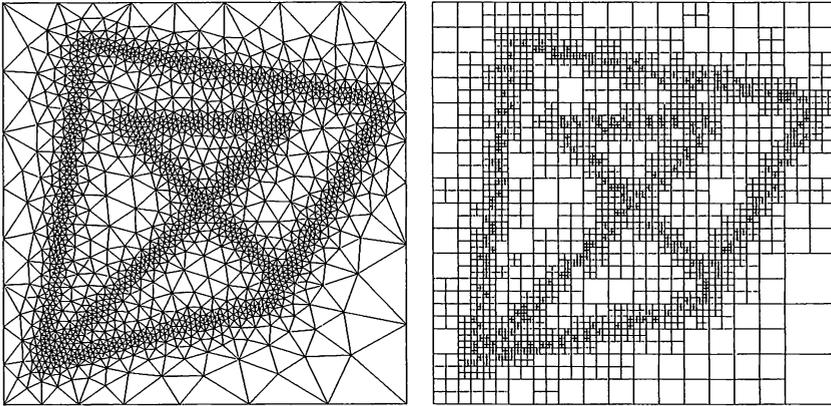


Рис. 4. Неструктурированная и иерархическая сетки

ную [62], [67]. При этом построение локально измельченной иерархической сетки и оператор интерполяции зависят от применяемого многосеточного метода.

2.3. Методы декомпозиции области

Методы декомпозиции области сводят процесс решения краевой задачи к последовательности задач в подобластях, на которые разбита расчетная область. Разбиение на подобласти может осуществляться разными способами и с разными целями. С точки зрения вычислительных технологий метод декомпозиции может быть применим на разных этапах технологической цепочки. Ниже мы рассмотрим некоторые алгоритмы как элементы технологии, группируя их по достигаемым целям.

Одна из целей декомпозиционных алгоритмов — облегчение численного решения, или вычислительно более эффективная процедура решения, сеточной краевой задачи. Рассмотрим несколько примеров. Предположим, что в диффузионной задаче коэффициент диффузии терпит сильные скачки при переходе от подобласти к подобласти, а нам доступны эффективные решатели для сеточных задач с плавно меняющимися коэффициентами. Построение экономичного переобуславливателя для сеточного оператора

во всей области, обеспечивающего скорость сходимости, не зависящую от величины и топологии скачка, возможно как для конформных сеток, так и для сеток, нестыкующихся на границе между подобластями (интерфейсами). Технологическая основа большинства методов декомпозиции — использование решателей сеточных задач в подобластях и разработка интерфейсного переобуславливателя для дополнения Шура для степеней свободы на интерфейсах. Интерфейсный переобуславливатель может быть либо результатом внутреннего итерационного процесса [103], [85]⁺, либо комбинацией сеточных краевых задач в подобластях, либо аддитивным наложением локальных интерфейсных переобуславливателей, ассоциированных с разными топологическими частями интерфейса [112], [105]. В последнем случае накладываются сильные ограничения на структуру интерфейса и, соответственно, на топологию скачков коэффициента диффузии. Некоторые декомпозиционные переобуславливатели [24], [132]⁺ не используют интерфейсное переобуславливание, ограничиваясь исключительно решателями в подобластях. При этом топология распределения коэффициентов по подобластям не может быть произвольной. Отметим, что аддитивный или мультипликативный метод Шварца предлагают простые переобуславливатели в случае разбиения на перекрывающиеся подобласти. Обмен данными между подобластями осуществляется только на границах подобластей [112], [105]. Однако подобный подход эффективен в случае мало меняющихся коэффициентов диффузии, за исключением случаев с простой топологией [72].

Второй пример. Предположим, что единственный доступный нам решатель сеточных задач — LU -разложение для разреженных матриц. Факторизация является весьма эффективным (как по производительности, так и по памяти) методом для небольших матриц (порядка нескольких тысяч), поэтому можно разбить все множество степеней свободы из расчетной области на подмножества мощности $\sim 10^3 - 10^4$, ассоциированные с подобластями. При этом количество подобластей может стать весьма большим, что значительно ухудшает скорость сходимости итерационных процессов с вышеупомянутыми переобуславливателями. Причиной этому является слабое уменьшение итерационной ошибки без подавления низкочастотных компонент за счет решения задачи во всей области. Для того чтобы избежать зависимости скорости сходимости от числа подобластей, применяется технология каркасного пространства, обеспечивающая подавление низкочастотных компонент ошибки, по аналогии с многосеточными методами. При этом используется либо решатель, либо переобуславливатель для исходного оператора в каркасном пространстве, число степеней свободы в котором, как правило, порядка числа подобластей. Отметим, что выбор эффективного решателя сеточных подзадач зависит от типа сетки: на неструктурированных сетках технологически предпочтительно выглядят методы фак-

торизации [136]; на прямоугольных сетках имеет смысл воспользоваться быстрыми прямыми методами, формируя прямоугольные подобласти; на иерархических сетках удобны многосеточные решатели [85], [140].

Процедура разбиения области на пересекающиеся или непересекающиеся подобласти может осуществляться либо на уровне матриц, либо на уровне сетки, либо на уровне геометрического моделирования. В последнем случае используют данные о геометрической модели, зачастую представляемой объединением и/или пересечением областей простой формы, а также данные о зонах с одинаковым значением коэффициентов дифференциального оператора. Разбиение на уровне сетки подразумевает разделение узлов или ячеек на непересекающиеся подмножества (которые могут быть впоследствии расширены для получения перекрывающихся подобластей). Разбиение может осуществляться на разных принципах [139], учитывающих или не учитывающих ребра сетки и физическое расположение узлов. Разбиение на уровне матриц является разбиением графа матрицы, представляющего структуру разреженности матрицы. В этом случае невозможно гарантировать не только гладкость интерфейса, но даже односвязность подобластей. Зачастую негладкость интерфейса и топологическая несогласованность интерфейса и скачка коэффициентов приводят к замедлению скорости сходимости итерационного процесса, поэтому важны методы, не чувствительные к побочным эффектам автоматического разбиения на подобласти.

Другой целью декомпозиционных алгоритмов может быть параллелизация вычислений. Разделение данных по принадлежности подобластям очень удобно при распределении данных по процессорам. При этом наряду со скоростью сходимости ключевую роль играет равномерное распределение нагрузки по процессорам для получения высокой параллельной эффективности. Иногда геометрические соображения (распределение коэффициента, гладкость интерфейса) могут приноситься в жертву удобству и эффективности параллелизации. Поэтому в параллельных расчетах могут использоваться методы, не обеспечивающие (почти) независимость скорости сходимости от шага сетки, но не чувствительные к гладкости интерфейса или поведению скачка коэффициента [133]. Поскольку в параллельных вычислениях количество процессоров и подобластей может быть большим, использование переобуславливателя или решателя на каркасном пространстве является неизбежным. Каркасная задача, как правило, не является большой и решается на одном процессоре. Иногда каркасная задача решается приближенно на всех процессорах с помощью внутренних итераций. В последнем случае каркасный решатель входит в переобуславливатель мультипликативно, после независимой (аддитивной) обработки каждым процессором своей подзадачи и вычисления невязки после аддитивного переобуславливания [133], [96]. Подобная комбинация аддитивного

и мультипликативного метода Шварца порождает гибридный метод декомпозиции [112]. В случае, когда каркасный решатель входит в переобуславливатель аддитивно, т.е. независимо от переобуславливания в подобластях, он применяется на дополнительном процессоре, при этом желательна балансировка его нагрузки с нагрузкой других процессоров. Примером технологических решений параллельных методов декомпозиции могут служить программные комплексы [140], [136].

Сингулярно возмущенные операторы предоставляют дополнительные возможности для эффективной параллелизации решателей на основе методов декомпозиции с перекрывающимися подобластями. Так, неявная аппроксимация по времени с шагом τ конечно-разностного параболического уравнения на сетке с шагом h порождает оператор с сингулярно возмущенным членом нулевого порядка. При этом глубина перекрытия порядка $\sqrt{\tau} \ln(\epsilon h^d / \tau)$ обеспечит точность ϵ приближения дискретного решения в каждой подобласти вне зоны перекрытия, даже если граничные значения для подзадач берутся с предыдущего временного слоя [93]. Таким образом, независимое решение подзадач с относительно небольшим перекрытием обеспечивает хорошее приближение к неявному решению. Аналогичным образом можно использовать свойства оператора конвекции-диффузии с малым диффузионным коэффициентом. Двухуровневое разбиение с малыми перекрытиями и специальный вариант мультипликативного метода Шварца обеспечивают высокую скорость сходимости (не зависящую от шага сетки и числа процессоров) без введения каркасного пространства и хорошую параллелизацию алгоритма [66].

Композиция подзадач с разными типами операторов — третья цель методов разбиения области. Подобные разбиения возникают, например, когда сложный оператор в большей части области заменяется на упрощенный, сохраняя свой исходный вид лишь в меньшей подобласти, где использование сложной модели физически оправдано. Другой причиной композиции разных уравнений может быть моделирование взаимодействия различных физических процессов, происходящих в разных зонах расчетной области. Для уравнений второго порядка их склейка на границе между подобластями означает непрерывность решения и его потоков и может осуществляться явно, через выписывание интерфейсного уравнения в терминах операторов Пуанкаре-Стеклова, либо неявно, через выполнение некоей итерационной процедуры. Общий функциональный подход формирования уравнения Пуанкаре-Стеклова и возможные способы его итерационного решения были рассмотрены в [28], а приложения этого подхода подробно изучаются в [105]. Преимуществом явного выписывания интерфейсного уравнения является возможность вычислить итерационную невязку, что позволяет применить итерационные алгоритмы с автоматическим выбором параметров (CG, GMRES, BiCGstab и т.п.). Неявная итерационная склейка использует

гетерогенный метод декомпозиции типа Нейман-Дирихле [105], при котором последовательное решение краевых подзадач гарантирует непрерывность решения (или потока), а итерационная релаксация склеивает потоки (или решения). Преимуществом неявной склейки оказывается простейшая комбинация двух независимых решателей разных подзадач, а его недостатком является необходимость выбора релаксационного параметра.

2.4. Технологии "черного ящика"

Отличительной особенностью методов декомпозиции, фиктивных областей и многосеточных методов является дополнительная информация о сеточной системе, которую необходимо решить. С формальной точки зрения линейные системы могут быть заданы только через элементы матрицы и вектор правой части. Остальные данные носят вспомогательный характер и могут быть использованы для ускорения процесса решения. При создании программных комплексов, реализующих те или иные решения краевых задач, очень важным является представление о том, какими методами будут решаться линейные системы. Разрабатываемая структура данных должна учитывать требования используемых методов. В случае, когда программный комплекс не предоставляет никакой дополнительной информации о сеточной системе, приходится использовать решатель, использующий только минимальную информацию о системе, т.е. матричные элементы и вектор правой части. Подобные подходы называются технологиями "черного ящика", поскольку процесс получения решения скрыт от пользователя, роль которого сводится к вводу матрицы и вектора правой части и выводу полученного решения. С одной стороны, это очень удобный способ интегрирования решателя в программный комплекс в силу минимизации обмена данными; с другой стороны, это затрудняет управление процедурой решения со стороны пользователя. Кроме того, технологии "черного ящика" зачастую используют некоторые предположения о матрице, которые могут не удовлетворяться на практике. Тем не менее именно эти технологии позволяют решать многие разреженные системы, не обладающие никакой дополнительной информацией.

Наиболее общими методами решения систем линейных уравнений являются методы факторизации, в которых матрица системы разлагается на треугольные множители: $A = LU$, где L и U — нижне- и верхнетреугольные матрицы, соответственно, а затем решаются последовательно две системы с треугольными матрицами. Для сеточных систем, возникающих при аппроксимации краевых задач, матрица A часто является разреженной, поэтому арифметическая цена факторизации матрицы порядка n оказывается значительно меньше оценки $n^3/3$, верной для плотных матриц. Теоретические оценки треугольного разложения для разрежен-

ных матриц [37] не дают четкого представления об области его практической применимости. Появившиеся в последние годы алгоритмы и их реализации [20, 21, 60, 75] позволяют существенно расширить границы этой области. Этому удалось добиться за счет усовершенствования методики факторизации и применения параллельных компьютеров. Так, например, конечно-разностная аппроксимация оператора Лапласа на трехмерной сетке $50 \times 50 \times 50$ ($n = 125000$) может быть успешно факторизована на персональном компьютере, а на трехмерной сетке $100 \times 100 \times 100$ ($n = 1000000$) — на 32-процессорном параллельном компьютере. Тем не менее приемлемыми размерностями для решения неструктурированных трехмерных задач на однопроцессорном компьютере, по-видимому, обладают разреженные матрицы порядка десятков тысяч. Основные этапы разложения матрицы на треугольные множители — переупорядочивание, символическая факторизация, и численная факторизация. Параллелизация, как правило, применяется к последнему этапу, который основан на мультифронтальном алгоритме. Стоит отметить, что современный стандарт пакетов факторизации требует механизма контроля ошибки (оценка числа обусловленности, невязки), а также улучшения точности решения (итерационное уточнение).

Широко применяемой альтернативой методам точной факторизации является технология неполного LU -разложения, при которой произведение найденных треугольных множителей формирует переобуславливатель, используемый при последующем итерационном решении. Достоинствами этого подхода можно назвать быстроту факторизации и умеренные затраты на машинную память, а недостатками — относительно медленная скорость сходимости и возможное появление нулевых пивотов, т.е. тех матричных элементов, на которые необходимо осуществлять деление. Построчное формирование треугольных множителей при неполной факторизации представляет собой усеченную процедуру точной факторизации: записываются только те элементы, которые удовлетворяют определенным критериям, остальные либо отбрасываются, либо используются каким-либо образом при формировании текущей строки/столбца у множителей. Критерии записи элементов делятся на критерии по позиции и критерии по значению, когда записываемое значение должно быть не меньше некоторого порога [14], [50], [109]. Позиционные критерии используют заранее определенную структуру разреженности треугольных множителей. Например, это может быть структура разреженности исходной матрицы (ILU0). Разработаны параллельные версии методов неполной факторизации, причем для компьютеров как с общей, так и распределенной (по процессорам) памятью [109]. Одно из направлений развития методики — построение разложений, обеспечивающих высокую скорость сходимости итерационных процессов за счет более аккуратной (а следовательно, вычислительно более дорогой) построчной процедуры факторизации [90].

Алгебраические многосеточные методы [58]⁺, подобно методам неполной факторизации, формируют переобуславливатель для исходной матрицы, однако часто обеспечивают его спектральную эквивалентность с исходной матрицей, т.е. независимость скорости сходимости от порядка матрицы. Будучи многосеточными методами, они наследуют все основные элементы многосеточной технологии; будучи технологиями "черного ящика", они используют только информацию об элементах матрицы. Поэтому необходимые данные об иерархии алгебраических "сеточных уровней" восстанавливаются искусственно, даже в том случае, если сетка не обладает никакой иерархией. На первом этапе все степени свободы разделяются на мелкие и грубые, причем последние в свою очередь разделяются на мелкие и грубые степени свободы следующего уровня, и так далее рекурсивно.

Далее определяются операторы продолжения векторов с грубых на мелкие степени свободы каждого уровня, операторы интерполяции на более грубые уровни и матрицы каждого уровня. После этого организация вычислений аналогична стандартным многосеточным методам. Обычно операторы интерполяции выбираются как транспонированные матрицы продолжения, а матрицы более грубого уровня являются Галеркинскими проекциями матриц более мелкого уровня (матрица самого мелкого уровня есть исходная матрица). Разделение степеней свободы на мелкие и грубые и построение операторов продолжения подчинены главной задаче — ошибка на более грубом уровне должна быть более гладкой сеточной функцией. Технически эти процедуры допускают несколько реализаций, причем каждая из них использует свое понятие "гладкости" ошибки. Последнее, в свою очередь, опирается на определенные свойства сеточной системы, которыми, как предполагается, она обладает. Классическим примером может быть требование первой программной реализации алгебраического многосеточного метода [116] — матрица системы должна быть M -матрицей. Таким образом, алгебраические многосеточные методы, в отличие от методов полной или неполной факторизации, предполагают наличие у матрицы определенных свойств, порождаемых дискретизациями дифференциальных операторов. В случаях, когда матрица не обладает этими свойствами, эти методы могут быть применены к подматрицам с этими свойствами в рамках блочных переобуславливателей [94], [95].

3. Примеры организации технологических процессов

Реализация высокопроизводительных вычислительных технологий сопряжена с разработкой больших объемов программных комплексов и осуществима практически только в коммерческих прикладных пакетах. Опи-

сание десятков таких продуктов представлены в Интернете, среди которых широкое распространение имеют, например, [138], [137], [135]. Их конечный успех зачастую определяется не столько математическими, сколько менеджерскими технологиями, но в целом рынок пакетов прикладных программ (ППП) содержит значительный ассортимент разработок, ориентированных на широкий круг задач, связанных с выполнением крупномасштабных вычислительных экспериментов.

3.1. Цели и назначение ППП БАЗИС-А

В качестве примера организации технологических процессов мы приведем краткое описание ППП БАЗИС-А [9], ориентированного на комплексное трехмерное моделирование следующих взаимосвязанных технологических процессов в алюминиевом электролизере [32]:

- моделирование температурных полей с учетом различных источников тепла, конвективных потоков и зависимостей теплофизических параметров от температуры;
- расчет электрических полей и токов с учетом температурной зависимости электропроводности и электрохимических процессов;
- расчет магнитных полей с учетом объемного распределения плотностей токов, внешних токоподводов и свойств ферромагнитных материалов, а также влияния близлежащих электролизеров;
- моделирование магнитогидродинамических течений в расплавах электролита и алюминия с учетом пондеромоторных сил и поверхностных волновых явлений, описываемых уравнениями Навье-Стокса;
- в рамках линейной теории термоупругости расчет деформаций и напряжений под воздействием нагрузок различного типа.

В целом электролизер представляет собой сложную физико-техническую систему, а проблемы расчета различных полей, необходимых для понимания и управления процессами производства алюминия, рассматривались многими авторами.

ППП БАЗИС-А основан на использовании современных эффективных алгоритмов и информационных технологий, включая средства геометрического и функционального моделирования, построение адаптивных сеток, поэлементные принципы аппроксимации краевых задач методами конечных объемов и/или конечных элементов, а также быстрые итерационные алгоритмы решения систем линейных и нелинейных алгебраических уравнений с разреженными матрицами высокого порядка, составляющих ядро вычислительных инструментариев для создания различных программных конфигураций и приложений.

Реализация алгоритмов и архитектура ППП базируется на гибких структурах данных: геометрических, функциональных, сеточных и алгебраических (ГСД, ФСД, ССД, АСД), — обеспечивающих взаимодействие различных программных компонент, а также развиваемость пакета в целом за счет добавления новых математических моделей и методов.

Данный пакет программ содержит развитый графический пользовательский интерфейс (ГПИ) и постпроцессор, обеспечивающие задание сложных входных геометрических и функциональных данных (произвольные конфигурации материальных сред, различные системы дифференциальных уравнений и граничные условия) и наглядную визуализацию результатов.

ППП БАЗИС-А может использоваться на стадиях НИР или ОКР для глубокого изучения характера физических эффектов в различных режимах электролиза алюминия или на этапах проектирования новых типов электролизеров с апробацией инженерных решений, или в составе заводского АСУТП типа "виртуальный электролизер" для оперативного управления технологическими процессами, оптимизации эксплуатационных режимов и выработки рекомендации по устранению нежелательных ситуаций. ППП БАЗИС-А может применяться и в других прикладных областях, связанных с необходимостью расчетов электростатических, магнитных и/или тепловых полей, а также магнитогидродинамических течений и термоупругих напряжений.

Основными предполагаемыми пользователями данного пакета являются научные работники и инженеры, связанные с проектированием, исследованием и эксплуатацией алюминиевых электролизеров или других электромагнитных и теплофизических устройств. ППП БАЗИС-А может использоваться также как обучающая программа для студентов и аспирантов физико-математических и технических специальностей, а также на курсах повышения квалификации инженерного персонала для освоения технологий математического моделирования.

3.2. Математические постановки задач моделирования процессов в алюминиевом электролизере

Рассматриваемые самосогласованные трехмерные задачи являются в общем случае нестационарными и описываются системой дифференциальных уравнений, нелинейные связи между которыми осуществляются через зависимость коэффициентов и/или правых частей от соответствующих решений. Электромагнитные поля при этом предполагаются квазистационарными, т.е. описываемыми формально стационарными уравнениями, но с параметрической зависимостью входных данных от времени. С учетом основных электромагнитных, теплофизических и гидродинамических фак-

торов краевые задачи для различных полей могут быть сформулированы следующим образом.

3.2.1. Электростатическая задача

Электрический потенциал $\varphi(\vec{x})$ описывается в ограниченной расчетной области Ω_φ с внешней границей Γ_φ уравнением

$$\operatorname{div}(\sigma \operatorname{grad} \varphi) = s, \quad \vec{x} = (x, y, z) \in \Omega_\varphi, \quad (3.1)$$

где проводимость среды $\sigma = \sigma(T)$ есть функция от температуры, различающаяся в каждой из физических подобластей. Правая часть в уравнении (3.1) определяется индуцированными токами в соответствии с законом Фарадея:

$$\vec{j}_f = \sigma[\vec{u} \times \vec{B}], \quad s = \operatorname{div} \vec{j}_f,$$

где \vec{u} , \vec{B} суть векторы скорости движущейся проводящей среды (расплава) и магнитной индукции.

На различных участках внешней границы Γ_φ расчетной "электростатической" области могут задаваться или условия Дирихле (заданный потенциал), или однородные условия Неймана на плоскостях симметрии, или определяемые из качественных физических соображений условия третьего рода (Ньютона), или заданное (из каких-то дополнительных физических соображений) распределение плотности входящих и (или) выходящих токов.

Главными результатами решения данной задачи являются векторы напряженности электрического поля и плотности электрического тока

$$\vec{E} = -\operatorname{grad} \varphi, \quad \vec{j} = \sigma \vec{E} + \vec{j}_f. \quad (3.2)$$

3.2.2. Магнитостатическая задача

Данную краевую задачу формулируем относительно функции $\psi(\vec{x})$, которая в "магнитостатической" расчетной области Ω_ψ с границей Γ_ψ удовлетворяет уравнению

$$\operatorname{div}(\mu \operatorname{grad} \psi) = 0, \quad (3.3)$$

где функция μ (магнитная проницаемость) в ферромагнитной среде Ω_f зависит от магнитного поля (а также от температуры T) и является постоянной в неферромагнитной среде Ω_0 , $\Omega_\psi = \Omega_f \cup \Omega_0$. Традиционно функция ψ в Ω_f называется полным потенциалом, а в Ω_0 — неполным потенциалом, поскольку в соответствующих подобластях вектор напряженности \vec{H} определяется различным образом:

$$\vec{H} = \begin{cases} -\operatorname{grad} \psi, & \vec{x} \in \Omega_f, \\ -\operatorname{grad} \psi + \vec{H}_0, & \vec{x} \in \Omega_0, \end{cases} \quad (3.4)$$

где \vec{H}_0 есть вектор напряженности магнитного поля, создаваемого определяемыми в пункте 3.2.1 внешними токами в однородном пространстве с $\mu = 1$; этот вектор описывается формулой Био-Савара

$$\vec{H}_0(\vec{x}) = \frac{1}{4\pi} \int_V \frac{[\vec{j}, \vec{r}]}{r^3} d\vec{x}', \quad (3.5)$$

где V — ограниченный объем, включающий ненулевые учитываемые токи $\vec{j}(\vec{x}')$, $\vec{r}' \in V$, а \vec{r} — радиус-вектор, соединяющий точку проводника \vec{x}' с точкой наблюдения \vec{x} .

На границе Γ_f ферромагнитной среды решение уравнения (3.3) удовлетворяет условиям сопряжения со скачками

$$\mu \frac{\partial \psi}{\partial n} \Big|_{\Gamma_f^+} = \frac{\partial \psi}{\partial n} \Big|_{\Gamma_f^-} + (\vec{n}, \vec{H}_0), \quad \psi|_{\Gamma_f^+} = \psi|_{\Gamma_f^-} + \psi^0, \quad \vec{x} \in \Gamma_f, \quad (3.6)$$

где знак "—" относится к неферромагнитной среде, а ψ^0 есть решение вспомогательной краевой задачи

$$\Delta \psi^0 = 0, \quad \vec{x} \in \Omega_0, \quad \frac{\partial \psi^0}{\partial n} \Big|_{\Gamma_f} = (\vec{n}, \vec{H}), \quad \vec{x} \in \Gamma_f. \quad (3.7)$$

На внешней границе ограниченной области Ω_ψ ставятся краевые условия из качественных соображений (в строгой постановке магнитное поле стремится к нулю на бесконечности). Например, над электролизером поле может задаваться по закону Био-Савара без учета влияния стального корпуса, а при учете влияния магнитных полей соседних электролизеров на плоскостях симметрии между ними ставится однородное условие Неймана.

Функция магнитной проницаемости $\mu(H = |\vec{H}|)$ ферромагнитного материала определяется соотношением $\vec{B} = \mu\mu_0\vec{H}$, где μ_0 — магнитная проницаемость вакуума, а \vec{B} — вектор магнитной индукции. Она для разных материалов или задается приближенно эмпирическими формулами, или определяется из аппроксимации экспериментальных данных.

3.2.3. Тепловая задача

Рассмотрим для простоты стационарное уравнение теплопроводности без учета конвективного теплопереноса в расплаве, а также эффектов влияния электромагнитного поля на перенос тепла:

$$-\text{div}(\lambda \text{grad } T) = f \equiv \sigma |\text{grad } \psi|^2, \quad \vec{x} \in \Omega_T, \quad (3.8)$$

где коэффициент теплопроводности λ зависит от температуры, а правая часть представляет источник тепла, обусловленный джоулевым нагревом.

На внешней границе Γ_T "тепловой" расчетной области Ω_T ставятся краевые условия теплообмена с внешней средой.

3.2.4. Модель описания магнитогидродинамических течений

Рассмотрим стационарное течение алюминиевого расплава, в котором вектор скорости $\vec{u} = (u, v, w)$ и давление p не зависят от времени и описываются в приближении Буссинеска системой уравнений типа магнитогидродинамики Навье-Стокса для сжимаемой жидкости:

$$\begin{aligned} \rho(\vec{u} \cdot \nabla)\vec{u} &= -\text{grad } p + \nu \Delta \vec{u} + [\vec{j} \times \vec{B}] + \rho \vec{g}, \quad \vec{x} \in \Omega_m, \\ \text{div } \rho \vec{u} &= 0, \end{aligned} \quad (3.9)$$

где ν — коэффициент динамической вязкости, $\vec{g} = (0, 0 - g)^t$, g — величина ускорения свободного падения, ρ — зависящая от температуры, в силу объемного расширения, функция плотности, а член с векторным произведением определяет действие пондеромоторных (лоренцовых) сил.

На границах жидкости с твердыми стенками ставится условие прилипания $\vec{u} = 0$, а на верхней свободной поверхности расплава формулируются кинематические условия $d\vec{x}/dt = \vec{u}$, а также задаются условия равновесия.

Соответствующие нестационарные течения в расплаве описываются также системой уравнений (3.9), в которой к левой части первого из них добавляется член с производной по времени $\rho \frac{\partial \vec{u}}{\partial t}$, а также ставятся дополнительно начальные условия при $t = 0$.

3.2.5. Уравнения термоупругости

Рассматривается трехмерная статическая задача линейной теории упругости с учетом термических нагрузок в области, состоящей из параллелепипедов с разными, вообще говоря, прочностными характеристиками.

Данная задача для изотропных сред описывается уравнением Ламе

$$-2 \text{div} (\mu \varepsilon(\vec{u})) - \text{grad} (\lambda \text{div } \vec{u}) = \vec{f} - \text{grad} (\alpha(2\mu + 3\lambda)T), \quad (3.10)$$

где используются следующие обозначения: $\vec{u} = (u_1, u_2, u_3)^T$ — вектор перемещений, задающий векторное поле $\vec{u} = \vec{u}(\vec{x})$ в точках $\vec{x} = (x_1, x_2, x_3)^T$ недеформированного тела, $T = T(\vec{x})$ — скалярное температурное поле, \vec{f} — вектор объемной нагрузки, α — коэффициент термического расширения, $\lambda > 0$ и $\mu > 0$ — коэффициенты Ламе, а $\varepsilon(\vec{x}) = \varepsilon_{i,j}$ есть тензор деформаций с компонентами

$$\varepsilon_{i,j} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad i, j = 1, 2, 3. \quad (3.11)$$

Относительно краевых условий предполагается, что часть тела закреплена, часть свободна, и при этом какие-либо силы на границе не действуют.

При известных перемещениях симметричный тензор напряжений $\sigma(\vec{x})$ определяется "температурной модификацией" закона Гука

$$\sigma = 2\mu\varepsilon + \lambda tr(\varepsilon)\delta - \alpha(2\mu + 3\lambda)T\delta, \quad (3.12)$$

где δ — единичный тензор.

3.3. Краткое описание алгоритмов

Мы рассмотрим принципы дискретизации и аппроксимации методами конечных объемов и алгебраизации описанных выше задач на примере одного дифференциального уравнения вида (3.1), (3.3), или (3.8) относительно одной неизвестной (скалярной) функции u .

На неравномерной параллелепипеодоидальной сетке $\Omega^h = \{x_i, y_j, z_k\}$,

$$i = 1, \dots, I, \quad j = 1, \dots, J, \quad k = 1, \dots, K,$$

покрывающей всю расчетную область, то есть $\Omega \in [x_1, x_{I+1}] \times [y_1, y_{J+1}] \times [z_1, z_{K+1}]$, определяются конечные объемы (ячейки Дирихле-Вороного)

$$V_{i,j,k} = \{x_{i-1/2} < x < x_{i+1/2}, \quad y_{j-1/2} < y < y_{j+1/2}, \quad z_{k-1/2} < z < z_{k+1/2}\}$$

с поверхностями $S_{i,j,k}$, для которых выписываются интегральные соотношения баланса

$$\int_{S_{i,j,k}} \lambda \frac{\partial u}{\partial n} ds = \int_{V_{i,j,k}} f dv. \quad (3.13)$$

Уравнения (3.13) выписываются для всех внутренних узлов $(x_i, y_j, z_k) \in \Omega$, а также для узлов, лежащих на части границы Γ , содержащей краевые условия 2-го или 3-го рода, после чего они аппроксимируются с учетом условий на внешних и внутренних границах. Предполагается, что все граничные поверхности проходят по координатным плоскостям сетки Ω^h . Если узел (i, j, k) лежит на границе расчетной области, то баланс субстанции выписывается для "внутренней" части ячейки Дирихле-Вороного $V_{i,j,k} \cap \Omega$.

В результате аппроксимации соотношений (3.13) формируется система алгебраических уравнений с разреженной матрицей

$$\begin{aligned} (Au)_{i,j,k} \equiv & p_{i,j,k}^0 u_{i,j,k} - p_{i,j,k}^1 u_{i-1,j,k} - p_{i,j,k}^2 u_{i,j-1,k} - \\ & - p_{i,j,k}^3 u_{i+1,j,k} - p_{i,j,k}^4 u_{i,j+1,k} - p_{i,j,k}^5 u_{i,j,k-1} - \\ & - p_{i,j,k}^6 u_{i,j,k+1} = f_{i,j,k}, \quad (i, j, k) \in \Omega^h \cup \Gamma_2^h \cup \Gamma_3^h, \end{aligned} \quad (3.14)$$

где коэффициенты и правые части вычисляются с учетом краевых условий.

Сеточное решение $u^h = \{u_{i,j,k}\}$ аппроксимирует решение исходной дифференциальной задачи на равномерной сетке с погрешностью $O(h^2)$, а в общем случае ошибка имеет первый порядок по h .

Если уравнение (3.13) является нелинейным, то коэффициенты системы (3.14) зависят от искоемых $u_{i,j,k}$ и для ее решения применяется какой-либо итерационный метод квазилинеаризации, на каждом шаге которого решается СЛАУ того же вида (3.14) с симметричной положительно определенной матрицей. Ее решение, в свою очередь, осуществляется быстрым итерационным методом неполной факторизации с ускорением сопряженными алгоритмами.

При решении уравнений Навье-Стокса, в предположении заданности электромагнитных и теплового полей, после аппроксимации получаем систему следующего вида:

$$\begin{aligned} & \begin{bmatrix} \bar{A}_{11} & \nabla \\ \nabla^t & 0 \end{bmatrix} \begin{bmatrix} \vec{u} \\ p \end{bmatrix} = \\ & = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ p \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ 0 \end{bmatrix} = \begin{bmatrix} f \\ 0 \end{bmatrix}, \quad (3.15) \end{aligned}$$

где коэффициенты матричных блоков в \bar{A}_{11} зависят от искомого решения. Применяя процесс квазилинеаризации, на каждом его шаге будем получать систему несимметричных линейных уравнений того же вида (3.15), для решения которой используем "внутренние" итерации с быстрыми алгоритмами неполной факторизации и обобщенными сопряженными градиентами.

Блочная система уравнений (3.9) представляет сложную задачу с седловой точкой со знаконеопределенной матрицей, алгоритмам решения которой посвящено большое количество работ отечественных и зарубежных авторов [47]⁺, [76]⁺.

Для определения вектора перемещений после аппроксимации уравнения Ламе (3.10) решается блочная система с матрицей вида \bar{A}_{11} из (3.15) [78].

При расчете различных взаимодействующих полей, описанных в пунктах 3.2.1–3.2.5, реализуется дополнительный внешний итерационный процесс типа блочного алгоритма Зейделя.

3.4. Вычислительно-информационные технологии

Вычислительные инструментари и технологии основываются на современных подходах, включая следующие компоненты.

- Геометрическое и функциональное моделирование, обеспечивающее описание и модификацию сложных трехмерных геометрических объектов и математических моделей, а также автоматизацию построения алгоритмов и оптимизацию геометрических и/или физических параметров задачи. Система геометрического и функционального моделирования (СГФМ) призвана обеспечить, с одной стороны, автономные реализации графических пользовательских интерфейсов для конкретных прикладных областей, а с другой — представление межмодульных интерфейсов и внутренних структур данных для управления вычислительным процессом. Результатом работы СГФМ являются геометрическая и функциональная структуры данных (ГФСД).
- Адаптивные локально-модифицированные сетки, конструируемые с учетом геометрических и функциональных особенностей задачи. Алгоритмы генерации, анализа и обработки сеток поддерживаются сеточными структурами данных (ССД), описывающими, соответственно, сеточные подобласти (макросетка) и элементарные сеточные объекты (вершины, ребра, грани и объемы), а также их взаимосвязи с ГФСД.
- Конечно-объемные и конечно-элементные аппроксимации повышенной точности с учетом сингулярностей решений дифференциальных уравнений. Формально блок аппроксимации переводит ССД в алгебраическую структуру данных (АСД), которая отображает исходную математическую постановку на дискретном уровне. Алгоритмы построения систем алгебраических уравнений базируются на поэлементных технологиях МКО и МКЭ, заключающихся в вычислении локальных матриц баланса или жесткости, соответственно, и сборки глобальных матриц. Данный подход значительно повышает уровень технологии программирования и идеально приспособлен для распараллеливания.
- Быстрые методы решения систем линейных и нелинейных алгебраических уравнений (СЛАУ и СНАУ) высокого порядка с разреженными матрицами на разных этапах моделирования стационарных и нестационарных процессов. Для нелинейных систем решение осуществляется путем квазилинеаризации с итерационным решением СЛАУ на каждой внешней "нелинейной" итерации. Реализация итерационных алгоритмов проводится с использованием экономичных разреженных строчных форматов для симметричных или несимметричных матриц. Быстрые итерационные процессы выполняются с помощью предобусловленных методов неполной факторизации и ускорения обобщенными методами сопряженных градиентов.

3.5. Общая структура пакета программ БАЗИС-А

ППП БАЗИС-А предназначен для функционирования в операционных средах Windows или Unix. Его три составные части следующие.

- Комплекс вычислительных модулей, включающий предварительную обработку исходной информации и формирования внутренних структур данных, процедуры дискретизации и аппроксимации решаемых краевых задач для систем дифференциальных уравнений, решения получаемых алгебраических систем, а также средства управления общим вычислительным процессом.
- Постпроцессор для обработки и подготовки к визуализации результатов расчета трехмерных полей, включая построение изолиний решений в заданных сечениях, изоповерхностей, силовых линий, двумерных и одномерных графиков трехмерных функций, а также формирование векторных полей в плоскостях и в объеме.
- Графический пользовательский интерфейс, включающий цветные изображения трехмерных объектов с возможностями аксонометрических и полутонных представлений, их модификаций и редактирования. Целью ГПИ является полностью интерактивный ввод исходных геометрических и функциональных данных, визуализация результатов расчетов и управление вычислительным сеансом. Вычислительные модули ППП БАЗИС-А запрограммированы на языке ФОРТРАН 77, а интерфейсный блок со средствами трехмерного моделирования реализован с помощью объектно-ориентированных и графических средств JAVA..
- Структура ГПИ представлена неизменным ядром и пополняемым набором интерфейсных компонент, разрабатываемых по мере настройки ППП на новые приложения. Пользователю предоставляется набор окон с панелями инструментов, меню и иконками управления, обеспечивающими ввод исходных данных в числовом виде и их графическое отображение со стандартным набором средств визуализации.

Работа выполнена при финансовой поддержке РФФИ (коды проектов 02-01-00590, 02-01-1176, 04-07-90336) и гос. программы "Матричные методы в интегральных и дифференциальных уравнениях".

Литература

- [1] *Бахвалов Н.С. и др.* Численные методы решения задач математической физики // В настоящем сборнике, с. 7–97.
- [2] *Бахвалов Н.С.* Эффективные методы решения жестких многомерных многопараметрических задач // ЖВМ и МФ. 1999. Т. 39. С. 2019–2049.

- [3] *Воеводин В.В.* Об одном методе понижения порядка матриц при решении интегральных уравнений // Численный анализ на ФОРТРАНе. М.: МГУ, 1979. С. 21–26.
- [4] *Валиуллин А.Н., Ганюжа В.Г., Ильин В.П., Шапеев В.П., Яценко Н.Н.* Задача автоматического построения и исследования на ЭВМ разностных схем в аналитическом виде // ДАН СССР. 1984. Т. 275, № 3. С. 528–532.
- [5] *Воеводин В.В., Тыртышников Е.Е.* Вычисления с теплицевыми матрицами // Вычислительные процессы и системы. М.: Наука, 1983, вып. 1. С. 124–266.
- [6] *Воеводин В.В., Тыртышников Е.Е.* Вычислительные процессы с теплицевыми матрицами. М.: Наука, 1987.
- [7] *Воеводин В.В., Тыртышников Е.Е.* Матричные методы и алгоритмы. М.: ИВМ РАН, 1993.
- [8] *Годунов С.К.* Современные аспекты линейной алгебры. Новосибирск: Научная книга, 1997.
- [9] *Голосов И.С., Горбенко Н.И., Гурьева Я.Л., Ильин В.П. и др.* Пакет программ БАЗИС-А для моделирования технологических процессов в алюминиевом электролизере // Труды докл. междунар. конф. "Алюминий Сибири-2002". Красноярск: ИВМ СО РАН, 2002. С. 187–194.
- [10] *Горейнов С.А., Замарашкин Н.Л., Тыртышников Е.Е.* Псевдоскелетные аппроксимации матриц // Доклады РАН. 1995. Т. 343, № 2. С. 151–152.
- [11] *Гохберг И., Семенцул А.А.* Об обращении конечных теплицевых матриц и их континуальных аналогов // Матем. иссл. 1972. Т. 7, № 24. С. 201–224.
- [12] *Ершов А.П., Ильин В.П.* Пакеты программ как методология решения прикладных задач // Пакеты прикладных программ. Проблемы и перспективы. М.: Наука, 1982. С. 4–18.
- [13] *Ильин В.П.* Численные методы решения задач электрофизики. М.: Наука, 1985.
- [14] *Ильин В.П.* Методы неполной факторизации для решения алгебраических систем. М.: Наука, 1995.
- [15] *Ильин В.П.* Вычислительная информатика: открытие науки. Новосибирск: Наука, 1991.
- [16] *Ильин В.П.* Вычислительно-информационные технологии математического моделирования // Автометрия. 2000. № 1. С. 3–16.
- [17] *Ильин В.П.* Методы конечных разностей и конечных объемов для эллиптических уравнений. Новосибирск: ИМ СО РАН, 2000.
- [18] *Ильин В.П.* Геометрическое и функциональное моделирование в задачах математической физики // Вычислительные технологии. 2001. Т. 6, ч. 2. С. 315–321.
- [19] *Ильин В.П., Ицкович Е.А.* Трехсеточный метод решения двумерных краевых задач. Препринт № 1079. Новосибирск: ВЦ СО РАН, 1996.
- [20] *Ильин В.П., Карначук В.И., Ларин М.Р.* Древоподобный подход к организации структуры данных для разложения Холесского // ЖВМ и МФ. 1994. Т. 34, № 12. С. 1747–1756.
- [21] *Ильин В.П., Нефедов В.Л.* Метод диагональных переносов (TS-алгоритм) для решения пятиточечных уравнений // ЖВМ и МФ. 2001. Т. 41, № 2. С. 186–199.

- [22] Кузнецов Ю.А. Итерационные методы в подпространствах // Вычислительные процессы и системы. Т.2. М.: Наука, 1984. С. 265–350.
- [23] Кузнецов Ю.А. (ред.) Методы математической физики. М.: ОВМ АН СССР, 1988.
- [24] Кузнецов Ю.А. (ред.) Численный анализ и математическое моделирование. М.: ОВМ АН СССР, 1989.
- [25] Кузнецов Ю.А. (ред.) Численный анализ и математическое моделирование. М.: ОВМ АН СССР, 1990.
- [26] Кузнецов Ю.А. (ред.) Численные методы и математическое моделирование. М.: ИВМ РАН, 1992.
- [27] Кузнецов Ю.А. (ред.) Численные методы и приложения // Деп. в ВИНТИ. 1995. №. 394–В95.
- [28] Лебедев В.И. Методы композиции. М.: ОВМ АН СССР, 1986.
- [29] Лифанов И.К., Тыртышников Е.Е. Теплицевы матрицы и сингулярные интегральные уравнения // Вычисл. процессы и системы. Вып.7. М.: Наука, 1990. С. 94–278.
- [30] Марчук Г.И., Кузнецов Ю.А. Итерационные методы и квадратичные функционалы // Методы вычислительной математики. Новосибирск: Наука, 1975. С. 4–143.
- [31] Мацюкин А.М. Методы фиктивных компонент и альтернирования по подпространствам // Диссертация д.ф-м.н. Новосибирск: ВЦ СО АН СССР, 1988.
- [32] Минцис М.Я., Поляков П.В., Сиразутдинов Г.А. Электрометаллургия алюминия. Новосибирск: Наука, 2001.
- [33] Нечепуренко Ю.М. Быстрые численно устойчивые алгоритмы для широкого класса линейных дискретных преобразований. Препринт № 92. М.: ОВМ РАН, 1985.
- [34] Ольшанский М.А. Анализ многосеточного метода для уравнений конвекции-диффузии с краевыми условиями Дирихле // ЖВМ и МФ. 2004. Т. 44. С. 1462–1491.
- [35] Оселедец И.В. Применение разделенных разностей и *B*-сплайнов для построения быстрых дискретных преобразований вейвлетовского типа на неравномерных сетках // Матем. заметки. 2004 (в печати).
- [36] Оселедец И.В., Тыртышников Е.Е. Приближенное обращение матриц при решении гиперсингулярного интегрального уравнения // ЖВМ и МФ. 2004 (в печати).
- [37] Писсанецки С. Технология разреженных матриц. М.: Мир, 1988.
- [38] Попов Ю.П., Самарский А.А. Вычислительный эксперимент. Новое в жизни, науке, технике // Сер. Мат. кибернетика, вып. 11. М.: Знание, 1983.
- [39] Тыртышников Е.Е. Теплицевы матрицы, некоторые их аналоги и приложения. М.: ОВМ РАН, 1989.
- [40] Тыртышников Е.Е. Краткий курс численного анализа. М.: ВИНТИ, 1994.
- [41] Тыртышников Е.Е. Параллельные алгоритмы для обобщенно-теплицевых систем // ЖВМ и МФ. 1996. Т. 36, № 6. С. 5–19.
- [42] Тыртышников Е.Е. (ред.) Матричные методы и вычисления. М.: ИВМ РАН, 1999.
- [43] Тыртышников Е.Е. Некоторые применения матричного признака равномерности // Матем. сб. 2001. Т. 192, № 12. С. 145–156.

- [44] *Тыртышников Е.Е.* Тензорные аппроксимации матриц, порожденных асимптотически гладкими функциями // Матем. сб. 2003. Т. 194, № 6. С. 147–160.
- [45] *Чижонков Е.В.* Релаксационные методы решения седловых задач. М.: ИВМ РАН, 2002.
- [46] *Шайдуров В.В.* Многосеточные методы конечных элементов. М.: Наука, 1989.
- [47] *Яценко Н.Н., Карначук В.И., Коновалов А.Н.* Проблемы математической технологии // Численные методы механики сплошной среды. 1977. Т. 8, № 3. С. 129–157.
- [48] *Яценко Н.Н., Преображенский Н.Г., Разумовский О.С.* Методологические проблемы математической физики. Новосибирск: Наука, 1986.
- [49] *Andreeva M.Yu, P'in V.P., Itskovich E.A.* Two solvers for nonsymmetric SLAE, Bull. of the NCC // Num. Anal. 2003. V. 12. P. 1–16.
- [50] *Axelsson O.* Iterative Solution Methods. N.Y.: Cambridge Univ. Press, 1996.
- [51] *Banegas A.* Fast Poisson solvers for problems with sparsity // Math. Comp. 1978. V. 32. P. 441–446.
- [52] *Bespalov A., Finogenov S., Kuznetsov Yu., Lipnikov K., Supalov A.* Generation of three-dimensional locally fitted meshes. Algorithms and software // Tech. Report Jyvääskylä University. Finland. 1993.
- [53] *Beylkin G., Mohlenkamp M.M.* Numerical operator calculus in higher dimensions // Proc. Natl. Acad. Sci. USA. 2002. V. 99. P. 10246–10251.
- [54] *Bramble J., Pasciak J., Xu J.* Parallel multilevel preconditioners // Math. Comp. 1990. V. 55. P. 1–22.
- [55] *Brandt A.* Multilevel computations of integral transforms and particle interactions with oscillatory kernels // Computer Physics Communications. 1991. V. 65. P. 24–38.
- [56] *Chan R.H., Ng M.* Conjugate gradient method for Toeplitz systems // SIAM Review. 1996. V. 38. P. 427–482.
- [57] *Chugunov V., Svyatski D., Tyrtyshnikov E., Vassilevski Yu.* Parallel iterative multilevel solution of mixed finite element systems for scalar equations // Concurrency Computat. 2005 (в печати).
- [58] *Cleary A., Falgout R., Henson V., Jones J., Manteuffel T., McCormick S., Miranda J., Ruge J.* Robustness and scalability of algebraic multigrid // SIAM J. Sci. Comp. 2000. V. 21. P. 1886–1908.
- [59] *Coifman R., Rokhlin V., Wandzura S.* The fast multipole method for the wave equation: a pedestrian prescription // IEEE Antennas Propagat. Mag. 1993. V. 35, № 3. P. 7–12.
- [60] *Davis T., Duff I.* An unsymmetric-pattern multifrontal method for sparse LU factorization // SIMAX. 1997. V. 18. P. 140–158.
- [61] *Daubechies I.* Ten lectures on wavelets. Philadelphia: SIAM, 1992.
- [62] *Dyadechko V., Iliash Yu., Vassilevski Yu.* Structuring preconditioners for unstructured meshes // Russ. J. Numer. Anal. Math. Modelling. 1996. V. 11. P. 139–154.
- [63] *Finogenov S., Kuznetsov Yu.* Two-stage fictitious components method for solving the Dirichlet boundary value problem // Sov. J. Numer. Anal. Math. Modelling. 1988. V. 3. P. 301–323.

- [64] *Ford J.M., Tyrtysnikov E.E.* Combining Kronecker product approximation with discrete wavelet transforms to solve dense, function-related systems // *SIAM J. Sci. Comp.* 2003. V. 25, № 3. P. 961–981.
- [65] *Ford J.M., Oseledets I.V., Tyrtysnikov E.E.* Matrix approximations and solvers using tensor products and non-standard wavelet transforms related to irregular grids // *Rus. J. Numer. Anal. and Math. Modelling.* 2004. V. 19, № 2. P. 185–204.
- [66] *Garbey M., Kuznetsov Yu., Vassilevski Yu.* Parallel Schwarz method for a convection-diffusion problem // *SIAM J. Sci. Comp.* 2000. V. 22. P. 891–916.
- [67] *Globisch G., Nepomnyaschikh S.* The hierarchical preconditioning on unstructured grids // *Computing.* 1998. V. 61. P. 307–330.
- [68] *Goreinov S.A., Tyrtysnikov E.E.* The maximal-volume concept in approximation by low-rank matrices // *Contemporary Mathematics.* 2001. V. 208. P. 47–51.
- [69] *Goreinov S.A., Tyrtysnikov E.E., Yeregin A.Yu.* Matrix-free iteration solution strategies for large dense linear systems // *Numer. Linear Algebra with Appl.* 1997. V. 4, № 4. P. 273–294.
- [70] *Goreinov S.A., Tyrtysnikov E.E., Zamarashkin N.L.* A theory of pseudo-skeleton approximations // *Linear Algebra Appl.* 1997. V. 261. P. 1–21.
- [71] *Golub G.H., Van Loan Ch.F.* *Matrix Computations.* Baltimore and London: The John Hopkins Univ. Press, 1989. (Рус. пер.: Голуб Джс., Ван Лоун Ч. Матричные вычисления. М.: Мир, 1999.)
- [72] *Graham I., Hagger M.* Unstructured additive Schwarz-CG method for elliptic problems with highly discontinuous coefficients // *SIAM J. Sci. Comp.* 1999. V. 20. P. 2041–2066.
- [73] *Grauschopf T., Griebel M., Regler H.* Additive multilevel preconditioners based on bilinear interpolation, matrix dependent geometric coarsening and algebraic multigrid coarsening for second order elliptic PDEs // *Applied Numerical Mathematics.* 1997. V. 23. P. 63–96.
- [74] *Greengard L., Rokhlin V.* A fast algorithm for particle simulations // *J. Comput. Physics.* 1987. V. 73. P. 325–348.
- [75] *Gupta A.* Improved symbolic and numerical factorization for unsymmetric sparse matrices // *SIAM J. Matr. Anal. Appl.* 2002. V. 24. P. 529–552.
- [76] *Gurieva Y.L.* On one numerical solution of 3D Navier–Stokes equations // *Proc. ICCM-2004.* Novosibirsk: NCC Publ., 2004. P. 852–857.
- [77] *Gurieva Y.L., Il'in V.P.* On the finite volume technology for mixed boundary value problems // *Proc. AMCA-95.* Novosibirsk: NCC Publ., 1995. P. 650–655.
- [78] *Gurieva Y.L., Il'in V.P.* On the finite volume solution of 3D thermoelasticity problem // *Proc. ICCM-2004.* Novosibirsk: NCC Publ., 2004. P. 858–863.
- [79] *Hackbusch W.* *Iterative Solution of Large Sparse Systems of Equations.* N.Y.: Springer-Verlag, 1994.
- [80] *Hackbusch W., Khoromskij B.N., Tyrtysnikov E.E.* Hierarchical Kronecker tensor-product approximations // *Max-Planck-Institut für Mathematik in den Naturwissenschaften.* Preprint № 35. Leipzig, 2003.
- [81] *Hackbusch W., Nowak Z.P.* On the fast matrix multiplication in the boundary elements method by panel clustering // *Numer. Math.* 1989. V. 54, № 4. P. 463–491.

- [82] *Hakopian Yu., Kuznetsov Yu.* Algebraic multigrid substructuring preconditioners // *Sov. J. Numer. Anal. Math. Modelling.* 1991. V. 6. P. 453–383.
- [83] *Harten A.* Multiresolution representation and numerical algorithms: a brief review // *ICASE Report.* October 1994. P. 94–59.
- [84] *Heinig G., Rost K.* Algebraic Methods for Toeplitz-like Matrices and Operators. Berlin: Akademie-Verlag, 1984.
- [85] *Hoppe R., Iliash Yu., Kuznetsov Yu., Vassilevski Yu., Wohlmuth B.* Analysis and parallel implementation of adaptive mortar element methods // *East-West J. Numer. Math.* 1998. V.6. P. 223–248.
- [86] *Iliash Yu., Kuznetsov Yu., Vassilevski Yu.* On application of strengthened AMG for partially unstructured meshes to unsteady fully potential flow problem with moving boundaries // *Experimentation, Modelling and Computation in Flow, Turbulence and Combustion.* 1997. V. 2. P. 71–86.
- [87] *Il'in V.P., Itskovich E.A.* Two explicit incomplete factorization methods // *Bullet. of NCC. Num. Anal.* 2002. № 11. P. 51–60.
- [88] *Il'in V.P., Laevsky K.Yu.* Generalized compensation principle in incomplete factorization methods // *Rus. J. Num. Anal. Math. Mod.* 1997. V. 12, № 5. P. 399–420.
- [89] *Il'in V.P., Laevsky Yu.M.* On high-order compact schemes in the finite element method // *Rus. J. Num. Anal. Math. Mod.* 2002. V. 17, № 1. P. 25–39.
- [90] *Kaporin I.* High quality preconditioning of a general symmetric positive definite matrix based on its $U^T U + U^T R + R^T U$ -decomposition // *Numer. Linear Algebra Appl.* 1998. V. 5. P. 483–509.
- [91] *Kobelkov G.* Fictitious domain method and the solution of elliptic equations with highly varying coefficients // *Sov. J. Numer. Anal. Math. Modelling.* 1987. V. 2. P. 407–418.
- [92] *Kuznetsov Yu.* Algebraic multigrid domain decomposition method // *Sov. J. Numer. Anal. Math. Modelling.* 1989. V. 4. P. 351–380.
- [93] *Kuznetsov Yu.* New algorithms for approximate realization of implicit difference schemes // *Sov. J. Numer. Anal. Math. Modelling.* V.3. P. 99–114.
- [94] *Lacroix S., Vassilevski Yu., Wheeler M., Wheeler J.* Iterative solution methods for modeling multiphase flow in porous media fully implicitly // *SIAM J. Sci. Comp.* 2003. V. 25. P. 905–926.
- [95] *Lipnikov K., Vassilevski Yu.* Parallel adaptive solution of the Stokes and Oseen problems on unstructured 3D meshes // *Proceedings of Int. Conf. ParCFD 2003.* Berlin: Elsevier, 2004. P. 153–162.
- [96] *Lipnikov K., Vassilevski Yu.* Parallel adaptive solution of 3D boundary value problems by Hessian recovery // *Comp. Methods Appl. Mech. Engrn.* 2003. V. 192. P. 1495–1513.
- [97] *Marchuk G.I., Ershov A.P.* Man-machine interaction in solving a certain class of differential equations // *Proceed. of the IFIP Congress.* N.Y., 1965. V. 2. P. 550–551.
- [98] *Marchuk G.I., Il'in V.P.* Parallel computations in grid methods physics problems // *Proceed. of the IFIP Congress.* Tokyo, 1980. P. 671–676.
- [99] *Marchuk G.I., Kuznetsov Yu.I., Matsokin A.M.* Fictitious domain and domain decomposition methods // *Sov. J. Numer. Anal. Math. Modelling.* 1986. V. 1. P. 3–35.

- [100] *Myagchilov S.V., Tyrtysnikov E.E.* A fast matrix-vector multiplier in discrete vortex method // Rus. J. Numer. Anal. Math. Modelling. 1992. V. 7, № 4. P. 325–342.
- [101] *Nechepurenko Yu.M., Tyrtysnikov E.E.* Multi-dimensional discrete Fourier transform and its optimization // SNAMM. 1990. V. 5, № 3. P. 241–250.
- [102] *Nepomnyaschikh S.* Mesh theorems of traces, normalizations of function traces and their inversion // Sov. J. Numer. Anal. Math. Modelling. 1991. V. 6. P. 223–242.
- [103] *Nepomnyaschikh S.* Application of Domain Decomposition to Elliptic Problems on with Discontinuous Coefficients // 4th Int. Symp. on DDM for PDEs. (Eds. R.Glowinski et al.). Philadelphia: SIAM, 1991. P. 242–251.
- [104] *Nepomnyaschikh S.* Fictitious space method on unstructured meshes // East-West J. Numer. Math. 1995. V. 3. P. 71–79.
- [105] *Quarteroni A., Valli A.* Domain Decomposition Methods for Partial Differential Equations. Oxford: Science Publications, 1999.
- [106] *Rjasanow S., Bebendorff M., Tyrtysnikov E.* Approximation using diagonal-plus-skeleton matrices // Mathematical Aspects of Boundary Element Methods. N.Y.: Chapman&Hall/CRC, 1999. P. 45–52.
- [107] *Rokhlin V.* Rapid solution of integral equations of classical potential theory // J. Comput. Physics. 1985. V. 60. P. 187–207.
- [108] *Rossi T., Toivanen J.* A nonstandard cyclic reduction method, its variants and stability // SIAM J. Matrix Anal. Appl. 1999. V. 20. P. 628–645.
- [109] *Saad Y.* Iterative Methods for Sparse Linear Systems. Second Edition. N.Y.: SIAM, 2003.
- [110] *Serra Capizzano S., Tyrtysnikov E.* Any circulant-like preconditioner for multilevel Toeplitz matrices is not superlinear // SIAM J. Matrix Analysis and Appl. 1999. V. 21, № 2. P. 431–439.
- [111] *Serra Capizzano S., Tyrtysnikov E.* How to prove that a preconditioner cannot be superlinear // Math. Comp. 2003. V. 72, № 243. P. 1305–1316.
- [112] *Smith B., Bjorstad P., Gropp W.* Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations. Cambridge: University Press, 1996.
- [113] *Strang G.* Wavelets and dilation equations: a brief introduction // SIAM Review. 1989. V. 31. P. 614–627.
- [114] *Strang G.* Wavelet transforms versus Fourier transforms // Bulletin of the Amer. Math. Soc. 1993. V. 28, № 2. P. 288–305.
- [115] *Strela V.V., Tyrtysnikov E.E.* Which circulant preconditioner is better? // Math. Comp. 1996. V. 65, № 213. P. 137–150.
- [116] *Stüben K.* Algebraic multigrid (AMG): experiences and comparisons // Appl. Math. Comput. 1983. V. 13. P. 419–452.
- [117] *Sun X., Pitsianis N.P.* A matrix version of the fast multipole method // SIAM Review. 2001. V. 43, № 2. P. 289–300.
- [118] *Tyrtysnikov E.E.* Cauchy-Toeplitz matrices and some applications // Linear Algebra Appl. 1991. V. 149. P. 1–18.
- [119] *Tyrtysnikov E.E.* Optimal and superoptimal circulant preconditioners // SIAM J. Matrix Analysis and Appl. 1992. V. 13, № 2. P. 459–473.
- [120] *Tyrtysnikov E.E.* A fast and parallel inertia finder for Toeplitz expanded matrices // East-West J. Numer. Math. 1995. V. 4. P. 301–316.

- [121] *Tyrtysnikov E.E.* Circulant preconditioners with unbounded inverses // *Linear Algebra Appl.* 1995. V. 216. P. 1–23.
- [122] *Tyrtysnikov E.E.* A unifying approach to some old and new theorems on distribution and clustering // *Linear Algebra Appl.* 1996. V. 232. P. 1–43.
- [123] *Tyrtysnikov E.E.* Mosaic-skeleton approximations // *Calcolo.* 1996. V. 33, № 1–2. P. 47–57.
- [124] *Tyrtysnikov E.E.* *A Brief Introduction to Numerical Analysis.* Boston: Birkhauser, 1997.
- [125] *Tyrtysnikov E.E.* Mosaic ranks for weakly semiseparable matrices // *Notes on Numerical Fluid Mechanics.* 2000. V. 73. P. 36–41.
- [126] *Tyrtysnikov E.E.* Incomplete cross approximation in the mosaic-skeleton method // *Computing.* 2000. V. 64, № 4. P. 367–380.
- [127] *Tyrtysnikov E.E.* Kronecker-product approximations for some function-related matrices // *Linear Algebra Appl.* 2004. V. 379. P. 423–437.
- [128] *Tyrtysnikov E.E., Chan R.* Spectral equivalence and proper clusters for boundary element method matrices // *Int. J. Numer. Meth. Engrn.* 2000. V. 49. P. 1211–1224.
- [129] *Tyrtysnikov E.E., Vassilevski Yu.* A mosaic preconditioner for a dual schur complement // *Numerical Mathematics and Advanced Applications. Proceedings of ENUMATH 2001.* Milano: Springer-Verlag, 2003. P. 867–880.
- [130] *Tyrtysnikov E.E., Vulkov L. et al. (eds.)* Mosaic ranks and skeletons. *Lecture Notes in Computer Science 1196 / Numerical Analysis and Its Applications // Proceedings of WNAA-96.* N.Y.: Springer-Verlag, 1996. P. 505–516.
- [131] *Tyrtysnikov E.E., Zamarashkin N.L., Yeremin A.Yu.* Clusters, preconditioning, convergence // *Linear Algebra Appl.* 1997. V. 263. P. 25–48.
- [132] *Vassilevski Yu.* On application of global coarse space for two-level domain decomposition method in the case of a large number of subdomains // *Russ. J. Numer. Anal. Math. Modelling.* 1993. V. 8. P. 59–82.
- [133] *Vassilevski Yu.* A hybrid domain decomposition method based on aggregation // *Numer. Linear Algebra Appl.* 2004. V. 11. P. 327–341.

Интернет-адреса программных комплексов

- [134] ALBERT: <http://www.mathematik.uni-freiburg.de/IAM/Research/projectsdz/albert/>
- [135] ANSYS: <http://www.ansys.com/>
- [136] DOUG: <http://www.maths.bath.ac.uk/~parsoft/doug>
- [137] FEMLAB: <http://www.comsol.com/products/femlab/>
- [138] MATLAB: <http://www.mathworks.com/products/matlab/>
- [139] Metis: <http://www.netlib.org/linalg/metis-4.0.tar.gz>
- [140] PETSc: <http://www-unix.mcs.anl.gov/petsc/petsc-2>
- [141] PLTMG: <http://www.scicomp.ucsd.edu/~reb/>
- [142] UG: <http://cox.iwr.uni-heidelberg.de/~ug/>

Стохастическое моделирование и метод Монте-Карло

*Г.А. Михайлов, Б.А. Каргин, С.М. Пригарин,
В.С. Антюфеев, В.А. Огородников, К.К. Сабельфельд,
С.С. Артемьев, А.В. Войтишек*

Содержание

Введение	150
1. Весовые методы Монте-Карло	151
1.1. Весовые оценки	151
1.2. Модификация базового пространства	154
1.3. Оптимизация моделирования по части переменных	156
1.4. Оптимизация моделирования процесса переноса	157
1.5. Векторные оценки для треугольных систем интегральных уравнений	159
1.6. Вычисление параметрических производных и собственных чисел	161
1.7. Примеры использования оценок параметрических производных	163
2. Статистическое моделирование в оптике атмосферы и океана	164
2.1. Задачи переноса оптического излучения в сферической аэрозольной атмосфере	164
2.2. Задачи лазерного зондирования рассеивающих сред	165
2.3. Стохастические задачи теории переноса	166
2.4. Моделирование радиационного поля в плоской облачной атмосфере	167
2.5. Перенос оптического излучения в системе океан–атмосфера	173
2.6. Решение обратных задач атмосферной оптики с помощью метода Монте-Карло	174
	149

3. Моделирование случайных полей в связи с решением задач статистической метеорологии	179
3.1. Моделирование гауссовских последовательностей и полей	179
3.2. Моделирование негауссовских процессов и полей	182
3.3. Приложение в метеорологии	184
4. Решение стохастических задач математической физики	185
4.1. Методы случайного блуждания для решения многомерных краевых задач математической физики	186
4.2. Эйлеровы и лагранжевы стохастические модели переноса частиц в турбулентных потоках и пористых средах	195
5. Методы Монте-Карло для численного решения СДУ и в финансовой математике	199
5.1. Численные методы решения СДУ	200
5.2. Математические модели цен финансовых инструментов	204
6. Дискретно-стохастические численные методы	207
6.1. Смешанные методы моделирования случайных величин	207
6.2. Дискретно-стохастические схемы численного интегрирования	209
6.3. Функциональные оценки метода Монте-Карло	210
6.4. Приложения теории дискретно-стохастических методов	213
Литература	213

Введение

Статья посвящена разработке стохастических вычислительных моделей и соответствующих алгоритмов метода Монте-Карло для решения задач математической физики, а также индустриальной и финансовой математики. Разделы текста объединены общей методологией интегральных и вероятностных представлений, которые позволяют строить статистические оценки решения с конечной дисперсией. Представленные подходы особенно эффективны для решения задач со случайными функциональными параметрами, в связи с чем разработано численное моделирование случайных процессов и полей. Тексты разделов подготовлены авторами согласно перечню: 1. Михайлов Г.А.; 2. Каргин Б.А.; Пригарин С.М.; Антюфеев В.С.; 3. Огородников В.А.; 4. Сабельфельд К.К.; 5. Артемьев С.С.; 6. Войтишек А.В.

1. Весовые методы Монте-Карло

Статистическое моделирование наиболее часто используется для решения задач физики и техники, в основе которых лежат вероятностные модели, связанные с некоторыми цепями Маркова. В принципе такие задачи можно решать непосредственно, численно моделируя траектории этих цепей. На основе исходного феноменологического описания проблемы можно даже строить весовые модификации алгоритма, домножая вспомогательный "вес" после каждого элементарного перехода на отношения соответствующих (может быть, обобщенных) плотностей исходного и моделируемого распределений. Однако для детального изучения и построения различных модификаций весовых алгоритмов необходимо использовать интегральные уравнения второго рода, ядра которых совпадают с плотностями перехода базовых цепей Маркова. В частности, на основе таких уравнений можно разрабатывать "ценностные" алгоритмы с малыми вероятностными погрешностями, что особенно важно для оценок малых вероятностей (см. 1.2–1.4). С помощью интегральных уравнений и представлений разработаны теория и алгоритмы оценки параметрических производных и градиентов решений рассматриваемых задач (см. 1.6–1.7). В последнее время было выяснено, что для построения весовых модификаций может быть особенно эффективным увеличение размерности фазового пространства путем включения в число фазовых координат моделируемых вспомогательных случайных величин. Соответствующая факторизация ядра базового интегрального уравнения и вспомогательной плотности перехода иногда дает требуемую весовую модификацию (см. 1.2).

1.1. Весовые оценки

Математическая модель ряда прикладных задач строится на основе рассмотрения некоторого скачкообразного, обрывающегося с вероятностью единица однородного марковского процесса (см., например, [14]). При этом траектория процесса вполне определяется ее состояниями в моменты скачков, т.е. фактически можно рассматривать обрывающуюся однородную цепь Маркова с заданной переходной функцией $P(x, S)$, где $x \in X$, X — m -мерное евклидово пространство, $S \subset X$ — измеримое по Лебегу множество. Для построения весовых алгоритмов моделирования целесообразно рассматривать соответствующую условной мере $P(x', S)$ обобщенную субстохастическую плотность перехода $k(x', x)$. Обобщенная плотность распределения $k(x', x)$ определяется $\forall x' \in X$ равенством

$$\int_X h(x)P(x', dx) = \int_X k(x', x)h(x) dx \quad \forall h \in C_0(X),$$

где $C_0(X)$ — множество непрерывных ограниченных функций. Отметим, что здесь и далее рассматриваются и ненормированные (т.е. не обязательно вероятностные) распределения.

Настоящая работа ориентирована, в частности, на приложения в теории переноса частиц, в которой кроме измеримых плотностей, соответствующих абсолютно непрерывным распределениям, используются также "дельта-функции", означающие интегрирование по некоторым многообразиям меньшей сравнительно с m размерности (см., например, [14], раздел 2.1.1). Использовать обобщенные плотности (вместо интегрирования по соответствующим мерам) в теории статистического моделирования предложил Н.Н. Ченцов [54] в связи с тем, что такой подход упрощает построение и реализацию модификаций моделирования. Это важно особенно в тех случаях, когда вводятся дополнительные фазовые переменные, причем базовые переменные — координаты и скорости — связаны с дополнительными функционально, т.е. их условные распределения определяются соответствующими "дельта-функциями".

Предполагается, что

$$\int k(x', x) dx = q(x') \leq 1 - \delta < 1.$$

Величина $q(x')$ имеет смысл вероятности необрыва траектории в заданной точке x' . Вследствие последнего неравенства цепь обрывается с вероятностью единица, и среднее число состояний конечно.

Рассматривается однородная обрывающаяся цепь Маркова x_0, x_1, \dots, x_N , определяемая плотностью $f(x)$ распределения начального состояния x_0 и субстохастической обобщенной плотностью перехода $k(x', x)$. Здесь N — номер состояния, в котором реализуется обрыв траектории (иначе, *момент остановки*). Обобщенная плотность распределения состояний, непосредственно следующих за начальным, выражается равенством

$$\varphi_1(x) = \int f(x')k(x', x) dx' = [Kf](x).$$

Следовательно, обобщенная плотность распределения фазовых состояний цепи $\varphi(x) = \sum_{n=0}^{\infty} \varphi_n(x)$, где $\varphi_n(x)$ — плотность распределения состояний номера n , представляет собой ряд Неймана для следующего интегрального уравнения 2-го рода:

$$\varphi(x) = \int_X k(x', x)\varphi(x') dx' + f(x), \quad (1.1)$$

или $\varphi = K\varphi + f$. Это уравнение можно рассматривать в пространстве $N_1(X)$ обобщенных плотностей мер ограниченной вариации, так как

ряд $\sum_{n=0}^{\infty} (K^n f, h)$, в силу условия $q(x') \leq 1 - \delta$, сходится $\forall h \in C_0(X)$. Для заданной функции $h \in C_0(X)$ рассмотрим также сопряженное уравнение $\varphi^* = K^* \varphi^* + h$, где

$$[K^* \varphi^*](x') = \int_X k(x', x) \varphi^*(x) dx.$$

Предполагается, что $K^* \in [C_0(X) \rightarrow C_0(X)]$. Справедливо представление

$$\varphi^* = \sum_{n=0}^{\infty} K^{*n} h = \sum_{n=0}^{\infty} (K^n \delta, h),$$

где $\delta(\cdot)$ — обобщенная плотность локализованного в соответствующей точке источника.

Методы Монте-Карло обычно используются для оценки линейных функционалов вида

$$I_h = (\varphi, h) = \int_X \varphi(x) h(x) dx = \sum_{n=0}^{\infty} (K^n f, h), \quad h \in C_0(X).$$

Если реализуется прямое моделирование исходной цепи Маркова, то для оценки величины I_h используется соотношение $I_h = E \left[\sum_{n=0}^N h(x_n) \right]$. Однако можно моделировать и другую, вспомогательную, цепь Маркова с плотностью перехода $p(x', x)$, взаимно регулярной с $k(x', x)$, и начальной плотностью $\pi(x)$, взаимно регулярной с $f(x)$; кроме того, предполагается, что $p(x', x) \neq 0$ и $\pi(x) \neq 0$ на носителях функций $k(x', x)$ и $f(x)$, соответственно. При выполнении этих "общих условий несмещенности" отношения $k(x', x)/p(x', x)$ и $f(x)/\pi(x)$ имеют смысл и определяются отношениями измеримых сомножителей рассматриваемых функций. Это позволяет ввести *вспомогательные веса* по формулам:

$$Q_0(x_0) = \frac{f(x_0)}{\pi(x_0)}, \quad Q_n = Q_{n-1} \frac{k(x_{n-1}, x_n)}{p(x_{n-1}, x_n)}. \quad (1.2)$$

Полагая $\xi = \sum_{n=0}^N Q_n h(x_n)$, имеем $I_h = E\xi$ (см., например, [14]). Величина ξ представляет собой весовую "оценку по столкновениям". Введем величину $\rho(K)$ по формуле

$$\rho(K) = \rho(K^*) = \inf_n \|K^{*n}\|^{1/n}.$$

Обозначим символом K_p оператор с ядром $k^2(x', x)/p(x', x)$. Известно [14, 23], что если $\rho(K_p) < 1$ и $f^2/\pi \in N_1(X)$, то $D\xi < \infty$. Известно также, что если $h(x) \geq 0$ и

$$p(x', x) = \frac{k(x', x)\varphi^*(x)}{[K^* \varphi^*](x')}, \quad \pi(x) = \frac{f(x)\varphi^*(x)}{(f, \varphi^*)},$$

то $D\xi = 0$ [36, 63]. Если вместо $\varphi^*(x)$ здесь использовать функцию $g(x) = C\varphi^*(x)[1 + \epsilon(x)]$, $|\epsilon(x)| \leq \epsilon$, то при малом ϵ величина $D\xi$ мала [36]. Соответствующие алгоритмы объединяются под названием "моделирование по ценности" [14, 36, 65]. Функцию $\varphi^*(\cdot)$ в теории весовых методов Монте-Карло принято называть "функцией ценности" в связи с вероятностным представлением, отражающим ее физический смысл [28]:

$$\varphi^*(x) = h(x) + E \sum_{n=1}^N Q_n h(x_n) \quad \text{при } x_0 \equiv x. \quad (1.3)$$

Допустимы кусочно-непрерывные функции $h(\cdot)$, если интегралы $(K^n f, h)$ определены.

1.2. Модификация базового пространства

Как правило, переход $x' \rightarrow x$ осуществляется в результате выбора совокупности значений вспомогательных случайных величин (может быть, векторных): $\mathbf{t} = (t_1, \dots, t_m) \in T$, т.е. справедливо представление:

$$k(x', x) = \int_{\tau} k_1(x', t_1) k_2((x', t_1), t_2) \dots k_m((x', t_1, \dots, t_{m-1}), t_m) \times \\ \times \delta(x - x(x', \mathbf{t})) dt.$$

Здесь $x(x', \mathbf{t})$ — функция, определяющая новое значение стандартных евклидовых координат через x' и значения вспомогательных переменных \mathbf{t} .

Обычно построение весов непосредственно по формулам (1.2) для решения достаточно сложных задач оказывается практически невозможным. Для преодоления этого затруднения целесообразно перейти к модифицированному фазовому пространству $T \times X$, точками которого являются совокупности (\mathbf{t}, x) . Соответствующее субстохастическое ядро имеет вид:

$$\mathbf{k}((\mathbf{t}', x'), (\mathbf{t}, x)) = \delta(x - x(x', \mathbf{t})) \prod_{i=1}^m k_i((x', t_1, \dots, t_{i-1}), t_i), \quad (1.4)$$

причем (x', t_1, t_0) следует рассматривать как x' . Введем обозначения: \mathbf{K} — интегральный оператор с ядром (1.4); $\mathbf{f}(\mathbf{t}, x) = f_0(\mathbf{t}) \cdot f(x)$, где $f_0(\mathbf{t})$ — некоторая плотность вероятностей в T ; $\varphi_t = \mathbf{K}\varphi_t + \mathbf{f}$; $\mathbf{I}_h = (\varphi_t, h)$; $\varphi_t^* = \mathbf{K}^* \varphi_t^* + h$.

Справедливы следующие равенства:

$$\varphi(x) = \int_T \varphi_t(\mathbf{t}, x) dt, \quad \mathbf{I}_h = I_h, \quad \varphi_t^*(\mathbf{t}, x) \equiv \varphi^*(x).$$

Первое равенство получается почленным интегрированием модифицированного интегрального уравнения по \mathbf{t} , второе — частичным интегрированием по \mathbf{t} в интеграле (φ_t, h) , а третье вытекает из вероятностного представления функции ценности (1.3).

Для построения модифицированной оценки по столкновениям ξ_t определяются вспомогательная переходная плотность $\mathbf{p}(\cdot, \cdot)$ вида (1.4), начальная плотность $\pi_t(\mathbf{t}, x) = f_0(\mathbf{t})\pi(x)$ и соответствующие веса $\{\mathbf{Q}_n\}$, которые получаются домножением на весовые множители после каждого элементарного перехода. При выполнении указанных в пункте 1.1 общих условий несмещенности (для каждого элементарного перехода) имеем:

$$I_h = \mathbb{E}\xi_t, \quad \text{где} \quad \xi_t = \sum_{n=0}^N \mathbf{Q}_n h(x_n),$$

причем $D\xi_t < +\infty$, если $\rho(\mathbf{K}_p) < 1$, где \mathbf{K}_p — интегральный оператор с ядром $\mathbf{k}^2(\cdot, \cdot)/\mathbf{p}(\cdot, \cdot)$.

Введем для $i = 1, \dots, m$ вспомогательные функции ценности

$$\begin{aligned} \varphi_i^*(x', t_1, \dots, t_i) = & \underbrace{\int \dots \int}_{m-i+1} \delta(x - x(x', \mathbf{t})) \left[\prod_{j=i+1}^m k_j((x', t_1, \dots, t_{j-1}), t_j) \right] \times \\ & \times \varphi^*(x) dt_{i+1} \dots dt_m dx, \end{aligned}$$

полагая $\prod_{m+1}^m \equiv 1$. Нетрудно заметить, что $\varphi_i^*(x', t_1, \dots, t_{i-1}, t_i)$ представляет собой условное математическое ожидание величины ξ_t при условии, что траектория цепи начинается во вспомогательной точке (x', t_1, \dots, t_i) .

Теорема 1.1. Пусть $h(\cdot) \geq 0$. Если для $i = 2, \dots, m$

$$p_i((x', t_1, \dots, t_{i-1}), t_i) = \frac{k_i((x', t_1, \dots, t_{i-1}), t_i) \varphi_i^*(x', t_1, \dots, t_{i-1}, t_i)}{\varphi_{i-1}^*(x', t_1, \dots, t_{i-2}, t_{i-1})},$$

$$p_1(x', t_1) = \frac{k_1(x', t_1) \varphi_1^*(x', t_1)}{\varphi^*(x') - h(x')} \quad \text{и} \quad \pi(x) = \frac{f(x) \varphi^*(x)}{I_h},$$

то $D\xi_t = 0$.

Отметим, что, меняя порядок выбора величин t_1, \dots, t_m , т.е. способ факторизации субстохастического ядра, можно строить различные интегральные уравнения и на их основе весовые оценки заданного функционала I_h . В частности, можно осуществлять сдвиг вдоль цепочки элементарных переходов, т.е. фиксировать фазовое состояние после выбора t_i , $i < m$. Иногда это может упростить выражение вспомогательного веса и параметрический анализ результатов.

Как будет видно из дальнейшего, использование "ценностных" плотностей вида (1.4) лишь для части вспомогательных переменных может быть малоэффективным.

1.3. Оптимизация моделирования по части переменных

Здесь будут использованы две, вообще говоря, векторные вспомогательные случайные величины t_1, t_2 , т.е. ядро имеет вид

$$k((t, x), (t', x')) = \delta(x' - x'(x, t'))k_1(x, t'_1)k_2((x, t'_1), t'_2).$$

В дальнейшем вспомогательные переменные будут рассматриваться всегда, поэтому будут использоваться более простые обозначения, принятые в теории весовых оценок (см. [14]): $\xi, \varphi, \varphi^*, J_h, K, K_p$ и т.д. Выполняются соотношения

$$\begin{aligned} \xi &= \frac{f(x_0)}{\pi(x_0)}\xi_{x_0}, \quad \xi_x = h(x) + \sum_{n=1}^N Q_n h(x_n), \\ f &\in N_1(X), \quad h \in C_b(X) = C_0(X) \cap C_+(X), \\ E\xi_x &= \varphi^*(x), \quad E\xi_x^2 = \int_X \frac{f^2(x)}{\pi(x)} E\xi_x^2 dx. \end{aligned}$$

Поэтому целесообразно рассмотреть задачу равномерной относительно $x \in X$ минимизации величины $E\xi_x^2$. Согласно теореме 1.1, использование "ценностных" плотностей вида (1.6) для всех элементарных переходов дает абсолютный минимум: $E\xi_x^2 = (\varphi^*(x))^2 \forall x \in X$. Практически такая глобальная оптимизация моделирования весьма затруднительна, поэтому важно рассмотреть возможность уменьшения величины $E\xi_x^2$ путем оптимального подбора плотности распределения части вспомогательных случайных величин, например, t_1 .

Рассмотрим цепь Маркова с субстохастической плотностью перехода:

$$p((t, x), (t', x')) = \delta(x' - x'(x, t'))p_1(x, t'_1)p_2((x, t'_1), t'_2).$$

Пусть дополнительно

$$\int_{T_1} k_1(x, t'_1) dt'_1 \equiv \int_{T_1} p_1(x, t'_1) dt'_1 \equiv 1, \quad \int_{T_2} \frac{k_2^2((x, t'_1), t'_2)}{p_2((x, t'_1), t'_2)} dt'_2 \leq q < 1.$$

Теорема 1.2. [40]. *Уравнение*

$$g(x) = h(x)[2\varphi^*(x) - h(x)] + \left\{ \int_{T_1} k_1(x, t'_1) \left[\int_{T_2} \frac{k_2^2((x, t'_1), t'_2)}{p_2((x, t'_1), t'_2)} g(x') dt'_2 \right]^{1/2} dt'_1 \right\}^2$$

имеет единственное решение, и плотность

$$p_1(x, t'_1) = \frac{k_1(x, t'_1) \left[\int_{T_2} \frac{k_2^2((x, t'_1), t'_2)}{p_2((x, t'_1), t'_2)} g(x') dt'_2 \right]^{1/2}}{\int_{T_1} k_1(x, t'_1) \left[\int_{T_2} \frac{k_2^2((x, t'_1), t'_2)}{p_2((x, t'_1), t'_2)} g(x') dt'_2 \right]^{1/2} dt'_1}$$

даст равномерно наименьшее значение $E\xi_x^2 \in C_b(X)$.

1.4. Оптимизация моделирования процесса переноса

При решении задач теории частиц переход $x' \rightarrow x$ осуществляется следующим образом [14] ($x = r, v$):

- 1) реализуется поглощение с вероятностью $\sigma_c(x)/\sigma(x)$ или номер ν типа рассеяния, причем $P(\nu = i) = \sigma_s^{(i)}(x)/\sigma(x)$;
- 2) выбирается новая скорость \mathbf{v} соответственно индикатрису w_i ;
- 3) определяется длина χ свободного пробега в направлении $\omega = \mathbf{v}/v$ соответственно субстохастической плотности

$$p_\chi(l; \mathbf{r}', \mathbf{v}) = \sigma(\mathbf{r}' + \omega l, v) \exp(-\tau_{\text{оп}}(l; \mathbf{r}', \mathbf{v})), \quad l \leq l^*(\mathbf{r}', \omega),$$

где $\tau_{\text{оп}}(l; \mathbf{r}, \mathbf{v}) = \int_0^l \sigma(\mathbf{r}' + s\omega, v) ds$ — оптическая длина отрезка $[\mathbf{r}', \mathbf{r}' + l\omega]$, $l^*(\mathbf{r}', \omega)$ — расстояние от точки \mathbf{r}' вдоль направления ω до границы среды, которую можно считать выпуклой; здесь возможен обрыв траектории вследствие вылета частицы из среды;

- 4) определяется новая точка столкновения по формуле $\mathbf{r} = \mathbf{r}' + \chi\omega$.

Заметим, что обобщенная субстохастическая плотность вероятностей величины ν

$$p_\nu(z; x) = \sum_i \delta(z - i) \sigma_s^{(i)}(x)/\sigma(x).$$

В данном случае $\mathbf{t} = (i, l)$ и

$$\mathbf{k}((\mathbf{t}', x'), (\mathbf{t}, x)) = p_\nu(i, x') w_i(\mathbf{v}' \rightarrow \mathbf{v}, \mathbf{r}') p_\chi(l; \mathbf{r}', \mathbf{v}) \times \delta(\mathbf{r} - (\mathbf{r}' + l\omega)). \quad (1.5)$$

Это новый вид ядра интегрального уравнения переноса, соответствующий обычному, указанному выше, способу построения перехода от столкновения к столкновению. Соотношение $q(x') \leq 1 - \delta$ здесь выполняется и при $\sigma_c \equiv 0$ вследствие ограниченности среды. Если необходимо учесть зависимость от времени τ , то ядро (1.5) домножается на $\delta(\tau - \tau' - l/v)$. Если заменить интегрирование по значениям ν соответствующим суммированием, то в выражении (1.5) $p_\nu(i, x)$ заменяется на $\sigma_s^{(i)}(x)/\sigma(x)$.

Для построения весовой оценки вводится вспомогательная переходная плотность типа (1.5) таким образом, что выполняются указанные в пункте 1.1 условия несмещенности и весовые множители имеют смысл для каждого элементарного перехода. Особенно важно, что в число координат включается номер типа рассеяния, так как во многих задачах теории переноса индикатрисы для разных типов рассеяния содержат сингулярности разных типов. Именно с помощью такого включения в [41] были построены весовые алгоритмы статистического моделирования эволюции ансамблей взаимодействующих частиц для приближенного решения нелинейного уравнения Больцмана. Точнее говоря, в [41] было предложено включить в число координат фазового пространства ансамбля частиц номер пары частиц, взаимодействующих в очередной, случайно выбранный момент времени. Отметим, что для построения таких алгоритмов в пространственно-неоднородном случае целесообразно использовать субстохастическое ядро вида (1.5), рассматривая в качестве X фазовое пространство координат и скоростей всех взаимодействующих частиц.

Рассмотрим теперь примеры оптимизации весовых алгоритмов статистического моделирования односкоростного процесса переноса частиц.

Для улучшения оценки прохождения излучения через оптически толстый слой целесообразно [36] использовать асимптотическое приближение к функции ценности вида $(\exp(z/L)a(\mu))$, где L — так называемая диффузионная длина; для простоты полагаем $\sigma = 1$. При этом длина пробега моделируется как при "экспоненциальном преобразовании" [14, 65] с параметром $1/L$, а рассеяние — соответственно функции $\omega(\mu', \mu)a(\mu)$. В [36] показано, что относительная погрешность оценки прохождения (а следовательно, и трудоемкость алгоритма) здесь растет пропорционально толщине слоя, а не экспоненциально, как при прямом моделировании.

Если же осуществляется модификация лишь моделирования длины пробега, то использование "ценностного" множителя $\exp(z/L)$, т.е. экспоненциального преобразования с параметром $1/L$, может быть неэффективным. Например, в случае "дельта-рассеяния" дисперсия соответствующей весовой оценки совпадает с дисперсией исходной "частотной" оценки прямого моделирования для вероятности прохождения [36] — это весьма нетривиальный факт. Поэтому для частичной оптимизации надо пытаться использовать Теорему 1.2, т.е. функцию ценности, определяемую нелинейным уравнением. Проведенные в [40] соответствующие "эвристически-асимптотические" оценки показали, что, в частности, для изотропного рассеяния асимптотически оптимальным является вариант экспоненциального преобразования с параметром $(1 - q)^{1/2}$ вместо $1/L$. Здесь $q = \sigma_s/\sigma$.

Отметим, что "ценностные" модификации диффузионных процессов рассмотрены в [25].

1.5. Векторные оценки для треугольных систем интегральных уравнений

Рассмотрим в L_∞ систему вида

$$\varphi_i = h_i + \sum_{j=1}^i K_{ij} \varphi_j, \quad i = 1, \dots, m, \quad (1.6)$$

где K_{ij} — интегральный оператор с ядром $k_{ij}(x, y)$: $[K_{ij} \varphi](x) = \int_X k_{ij}(x, y) \varphi(y) dy$. В операторной форме система (1.6) есть $\Phi = H + \mathbf{K}\Phi$.

Здесь $\|\Phi\|_{L_\infty} = \text{vrai sup}_{i,x} |\varphi_i(x)|$.

Теорема 1.3 [38]. *Если K_{ij} — ограниченные операторы, то*

$$\rho(\mathbf{K}) \leq \max_i \rho(K_{ii}) = \rho_0.$$

Если же $k_{ij}(x, y) \geq 0$, то $\rho(\mathbf{K}) = \rho_0$.

Для построения векторной оценки решения системы интегральных уравнений $\Phi = \mathbf{K}\Phi + H$ определяется обрывающаяся цепь Маркова $x = x_0, x_1, \dots, x_N$ с переходной плотностью $p(x, y)$, отличной от нуля на носителе матричного ядра $K(x, y)$, т.е. на объединении носителей ядер $k_{ij}(x, y)$. Вводится также случайный матричный вес

$$\tilde{Q}_0 = \{\delta_{ij}\}, \quad \tilde{Q}_n = \tilde{Q}_{n-1} K(x_{n-1}, x_n) / p(x_{n-1}, x_n), \quad n = 1, 2, \dots,$$

где δ_{ij} — символ Кронекера. При дополнительном условии $\rho(\mathbf{K}_1) < 1$ (где \mathbf{K}_1 — оператор, получаемый из \mathbf{K} заменой ядер k_{ij} на их модули) имеем

$$\mathbf{E} \xi_x = \Phi(x), \quad \xi_x = \sum_{n=0}^N \tilde{Q}_n H(x_n).$$

Для ковариационной матрицы $\Psi(x) = \mathbf{E}[\xi_x, \xi'_x]$ матрично-интегральное уравнение

$$\Psi(x) = [H\Psi' + \Psi H' - HH']_x + \int \frac{K(x, y)\Psi(y)K'(x, y)}{p(x, y)} dy, \quad (1.7)$$

или $\Psi = \mu + K_p \Psi$, было получено в [36] при условии, что сходится ряд Неймана для уравнения $\Psi = \mu_1 + K_{p,1} \Psi$, где матрично-интегральный оператор $K_{p,1}$ и функция μ_1 соответствуют задаче с модулями функциональных характеристик. Обозначим интегральный оператор с ядром $k_{ij}^2(x, y)/p(x, y)$ через $K_{p,ij}$.

Теорема 1.4 [38]. *Если $K_{p,ii}$ — ограниченные операторы, то для системы (1.6) имеет место неравенство*

$$\rho(\mathbf{K}_p) \leq \max_i \rho(K_{p,ii}) = \rho_0^{(p)}.$$

Если, дополнительно, $k_{ij} \geq 0$, то $\rho(\mathbf{K}_p) = \rho_0^{(p)}$.

Метод Монте-Карло обычно используется для оценки функционалов вида

$$I = (F, \Phi) = \int F'(x)\Phi(x) dx,$$

где $F'(x) = (f_1(x), \dots, f_m(x))$, причем $\|F\|_{L_1} = \sum_j \int |f_j(x)| dx < +\infty$.

Пусть точка x_0 распределена с плотностью $\pi(x)$ такой, что $\pi(x) \neq 0$, если $F'(x)\Phi(x) \neq 0$. Тогда

$$I = E \left[\frac{F'(x_0)}{\pi(x_0)} \xi_{x_0} \right] = E \sum_{n=0}^N \frac{F'(x_0)}{\pi(x_0)} \tilde{Q}'_n H(x_n) = E \sum_{n=0}^N H'(x_n) \tilde{Q}'_n \frac{F(x_0)}{\pi(x_0)}.$$

Случайный векторный вес $Q_n^{(l)} = \tilde{Q}'_n F(x_0)/\pi(x_0)$ вычисляется на основе формулы

$$Q_n^{(l)} = [K'(x_{n-1}, x_n)/p(x_{n-1}, x_n)] Q_{n-1}^{(l)}.$$

Полагая $\zeta = F'(x_0)\xi_{x_0}/\pi(x_0)$, имеем

$$E \zeta^2 = E [F'(x_0)\xi_{x_0}\xi_{x_0}^l(x_0)F(x_0)/\pi^2(x_0)].$$

Следовательно, дисперсия ζ определяется матричной функцией $\Psi(x) = E[\xi_x \xi_x^l]$, о которой речь шла выше. В силу соотношения

$$E \left[F'(x_0)\Psi(x_0)F(x_0)/\pi^2(x_0) \right] \leq \|\Psi\|_{L_\infty} \int_X \left[\left(\sum_{i=1}^m |f_i(x)| \right)^2 / \pi(x) \right] dx$$

для конечности $E \zeta^2$, в дополнение к условию теоремы 1.4, необходимо предположить, что

$$\int_X \left[\left(\sum_{i=1}^m |f_i(x)| \right)^2 / \pi(x) \right] dx < +\infty.$$

Если номер компоненты вектор-функции рассмотреть как вспомогательную координату фазового пространства (см. п. 1.2), то исходная система становится одним интегро-алгебраическим уравнением. Это позволяет использовать "моделирование по ценности" (см. п. 1.1) и упрощает построение оценок параметрических производных. С другой стороны, если используется плотность перехода, не зависящая от вспомогательной координаты, то дисперсия скалярной оценки больше дисперсии векторной оценки, которая может быть получена условным осреднением скалярной оценки для фиксированной траектории x_0, x_1, \dots, x_N .

Отметим, что на основе общей теории векторных оценок был построен критерий конечности дисперсии оценок функционалов теории переноса с учетом поляризации [38]. Другое важное приложение этой теории рассмотрено в следующем пункте.

1.6. Вычисление параметрических производных и собственных чисел

Рассмотрим в $L_\infty(X)$ интегральное уравнение с ядром, зависящим от параметра λ :

$$\varphi(x, \lambda) = \int_X k(x, y, \lambda) \varphi(y, \lambda) dy + h(x, \lambda). \quad (1.8)$$

Перепишем (1.8) в операторной форме: $\varphi = K\varphi + h$. Хотя интегральный оператор в (1.8) имеет сопряженный вид, он обозначается (для простоты) символом K . Рассмотрим задачу вычисления производных $\varphi^{(n)}(x, \lambda) = \partial^n \varphi(x, \lambda) / \partial \lambda^n$.

Введем следующее обозначение: $K^{(n)}$ — интегральный оператор с ядром $k^{(n)}(x, y, \lambda)$. Предполагается, что функции $k^{(n)}$ и $h^{(n)}$ измеримы по x . Путем формального дифференцирования уравнения (1.8) n раз по λ получаем треугольную систему интегральных уравнений

$$\varphi^{(n)} = \sum_{i=0}^n C_n^i K^{(n-i)} \varphi^{(i)} + h^{(n)}, \quad n = 0, 1, \dots, m. \quad (1.9)$$

Далее дается обоснование использования системы (1.9) для вычисления производных $\varphi^{(n)}$.

Теорема 1.5 [38]. *Предположим, что неравенства*

$$|k^{(n)}(x, y; \lambda')| \leq k_n(x, y), \quad |h^{(n)}(x, \lambda)| \leq h_0(x), \quad h_0 \in L_\infty,$$

выполняются при $\lambda - \varepsilon \leq \lambda' \leq \lambda + \varepsilon$ для некоторого $\varepsilon > 0$, интегральные операторы K_n с ядрами $k_n(x, y)$ ограничены, $n = 0, \dots, m$, и $\rho(K_0) < 1$. Тогда $\rho(\mathbf{K}) < 1$ и функции $\varphi^{(n)}$, $n = 0, 1, \dots, m$, удовлетворяют системе (1.9).

Рассмотрим методы Монте-Карло для решения построенных треугольных систем интегральных уравнений и для непосредственного вычисления производных. Теорема 1.4 показывает, что дисперсии векторных алгоритмов конечны, если спектральный радиус оператора с ядром $k^2(x, y, \lambda)/p(x, y)$ меньше единицы ($n = 1, \dots, m$), т.е. если выполняется стандартное условие конечности дисперсий скалярных оценок метода Монте-Карло для уравнения (1.8).

Пусть $\xi_x(\lambda)$ — стандартная оценка по столкновениям (см. п. 1.1) для уравнения (1.8), причем выполняются “общие условия несмещенности” (см. п. 1.1).

Теорема 1.6 [38]. *В условиях теоремы 1.5 выполняются следующие утверждения:*

$$1) \mathbf{E} \xi_x^{(m)}(\lambda) = \mathbf{E} \sum_{n=0}^m [Q_n(\lambda) h(x_n, \lambda)]^m = \varphi^{(m)}(x, \lambda);$$

- 2) если все функции в (1.8) и их необходимые производные неотрицательны, то можно использовать $\varepsilon = 0$ в условиях теоремы 1.5;
 3) если $\rho(K_p) < 1$, то $D\xi^{(m)} < +\infty$, $m = 0, 1, 2, \dots$

Теорема 1.6 легко распространяется на функционалы

$$(f, \varphi_\lambda) = E[(f(x_0)/\pi(x_0))\xi_{x_0}(\lambda)], \quad \varphi_\lambda \equiv \varphi(x, \lambda).$$

В частности, если $\rho(K_p) < 1$ и $f^2/\pi \in L_1$, то

$$D[(f(x_0)/\pi(x_0))\xi_{x_0}^{(m)}(x)] < +\infty.$$

Равенство

$$E[(f(x_0)/\pi(x_0))\xi_{x_0}(\lambda)] = (f, E\xi^{(m)}(\lambda)) = (f, \varphi_\lambda^{(m)})$$

является следствием теоремы Фубини, так как $f \in L_1$, $\varphi^{(m)} \in L_\infty$, а равенство $(f, \varphi_\lambda)^{(m)} = (f, \varphi_\lambda^{(m)})$ выполняется вследствие соотношения $\|\varphi_{\lambda'}^{(m)}\| < c < +\infty$ для $\lambda' \in [\lambda - \varepsilon, \lambda + \varepsilon]$. В условиях утверждения 2 из теоремы 1.6 значение $\varepsilon = 0$ допустимо и здесь.

Применим рассматриваемую методику для оценки итераций резольвенты. С этой целью в $L_\infty(X)$ рассмотрим уравнение (1.8) вида

$$\varphi(x, \lambda) = \lambda^{-1} \int_X k(x, y)\varphi(x, \lambda)dy + \lambda^{-1}h(x)$$

при $|\lambda| > \rho(K)$. Имеем:

$$\varphi_\lambda = [\lambda - K]^{-1}h, \quad \varphi_\lambda^{(m)} = \frac{d^m \varphi_\lambda}{d\lambda^m} = (-1)^m m! [\lambda - K]^{-(m+1)}h,$$

$$\varphi_\lambda^{(m)} = E\zeta_x^{(m)}, \quad \zeta_x^{(m)} = \sum_{n=0}^N (-1)^m \frac{(n+m)!}{n!} \lambda^{-(n+1+m)} Q_n h(x_n).$$

Несмещенность оценки $\zeta_x^{(m)}$ и соотношение $D\zeta_x^{(m)} < +\infty$ следуют из теоремы 1.6. Окончательно получаем

$$[\lambda - K]^{-(m+1)}h = (-1)^m \varphi_\lambda^{(m)} / m! = E\xi_x^{(m)},$$

$$\xi_x^{(m)} = \sum_{n=0}^N C_{n+m}^m Q_n h(x_n) \lambda^{-(n+1+m)}.$$

При достаточно общих условиях (см., например, [10]) имеем

$$\frac{([\lambda - K]^{-(m+1)}h, f)}{([\lambda - K]^{-m}h, f)} \rightarrow \frac{1}{\lambda - \lambda^*},$$

где λ^* — собственное число оператора K , ближайшее к λ . Если $f(x)$ — плотность вероятностей, то

$$([\lambda - K]^{-(m+1)}h, f) = E \xi_x^{(m)},$$

где x — случайная точка, распределенная с плотностью f .

Заметим, что алгоритм метода Монте-Карло для оценки собственного числа на основе итерации оператора предложен в [7].

1.7. Примеры использования оценок параметрических производных

В [39] решается задача оценки первого и второго моментов критических значений параметров процесса переноса в рассеивающей и размножающей частице среде со случайной плотностью. С этой целью на основе специального итерационного процесса строятся оценки производных первого и второго порядка от критических параметров по значениям плотности в различных подобластях среды. Использование "метода максимального сечения" существенно упрощает геометрические вычисления. Сформулированы достаточно широкие условия конечности дисперсий оценок. Кроме того, для оценки изучаемых моментов построена простая вероятностная модель процесса переноса в бесконечной среде с дополнительной вероятностью поглощения, зависящей от случайной плотности. Представленные в статье результаты вычислений показывают практическую эффективность построенных оценок. При выполнении этой работы было учтено, что Теорема 1.6 обеспечивает конечность дисперсий оценок кратных параметрических производных при выполнении стандартного условия конечности дисперсии оценки исходного функционала от потока частиц.

В [6] на основе весовой модификации схемы Эйлера построен и обоснован алгоритм оценки производной $\partial u / \partial \omega$ по заданному направлению ω от решения u краевой задачи для сопряженного диффузионного уравнения. Для случая изотропной диффузии разработан также альтернативный алгоритм на основе "блуждания по сферам и в шарах" для решения системы локальных интегральных уравнений относительно функций u и $\partial u / \partial \omega$, а также комбинированный алгоритм. Приведен численный пример для задачи с известным решением.

В [77] представлены результаты, полученные на основе построенного авторами вероятностного представления решения первой краевой задачи для полигармонического уравнения. Показано, что такое решение выражается параметрической производной от решения специально сконструированной задачи Дирихле для уравнения Гельмгольца. На этой основе разработаны новые алгоритмы "блуждания по решетке" и "блуждания по сферам"

для решения полигармонического уравнения. Это дало возможность построения легко реализуемого алгоритма метода Монте-Карло для оценки ковариационной функции решения бигармонического уравнения со случайными функциональными параметрами.

2. Статистическое моделирование в оптике атмосферы и океана

Начало широкомасштабной разработке и применению статистического моделирования для широкого спектра прямых и обратных задач оптики атмосферы и океана в СССР положили пионерские работы Г.И. Марчука и Г.А. Михайлова, выполненные в 1965–1967 годах [29], [30]. На их основе в последующем были разработаны эффективные вычислительные алгоритмы и пакеты прикладных программ для решения большого комплекса задач переноса излучения в атмосфере и океане с учетом неоднородности, стохастичности и реальной геометрии среды, многократного рассеяния, поляризации и рефракции излучения, поглощения в полосах атмосферных газов. Целая серия эффективных весовых алгоритмов была разработана на основе применения различных приближений для функции ценности при численном исследовании информативности оптического, в том числе лазерного, зондирования атмосферного аэрозоля, облачности и океана с искусственных спутников Земли. Были созданы численные модели радиационных полей и фонов рассеянной солнечной радиации в облачной и безоблачной атмосфере и системе океан-атмосфера с учетом стохастической структуры реальных оптических сред. С целью фильтрации атмосферных искажений при дистанционном измерении параметров среды из космоса разработан комплекс алгоритмов и программ для расчетов оптической передаточной функции атмосферы и системы океан-атмосфера.

Большая часть полученных результатов была опубликована в монографиях [31, 32, 19, 44]. В настоящем разделе дан краткий обзор основных вопросов, связанных с применением статистического моделирования в вышеуказанных задачах оптики атмосферы и океана.

2.1. Задачи переноса оптического излучения в сферической аэрозольной атмосфере

Одними из первых по инициативе Г.И. Марчука были рассмотрены задачи переноса оптического излучения в сферической аэрозольной атмосфере. Эти задачи актуальны в связи с необходимостью интерпретации оптических наблюдений земной атмосферы со спутников. Постановки по-

добных задач в наиболее общем виде сформулированы в [32]. В течение примерно 25 лет были разработаны эффективные алгоритмы моделирования полей оптического излучения в сферической аэрозольной атмосфере. К числу таких алгоритмов следует прежде всего отнести методику, основанную на использовании информации о решении сопряженного уравнения переноса, а также модификации метода зависимых испытаний, примененных для расчетов производных от характеристик поля оптического излучения по целому ряду оптических параметров среды. Были получены специальные локальные оценки характеристик рассеянного излучения в заданных точках пространства, основанные на использовании осевой симметрии системы планета-атмосфера-Солнце. Особо следует отметить методику расчетов поляризационных характеристик многократно рассеянного солнечного излучения в аэрозольной атмосфере, описываемых системой интегральных уравнений 2-го рода. На основе разработанных алгоритмов был создан пакет программ, который был использован для проведения большой серии численных экспериментов для целого ряда прямых и обратных задач спутникового оптического зондирования аэрозольной атмосферы.

2.2. Задачи лазерного зондирования рассеивающих сред

Одновременно Г.И.Марчуком и Г.А.Михайловым были поставлены задачи лазерного зондирования рассеивающих сред в приложениях, главным образом, к проблемам оптического дистанционного зондирования атмосферы и океана. Процесс переноса лазерного излучения описывается нестационарным уравнением переноса. В этих задачах, как правило, требуется определять такие тонкие характеристики поля излучения, как временные распределения интенсивности для локализованного коллимированного приемника при малой расходимости источника. Для расчетов временных распределений интенсивности лазерного излучения, рассеянного средой и приходящего в приемное устройство лидара, были разработаны специальные локальные оценки. Для оценки интенсивности отраженного излучения, проникшего в среду на большие оптические глубины и регистрируемого приемником на относительно больших временах был разработан эффективный алгоритм, основанный на использовании в качестве функции "ценности" приближенного асимптотического решения сферической проблемы Милка. Это позволило выполнить целую серию численных расчетов для задачи оптической лазерной локации.

Результаты, упомянутые в 2.1 и 2.2, были опубликованы в монографии [32]. Далее более подробно рассмотрены вопросы, связанные с решением методом Монте-Карло стохастических задач оптики атмосферы и океана, выполнявшимся в последние 10–15 лет.

2.3. Стохастические задачи теории переноса

Процесс переноса оптического излучения в приближении лучевой оптики в рассеивающих и поглощающих средах описывается интегральным уравнением (1.1), в котором $\varphi(x)$ является плотностью столкновений квантов света (фотонов) с элементами среды, $f(x)$ — плотностью распределения источников, X — фазовым пространством $R \times \Omega \times T$, $\vec{x}(\vec{r}, \vec{\omega}, t) \in X$ координат $\vec{r} \in R$, направлений $\vec{\omega} \in \Omega$ ($|\vec{\omega}| = 1$) и времени $t \in T = [0, \infty]$. Вид ядра уравнения (1.1) $k(\vec{x}', \vec{x})$ определяется типом задачи и граничными условиями. Задачи теории переноса состоят в оценке линейных функционалов I_h (см. раздел 1), где функция $h(\vec{x}) \geq 0$, так называемая "аппаратная функция", определяет вид исследуемой характеристики поля оптического излучения.

Рассматривается случай, когда один или несколько параметров задачи $\sigma = (\sigma_1, \dots, \sigma_s)$, от которых зависят ядро $k(\vec{x}', \vec{x})$ и плотность столкновений $\varphi(\vec{x})$, являются случайными функциями пространства и/или времени (случайными полями). Задача решается на основе вычисления некоторых случайных величин $\xi(\omega, \sigma)$, заданных на траекториях ω моделируемой марковской цепи и таких, что $E\xi[(\omega, \sigma)/\sigma] = I_h(\sigma)$. Траектории ω зависят от σ . Искомые функционалы определяются выражением $I = \langle I_h(\sigma) \rangle$, где символом $\langle \rangle$ обозначено математическое ожидание по распределению случайного поля σ . Такие задачи будем называть стохастическими. Кроме функционала I интерес, например, для решения обратной задачи может представлять величина $r_k = \langle I(\sigma)\sigma_k \rangle$, являющаяся коэффициентом взаимной корреляции $I(\sigma)$ и случайного поля σ_k ($k = \overline{1, s}$).

Решение методом Монте-Карло стохастических задач теории переноса основано на методе "двойной рандомизации" (см., например, [79]), который вытекает из следующих соотношений:

$$\begin{aligned} \langle I_h(\sigma) \rangle &= \langle E_\omega[\xi(\omega, \sigma)/\sigma] \rangle = E_{(\omega, \sigma)}\xi(\omega, \sigma), \\ \langle I(\sigma)\sigma_k \rangle &= \langle E_\omega[\xi(\omega, \sigma), \sigma_k] \rangle = E_{(\omega, \sigma)}[\xi(\omega, \sigma) \cdot \sigma_k]. \end{aligned} \quad (2.1)$$

Соотношения (2.1) показывают, что для вычисления величин I и r_k достаточно строить одну траекторию для каждой реализации случайного поля σ , а случайные оценки ξ и $\xi \cdot \sigma_k$ осреднять по реализациям σ . Способы построения траекторий ω в неоднородных средах, в том числе методы "максимальных сечений" и "минимальных пробегов", изложены, например, в [32], [44].

Из вышесказанного видно, что необходимым этапом решения стохастических задач теории переноса является построение реализаций случайных полей оптических параметров среды. Важную стимулирующую роль в этом направлении сыграли работы [33], [34], [35]. Они дали толчок развитию этого раздела статистического моделирования и позволили ставить и решать

качественно новые, более сложные и адекватные задачи не только атмосферной оптики, но и статистической метеорологии, диффузии частиц в турбулентных средах и т.д. Большое количество уже известных и целый ряд новых алгоритмов численного построения случайных полей отражены в монографиях [83], [88].

2.4. Моделирование радиационного поля в плоской облачной атмосфере

Статистическое моделирование процессов переноса солнечной радиации в плоской безоблачной атмосфере представляет собой сравнительно простую задачу. Применение метода Монте-Карло к решению этой задачи имеет смысл, в основном, для проведения тестовых расчетов, которые могут послужить основой для оценки точности других менее трудоемких численных и аналитических методов, которым присущи трудно контролируемые ошибки расчетов, связанные, в том числе, с идеализацией в той или иной мере реальных оптических параметров атмосферы. Примеры таких тестовых расчетов на основе метода Монте-Карло приведены, например, в [21].

Гораздо более важной задачей является моделирование процесса переноса солнечной радиации в облачной атмосфере. Важность этой задачи обусловлена двумя обстоятельствами. Во-первых, облака, всегда присутствующие над значительной частью земной поверхности, являются одним из главных факторов, определяющих перенос излучения в системе "атмосфера-подстилающая поверхность" и поступление солнечной энергии к поверхности Земли. И, во-вторых, любая облачность имеет стохастическую структуру, более или менее адекватное описание которой возможно лишь на основе статистического моделирования.

Итак, рассматривается стационарный процесс переноса частиц (фотонов) в плоском бесконечном слое $0 \leq z \leq H$. В этом случае $X = R \times \Omega$, $R = (-\infty, \infty) \times (-\infty, \infty) \times [0, H]$, $\Omega = [-1, 1] \times [0, 2\pi]$. Единичный вектор направления движения частицы $\vec{\omega}$ ($|\vec{\omega}| = 1$) задается величинами $\mu = \cos \theta$ и φ , θ — угол между вектором $\vec{\omega}$ и осью OZ , φ — азимутальный угол, т.е. угол между осью OX и вектором $\vec{\omega}_\perp$ — проекцией вектора $\vec{\omega}$ на плоскость $z = 0$; Ω_- и Ω_+ — пространства единичных векторов $\vec{\omega}$, для которых соответственно $\mu \in [-1, 0]$ и $\mu \in [0, 1]$. Воздействие рассеивающей и поглощающей среды на перенос частиц с длиной волны λ определяется заданием макроскопических сечений рассеяния $\sum_{s,\lambda}(\vec{r})$, поглощения $\sum_{a,\lambda}(\vec{r})$ и индикатрисы рассеяния $g_\lambda(\vec{r}, \tilde{\mu})$, где $\tilde{\mu} = (\vec{\omega}', \vec{\omega})$ — косинус угла между направлением движения частицы $\vec{\omega}'$ до столкновения и направлением $\vec{\omega}$ после рассеяния. Функцию $g_\lambda(\vec{r}, \tilde{\mu})$ будем задавать так, что $\int_{-1}^{+1} g_\lambda(\vec{r}, \tilde{\mu}) d\tilde{\mu} = 1$.

Далее будут использованы величины: $\sum_{\lambda}(\vec{r}) = \sum_{s,\lambda}(\vec{r}) + \sum_{a,\lambda}(\vec{r})$ — полное макроскопическое сечение ослабления, $q_{\lambda}(\vec{r}) = \sum_{s,\lambda}(\vec{r}) / \sum_{\lambda}(\vec{r})$ — альбеда однократного рассеяния (вероятность "выживания" частицы при столкновении) и $\tau_{\lambda}(\vec{r}', \vec{r}) = \int_0^l \sum_{\lambda}(\vec{r}' + \vec{\omega}s) ds$ — оптическая длина пути между точками \vec{r}' и \vec{r} , где $l = |\vec{r} - \vec{r}'|$, $\vec{\omega} = (\vec{r} - \vec{r}') / |\vec{r} - \vec{r}'|$.

Предполагается, что при попадании частиц на плоскость $z = 0$ они могут испытывать поглощение и отражение, которое будем характеризовать коэффициентом яркости $R_{\lambda}(\vec{r}_{\perp}, \vec{\omega}', \vec{\omega})$, где $\vec{\omega}'$ и $\vec{\omega}$ — направления соответственно падающего и отраженного в точке \vec{r}_{\perp} лучей.

Источником частиц (фотонов) в нашем случае с точностью, вполне достаточной для задач радиационной энергетики атмосферы, является мононаправленный бесконечно широкий поток фотонов, падающий на верхнюю границу атмосферы $z = H$. Такой источник описывается функцией $S(\lambda, \vec{r}, \vec{\omega}) = \pi S_{\lambda} \delta(z - H) \delta(\vec{\omega} - \vec{\omega}^0)$, где πS_{λ} — спектральная солнечная постоянная, $\vec{\omega}^0$ — единичный вектор направления падения солнечных лучей на границу атмосферы.

Поскольку процесс переноса излучения в атмосфере не сопровождается перераспределением энергии по длине волны, то для простоты можно рассматривать случай монохроматического излучения. Уравнение для этого случая построено в [19].

Переходя далее к стохастической задаче полагаем, что один или несколько (в любом сочетании) исходных оптических характеристик в выражениях (2.1), (2.2) $q(\vec{r})$, $\sum(\vec{r})$, $g(\vec{r}, \mu)$ и $R(\vec{r}_{\perp}, \vec{\omega}', \vec{\omega})$ являются случайными функциями пространства, которые обобщенно обозначены символом σ . Задачи, в которых существенную роль играет учет случайно-неоднородной структуры коэффициента яркости $R(\vec{r}_{\perp}, \vec{\omega}', \vec{\omega})$ земной поверхности, возникают в проблеме аэрокосмического землеведения при статистической обработке результатов дистанционного зондирования поверхности Земли, почв и растительности. Облачность в этих задачах является мешающим фактором, что требует применения статистических процедур фильтрации вносимых ею искажений. Такая же проблема возникает при исследовании возможности обнаружения очагов лесных пожаров из космоса. Учитывая все вышеизложенное в этом разделе ограничимся далее рассмотрением применения статистического моделирования для исследования радиационных процессов в изолированной стохастической облачности, пренебрегая процессами рассеяния и поглощения излучения в надоблачной и подоблачной атмосфере. Заметим, что имитационная модель изолированной облачности легко встраивается в радиационную модель аэрозольной атмосферы на основе уже внедренной в практику расчетов автоматизированной системы статистического моделирования атмосферно-оптических задач [22].

К настоящему времени разработано большое количество алгоритмов моделирования случайных полей, с успехом применяемых в атмосферно-оптических задачах [15, 48, 66, 83, 88]. В отечественных работах развиваются в основном три принципиально разных подхода к построению реализаций стохастической облачности, заложенных в работах [33, 34, 35]. Первый подход основан на моделировании точечных потоков Пальма на прямых и его частном варианте — пуассоновских потоках точек на прямых. Алгоритм моделирования трехмерного случайного поля на основе точечных потоков реализуется следующим образом. Пусть D — область трехмерного евклидова пространства, заключенная в прямоугольном параллелепипеде $[0, T_x] \times [0, T_y] \times [0, T_z]$. Получим в D случайное поле $\sigma(x, y, z, k_x, k_y, k_z)$, где $k_x(t)$, $k_y(t)$ и $k_z(t)$ — корреляционные функции вдоль осей OX , OY и OZ соответственно, следующим образом.

1. Строится поток Пальма $\tau_{x,k} = \tau_{x,k-1} + \eta_{x,k}$, $\tau_{x,0} = 0$ вдоль оси OX на отрезке $[0, T_x]$, где $\{\eta_{x,k}\}$ — независимые в совокупности неотрицательные случайные величины, распределенные с плотностью $p_{\eta_{x,1}}(t) = -k'_x(t)$, $p_{\eta_{x,i}} = -k''_x(t)/k'_x(0)$, $i = 1, 2, \dots$
2. Аналогичная процедура выполняется вдоль осей OY и OZ с плотностями, соответствующими корреляционным функциям $k_y(t)$ и $k_z(t)$.
3. Для каждой из прямоугольных ячеек $d_{i,j,k} \in D$, ограниченных плоскостями, перпендикулярными осям OX , OY и OZ в точках $\tau_{x,i}$, $\tau_{y,j}$ и $\tau_{z,k}$, выбираем свое независимое значение σ соответственно функциям одномерного распределения $F_\sigma(\cdot)$.

Пуассоновский поток точек реализуется в случае, когда в качестве корреляционных функций $k_x(t)$, $k_y(t)$, $k_z(t)$ выбираются функции $B_x(t) = \exp\{-A_x t\}$, $B_y(t) = \exp\{-A_y t\}$ и $B_z(t) = \exp\{-A_z t\}$, где A_x , A_y и A_z — средние количества точек, приходящихся на единицу длины вдоль осей OX , OY и OZ соответственно.

Данная модель случайного поля проста в реализации и эффективна в том смысле, что допускает оптимизацию процедуры построения траекторий частиц в такой среде в силу того, что ячейки d_{ijk} имеют простейшие формы прямоугольных параллелепипедов. Заметим, что данный алгоритм при $k_x(t) = k_y(t) = k_z(t)$ допускает модификацию, улучшающую реализацию поля $\sigma(\vec{r}, k)$ в смысле близости к непрерывным функциям. Такой алгоритм рассмотрен в [43], а соответствующая модель случайного поля названа "изотропной слоистой".

Свойства реализаций вышеописанных моделей подробно исследованы в [43], а ширококомасштабные расчеты на основе этих моделей радиационных характеристик сплошной стохастической облачности типа St выполнены в [40]. Примеры применения этих удобных и эффективных в реализации моделей для расчетов полей солнечной радиации в стохастической сплошной и с разрывами облачности (типа St и Sc) можно найти также

в самых последних работах (например, [48]). В [42] на основе существенного использования кусочно-слоистой структуры одномерных и трехмерных стохастических сред, получающейся при реализации описанной выше пуассоновской модели, выполнено аналитическое частичное осреднение соответствующих весовых оценок по распределению случайной плотности, что позволяет оценить изменение характеристик процесса переноса излучения при переходе от детерминированной среды к стохастической с той же средней плотностью.

Другой подход построения реализаций сплошной и разорванной облачности основан на спектральных моделях случайных полей [35]. Прежде всего отметим модели сплошной слоистообразной облачности со стохастическими верхними границами. Заметим, что нижняя граница облачности определяется уровнем конденсации и незначительно в сравнении с верхней границей варьирует в пространстве. Поэтому в моделях сплошной облачности вариациями нижней границы, как правило, пренебрегают и полагают ее плоской. Предположим, что нижняя граница облачности расположена на уровне $z = h_0$, а верхняя граница задается выражением

$$z = h_0 + h + \zeta(x, y), \quad (2.2)$$

где h — средняя толщина облака, а $\zeta(x, y)$ — некоторое случайное поле отклонений от нулевого среднего. Стохастическая структура поля $\zeta(x, y)$ определяется многими факторами: типом облачности, гидротермодинамическими процессами в атмосфере, обуславливающими процессы облакообразования и т. д. В большинстве расчетных моделей, основываясь на экспериментальных данных, полагают $\zeta(x, y)$ однородным гауссовским полем с нулевым средним, некоторой корреляционной функцией $k(|x_1 - x_2|, |y_1 - y_2|)$ и дисперсией $\sigma_\zeta^2 = k(0, 0)$; (x_1, y_1) и (x_2, y_2) — произвольные точки на плоскости $z = h_0 + h$. Изотропные поля путем изменения масштаба на одной из осей координат могут быть использованы для получения неизотропных структур. Поэтому чаще всего в прикладных расчетах применяются модели гауссовских однородных изотропных полей, вариации которых определяются методом рандомизации спектра [35]. В [66], например, для изотропного поля, спектральная мера которого обладает круговой симметрией, а корреляционная функция представима в виде

$$k(r) = \sigma_\zeta^2 \int_0^\infty J_0(\rho r) \mu(d\rho), \quad (2.3)$$

реализован следующий алгоритм:

$$\zeta(x, y) = \sigma_\zeta \sum_{k=1}^N a_k \sqrt{-2 \ln \alpha_k} \cdot \cos [(x \rho_k \cos \omega_k + y \rho_k \sin \omega_k) + 2\pi\beta_k]. \quad (2.4)$$

В (2.3) $r = (x^2 + y^2)^{1/2}$, J_0 — функция Бесселя первого рода, $\mu(d\rho)$ — радиальная спектральная мера на $[0, +\infty)$, а в (2.6) ρ_k , ω_k — полярные координаты точек спектра, a_k — коэффициенты, связанные с разбиением спектрального пространства, N — число участков разбиения, ξ_k и η_k — независимые стандартные нормальные величины, α_k и β_k — независимые случайные величины, равномерно распределенные на $[0, 1]$. Данный алгоритм может быть использован для моделирования разорванной кучевой облачности, если в качестве $h + \zeta(x, y)$ в (2.2) взять $\max(|\zeta(x, y)| - c, 0)$, где $c > 0$. Балл облачности n_0 в этом случае определяется соотношением

$$n_0 = 2 \left(1 - \Phi \left(\frac{c}{\sigma_\zeta} \right) \right), \quad (2.5)$$

где Φ — функция стандартного нормального распределения. Среднее количество облаков m_0 на единицу площади в этом случае определяется выражением

$$m_0 = 2c(2\pi\sigma_\zeta)^{-3/2} k \exp \left(-\frac{c^2}{2\sigma_\zeta^2} \right),$$

где

$$k = -\frac{\partial K(x, y)}{\partial x^2} \Big|_{x=y=0} = -\frac{\partial K(x, y)}{\partial y^2} \Big|_{x=y=0}.$$

Выбирая требуемое значение n_0 из (2.5) можно моделировать случайные реализации кучевой облачности с заданным баллом облачности. Поскольку корреляции между формой отдельных облаков, нижними и верхними границами сплошной облачности, с одной стороны, и внутренней стохастической структурой, с другой стороны, по-видимому, отсутствуют (по крайней мере, свидетельств обратного в литературе нет), то описанные выше два подхода моделирования случайных полей можно комбинировать для построения наиболее адекватных реализаций стохастической облачности.

Третий подход построения реализаций разорванной облачности, широко и успешно используемый в расчетных моделях, основан на моделировании индикаторных полей на основе пуассоновских потоков точек в пространстве [15, 48]. Подобные модели просты в реализации. В некотором объеме V пространства, ограничивающем облачное поле, по закону Пуассона

$$P_{(m)} = \frac{(\bar{m})^m}{m!} \exp(-\bar{m})$$

случайным образом выбираются m точек $\vec{r}_1, \dots, \vec{r}_m$, которые статистически независимо и равномерно распределены в V . Здесь \bar{m} — среднее значение m . Точки \vec{r}_i являются геометрическими центрами оснований некоторых фигур, аппроксимирующих форму отдельных облаков. Чаще всего

в качестве таких фигур выбираются сферы или "опрокинутые" усеченные параболоиды. Обозначим через $\chi(\vec{r})$ индикаторное поле $\prod_{i=0}^m g(\vec{r} - \vec{r}_i)$, где

$$g(\vec{r} - \vec{r}_i) = \begin{cases} 1, & r \in \tilde{A}_i \\ 0, & r \notin \tilde{A}_i, \end{cases}$$

$\tilde{A}_i = V/A_i$, A_i — множество точек, принадлежащих указанной выше выпуклой геометрической фигуре с центром в точке \vec{r}_i . В такой модели усреднение поля по числу точек m и пространственным точкам \vec{r}_i дает

$$\langle \chi(\vec{r}) \rangle = \exp \left\{ - \frac{\bar{m}}{V} V_0(\vec{r}) \right\},$$

где $V_0(\vec{r}) = V - \int_V g(\vec{r} - \vec{r}') d\vec{r}'$ — объем фигуры с центром в точке \vec{r} , принадлежащий V . Корреляционная функция в этом случае имеет вид

$$k_\chi(\vec{r}_1, \vec{r}_2) = \exp \left\{ - \frac{\bar{m}}{V} [V_0(\vec{r}_1) + V_0(\vec{r}_2) - V_s(\vec{r}_1, \vec{r}_2)] \right\},$$

где V_s — общий объем, принадлежащей V части пересечения фигур с центрами в точках \vec{r}_1 и \vec{r}_2 . Если отдельные облака имеют одинаковую форму и их проекция на горизонтальную плоскость является кругом с диаметром D , то балл облачности

$$n_0 = 1 - \exp \left(- \frac{\bar{m}}{V} h\pi \bar{D}^2 / 4 \right),$$

где $\bar{D}^2 = \int_0^{D_{\max}} D^2 f(D) dD$, $f(D)$ — плотность распределения диаметров оснований облаков, h — геометрическая толщина облачного слоя. Примеры реализаций кучевой облачности для данной модели приведены в [15], [48]. Там же даны некоторые оценки статистических характеристик построенных таким образом случайных индикаторных полей. Заметим, что данная модель позволяет учесть вертикальные корреляции облаков.

Четвертый подход к моделированию разорванных облачных полей, развиваемый в основном за рубежом, основан на использовании методов фрактальных поверхностей или n -шаговых каскадных процессов [88]. Достоинством таких моделей является возможность построения облаков заданной фрактальной размерности, получаемой из анализа реальных спутниковых данных.

На основе описанных в предыдущем разделе имитационных моделей стохастической облачности выполнено большое количество численных исследований радиационных полей. Подробный анализ этих результатов выходит за рамки настоящей работы.

2.5. Перенос оптического излучения в системе океан–атмосфера

Задача расчета поля электромагнитного излучения, отраженного и преломленного морской поверхностью, возникает при разработке и оптимизации методов дистанционного зондирования моря, а также при решении целого ряда научных и прикладных проблем оптики моря. Математическая суть этой задачи заключается в вычислении некоторых функционалов от решения уравнения (1.1), заданных на случайном поле, каковым является взволнованная морская поверхность.

В задачу, сформулированную в разделе 2.4, введем дополнительное условие — наличие в слое $0 \leq z \leq H$ некоторой поверхности S , заданной уравнением $z = h + \zeta(x, y)$, отражающей и преломляющей излучение. Предполагаем, что $\zeta(x, y)$ является случайной функцией, описывающей отклонения взволнованной морской поверхности от среднего уровня $z = H$. Относительно функции $\zeta(x, y)$ будем предполагать, что она однозначна и непрерывно дифференцируема по координатам x и y . В дальнейшем функция $\zeta(x, y)$ будет рассматриваться как одна из реализаций некоторого случайного поля (случайной поверхности).

Для описания воздействия морской поверхности на излучение примем модель, в которой отражение и преломление происходят по законам лучевой оптики на элементарной площадке, касательной к поверхности в точке воздействия. В таком случае при попадании луча, имеющего направление $\vec{\omega}'$, на поверхность S в некоторой точке $\vec{r} = (\vec{r}_\perp, h + \zeta(\vec{r}_\perp))$ с вероятностью $R(\vec{\omega}', \vec{s}(\vec{r}_\perp))$ происходит зеркальное отражение в направлении $\vec{\omega}' - 2(\vec{\omega}', \vec{s})\vec{s}$, а с вероятностью $1 - R(\vec{\omega}', \vec{s}(\vec{r}_\perp))$ происходит преломление в направлении $\nu\vec{\omega}' + \gamma\vec{s}$. Здесь $R(\vec{\omega}, \vec{s})$ — френелевский коэффициент отражения,

$$\gamma = (\vec{\omega}, \vec{s}) - \text{sign}((\vec{\omega}, \vec{s}))\sqrt{\nu^2 - 1 + (\vec{\omega}, \vec{s})^2}, \quad \nu = \begin{cases} 1/n, & (\vec{\omega}', \vec{s}) \leq 0, \\ n, & (\vec{\omega}', \vec{s}) > 0, \end{cases}$$

n — коэффициент преломления воды относительно воздуха; $\vec{s}(\vec{r}_\perp)$ — единичный вектор внешней нормали к поверхности S в точке $\vec{r} = (\vec{r}_\perp, h + \zeta(\vec{r}_\perp))$. В силу закона полного внутреннего отражения при $0 \leq (\vec{\omega}', \vec{s}) \leq \sqrt{n^2 - 1/n}$ полагаем $R(\vec{\omega}', \vec{s}) \equiv 1$.

Для сформулированной таким образом задачи в [67] было получено выражение для ядра уравнения (1.1). Предложенная там постановка исчерпывающе описывает процесс переноса в системе океан–атмосфера в том смысле, что здесь учтены все эффекты, связанные со стохастической структурой морской поверхности, а именно, переотражение излучения из одной точки поверхности в другую, затенение одних участков поверхности другими, преломление и отражение излучения поверхностью, а также полное

внутреннее отражение под поверхностью. Введем функцию

$$\tilde{\Sigma}(\vec{x}) = \tilde{\Sigma}(\vec{r}, \vec{\omega}) = \begin{cases} \Sigma(\vec{r}), & z \in (h + \zeta(\vec{r}_{\perp}), H], \\ z \in (0, h + \zeta(\vec{r}_{\perp})), \\ |(\vec{s}(\vec{r}_{\perp}, \vec{\omega}))|, & z = h + \zeta(\vec{r}_{\perp}), \vec{\omega} \in \Omega. \end{cases}$$

Тогда искомым функционал I определяется выражением

$$I = \langle I_h(\zeta) \rangle_{\zeta} = \int \left[\int_X \varphi(\vec{x}/\zeta) \tilde{\varphi}(\vec{x}) d\vec{x} \right] dp(\zeta), \quad (2.6)$$

где $\tilde{\varphi}(\vec{x}) = h(\vec{x})/\tilde{\Sigma}(\vec{x})$. Известно, что ветровое волнение морской поверхности с большой точностью описывается однородным гауссовским полем возвышений относительно среднего уровня. В [17] представлены спектральные модели гауссовской однородной поверхности, соответствующие различным спектральным участкам ветрового волнения. Представленные там алгоритмы позволяют строить наиболее адекватные численные реализации взволнованной поверхности океана. В [67] в предположении гауссовской однородной модели ветрового волнения на основе специального представления меры $P(\zeta)$ из (2.6) осуществлена приближенная замена непрерывного случайного поля $\zeta(\vec{r}_{\perp})$ конечно-мерными нормальными векторами $(\zeta(\vec{r}_{\perp,1}), \zeta(\vec{r}_{\perp,2}), \dots, \zeta(\vec{r}_{\perp,k}))$ и $(\vec{s}(\vec{r}_{\perp,1}), \vec{s}(\vec{r}_{\perp,2}), \dots, \vec{s}(\vec{r}_{\perp,k}))$, длины которых случайны и построение которых взаимосвязано с моделированием случайных траекторий фотонов. Это позволило избежать трудоемкой процедуры моделирования реализаций случайного поля $\zeta(\vec{r}_{\perp})$ и существенно уменьшить трудоемкость расчетов. Случайные оценки функционала (2.6) основаны на вычислении специальных весовых множителей. В настоящее время эта модель является, по-видимому, наиболее полным и эффективным описанием процесса переноса излучения в системе океан–атмосфера. Она вполне допускает ее сочетание со стохастическими моделями аэрозольной облачной атмосферы.

2.6. Решение обратных задач атмосферной оптики с помощью метода Монте-Карло

Наземные, атмосферные и спутниковые наблюдения дают информацию об угловом и пространственном распределении яркости рассеянного солнечного излучения. Основываясь на этой информации можно ставить и решать обратные задачи об определении оптических параметров атмосферы. Постановка обратных задач выглядит следующим образом: по данным измерения поля рассеянного солнечного излучения в атмосфере определить оптические параметры атмосферы (распределение коэффициентов рассеяния, поглощения, индикатрису рассеяния и т.д.).

Метод Монте-Карло позволяет оценивать интенсивность многократно рассеянного излучения. Поэтому им удобно пользоваться при построении итерационных алгоритмов для решения обратных задач атмосферной оптики. Рассмотрим несколько таких обратных задач.

1. Восстановление высотного хода коэффициента рассеяния по наблюдениям яркости дневного горизонта Земли.

На внешнюю поверхность атмосферы, окружающей сферическую поверхность Земли, падает параллельный поток солнечного излучения. Атмосфера разбита по высоте на n слоев с постоянными значениями коэффициента рассеяния внутри них. В некоторой точке C над верхней границей атмосферы расположен приемник излучения, измеряющий яркость дневного горизонта Земли. Ставится задача: зная измеренные интенсивности I_1, I_2, \dots, I_N , восстановить значения коэффициентов рассеяния $\sigma_1, \sigma_2, \dots, \sigma_n$ во всех слоях. Значения остальных оптических параметров предполагаются известными. Математически задача сводится к решению системы из N нелинейных уравнений относительно n неизвестных

$$\begin{aligned} I_1(\sigma_1, \dots, \sigma_n) &= I_1^*, \\ I_2(\sigma_1, \dots, \sigma_n) &= I_2^*, \\ &\dots\dots\dots, \\ I_N(\sigma_1, \dots, \sigma_n) &= I_N^*. \end{aligned}$$

Эта система уравнений решается итерационным методом Ньютона: на k -ом шаге итерации решается система линейных уравнений

$$\begin{aligned} D_{11}\delta_1 + \dots + D_{1n}\delta_n &= I_1^* - I_1^{(k)}, \\ D_{21}\delta_1 + \dots + D_{2n}\delta_n &= I_2^* - I_2^{(k)}, \\ &\dots\dots\dots, \\ D_{N1}\delta_1 + \dots + D_{Nn}\delta_n &= I_N^* - I_N^{(k)}, \end{aligned} \tag{2.7}$$

где

$$\delta_n = \sigma_m^{(k)} - \sigma_m^{(k-1)}, \quad I_m^{(k)} = I_m(\sigma_1^{(k)}, \dots, \sigma_N^{(k)}), \quad D_{ij} = \partial I_i / \partial \sigma_j,$$

а $\sigma_m^{(k)}$ — k -ое приближение к искомому коэффициенту σ_m^* .

Такая постановка задачи впервые была предложена Г.И.Марчуком [28]. Алгоритм вычисления производных D_{ij} интенсивности I_i по оптическим параметрам σ_j был предложен Г.А.Михайловым [32].

2. Восстановление индикатрисы углового рассеяния света по наблюдениям яркости в альмукантарате Солнца (как пример общего итерационного метода решения обратных задач путем последовательного выделения вклада однократного рассеяния).

Для выделения индикатрисы рассеяния необходимы точные значения компоненты кратного рассеяния и ее угловой зависимости. Метод Монте-Карло позволяет проводить соответствующие вычисления [2]. Как обычно, для решения задачи необходимо использовать метод последовательных приближений.

Рассматривается модель атмосферы в виде плоского однородного слоя. Приемник излучения находится на поверхности Земли и измеряет интенсивность рассеянной солнечной радиации. Наблюдения ведутся в нескольких направлениях с одинаковыми зенитными углами $\theta = \theta_{\odot}$ (θ_{\odot} — зенитный угол солнца) и различными азимутальными углами φ (φ — угол между вертикальной плоскостью наблюдения и плоскостью вертикала Солнца, $0^{\circ} < \varphi < 180^{\circ}$). Угол однократного рассеяния в этой схеме равен углу между направлением наблюдения и направлением на Солнце. Он может изменяться в пределах от нуля до $2\theta < 180^{\circ}$. Значения остальных оптических параметров модели атмосферы — коэффициенты молекулярного, аэрозольного рассеяния, поглощения, альбедо и индикатриса рассеяния подстилающей поверхности — считаются известными.

Ставится следующая задача. Зная измеренные величины I_1^*, \dots, I_m^* , восстановить индикатрису рассеяния света. Индикатриса полностью определяется своими значениями g_1, g_2, \dots, g_n в узловых точках, поэтому математически задача сводится к определению величин g_1, g_2, \dots, g_n . Отметим, что из-за геометрии задачи $2\theta < 180^{\circ}$ и число наблюдений m (число данных) меньше, чем число неизвестных n .

Алгоритм является итерационным и основан на приближенном выделении доли однократно рассеянного излучения из наблюдаемого потока рассеянной солнечной радиации. Пусть I_1, \dots, I_m — значения интенсивности, наблюдаемые в точке приемника: $I_k = I_k(g_1, \dots, g_n)$. Соответствующая интенсивность однократно рассеянного излучения для заданной модели определяется по простой формуле: $J_k = C \cdot g_k$, где $C = \exp(-\tau) \cos \theta_{\odot} / 2\pi$, τ — оптическая толщина атмосферы в направлении наблюдения. Она одинакова для всех направлений наблюдения, так как все направления наблюдения имеют одинаковые зенитные углы, равные θ_{\odot} . Поэтому τ и C не зависят от номера соответствующего направления.

Полагая приближенно

$$I_k^* - I_k \approx J_k^* - J_k = C(g_k^* - g_k),$$

получаем итерационный процесс: на r -ом шаге итерации решается тривиальная (диагональная) недоопределенная система линейных уравнений

$$C\delta_k = I_k^* - I_k^{(k)}, \quad m = 1, \dots, n,$$

где $\delta_m = g_m^{(k)} - g_m^{(k-1)}$, $I_m^{(k)} = I_m(g_1^{(k)}, \dots, g_n^{(k)})$, а $g_m^{(k)}$ — k -ое приближение к m -ой компоненте g^* искомого вектора $g = (g_1^*, \dots, g_n^*)$. Так как $r < n$, то будут заново пересчитаны лишь r компонент из n . Оставшиеся $(n - r)$ компонент — "хвост" индикатрисы — можно найти так. Возьмем некую "среднюю" индикатрису, умножим ее компоненты с r -ой по n -ую на подходящий коэффициент, так, чтобы r -ая компонента сравнялась с r -ой компонентой, восстановленной на очередной итерации. Затем приравниваем неизвестные компоненты (с r -ой по n -ую) восстановленной индикатрисы к соответствующим компонентам этой "усредненной" индикатрисы ("пришиваем хвост"). Конечно, эти компоненты нельзя считать достоверно восстановленными. Ниже показано, как регуляризация позволяет избавиться от этого дефекта алгоритма. Далее полученный вектор $g^* = (g_1^*, \dots, g_n^*)$ нормируется согласно условию

$$\sum_{i=1}^{n-1} \frac{g_i + g_{i+1}}{2} (\mu_i - \mu_{i-1}) = 1. \quad (2.8)$$

С увеличением оптической толщины атмосферного слоя качество восстановления падает. На графике восстановленной функции появляются пики, не типичные для физической функции. Это объясняется ухудшением обусловленности задачи. Были испробованы различные способы преодолеть возникшие затруднения.

Например, можно восстанавливать индикатрису в меньшем числе узлов, а затем интерполировать ее в промежутках. Однако сглаживание функции без учета ее природы не гарантирует качества восстановления.

Рассмотрим другой способ сглаживания. Известно, что график типичной индикатрисы имеет характерную форму. Она резко возрастает при $\mu \rightarrow 1$, а в целом похожа на выпуклую (вниз) функцию. При этом графики близких между собой экземпляров индикатрис имеют примерно одинаковый вид. Воспользуемся этим. Возьмем некоторую "среднюю" индикатрису \bar{g} , а искомую индикатрису представим в виде "композиции" этой усредненной и некоторой "добавки" и "сгладим" именно эту "добавку". Однако если взять упомянутую "композицию" в виде суммы

$$g = \bar{g} + \tilde{g},$$

"добавка" \tilde{g} снова будет "нерегулярной" функцией, ее график будет иметь форму, не поддающуюся сглаживанию. Существенно то, что "композицию" следует взять в виде произведения:

$$g = \bar{g} \cdot (1 + \tilde{g}).$$

Тогда "добавка" \tilde{g} будет "регулярной" функцией, которую можно сгладить.

Разложим индикатрису в ряд Фурье по полиномам Лежандра и отбросим несколько последних членов, которые отвечают за сильные флуктуации. Такой способ дает лучшие результаты, чем обычное сглаживание. Однако возникает вопрос, чем руководствоваться, отбрасывая члены разложения.

Ниже дается ответ на этот вопрос. Здесь мы рассмотрим специальное разложение для индикатрисы (а не для "добавки"). Это разложение тесно связано с природой индикатрисы рассеяния.

Мы рассматриваем кусочно-линейную аппроксимацию для индикатрис, т.е. разложение по базису из "пиковых" кусочно-линейных функций. Количество этих функций равно числу узлов, а коэффициенты разложения равны значениям функции в узлах. Решая обратную задачу, мы должны восстановить n компонент искомого вектора $\mathbf{g} = (\mathbf{g}_1, \dots, \mathbf{g}_n)$. Регуляризация, основанная на методе выделения главных компонент [64], позволяет уменьшить это число. Мы располагаем статистикой индикатрис. Поэтому восстанавливаемый вектор можно считать случайным вектором из статистического ансамбля с известной ковариационной матрицей C . Каждая компонента g_i имеет определенную дисперсию. Если рассмотреть представление вектора g в другой системе координат, то дисперсии проекций вектора g на новые оси будут другими. Метод главных компонент состоит в поочередном выделении таких взаимно-ортогональных направлений, для которых эти дисперсии будут иметь наибольшее значение. Известно [64], что эти направления являются собственными для ковариационной матрицы C случайного вектора g . Соответствующий базис называется базисом статистически ортогональных функций. Для последовательно остающихся направлений соответствующие дисперсии будут убывать. Представим вектор g в виде суммы $g = \bar{g} + \tilde{g}$, где \bar{g} — вектор, равный математическому среднему вектора g (этот детерминированный вектор нам известен). Мы ищем вектор \tilde{g} . Пусть h_1, \dots, h_n — векторы статистически ортогонального базиса. Рассмотрим разложение вектора g в этом базисе и обратим внимание на его последние (по величине дисперсии) компоненты. Допустим, что k из них пренебрежимо малы ($k < n$). Положим их равными нулю. Учитывая ортогональность базиса, это предположение запишется в виде системы из n равенств:

$$(\tilde{g}, h_i) = 0, \quad i = n - k + 1, \dots, n. \quad (2.9)$$

Соотношения (2.9) вместе с условием нормировки (2.8) для вектора g дают систему линейных уравнений. Эти уравнения следует рассматривать как дополнительные к уравнениям, которые мы решаем на очередном шаге итерации. Совокупное число уравнений превышает число неизвестных параметров. Поэтому можно отбросить искусственное условие, по которому к индикатрисе "пришивается хвост".

3. Моделирование случайных полей в связи с решением задач статистической метеорологии

Методы численного моделирования случайных процессов и полей находят широкое применение при решении теоретических и прикладных задач в различных областях науки и техники, причем область применения этих методов и сложность решаемых на их основе задач постоянно увеличиваются. В статистической метеорологии, климатологии, океанологии, гидрологии применение этих методов давно уже стало традиционным и показало их высокую эффективность и перспективность при решении широкого класса задач, включающего задачи усвоения гидрометеорологической информации, задачи, связанные с исследованием экстремальных свойств реально наблюдаемых процессов, вероятностным прогнозированием, исследованием свойств статистических оценок, синтезом динамических и вероятностных методов описания реальных процессов, решением экологических задач и т. д. [49, 83, 84].

3.1. Моделирование гауссовских последовательностей и полей

Эффективное решение современных задач из этого класса во многом определяется качеством используемых для этих целей вероятностных моделей, высокой точностью воспроизведения в этих моделях основных свойств реально наблюдаемых процессов. Один из путей построения таких моделей основан на использовании алгоритмов "метода условных математических ожиданий" [14] для моделирования гауссовских векторов, который в сочетании с различными методами учета негауссовости позволяет моделировать широкий класс скалярных и векторных гидрометеорологических процессов и полей большой размерности с корреляционными матрицами как общего, так и специального вида, например теплицева и блочно-теплицева вида, полученных путем соответствующей обработки данных наблюдений.

Класс процессов с корреляционными матрицами теплицева и блочно-теплицева типа достаточно широк, он включает в себя стационарные скалярные последовательности с произвольной корреляционной функцией в любом фиксированном числе точек, стационарные временные последовательности евклидовых векторов в заданной точке области, комплексы евклидовых векторов на заданной сетке, которые при решении прикладных гидрометеорологических задач могут интерпретироваться, например, как векторные ряды скорости ветра, океанических течений либо векторные поля на регулярной или нерегулярной сетке; стационарные временные последовательности аффинных векторов, которые могут интерпретироваться как комплексы метеорологических или океанологических процессов (вре-

менные ряды температуры, солнечной радиации, осадков); периодически коррелированные процессы и т. д [83].

Моделирование гауссовской стационарной последовательности векторов (которая может быть интерпретирована одним из перечисленных выше способов) $\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_n$ (или вектора $\vec{\xi}_{(n)} = (\vec{\xi}_1^T, \dots, \vec{\xi}_n^T)^T$, где $\vec{\xi}_k$ — вектор размерности p) с блочно-теплицевой ковариационной матрицей

$$R_{(n)} = \begin{bmatrix} R_0 & R_1 & \dots & R_{n-1} \\ R_1^T & R_0 & \dots & R_{n-2} \\ \dots & \dots & \dots & \dots \\ R_{n-1}^T & R_{n-2}^T & \dots & R_0 \end{bmatrix} \quad (3.1)$$

осуществляется последовательно по схеме [83]

$$\begin{aligned} \vec{\xi}_1 &= C_0 \vec{\varphi}_1, \\ \vec{\xi}_2 &= \vec{B}^T[1] J_{(1)} \vec{\xi}_{(1)} + C_1 \vec{\varphi}_2, \\ &\dots, \\ \vec{\xi}_n &= \vec{B}^T[n-1] J_{(n-1)} \vec{\xi}_{(n-1)} + C_{n-1} \vec{\varphi}_n, \end{aligned} \quad (3.2)$$

где $\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n$ — независимые гауссовские векторы размерности p , такие, что $M \vec{\varphi}_k \vec{\varphi}_k^T = I_p$, $\vec{B}[k] = (B_1^T[k], \dots, B_k^T[k])^T$, $k = 1, \dots, n-1$, $B_i[k]$ — матрицы $p \times p$, а C_i — нижние треугольные матрицы $p \times p$. Матричные векторы $\vec{B}[k]$ и остаточные ковариационные матрицы $Q_k = C_k C_k^T$ для $k = 1, \dots, n-1$ определяются уравнениями вида

$$\tilde{R}_{(k)} \vec{B}[k] = \vec{R}_k, \quad R_{(k)} \vec{B}[k] = \vec{R}_k.$$

Здесь $\tilde{R}_{(k)} = J_{(k)} R_{(k)} J_{(k)}$, $J_{(k)}$ — блочная матрица, на побочной блочной диагонали которой расположены единичные матрицы $p \times p$, а все остальные блоки — нулевые, $R_k = (R_1^T, \dots, R_k^T)^T$, $\vec{R}_k = (R_1, \dots, R_k)^T$. Алгоритм вычисления $\vec{B}[k]$ и Q_k сводится к векторному алгоритму Робинсона [83]

$$\begin{aligned} B_1^T[1] &= R_1^T, \tilde{B}_1^T[1] = R_1 R_0^{-1}, Q_0 = \tilde{Q}_0 = R_0, \\ (B_1^T[k+1], \dots, B_k^T[k+1]) &= \tilde{B}^T[k] - B_{k+1}^T[k+1] \vec{B}^T[k] J_{(k)}, \\ (\tilde{B}_1^T[k+1], \dots, \tilde{B}_k^T[k+1]) &= \vec{B}^T[k] - \tilde{B}_{k+1}^T[k+1] \vec{B}^T[k] J_{(k)}, \\ B_{k+1}[k+1] &= \tilde{Q}_K^{-1} (R_{k+1} - \tilde{R}_k^T J_{(k)} \vec{B}[k]), \\ \tilde{B}_{k+1}[k+1] &= Q_k^{-1} (\tilde{R}_{k+1} - \tilde{R}_k^T J_{(k)} \vec{B}[k]), \\ Q_k &= R_0 - \tilde{R}_k^T \vec{B}[k], \tilde{Q}_k = \tilde{R}_0 - \tilde{R}_k^T \vec{B}[k], \\ C_k C_k^T &= Q_k, \quad k = 1, \dots, n-1. \end{aligned} \quad (3.3)$$

Для построения вероятностных моделей гидрометеорологических процессов и полей используются различные модификации алгоритма (3.2), (3.3), учитывающие специфику рассматриваемых процессов (стационарность, однородность и изотропность, и т.д.). В ряде случаев оказывается приемлемым упрощенное представление корреляционных матриц, например, в виде ленточных матриц или в виде прямого произведения матриц сравнительно невысокой размерности [46]. В последнем случае алгоритм (3.2), (3.3) сводится к последовательности соответствующих скалярных алгоритмов.

В ряде случаев метод условных математических ожиданий используется для построения многомерных процессов авторегрессии [83]

$$\vec{\xi}_t = B_1^T [m] \vec{\xi}_{t-1} + \dots + B_m^T \vec{\xi}_{t-m} + \vec{u}_t,$$

где \vec{u}_t — последовательность взаимно-независимых p -мерных векторов таких, что $M\vec{u}_t\vec{u}_t^T = Q_m$, а $B_1^T [m], \dots, B_m^T [m]$ — матрицы $p \times p$. При этом вектор $\vec{\xi}_1, \vec{\xi}_2, \dots, \vec{\xi}_n$, построенный с помощью алгоритма (3.2), (3.3), используется в качестве начального.

Алгоритм (3.2), (3.3) может быть также использован для построения некоторых классов нестационарных процессов, в частности в гидрометеорологических приложениях он используется для построения скалярных периодически коррелированных рядов. Например, вероятностная модель периодически коррелированных рядов векторной скорости ветра основана на представлении соответствующих рядов в виде стационарного векторного гауссова процесса $\vec{\xi}_1, \dots, \vec{\xi}_m$ с ковариационной матрицей вида (1), где $\vec{\xi}_i$ — векторы размерности $2p$, $\vec{\xi}_i = ((\vec{\xi})_{i1}^T \dots \vec{\xi}_{ip}^T)^T$, $\vec{\xi}_{ik} = (u_{ik}, v_{ik})^T$, p — период коррелированности.

В работах [16, 45, 46, 83] для гауссова случая получены алгоритмы моделирования условных реализаций процессов и полей при заданных значениях в фиксированных точках, которые являются модификацией известного алгоритма моделирования гауссовских векторов $\vec{\xi} = (\vec{\xi}_1^T, \vec{\xi}_2^T)^T$ с нулевым средним и ковариационной матрицей R , который имеет вид

$$\vec{\xi}_1 = R_{12}R_{22}^{-1}\vec{\xi}_2 + A_1\vec{\varphi}_1, \quad (3.4)$$

где $\vec{\varphi}_1$ — вектор из независимых между собой и от $\vec{\xi}_2$ стандартных нормальных величин, а нижняя треугольная матрица A_1 такая, что

$$A_1A_1^T = R_{11.2} = R_{11} - R_{12}R_{22}^{-1}R_{21},$$

где блоки R_{11} и R_{22} ковариационной матрицы R — ковариационные матрицы векторов $\vec{\xi}_1$ и $\vec{\xi}_2$, а блок R_{12} — соответствующая взаимная матрица.

Предлагаемый в работе алгоритм не требует разложения $R_{11,2}$ на произведение двух треугольных матриц и сводится к преобразованиям:

- 1а. Независимо от $\vec{\xi}$ моделируется нормальный вектор $\vec{\eta} = (\vec{\eta}_1^T, \vec{\eta}_2^T)^T$ с нулевым средним и корреляционной матрицей R , причем размерности его подвекторов такие же, как и у вектора $\vec{\xi}$.
- 1б. Вектор $\vec{\xi}_1$ при фиксированном $\vec{\xi}_2$ строится в виде $\vec{\xi}_1 = R_{12}R_{22}^{-1}(\vec{\xi}_2 - \vec{\eta}_2) + \vec{\eta}_1$.

В отличие от (3.4), алгоритм (1а)–(1б) позволяет строить условные реализации стационарного процесса (либо однородного поля) в узлах регулярной сетки при фиксированных значениях ξ_2 в узлах опорной регулярной сетки рекурсивно с использованием алгоритма (3.2), (3.3) в его скалярном либо векторном варианте. Если опорная сетка нерегулярна, преобразования (1а)–(1б) реализуются приближенно с использованием методов оптимальной интерполяции.

Предложенный подход используется в [45] для построения условных пространственных полей метеоэлементов в узлах регулярной сетки при заданных значениях на нерегулярной сети метеорологических станций. С использованием этих алгоритмов построен метод оценки влияния неопределенности в начальных данных, обусловленной ограниченностью и нерегулярностью сети метеорологических станций, на результаты численного моделирования атмосферных процессов на основе гидротермодинамических моделей. Метод сводится к моделированию ансамбля начальных условных полей, решению нелинейных прогностических уравнений для каждого элемента ансамбля с последующей статистической обработкой прогностических полей и не связан с проблемой замыкания бесконечной системы уравнений для моментов распределения.

3.2. Моделирование негауссовских процессов и полей

Рассмотренные гауссовы процессы могут быть использованы в качестве основы для моделирования негауссовских гидрометеорологических процессов и полей. Для этой цели наиболее часто используют хорошо известный метод обратных функций распределения [14] в соответствии с которым для получения негауссовской величины η с функцией распределения $F(x)$ используется преобразование

$$\eta = F^{-1}(\Phi(\xi)), \quad (3.5)$$

где $\Phi(x)$ — функция одномерного нормального распределения, а ξ — гауссова величина с нулевым средним и единичной дисперсией. Для построения негауссова поля $\vec{\eta}_{(n)} = (\vec{\eta}_1^T, \dots, \vec{\eta}_n^T)^T$ со скалярными элементами η

в качестве ξ используются скалярные элементы рассмотренных выше гауссовых полей $\vec{\xi}_{(n)}$ (распределение $F(x)$ может зависеть от пространственных или временной координат поля), а скалярные элементы g корреляционной матрицы $G_{(n)}$ гауссова поля $\vec{\xi}_{(n)}$ связаны с соответствующими скалярными элементами r корреляционной матрицы $R_{(n)}$ негауссова поля $\vec{\eta}_{(n)}$ соотношением

$$r = f(g). \quad (3.6)$$

Конкретный вид этого соотношения с учетом (3.5) приведен в [14].

При использовании метода обратных функций распределения корреляционная матрица $R_{(n)}$ считается заданной. Обычно ее получают путем соответствующей обработки данных наблюдений. Необходимо решить систему нелинейных алгебраических уравнений вида (3.6) относительно элементов матрицы $G_{(n)}$. При определенных сочетаниях вида одномерных распределений $F(x)$ и значений корреляций r уравнения (3.6) могут не иметь решения. В работе [83] приведены соответствующие условия совместности одномерных распределений и корреляций. Если уравнение (3.6) при заданной функции распределения $F(x)$ и заданном r не имеет решения, то ищется приближенное решение, т.е. в качестве r выбирается значение, ближайшее к заданному, но такое, чтобы решение уравнения (3.6) существовало. В силу этого обстоятельства, а также нелинейности уравнений (3.6), полученная матрица $G_{(n)}$ в ряде случаев может оказаться отрицательно определенной, т.е. задача не имеет решения. В этих условиях приходится искать приближенное решение задачи. Один из эффективных способов нахождения приближенной положительно определенной матрицы $\hat{G}_{(n)}$, в определенном смысле близкой к $G_{(n)}$, основан на соответствующем спектральном разложении этой матрицы.

Наряду с методом обратных функций распределения для моделирования негауссовских процессов и полей при решении задач статистической метеорологии используются различные модификации алгоритма моделирования негауссовских процессов и полей на точечных потоках [36]. Например, в работе [16] рассмотрены алгоритмы кусочно-постоянной стохастической интерполяции процессов и полей дискретного аргумента, основанные на использовании регулярных точечных потоков. Используемые преобразования сохраняют основные свойства исходных процессов и полей дискретного аргумента: значения процесса и корреляций в узлах сетки, одномерные распределения в произвольной точке области, для процессов — стационарность, для полей — однородность (либо изотропность). Алгоритмы предназначены для построения вероятностных моделей временных рядов и полей метеоэлементов с использованием реальных данных на сетках большого размера, а также в тех случаях, когда необходимо моделировать значение поля в произвольной точке рассматриваемой области.

3.3. Приложение в метеорологии

В качестве примеров использования методов моделирования случайных процессов для решения задач статистической метеорологии в работе [83] рассмотрен ряд вероятностных моделей временных рядов, пространственных и пространственно-временных полей различных метеоэлементов: температуры воздуха, вектора скорости ветра, суточных сумм жидких осадков и т.д. Одним из важных приложений этих моделей является исследование экстремальных погодных условий. В частности с помощью негауссовой вероятностной модели временных рядов температуры воздуха, учитывающей суточный ход параметров распределения в виде смеси двух нормальных распределений, а также реальные корреляционные связи, были исследованы статистические свойства сильных и длительных похолоданий для некоторых гидрометеорологических станций северного полушария [47]. На основе вероятностной модели временных рядов суточных сумм жидких осадков, учитывающей специфику и статистические свойства реальных рядов осадков, построены распределения длительностей периодов с осадками и без осадков, позволяющих рассчитать вероятности длительных дождливых и засушливых периодов [83].

Основная трудность при построении стохастических многомерных моделей атмосферных процессов состоит в том, что объем имеющейся информации, как правило, оказывается недостаточным для построения реалистических моделей. В связи с этим представляется перспективным использовать численные модели гидротермодинамики атмосферы для восполнения недостающей информации и фильтрации нефизических составляющих в рамках единой динамико-вероятностной модели атмосферных процессов. Эта модель основана на вариационном согласовании [84] стохастической модели с численной математической моделью гидротермодинамики атмосферы. В основе такого согласования лежат разработанные А.В. Протасовым [84] численные алгоритмы вариационного усвоения гидрометеорологической информации с помощью математических моделей атмосферных процессов. В рамках этой модели реализации случайного поля заменяются соответствующими наиболее близкими по некоторой норме полями, которые являются решениями соответствующей системы дифференциальных уравнений, положенных в основу численной модели гидротермодинамики атмосферы.

При этом используемая модель гидротермодинамики атмосферы обеспечивает физическое содержание и полноту пространственно-временного разрешения в рамках соответствующей постановки, а стохастическая модель — соответствующую вероятностную структуру и задает ансамбль независимых реализаций рассматриваемых полей. Использование вариационного метода усвоения информации позволяет оптимизировать процесс

объединения динамических и статистических методов численного моделирования, а также динамических моделей различного масштаба и физического содержания. Таким образом, совместно с А.В. Протасовым предложен новый метод динамико-вероятностного моделирования атмосферных процессов, суть которого заключается в следующем.

На основе имеющейся реальной информации с использованием рассмотренных выше алгоритмов строится ансамбль реализации этих полей. Точность воспроизведения в модели статистических свойств реального процесса определяется объемом соответствующей информации и выбором приемлемого приближения (однородность, изотропность и т.д.), необходимого для применения алгоритмов, рассмотренных в 1.1, 1.2. Реализации полей из этого ансамбля используются как входные для задачи вариационного усвоения информации с помощью динамической модели. В результате мы получаем новый ансамбль, в котором каждая реализация удовлетворяет динамической модели, а статистическая структура полученного нового ансамбля — близка к исходной в пределах точности задачи усвоения.

С использованием этого ансамбля решен ряд практических задач статистической метеорологии, в частности предложен многошаговый метод моделирования климатического переноса пассивной примеси для локальной области, и на основе метода индикаторных функций исследуются некоторые пространственно-временные статистические свойства полученных климатических траекторий. На основе вариационного принципа предложен и численно апробирован метод численного стохастического моделирования области влияния для полей концентрации примеси на фоне климатических реализаций полей скорости ветра [49].

4. Решение стохастических задач математической физики

В настоящем разделе дан обзор результатов, полученных лабораторией стохастических задач математической физики ИВММГ СО РАН, г. Новосибирск; здесь представлены три основные области исследования: (1) — разработка методов случайного блуждания для решения многомерных краевых задач математической физики; (2) — создание эйлеровых и лагранжевых стохастических моделей и алгоритмов для решения задач переноса в стохастических полях скоростей, таких как турбулентные потоки и течения в пористой среде; (3) — моделирование процессов формирования и трансформации спектра размеров коагулирующих частиц на основе решения нелинейного уравнения Смолуховского.

Представленные исследования касаются в основном теоретических основ стохастических методов, однако приводятся также отдельные при-

меры решения некоторых прикладных задач, таких как задачи электростатики, упругости, задачи о дисперсии аэрозольных частиц в атмосфере, моделирование процессов роста капель в турбулентных потоках.

Предложен новый подход к построению стохастических методов решения многомерных краевых задач математической физики. В основе метода лежит переформулировка исходной дифференциальной задачи в систему интегральных уравнений и построение специальных рандомизированных итерационных методов для численного решения этих систем. В качестве одного из таких методов построен алгоритм случайного блуждания по специальным сферическим сеткам, а также бессеточный метод случайного блуждания по фиксированным детерминированным сферам. Данный подход позволил построить стохастические методы численного решения ряда краевых задач, которые ранее не могли быть решены в рамках традиционного вероятностного подхода: (1) — решение задачи Неймана без привлечения случайных процессов с отражением, (2) — бигармоническая задача с заданием на границе решения и нормальной производной, (3) — первая краевая задача для системы уравнений теории упругости.

4.1. Методы случайного блуждания для решения многомерных краевых задач математической физики

Здесь мы приводим обзор основных методов случайного блуждания и подробнее останавливаемся на методе случайного блуждания по фиксированным сферам. Существует три основных подхода к решению краевых задач.

I. Первый подход основан на использовании классических вероятностных представлений в виде математического ожидания в пространстве траекторий диффузионных случайных процессов. Строятся конечно-разностные аппроксимации решений соответствующих стохастических дифференциальных уравнений и затем производится усреднение по таким приближенным траекториям, которые сосредоточены во всей области поля решений [68]. Данный подход применим лишь к скалярным уравнениям второго порядка, а практически решаться могут лишь внутренние краевые задачи с условиями Дирихле. Имеются теоретические результаты по обобщению методов на случай задачи Неймана, однако моделирование диффузионных траекторий с отражением оказывается весьма дорогостоящей процедурой. Более того, и в случае задачи Дирихле эта трудность — моделирование случайного процесса при его приближении к границе — остается непростой проблемой.

II. Второй подход основан на использовании явных представлений функций Грина для некоторых стандартных областей, таких как сфера, полупространство, цилиндр и т.д. (см., например, [12, 62, 90]). Этот подход

по-существу близок первому подходу, он специально разработан в основном для решения уравнений с постоянными коэффициентами, поскольку именно для таких уравнений известны функции Грина. Здесь существуют два различных подхода к построению и обоснованию методов — один основан на интерпретации случайного блуждания как мартингала [62], второй — на использовании представления решения в виде сходящегося ряда Неймана и стандартной техники его вычисления с помощью цепей Маркова [12, 90].

III. Третий подход, впервые предложенный в [50] и подробно представленный в монографии [100], основан на использовании граничных интегральных уравнений теории потенциала. Здесь марковские процессы строятся на границе области, что и дало название методам — алгоритмы блуждания по границе. Этот класс методов также основан на построении подходящих марковских процессов, однако он существенно отличается от методов второго подхода прежде всего тем, что ряд Неймана для соответствующих граничных интегральных уравнений как правило расходится, и поэтому приходится строить специальные итерационные методы для их решения. При этом привлекается различная информация о структуре граничных интегральных уравнений; например, при построении изотропного блуждания по границе использовалась информация о границах спектра граничного интегрального оператора. Существенным является то обстоятельство, что алгоритмы блуждания по границе позволяют решать как внутренние, так и внешние краевые задачи, причем все классические задачи — с условием Дирихле, Неймана и с заданием линейных комбинаций решения и нормальной производной. При этом размерность задачи фактически уменьшается на единицу, поскольку фазовым пространством траекторий является граница области.

Методы блуждания по границе идеально приспособлены для решения задач в выпуклых областях. В случае же областей, которые существенно невыпуклы (например, область, состоящая из множества пересекающихся или непересекающихся сфер), они теряют свою эффективность в силу возрастания статистической погрешности.

В последних работах [99], [96] предлагается новый подход, который объединяет преимущества второго и третьего подходов. Идея заключается в том, чтобы, с одной стороны, эффективно использовать информацию о функции Грина для стандартных областей, как в подходе II, а с другой стороны — иметь дело с интегральной формулировкой задачи, с тем чтобы далее строить оптимальные стохастические итерационные процедуры их решения. Заметим, что этот подход, на первый взгляд близкий к традиционному методу блуждания по сферам, существенно отличается от него тем, что он основан на эквивалентной записи исходной задачи в виде системы интегральных уравнений с вполне определенным фиксированным фазовым пространством, что открывает совершенно новые возможности при постро-

ении стохастических алгоритмов для их решения. В частности эффективными оказываются нестационарные двушаговые итерационные методы с набором случайных параметров, которые в определенном смысле служат стохастическим аналогом известного метода с чебышевским набором параметров [27]. Другой вариант численного метода, оказавшийся чрезвычайно эффективным для широкого класса областей — стохастический вариант метода последовательной верхней релаксации [27].

Данный подход позволил построить стохастические методы численного решения ряда краевых задач, которые ранее не могли быть решены в рамках традиционного вероятностного подхода: (1) — решение задачи Неймана без привлечения случайных процессов с отражением, (2) — бигармоническая задача с заданием на границе решения и нормальной производной, (3) — первая краевая задача для системы уравнений теории упругости.

Отметим, что в этом списке представлены наиболее яркие примеры краевых задач, которые давно привлекают внимание специалистов по статистическому моделированию, однако предложенный метод является достаточно общим, он может быть применен к решению самых разных краевых задач как для систем уравнений, так и для уравнений высокого порядка. В частности он применим к решению задач теории упругости в напряжениях, к уравнению Стокса, системе уравнений Максвелла.

Алгоритм случайного блуждания по фиксированным сферам. Для простоты изложения мы рассмотрим вначале задачу Дирихле для уравнения Лапласа

$$\Delta u(x) = 0, \quad x \in G, \quad u(y) = \varphi, \quad y \in \Gamma = \partial G, \quad (4.1)$$

в некоторой ограниченной области G .

Решение этой задачи имеет вероятностное представление в виде математического ожидания в пространстве винеровских траекторий:

$$u(x) = \langle \varphi(X_\Gamma) \rangle, \quad (4.2)$$

где X_Γ — случайная точка первого достижения винеровским процессом X границы области Γ , при условии, что процесс стартовал в точке x . Это представление используется для построения различных вычислительных алгоритмов. Самый простой способ — построение конечно-разностной аппроксимации винеровского процесса. Для этого выбирается дискретизация по времени, и винеровский процесс аппроксимируется, например, с помощью стохастического аналога метода Эйлера. При этом затраты могут быть весьма внушительны, поскольку шаг интегрирования должен быть достаточно малым, и, соответственно, случайная длина траекторий до достижения ими границы может оказаться очень большой. Кроме того, при приближении к границе этот шаг должен становиться все меньше, и вычислительные затраты повышаются дополнительно.

Сильной стороной метода является его универсальность в том смысле, что представление (4.2) в точности сохраняется для общих эллиптических уравнений с переменными коэффициентами — в этом представлении нужно лишь винеровский процесс заменить на диффузионный, который однозначно записывается в виде системы стохастических дифференциальных уравнений Ито, которая в свою очередь легко строится с помощью коэффициентов исходной дифференциальной задачи. Моделирование же таких диффузионных процессов труда не представляет, остаются лишь те же трудности, что и в случае винеровского процесса.

Для уравнений с постоянными коэффициентами алгоритмы, основанные на таком вероятностном представлении, могут быть существенно улучшены на основе прямого моделирования точек выхода траекторий на границы стандартных областей, например, сферы. Данный класс методов известен как методы случайного блуждания по сферам. Этот метод был детально исследован в пионерской работе [81] и получил ряд обобщений на уравнения высокого порядка и системы эллиптических уравнений [50, 90]. Следует отметить, что методы исследования в [62] и в [12, 90] приводят к одним и тем же алгоритмам, однако они существенно отличаются: если в [62] ведется прямое доказательство сходимости на основе мартингалного подхода, то в [12, 90] делается формальный переход к интегральному уравнению и затем применяется стандартная техника вычисления его решения с помощью цепей Маркова. Последний подход является фактически рандомизированным вычислением ряда Неймана для интегрального уравнения, что эквивалентно применению метода простой итерации, скорость сходимости которого далеко не оптимальна, а условия применимости — весьма ограничительны. Именно это обстоятельство являлось серьезным препятствием на пути дальнейшего обобщения методов случайного блуждания и одновременно послужило мотивацией автору данного обзора для построения стохастических методов решения интегральных уравнений с помощью других итерационных методов, что и было впервые сделано в [50] и подробно представлено в монографиях [90, 100, 97]. Данный подход позволил построить алгоритмы случайного блуждания по границе, впервые представленные в той же публикации [50], которые оказались весьма эффективными для выпуклых областей. Следует отметить, что область применимости существенно расширилась — стало возможным решать все основные краевые задачи, причем как внутренние, так и внешние; кроме того, появилась возможность ускорения сходимости метода на основе информации о спектральных свойствах интегрального оператора. Некоторые примеры приложений можно найти в [90, 24, 78].

Важно отметить, что данный подход стал возможным в силу того, что исходная дифференциальная задача оказалась переформулированной в виде интегрального уравнения с фиксированным фазовым пространством (в

данном случае — граница области) с вполне определенными свойствами ядра и спектра. В связи с этим естественным было попытаться подойти с этой точки зрения и к традиционному методу блуждания по сферам. Это удалось сделать для класса областей, представляющих собой объединение шаров с непустым пересечением [99]. В результате мы получили алгоритм блуждания по детерминированным, фиксированным заранее сферам, область сходимости которого существенно расширилась, а скорость сходимости — увеличилась. В частности данный алгоритм позволил решить такие краевые задачи как задача Неймана, первая краевая задача для системы уравнений теории упругости, бигармоническая задача с условиями задания решения и его нормальной производной на границе. Для всех этих задач традиционный алгоритм блуждания по сферам расходится. Опыт расчетов показал, что новый метод хорошо работает не только для областей, состоящих из сфер (для таких областей он не имеет детерминированной погрешности, поскольку построенные оценки не имеют статистического смещения), но и для достаточно общих областей, которые могут быть покрыты системой пересекающихся шаров. При этом смещения, связанные с такой аппроксимацией, носят тот же характер, что и погрешности, возникающие при введении ϵ -границы в стандартном алгоритме блуждания по сферам, а скорость сходимости нового метода — значительно выше.

Отдельно следует остановиться на дискретном варианте этого метода. Поскольку здесь мы имеем дело с фредгольмовыми интегральными операторами, то возможно применение одной из дискретных аппроксимаций в виде системы линейных алгебраических уравнений с последующим построением дискретных случайных блужданий. Данный подход оказался весьма эффективным, он носит общий характер, и при его реализации легко строятся различные варианты итерационных методов. Для метода блуждания по фиксированным сферам особенно эффективными оказались два метода: метод последовательной верхней релаксации — для определенного класса областей, когда применимо так называемое черно-красное разбиение, — в [99] мы назвали такие области S_2 -областями; для общего случая также весьма эффективным оказался нестационарный итерационный процесс со случайным выбором параметров, являющийся в определенном смысле стохастическим аналогом чебышевского итерационного процесса.

Метод случайного блуждания по фиксированным сферам является достаточно новым, поэтому мы представим его в данном обзоре подробнее. Для простоты изложения мы начнем с плоской задачи Дирихле для уравнения Лапласа (4.1) в области, представляющей собой объединение двух пересекающихся кругов радиусов R_1 и R_2 :

$$G = K(x_0^{(1)}, R_1) \cup K(x_0^{(2)}, R_2); \quad K(x_0^{(1)}, R_1) \cap K(x_0^{(2)}, R_2) \neq \emptyset. \quad (4.3)$$

Введем обозначения: γ_1 — та часть окружности $S(x_0^{(1)}, R_1)$, что лежит во втором круге, и Γ_1 — часть окружности $S(x_0^{(1)}, R_1)$, не принадлежащая второму кругу; аналогично определяются дуги γ_2 и Γ_2 . Таким образом, граница области G состоит из Γ_1 и Γ_2 .

Регулярное решение уравнения Лапласа в любой внутренней точке x каждого из кругов представляется формулой Пуассона:

$$u(x) = \frac{R^2 - r^2}{2\pi R} \int_{S(x_0, R)} \frac{u(y)}{|x - y|^2} dS_y, \quad (4.4)$$

где $R = R_1$ в первом круге, и $R = R_2$ — во втором; $r = |x - x_0^{(i)}|$ — расстояние от точки x до центра круга, $i = 1, 2$. Пусть

$$p(y; x) = \frac{R^2 - |x - x_0^{(1)}|^2}{2\pi R} \cdot \frac{1}{|x - y|^2}.$$

Наша задача — дать интегральную формулировку задачи Дирихле для области G . Для произвольных точек $x \in \gamma_2$ и $y \in \gamma_1$ мы можем записать

$$\begin{aligned} u(x) &= \frac{R_1^2 - |x - x_0^{(1)}|^2}{2\pi R_1} \int_{S(x_0^{(1)}, R_1)} \frac{u(y)}{|x - y|^2} dS_y, \\ u(y) &= \frac{R_2^2 - |y - x_0^{(2)}|^2}{2\pi R_2} \int_{S(x_0^{(2)}, R_2)} \frac{u(x')}{|y - x'|^2} dS_{x'}. \end{aligned} \quad (4.5)$$

Введем обозначения $v_1(x) = u(x)$ для $x \in \gamma_2$, и $v_2(x) = u(x)$ для $x \in \gamma_1$. Тогда (4.5) можно записать так:

$$\begin{aligned} v_1(x) &= \int_{\gamma_1} p(y; x) v_2(y) dS_y + f_1(x), \\ v_2(y) &= \int_{\gamma_2} p(x'; y) v_1(x') dS_{x'} + f_2(y), \end{aligned} \quad (4.6)$$

где

$$\begin{aligned} f_1(x) &= \int_{\Gamma_1} p(y; x) \varphi(y) dS_y, \\ f_2(y) &= \int_{\Gamma_2} p(x'; y) \varphi(x') dS_{x'} + f_2(y), \end{aligned} \quad (4.7)$$

или в матричной форме

$$v = K v + F, \quad (4.8)$$

где $v = (v_1, v_2)^T$, $F = (f_1, f_2)^T$ и K является матрично-интегральным оператором, действующим на v следующим образом:

$$Kv(x, y) = \begin{pmatrix} 0 & \int_{\gamma_1} p(y; x)v_2(y) dS_y \\ \int_{\gamma_2} p(x'; y)v_1(x') dS_{x'} & 0 \end{pmatrix}.$$

Таким образом, мы приходим к системе интегральных уравнений (4.8) с интегральным оператором K , который имеет ряд приятных свойств. Прежде всего, L_1 — норма оператора K меньше 1 для любых расположений двух кругов с непустым пересечением, поскольку $\int_{S(x,R)} p(y; x)dS_y = 1$. Следовательно, $(E - K)^{-1}$ существует и представляется в виде сходящегося ряда Неймана; здесь E — единичный оператор.

Более точная оценка нормы следует из следующего утверждения.

Лемма 4.1. *Для любых $x \in \gamma_2$ и $y \in \gamma_1$*

$$\int_{\gamma_1} p(y; x) dS_y = \int_{\gamma_2} p(y'; y) dS_{y'} = 1 - \frac{\theta_1^*}{\pi} - \frac{\theta_2^*}{\pi},$$

где углы θ_1^* и θ_2^* определяются следующим образом: $2\theta_1^*$ — угол видимости дуги γ_1 из центра первого круга, а $2\theta_2^*$ — угол видимости дуги γ_2 из центра второго круга.

Доказательство. Для произвольной точки $x \in \gamma_2$ имеем

$$p(y; x) = \frac{\cos(\psi)}{\pi|x - y|} - \frac{1}{2\pi R_1},$$

что следует из

$$R_1^2 - r_1^2 + |x - y|^2 = 2R_1|x - y|\cos(\psi),$$

где ψ — угол между векторами $x - y$ и $x - x_0^{(1)}$. Отсюда получаем

$$\int_{\gamma_1} p(y; x) dS_y = \frac{1}{\pi} \int_{\gamma_1} \frac{\cos(\psi)}{|x - y|} dS_\psi - \frac{1}{2\pi R_1} \int_{\gamma_1} dS.$$

Первый интеграл в правой части есть потенциал двойного слоя, который, как известно [100], равен углу видимости дуги γ_1 из точки x , т.е. равен $(2\pi - 2\theta_2^*)/2$. Второй интеграл равен θ_2^*/π . Что и завершает доказательство, поскольку мы получим в точности тот же результат и для произвольной точки второго круга, когда $y \in \gamma_1$. Этот результат позволяет доказать следующую теорему:

Теорема 4.1. *Интегральный оператор K является фредгольмовым с ядрами $p(y; x)$, $p(x'; y)$, которые непрерывны на $x \in \gamma_2$, $y \in \gamma_1$, а при приближении фазовой точки к пересечению окружностей γ_1 и γ_2 сингулярность ядер носит полярный характер типа $p(y; x) \sim \frac{\sin(\theta_1^* + \theta_2^*)}{\pi |x - y|}$ при $x \rightarrow y$. Все собственные числа λ_i интегрального оператора K вещественны и имеют вид $\lambda_i = \pm \sigma_i \rho(K)$, где $\rho(K)$ — спектральный радиус оператора K ($\sigma_i \leq 1$ — некоторые положительные константы), причем спектральный радиус имеет вид*

$$\rho(K) = 1 - \frac{\theta_1^*}{\pi} - \frac{\theta_2^*}{\pi}.$$

Интегральное уравнение (4.8) имеет единственное решение, которое одновременно является решением задачи Дирихле (4.1) на дугах γ_1 , γ_2 , а решение в произвольных точках вычисляется с помощью интеграла Пуассона (4.4).

Обобщение на области G , состоящие из n шаров, каждый из которых имеет хотя бы одно пересечение с одним из остальных шаров, осуществляется достаточно просто. В результате мы имеем систему интегральных уравнений Фредгольмова типа, с теми же ядрами, а структура матрично-интегрального оператора — ленточная, с шириной ленты, зависящей от конфигурации зацеплений в системе шаров. Особенно простую структуру матрица будет иметь для S_2 -областей, которые определяются следующим образом: в определенных выше областях G недопустимы лишь такие конфигурации, в которых: (1) — имеется одновременное пересечение более чем двух кругов, (2) — в замкнутых системах кругов могут участвовать лишь четное число кругов. Можно показать, что соответствующая матрица для S_2 -областей является циклической индекса 2, что определяет ряд свойств, полезных при построении итерационных методов.

Имея таким образом определенные системы интегральных уравнений, можно построить самые различные варианты алгоритмов случайного блуждания для их решения. Во-первых, можно строить стандартное случайное блуждание по плотности $p(y; x)$ с обрывом на дугах Γ_1 , Γ_2 , как это было сделано в [98], поскольку в данном простом случае ряд Неймана сходится. Однако в более общих ситуациях, например, в случае системы уравнений Ламе, такой способ не работает, поэтому мы применяем специальные итерационные процессы, причем строим n -е приближение, моделируя n шагов марковского процесса без поглощения.

Рассмотрим систему уравнений теории упругости, — в случае плоской упругой среды, систему уравнений Ламе:

$$\mu \Delta \mathbf{u} + (\lambda + \mu) \nabla \operatorname{div} \mathbf{u} = 0, \quad \mathbf{u} = (u_1, u_2),$$

где константы λ и μ характеризуют упругую среду, и пусть $\sigma = 1 + 2\mu/(\lambda + \mu)$.

Построение эквивалентной системы интегральных уравнений осуществляется в точности по описанной схеме, исходя из обобщенной формулы Пуассона, выведенной нами в [97, 98]:

$$u_i(x) = \frac{R^2 - |x - x_0|^2}{2\pi R} \sum_{j=1}^2 \int_{S(x_0, R)} \frac{b_{ij} u_j(y)}{|x - y|^2} dS_y, \quad i = 1, 2, \quad (4.9)$$

где $b_{ij}(x, y)$ — элементы матрицы

$$B = \frac{1}{\sigma} \begin{pmatrix} (\sigma - 1) + 2 \cos^2 \varphi + & 2 \cos \varphi \sin \varphi + \\ + \frac{|x - y|}{R} \cos(\theta + \varphi) & + \frac{|x - y|}{R} \sin(\theta + \varphi) \\ 2 \cos \varphi \sin \varphi + & (\sigma - 1) + 2 \sin^2 \varphi - \\ + \frac{|x - y|}{R} \sin(\theta + \varphi) & - \frac{|x - y|}{R} \cos(\theta + \varphi) \end{pmatrix}.$$

Здесь углы θ и φ определяются так: угол θ есть полярный угол точки y в исходной системе координат, а угол φ — полярный угол той же точки y , но в системе координат с центром в точке x .

Удобно переписать эту систему уравнений в матричной форме, сохраняя обозначение $p(y; x)$ для ядра Пуассона:

$$\mathbf{u}(x) = \int_{S(x_0, R)} p(y; x) B \mathbf{u}(y) dS(y), \quad (4.10)$$

а результирующую систему интегральных уравнений, сосредоточенную на дугах кругов, запишем как $v = K v + F$.

Как видно, структура системы интегральных уравнений сохраняется, поскольку элементы матрицы B — непрерывные и ограниченные функции своих аргументов, так что основные свойства, такие, например, как фредгольмовость, сохраняются. Однако существенно новым моментом является то, что норма интегрального оператора здесь больше единицы, и традиционный метод блуждания по сферам не работает, поскольку ряд Неймана, представляющий дисперсию соответствующей случайной оценки, расходится. Поэтому мы используем подходящие итерационные методы, один из которых здесь идеально подходит — метод последовательной верхней релаксации. Для простоты приведем результат для случая уравнения (4.8).

Интегрально-матричное ядро K запишем в виде $K = L + U$, где L и U — нижне- и верхнетреугольный матрично-интегральные операторы

$$Lv = \begin{pmatrix} 0 & 0 \\ \int_{\gamma_2} p(x', y) v_1(x') dS_{x'} & 0 \end{pmatrix}, \quad Uv = \begin{pmatrix} 0 & \int_{\gamma_1} p(y, x) v_2(y) dS_y \\ 0 & 0 \end{pmatrix}.$$

Вводя произвольный скалярный параметр ω , перепишем наше уравнение $v = Kv + F$ в виде:

$$v = (E - \omega L)^{-1}[(1 - \omega)E + \omega U]v + \omega(E - \omega L)^{-1}F. \quad (4.11)$$

Теорема 4.2. *Для S_2 -областей матрица ядра является циклической индекса 2, причем $(E - \omega L)^{-1} = E + \omega L$, и интегральное уравнение (4.11) приобретает особенно простой вид:*

$$v = Tv + d,$$

где

$$T = (E + \omega L)[(1 - \omega)E + \omega U], \quad d = \omega(E + \omega L)F.$$

Параметр ω выбирается из следующих соображений. Во-первых, скорость сходимости итераций должна быть по-возможности максимальной, во-вторых, дисперсия соответствующей случайной оценки должна быть если не минимальной, то хотя бы конечной. Для рассматриваемой задачи наиболее хорошие результаты получены при $\omega = 0.97$, что близко значению этого параметра в случае метода Зейделя [27]. Детально результаты метода блуждания по фиксированным сферам изложены в [99, 96]. Отметим в заключение, что в методе могут использоваться и другие стандартные области, например, прямоугольники или эллипсы, для которых известны функции Грина.

4.2. Эйлеровы и лагранжевы стохастические модели переноса частиц в турбулентных потоках и пористых средах

Перенос частиц в турбулентных потоках или течениях в пористой среде является актуальной и перспективной задачей, привлекающей чрезвычайно большое внимание во многих экологических (рассеяние примеси в атмосфере и перенос в почвах) и промышленных задачах, таких как производство новых материалов и биотехнологии [82]. Описание взаимодействия частиц с помощью процессов агрегации–дисагрегации, в результате чего формируются кластеры различного размера, а также привлечение лагранжевых траекторий таких частиц для построения пространственных концентраций являются ключевыми методами исследования в аэрозольной технологии (например, фотохимическое образование частиц [85], формирование облаков [95], перенос аэрозольных частиц в пограничном слое атмосферы [89], [74], образование сажистых частиц в процессах горения, физика полимеров и химия поверхности в процессах формирования частиц силикона [85]. В [91, 95] нам удалось обнаружить с помощью расчетов эффект сильного ускорения роста частиц за счет перемежаемой турбулентности.

С физической точки зрения эта задача включает в себя огромный спектр масштабов: достаточно упомянуть только "простейшую" составляющую — сам поток: для развитой турбулентности с числом Рейнольдса порядка десятков тысяч традиционные решатели уравнения Навье-Стокса потребовали бы таких объемов машинной памяти и времени, которые еще далеко не достигнуты в самых современных вычислительных системах. Течения в пористой среде, вообще говоря, имеют малые числа Рейнольдса и не являются турбулентными, однако они имеют чрезвычайно сложную пространственную структуру в силу сложности пористой структуры среды. Так что если учесть, что частицы в этих потоках еще и взаимодействуют между собой, в частности коагулируют и распадаются, вследствие чего спектр размеров меняется от одного мономера до кластеров и частиц, состоящих из миллионов мономеров, то станет ясно, что вероятностный подход к таким задачам является вполне естественным. Возрастающий интерес к бессеточным численным методам отражает рост сложности исследуемых проблем на фоне одновременного прогресса в мощностях ЭВМ. Это придало новый импульс в развитии методов частиц и стохастических алгоритмов моделирования. В частности, для расчета сложных многомерных течений с успехом применяется подход, известный под названием "smoothed particle hydrodynamics method".

Математически мы имеем дело с двумя фундаментальными задачами: (1) — построение стохастических моделей турбулентных потоков и течений в пористой среде, (2) — развитие стохастических моделей для динамики частиц, движущихся в турбулентных или других сложных потоках.

Для построения стохастических моделей несущего поля скоростей мы используем три основных подхода.

1. В случае пористой среды строится стохастическое поле гидравлической проницаемости $K(x)$ в условиях заданной разности давления (например, определяемой через гравитацию), а затем поле скоростей находится с помощью уравнения Дарси [76, 93]:

$$q(x) = \theta v(x) = -K(x) \nabla \varphi(x).$$

Функция $\varphi(x)$ удовлетворяет уравнению

$$\sum_{i=1}^3 \frac{\partial}{\partial x_i} \left(K(x) \frac{\partial \varphi}{\partial x_i} \right) = 0,$$

где θ — пористость, $q(x)$ — скорость Дарси, а $v(x)$ — скорость в порах.

2. В турбулентных потоках скорость в принципе описывается уравнением Навье-Стокса, однако его численное решение при больших числах Рейнольдса практически невозможно. Мы используем принятое в данной области статистическое описание, когда поле скоростей строится как случай-

ное поле $v(x)$, обладающее некоторыми фундаментальными статистическими свойствами. Такой подход для численного моделирования был впервые предложен в работе [71] и затем развит новосибирской группой [79, 51, 90].

3. Как в турбулентных потоках, так и в течениях в пористой среде, моделирование динамики частиц требует определения скорости лишь вдоль траекторий частиц, не рассчитывая его во всей области течения. Лагранжевы стохастические модели строятся именно на этом принципе, на основе стохастических уравнений для траекторий частиц, без привлечения какой-либо интерполяции. Для задач турбулентного переноса эти модели имеют вид стохастического дифференциального уравнения типа Ланжевена [74, 87, 94]:

$$dv(t) = a(X(t), t)dt + b(X(t), t)dW(t), \quad X(0) = x, \quad (4.12)$$

где $W(t)$ — винеровский процесс.

В этой системе стохастической дифференциальной системе уравнений Ито тензор $b(x, t)$ выводится из эйлеровой стохастической модели потока, а функция сноса $a(x, t)$ может быть получена из экспериментов или из традиционных численных методов. В [104, 74] построен класс стохастических моделей, известных как "well-mixed stochastic models", который открыл существенно новые перспективы в развитии лагранжевых стохастических моделей и был успешно применен к решению задач переноса в погранслое турбулентной атмосферы [74, 89, 72], а также к исследованию процессов перемешивания в реках [61]. Что касается лагранжевых стохастических моделей для задач переноса в пористой среде, то они только начинают развиваться. Первых два подхода, описанных выше, принято называть эйлеровыми методами, поскольку в них строится вначале модель всего поля течения в фиксированной эйлеровой системе координат, в отличие от лагранжевого подхода, где этот первый шаг отсутствует, и описание сразу ведется вдоль лагранжевых траекторий. Эйлеровы модели более надежны и строги, поскольку в них может быть непосредственно учтена любая статистическая информация о потоке (или гидравлической проницаемости пористой среды), например, одноточечная плотность распределения, корреляционная структура, статистическое подобие, анизотропия. В этой области группам из Новосибирска, Принстона и Института Куранта удалось построить два эффективных метода моделирования таких эйлеровых случайных полей: один метод основан на рандомизированном спектральном представлении [90, 51], другой — на взрывлет-разложении в пространстве Фурье [60]. Лагранжевы же методы могут быть гораздо более эффективными с численной точки зрения, поскольку они не связаны с построением всего поля скоростей и требуют малого числа коэффициентов в представлении скорости локально в окрестности положения частицы. Тонкость здесь заключается в том, чтобы эти коэффициенты отражали су-

щественные характеристики течения. Это хорошо известная проблема взаимосвязи эйлеровых и лагранжевых статистических характеристик [87]. В лагранжевых стохастических моделях, описываемых уравнениями типа (4.12), коэффициент $b(x, t)$ выводится довольно просто, тогда как вывод сноса $a(x, t)$ более сложен при наличии неоднородности и требует привлечения дополнительной экспериментальной информации, а также известного свойства "well-mixed" [104]. При выбранных коэффициентах уравнение Ито далее решается каким-либо из стандартных методов, таких, например, как явно-неявная схема Эйлера или схема Мильштейна [68]. Существуют также гибридные эйлерово-лагранжевы модели в духе наших работ [73, 75], где некоторые относительно крупные масштабы моделируются в рамках эйлерового подхода, а локальные масштабы — на основе лагранжевой модели. Такой подход позволяет учесть особенности течения, которые в рамках лагранжевого подхода не могут быть учтены в принципе, но одновременно он является весьма экономичным с численной точки зрения. Таким образом, идея состоит в упрощении тех особенностей течения, которые могут быть описаны малым числом параметров.

До сих пор практически единственным методом моделирования переноса в стохастических полях скоростей являлся RDM-метод (метод случайных смещений). Этот метод, хотя он и достаточно эффективен, имеет ряд существенно слабых сторон: в его описании не используется вектор скоростей, что сказывается на нарушении многих существенных свойств [76]. В связи с этим мы предприняли разработку лагранжевых стохастических моделей на основе стохастического моделирования эйлерова поля скоростей и использования критерия "well-mixed" [76].

Чтобы смоделировать динамику взаимодействующих частиц, переносимых потоком, необходимо построить вначале поле скоростей $v(t, x)$, а затем решать уравнение Смолуховского. Система уравнений для концентрации частиц записывается в виде [69, 70]:

$$\begin{aligned} & \frac{\partial n_l(t, x)}{\partial t} + \mathbf{v}(t, x) \cdot \nabla_x n_l(t, x) = \\ & = D_l \Delta n_l(t, x) \frac{1}{2} \sum_{i+j=l} K_{ij} n_i n_j - n_l \sum_{i=1}^{\infty} K_{li} n_i + F_l(t, x) \end{aligned}$$

при заданных начальных распределениях $n(0, x) = n_l^{(0)}(x)$, $l = 1, 2, \dots$. Здесь мы используем обозначения: l -кластер — частица, содержащая l монономеров (или структурных единиц); n_l — плотность числа l -кластеров; K_{ij} — коэффициент коагуляции, характеризующий частоту столкновений между i -кластером и j -кластером; D_l — коэффициент диффузии l -кластера и $F_l(t, x)$ — интенсивность источника l -кластера. Литература, посвященная численным методам решения для уравнения Смолуховского весьма

обширна: это уравнение привлекает все большее внимание, что объясняется его универсальностью во многих прикладных областях, а также тем, что оно было слабо исследовано с математической точки зрения. Подробный обзор литературы по стохастическим моделям для решения уравнения Смолуховского дан в [70]; там же разработан алгоритм типа Нанбу для однородного случая, а в [92] — лагранжев стохастический метод для класса неоднородных уравнений Смолуховского, где наряду со средним потоком учитывается пространственная диффузия частиц. Следует отметить, что для случая, когда коэффициент диффузии зависит от размера частицы, обобщение данного метода остается открытой проблемой. Другая интересная проблема — так называемый *gelation phenomenon* — феномен желирования. Скорость столкновения частиц в этом случае настолько велика (например, как в случае $K = ij$), что решение взрывается — и образуется суперкластер бесконечной массы. Этот эффект хорошо известен в полимерных исследованиях как эффект желирования; он практически не поддается моделированию детерминированными численными методами, но довольно естественно анализируется с помощью стохастического моделирования. В области переноса в пористой среде этот эффект находит применение при моделировании течений, содержащих глину: растущие глинистые частицы начинают забивать поры, и течение существенно перестраивается.

5. Методы Монте-Карло для численного решения СДУ и в финансовой математике

Начало создания теории стохастических дифференциальных уравнений (СДУ) относится к середине XX в. и связано с именами японского математика Ито и российского — Р.Л. Стратоновича. СДУ являются обобщением обыкновенных дифференциальных уравнений путем добавления в правую часть слагаемого, включающего винеровский процесс (гауссовский процесс с независимыми приращениями, с нулевым математическим ожиданием и дисперсией, растущей пропорционально времени).

Пусть заданы q -мерный стандартный винеровский процесс $w(\cdot)$, p -мерная вектор-функция $f(t, y)$, матричная функция $\sigma(t, y)$ размера $p \times q$, случайный вектор y_0 , независимый с $w(t)$ при $t \geq 0$, причем $E|y_0|^2 < \infty$. Тогда задача Коши для системы СДУ ставится следующим образом: найти непрерывный p -мерный случайный процесс $y(\cdot)$, для которого с вероятностью единица для всех $t \in [0, T_{\text{кон}}]$ одновременно выполняется равенство

$$y(t) = y_0 + \int_0^t f(s, y(s)) ds + \int_0^t \sigma(s, y(s)) dw(s). \quad (5.1)$$

В зависимости от способа определения стохастического интеграла $\int_0^t \sigma(s, y(s))dw(s)$ СДУ понимают в смысле Ито или в смысле Стратоновича.

Систему СДУ (5.1) часто записывают в эквивалентной дифференциальной форме:

$$dy(t) = f(t, y(t))dt + \sigma(t, y(t))dw(t). \quad (5.2)$$

Функцию $f(t, y)$ принято называть функцией сноса, а функцию $\sigma(t, y)$ — интенсивностью шума.

При выполнении условий

$$\begin{aligned} |f(t, y_1) - f(t, y_2)|^2 + \|\sigma(t, y_1) - \sigma(t, y_2)\|^2 &\leq C_1|y_1 - y_2|^2 \\ |f(t, y)|^2 + \|\sigma(t, y)\|^2 &\leq C_2(1 + |y|^2) \end{aligned}$$

равномерно по t решение системы СДУ в смысле Ито (5.1) существует и единственно (см., например, [9]).

В более общем определении СДУ винеровский процесс заменяют на процесс с независимыми приращениями, например, путем добавления процесса Пуассона, в результате чего решение СДУ становится разрывным со случайной величиной скачка и случайным временем скачка.

По характеру поведения вероятностных характеристик своих решений СДУ можно классифицировать как "жесткие", когда математическое ожидание имеет быстрые переходные участки, осциллирующие, когда математическое ожидание или дисперсия имеют квазипериодическое поведение, неустойчивые — с растущей со временем дисперсией и несимметричные — с растущей асимметрией одномерных распределений. Алгоритмы статистического моделирования траекторий решений СДУ каждой группы имеют свои специфические особенности по точности, устойчивости и трудоемкости и требуют индивидуального подбора численных методов решения СДУ и методов улучшения монтекарловских оценок рассчитываемых функционалов (см. [3, 58]).

5.1. Численные методы решения СДУ

В отличие от всех других методов анализа решений СДУ, имеющих ограниченную область применимости (теория мартингалов, системы ОДУ на моментные функции, параболическое уравнение Фоккера–Планка–Колмогорова на плотность распределения решения СДУ, линеаризация СДУ, разложение решения по малому параметру), численные методы решения СДУ являются универсальными и позволяют оценить любые функционалы от решений с точностью, ограниченной только возможностями используемой вычислительной техники.

Простейшим численным методом для решения систем СДУ в смысле Ито является обобщение явного метода Эйлера:

$$y_{n+1} = y_n + hf(t_n, y_n) + \sqrt{h}\sigma(t_n, y_n)\zeta_n, \quad n = 0, 1, \dots, N - 1, \quad (5.3)$$

где y_n — значение приближенного решения системы СДУ в узле сетки по времени t_n , h — шаг интегрирования в узле t_n , $\{\zeta_n\}$ — последовательность независимых между собой нормальных случайных векторов с независимыми в совокупности компонентами ζ_{jn} , $j = 1, \dots, q$, имеющими нулевое математическое ожидание и единичную дисперсию. Кроме того, предполагается независимость между собой векторов ζ_n и y_n для каждого n .

Численный метод слабо сходится на решении задачи Коши с порядком k , если для любой достаточно гладкой функции $g(y)$

$$|Eg(y(T_{\text{кон}})) - Eg(y_N)| = O(h^k), \quad h \rightarrow 0.$$

Обобщенный метод Эйлера слабо сходится с первым порядком. Более точным является метод Мильштейна:

$$y_{n+1} = y_n + hf(t_n, y_n) + \sqrt{h}\sigma(t_n, y_n)\zeta_n + \frac{h}{2} \frac{\partial \sigma}{\partial y} \sigma(t_n, y_n)(\zeta_n^2 - 1). \quad (5.4)$$

Методы высокого порядка сходимости в слабом смысле строятся, как и в случае ОДУ, путем сравнения разложений в ряд Тейлора точного и численного решений. Существует также понятие сходимости численных методов в среднеквадратическом смысле, однако это определение не является конструктивным, так как до настоящего времени не построено численных методов решения систем СДУ с порядком сходимости в среднеквадратическом смысле выше $1/2$ из-за сложности аппроксимации стохастических интегралов от одного винеровского процесса по другому.

Для проверки устойчивости численных методов решения СДУ необходимо выбрать тестовое СДУ. Например, в работе [3] для этой цели предлагается использовать линейное скалярное СДУ с аддитивным шумом, решение которого является стационарным гауссовским процессом:

$$dy(t) = \mu y(t)dt + \sigma dw(t), \quad (5.5)$$

где $\mu < 0$ и σ — постоянные вещественные коэффициенты. Свойство устойчивости метода согласуется с асимптотическим поведением распределения точного решения тестового СДУ, в частности с поведением математического ожидания и дисперсии точного решения при $t \rightarrow \infty$.

В работе [3] построено обобщение на системы СДУ полунявного метода Рунге-Кутты вида

$$y_{n+1} = y_n + \left[I - \frac{h}{2} \frac{\partial f}{\partial y}(t_n, y_n) \right]^{-1} [hf(t_n, y_n) + \sqrt{h}\sigma(t_n, y_n)\zeta_n], \quad (5.6)$$

где I — единичная матрица, $\partial f/\partial y$ — матрица Якоби. В работах [3, 58] доказано, что метод (5.6) имеет лучшие свойства устойчивости, чем методы Эйлера и Мильштейна, и лучше подходит для численного решения осциллирующих СДУ.

Наиболее сложными для численного анализа являются стохастические динамические системы, имеющие колебательные режимы движения. Необходимость анализа стохастических динамических систем с колебательными режимами возникает при построении математических моделей различных явлений в радиотехнике, турбулентности, движении летательных аппаратов, волновых изменениях биржевых цен различных финансовых инструментов, сезонных природных процессах и т.п. Зачастую аналитические исследования решений индивидуальной нелинейной системы СДУ с несколькими колебательными режимами требуют разработки специфических подходов, тем более проблемно говорить об аналитическом исследовании параметрических семейств систем СДУ. Наиболее часто используемый метод гауссовой аппроксимации здесь мало применим в силу нелинейной специфики уравнений. Аналогично, не всегда можно говорить о приближениях первого порядка, так как малые колебания в автоколебательных системах нетипичны. В то же время методы Монте-Карло являются универсальными и применимыми для статистического моделирования решений любых систем СДУ, в частности нелинейных систем СДУ большой размерности с колебательными решениями. С помощью статистического моделирования можно исследовать наличие у таких систем аттракторов, выявить точки бифуркации, определить области устойчивости и неустойчивости решений, т.е. фактически провести полный параметрический анализ систем СДУ. Анализ стохастических динамических систем связан с расчетом большого количества вероятностных характеристик случайных процессов. Это моменты, плотности распределений, корреляции, вероятности выходов или невыходов траекторий из заданных областей и т.д. Для колебательных систем в первую очередь интересны вероятностные характеристики, связанные с колебаниями, например, распределение амплитуд и периодов колебаний, вероятности перехода системы из одного колебательного режима в другой, вероятность нахождения фазовой траектории в заданной области. В работе [58] разработаны специальные алгоритмы переменного шага с контролем точности вычислений, предназначенные для численного решения осциллирующих СДУ.

Исключительно сложным для стохастических колебательных систем является анализ их устойчивости [55]. Такие исследования наиболее актуальны в теории и практике оптимального управления движениями механических систем, подверженных случайным возмущениям. Обычно колебательные явления связаны с ростом дисперсии решения соответствующей системы СДУ. Такие системы с растущей дисперсией возникают в финан-

совой математике при моделировании ценовых рядов акций и финансовых фьючерсов, в радиотехнике при моделировании линейных и нелинейных колебательных контуров, а также при решении краевых задач математической физики с помощью вероятностного представления решения. Попытки использовать численные методы для статистического моделирования решения СДУ с растущей дисперсией заканчиваются одинаково неудачно из-за низкой точности оценки функционалов от решения. В этой ситуации использование методов высокого порядка сходимости в слабом смысле и моделирование огромного количества траекторий не приводит к успеху.

В работе [1] для решения СДУ с растущей дисперсией предлагается представлять решение исходной неустойчивой системы СДУ посредством двух составляющих, причем стохастическая составляющая задается системой СДУ с решением, близким к стационарному процессу, а неслучайная составляющая описывается неустойчивой системой ОДУ, численное решение которой не составляет труда. Система ОДУ включает в себя уравнения на первые и вторые моменты решения исходного СДУ, причем эти моменты могут быть представлены приближенно. Далее с использованием формулы дифференцирования Ито [9] получаются системы СДУ для стохастической составляющей как в случае линейной, так и нелинейной исходной системы СДУ. Дополнительно могут быть использованы некоторые классические способы уменьшения дисперсии монтекарловских оценок применительно к численному решению неустойчивых СДУ [13].

Например, для линейной системы СДУ в смысле Ито

$$dy(t) = Aydt + \sum_{j=1}^q S^{(j)} y dw_j(t) \quad (5.7)$$

системы ОДУ на математическое ожидание и матрицу вторых моментов имеют вид

$$\frac{dm(t)}{dt} = Am(t), \quad \frac{d\Gamma(t)}{dt} = A\Gamma(t) + \Gamma(t)A^* + \sum_{j=1}^q S^{(j)}\Gamma(t)^*.$$

Сделав в СДУ (5.7) замену переменных

$$x_i(t) = \frac{y_i(t) - m_i(t)}{\sqrt{d_i(t)}}, \quad d_i(t) = \gamma_{ii}(t) - m_i^2(t), \quad (5.8)$$

и применив формулу Ито, получим систему СДУ на процесс $x(t)$:

$$dx_i(t) = \left(-\frac{d'_i x_i}{2d_i} + \frac{1}{\sqrt{d_i}} \sum_{k=1}^p a_{ik} \sqrt{d_k} x_k \right) dt + \frac{1}{\sqrt{d_i}} \sum_{j=1}^q \sum_{k=1}^p s_{ik}^{(j)} (\sqrt{d_k} x_k + m_k) dw_j, \quad (5.9)$$

где

$$d_i'(t) = \sum_{k=1}^p \left(2a_{ik}\gamma_{ki}(t) + \gamma_{ik}(t)a_{ik} + \sum_{l=1}^p s_{il}^{(j)}\gamma_{lk}(t)s_{ik}^{(j)} \right) - 2m_i(t) \sum_{k=1}^p a_{ik}m_k(t).$$

Новые переменные $x_i(t)$ имеют нулевое математическое ожидание и единичную дисперсию.

Использование методов Монте-Карло для решения краевых задач математической физики имеет давние традиции [13]. В последние десятилетия все больший интерес для математиков представляет численный метод решения линейных эллиптических и параболических уравнений с помощью СДУ. Метод основан на вероятностном представлении решения краевой или начально-краевой задачи в виде функционала от решения соответствующей системы СДУ. В отличие от конечно-разностных методов, при решении уравнений с частными производными с помощью СДУ не требуется построения сеток по пространственным переменным, что значительно облегчает подготовку задачи к решению, а алгоритм решения и его программная реализация становятся предельно простыми. Наиболее сложным вопросом при таком подходе является точность определения момента пересечения моделируемой траектории решения СДУ границы области, особенно для случая осциллирующих и неустойчивых соответствующих систем СДУ. В работе [58] исследованы вопросы зависимости точности оценок решений эллиптических и параболических уравнений от размера шага интегрирования численного метода решения СДУ.

5.2. Математические модели цен финансовых инструментов

Математические модели в виде СДУ все шире используются в многочисленных приложениях. К классическим статистической механике и физике, радиотехнике, автоматическому управлению, химии, теории надежности, иммунологии добавилась финансовая математика. Теория расчета стоимости опциона, современная теория формирования портфеля ценных бумаг, теория ценообразования рискованного актива требуют использования СДУ в качестве моделей движения курсов ценных бумаг и валют, акционных индексов и процентных ставок.

Математическая модель должна, с одной стороны, быть адекватной историческим ценам, а с другой стороны, достаточно простой и гибкой при ее использовании по назначению. Особенно важна стохастическая математическая модель при проверке эффективности торговых стратегий и торговых компьютерных программ, так как она позволяет моделировать ансамбль траекторий, что невозможно в реальности, где всегда имеется лишь единственная реализация цены.

Основой для выбора структуры математической модели цены финансового инструмента является статистический анализ временных рядов из исторических значений цен. Статистический анализ позволяет выявить специфические закономерности поведения вероятностных характеристик цен, оценить закон распределения значений цен, выбрать финансовые инструменты, приемлемые для реальной прибыльной торговли. В работах [4, 5] предложены и исследованы различные линейные и нелинейные модели цен финансовых инструментов. Здесь же построен ряд численных алгоритмов для оценки неизвестных параметров моделей по дискретным наблюдениям за ценами, основанных на методе наименьших квадратов и методе максимального правдоподобия.

Структура стохастической модели цены финансового инструмента обычно выбирается в зависимости от решаемой с помощью этой модели задачи. Так, известная модель цены акции в виде СДУ Ито:

$$dP(t) = \mu P(t)dt + \sigma P(t)dw(t), \quad P(0) = P_0, \quad (5.10)$$

где μ, σ — постоянные вещественные коэффициенты, с успехом используемые для расчета стоимости опционов [59]. Точное решение (5.10) можно записать в виде

$$P(t) = P_0 \exp \left(\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma w(t) \right), \quad (5.11)$$

или в узлах равномерной сетки по времени $\{t_n\}$ с шагом h в рекуррентном виде

$$P_{n+1} = P_n \exp \left(\left(\mu - \frac{\sigma^2}{2} \right) h + \sqrt{h} \sigma \zeta_n \right), \quad n = 0, \dots, N - 1, \quad (5.12)$$

где $\{\zeta_n\}$ — последовательность независимых между собой стандартных нормальных случайных величин.

Для неслучайного положительного начального значения одномерная плотность распределения решения СДУ (5.10) при каждом фиксированном t является логарифмически нормальной.

К недостаткам модели (5.10) следует отнести то, что она совершенно не учитывает волновой характер движения цен акций и не может использоваться, например, для моделирования гипотетических рыночных ситуаций и тестирования компьютерных торговых программ. Волновое движение цен отмечено не только у акций, но и у курсов торгуемых валют и у цен разнообразных финансовых фьючерсов.

Хорошо известно, что волновые движения у решений ОДУ могут быть введены различными способами: через собственные колебания, параметрические колебания, автоколебания, вынужденные колебания. Эти же способы можно использовать в моделях случайных процессов, являющихся

решениями СДУ. Например, в модели (5.10) параметрические колебания можно ввести следующим образом [4]:

$$dP(t) = (\mu + A\cos(\omega t))P(t)dt + \sigma P(t)dw(t), P(0) = P_0. \quad (5.13)$$

Путем вариации параметров μ , σ , A , ω можно добиться нужного волнового движения вероятностных характеристик процесса $P(\cdot)$.

В работе [5] в качестве модели цены предложено использовать СДУ, обобщающее известное уравнение Ван-дер-Поля, описывающее поведение автоколебательной системы:

$$dP(t) = \varphi(t)dt, \\ d\varphi(t) = (\varepsilon(t)\varphi(t)(1 - r(t)P(t)^2) + b(t)P(t))dt + \sigma(t, P(t))dw(t),$$

где $\varepsilon(t)$, $r(t)$, $b(t)$, $\sigma(t, y)$ — заданные неслучайные функции, выбором которых задается определенное поведение вероятностных характеристик автоколебаний.

Использование моделей с разным типом колебаний позволяет в каждой конкретной задаче добиться нужных целей наиболее полно и точно. При выборе модели приходится учитывать ее сложность, так как многомерные нелинейные модели могут оказаться слишком сложными с точки зрения оценки параметров по историческим ценам и при выборе параметров при попытке задания определенного гипотетического движения цен.

Современное состояние мировой компьютерной сети Интернет дает возможность огромному количеству физических и юридических лиц участвовать в биржевой торговле ценными бумагами, в частности акциями. Профессиональная торговля доступна только высококвалифицированным опытным специалистам по торговле акциями и невозможна без компьютерной поддержки. В настоящее время имеется большое количество компьютерных торговых программ, которые выдают сигналы на покупку и продажу, основываясь на поступающей с биржи числовой информации в режиме реального времени. Каждая такая программа имеет некоторый набор параметров, варьируя которые можно добиваться определенных характеристик торговли, например, высокой годовой доходности, минимального риска и минимальной длительности ряда убыточных сделок. Обычно параметры подбираются путем тестирования программ на исторических ценах акции за определенный прошлый период времени. Однако практика торговли показывает, что зачастую оптимальная торговая программа дает совершенно убыточную реальную торговлю. Поэтому тестирование торговых программ желательно проводить на ансамбле цен, чего нет в действительности, но что может быть получено путем статистического моделирования цены акции на компьютере.

6. Дискретно-стохастические численные методы

Традиционно методы Монте-Карло рассматриваются в качестве альтернативных "детерминированным" численным методам (в частности конечно-разностным и конечно-элементным схемам). Однако во многих случаях эффективными оказываются смешанные алгоритмы, содержащие в себе элементы детерминированных и стохастических численных схем. Такие комбинированные алгоритмы можно назвать *дискретно-стохастическими численными методами*.

6.1. Смешанные методы моделирования случайных величин

Целесообразность применения комбинированных алгоритмов обнаруживается уже на уровне моделирования случайных величин. При получении выборочных значений непрерывных случайных величин (в частности, при использовании гистограмм, полигонов частот и других приближений плотностей, а также при построении методов исключения, в том числе двусторонних [8], [108]), требуется моделировать случайные векторы $\boldsymbol{\eta}$ с плотностями распределения вида

$$p_{\boldsymbol{\eta}}(\mathbf{u}) \approx HL_{(M)}f(\mathbf{u}) = H \sum_{i=0}^M w_i \chi_i(\mathbf{u}) \quad (6.1)$$

на компактном множестве $U \subseteq R^l$; здесь H — нормирующая константа. В формуле (6.1) $L_{(M)}f$ обозначает аппроксимацию (или интерполяцию) неотрицательной функции f на сетке $X^{(M)} = \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_M\}$. Базисные функции $\Xi^{(M)} = \{\chi_0, \chi_1, \dots, \chi_M\}$ и коэффициенты $W^{(M)} = \{w_0, w_1, \dots, w_M\}$ определенным образом связаны с узлами сетки $X^{(M)}$ (в частности, коэффициенты $W^{(M)}$ являются, как правило, комбинациями значений $\{f(\mathbf{x}_0), f(\mathbf{x}_1), \dots, f(\mathbf{x}_M)\}$). Базис $\Xi^{(M)}$ и коэффициенты выбираются таким образом, что функция $p_{\boldsymbol{\eta}}$ близка к функции Hf в некоторой функциональной норме, и аппроксимация (6.1) устойчива. Дополнительным требованием является наличие эффективного алгоритма реализации выборочных значений вектора $\boldsymbol{\eta}$ согласно распределению (6.1) (в этом случае мы будем называть базис $\Xi^{(M)}$ "моделируемым"). Пусть

$$\chi_i(\mathbf{u}) \geq 0 \text{ для } \mathbf{u} \in R^l \text{ и } w_i \geq 0, \quad i = 0, 1, \dots, M. \quad (6.2)$$

Тогда можно записать плотность (6.1) в виде

$$p_{\boldsymbol{\eta}}(\mathbf{u}) = \sum_{i=0}^M P_i \tilde{p}_i(\mathbf{u}); \quad \tilde{p}_i(\mathbf{u}) = \frac{\chi_i(\mathbf{u})}{Y_i}, \quad Y_i = \int_{R^l} \chi_i(\mathbf{v}) d\mathbf{v}, \quad P_i = H w_i Y_i. \quad (6.3)$$

Для плотностей вида (6.3) имеется алгоритм метода суперпозиции [14], состоящий в выборе номера m по вероятностям $P_i = \mathbf{P}\{m = i\}$ и получении выборочного значения вектора $\boldsymbol{\eta}$ согласно плотности \tilde{p}_m . Таким образом, помимо (6.2) следует требовать, чтобы для плотностей $\{\tilde{p}_i\}$ из (6.3) имелись алгоритмы численного статистического моделирования.

В работе [108] было проведено сравнение различных аппроксимационных базисов $\Xi^{(M)}$ с точки зрения сформулированных требований. Наилучшей, с точки зрения сформулированных выше требований, оказалась конечно-элементная аппроксимация Стренга-Фикса [26], которая строится следующим образом. Для простоты возьмем в качестве U прямоугольный параллелепипед

$$U = \{\mathbf{u} = (u^{(1)}, \dots, u^{(l)}) \in R^l \mid a_k \leq u^{(k)} \leq b_k, \quad k = 1, \dots, l\}.$$

Предположим, что в R^l задана равномерная прямоугольная сетка и каждому узлу \mathbf{x}_i из $X^{(M)}$ можно сопоставить мультииндекс $\bar{j}_{(i)} = (j_{(i)}^{(1)}, \dots, j_{(i)}^{(l)})$ так, что $\mathbf{x}_i = (j_{(i)}^{(1)}h, \dots, j_{(i)}^{(l)}h)$, где h — шаг сетки. Аппроксимация Стренга-Фикса определяется базисом

$$\chi_i(\mathbf{u}) = \chi_{(j_{(i)}^{(1)}, \dots, j_{(i)}^{(l)})}(u^{(1)}, \dots, u^{(l)}) = \chi_{j_{(i)}^{(1)}}(u^{(1)}) \times \dots \times \chi_{j_{(i)}^{(l)}}(u^{(l)}), \quad (6.4)$$

где $\chi_{j_{(i)}^{(m)}}(u^{(m)}) = \chi(u^{(m)}/h - j_{(i)}^{(m)})$, а χ — финитная, одинаковая для всех координат, производящая функция. Как правило, в качестве производящей функции выбирают сплайн $\beta^{(r)}$ порядка r [26]. Здесь уместно заметить, что использование сплайнов имеет большое преимущество с точки зрения эффективной численной реализуемости полученных аппроксимаций, поскольку для случайных величин, имеющих сплайн в качестве плотности, существуют эффективные моделирующие алгоритмы, основанные на том факте, что функция $\beta^{(r)}$ является плотностью распределения случайной величины $\alpha_1 + \dots + \alpha_r + \alpha_{r+1} - (r+1)/2$, где α_i — независимые стандартные (т.е. равномерно распределенные на $[0, 1)$) случайные числа.

В работах [8, 56, 57, 101, 102, 103, 106] в качестве производящей функции использованы сплайны первого порядка (или "функции-крышки"), в этом случае приближение $L_{(M)}f$ из (6.1) называется *мультилинейной аппроксимацией* функции f . Следующее утверждение отражает аппроксимационные свойства приближения (6.1) с базисом (6.4).

Лемма 6.1 [26]. Пусть $f \in C^{p+1}(U)$ и $\chi \in C^p(R)$, тогда найдутся такие коэффициенты w_i в (6.1), что справедлива оценка

$$\rho_{C^s(U)}(f, L_{(M)}f) \leq H_s h^{p+1-s} \|f\|_{C^{p+1}(U)}, \quad 0 \leq s \leq p,$$

где константы H_s не зависят от f и h .

Из этого утверждения можно сделать вывод, что для получения более высокого порядка по h оценки погрешности аппроксимации Стрэнга-Фикса следует выбирать более гладкие производящие функции. Для мультилинейной аппроксимации оптимальный порядок сходимости в лемме 6.1 дают коэффициенты $w_i = f(\mathbf{x}_i)$ [26], и тогда $L_{(M)}f(\mathbf{x}_i) = f(\mathbf{x}_i)$. При этом $\chi = \beta^{(1)} \in C^0(R)$ и для $f \in C^1(U)$ получаем $\rho_{C^0(U)}(f, L_{(M)}f) \leq H_0 h \|f\|_{C^1(U)}$. Справедливости ради отметим, что в [26] показано, что при $f \in C^2(U)$ погрешность мультилинейной аппроксимации $\rho_{C^0(U)}(f, L_{(M)}f)$ пропорциональна h^2 . В случае $r > 1$ выбор подходящих коэффициентов $\{w_i\}$ в (6.1) более сложен. Существуют алгоритмы построения интерполирующей сплайн-функции, т.е. сплайна, проходящего через значения функции в узлах [52]. Однако в многомерном случае не удается получить, как в случае мультилинейной аппроксимации, явных формул для вычисления коэффициентов интерполирующей сплайн-функции. Это затрудняет реализацию таких алгоритмов на ЭВМ и, кроме того, усложняет рассмотрение устойчивости погрешности аппроксимации к возможной ошибке задания значений функции в узлах сетки [107].

Теперь сформулируем свойство "сноса погрешности в узлы" для мультилинейной аппроксимации, которое обосновывает устойчивость мультилинейной аппроксимации к погрешности задания значений функции в узлах.

Лемма 6.2 [106]. Пусть заданы две функции $f, \tilde{f} \in C^0(U)$, тогда для мультилинейной аппроксимации имеет место неравенство

$$\sup_{\mathbf{u} \in U} \rho_{C^0(U)}(L_{(M)}f, L_{(M)}\tilde{f}) \leq \max_{i=0,1,\dots,M} |f(\mathbf{x}_i) - \tilde{f}(\mathbf{x}_i)|.$$

В работе [108] показано, что в одномерном случае при $U = [0, 1] \subset R$ неплохими свойствами аппроксимации и "моделируемости" обладает аппроксимация Бернштейна с базисными функциями

$$\chi_i(u) = C_n^i u^i (1-u)^{M-i}, \quad i = 0, 1, \dots, M; \quad 0 \leq u \leq 1.$$

6.2. Дискретно-стохастические схемы численного интегрирования

Стандартный алгоритм метода Монте-Карло для вычисления интеграла $I = \int_U f(\mathbf{u}) d\mathbf{u}$ (см., например, [14]) основан на представлении I в виде математического ожидания и приближении этого ожидания средним арифметическим:

$$I = \int_U f(\mathbf{u}) d\mathbf{u} = \mathbf{E}\xi \approx \frac{\xi_1 + \dots + \xi_n}{n}, \quad \xi = \frac{f(\boldsymbol{\eta})}{p\boldsymbol{\eta}(\boldsymbol{\eta})}, \quad \boldsymbol{\eta} \sim p\boldsymbol{\eta}; \quad (6.5)$$

здесь $\{\xi_j\}$ — численно реализуемые выборочные значения случайной величины ξ , а знак " \sim " означает распределение по соответствующей вероятностной плотности.

При решении проблемы выбора плотности p_η в алгоритме (6.5) следует минимизировать трудоемкость $S = t\mathbf{D}\xi$, где t — среднее время ЭВМ для получения одного выборочного значения ξ_j [14]. Большинство приемов уменьшения S связано с ограничением дисперсии $\mathbf{D}\xi$. В частности, в случае $f \geq 0$ выбор p_η в виде (6.1) может дать малую дисперсию (выбор плотности p_η , близкой к функции Hf , называется *выборкой по важности* [14]).

В работе [8] рассмотрена выборка по важности с плотностью (6.1) и мультилинейным базисом (6.4) и с помощью лемм 6.1 и 6.2 получена оценка сверху для дисперсии

$$\mathbf{D}\xi \leq \frac{\tilde{H}}{Q} h^4 \|f\|_{W_2^2(U)}. \quad (6.6)$$

Величина $Q = \min_i f(x_i)$ должна быть отделена от нуля. Другой способ применения приближения (6.1) для уменьшения дисперсии $\mathbf{D}\xi$ связан с алгоритмом *выделения главной части* [8], в котором интеграл представляется в виде

$$I = I_1 + I_2; \quad I_1 = \int_U (f(\mathbf{u}) - L_{(M)}f(\mathbf{u})) d\mathbf{u}; \quad I_2 = \int_U L_{(M)}f(\mathbf{u}) d\mathbf{u},$$

причем интеграл I_2 вычисляется аналитически, а для приближения величины I_1 используется алгоритм (6.5). В работе [8] показано, что при применении приближения $L_{(M)}f$ из (6.1) с мультилинейным базисом (6.4) алгоритм выделения главной части дает тот же порядок уменьшения дисперсии по шагу h сетки $X^{(M)}$, что и алгоритм выборки по важности (см. соотношение (6.6)). При вычислении величины I_2 можно использовать более простые плотности (например, плотность равномерного распределения в U) и соответствующие им алгоритмы моделирования, также более простые по сравнению с алгоритмом выборки по важности. В работе [8] предложена также комбинированная численная схема, использующая идеи построения алгоритмов выборки по важности и выделения главной части. Однако сравнительное тестирование, проведенное в работе [108], показало, что комбинированный алгоритм не дает существенного выигрыша по сравнению с алгоритмом выделения главной части, являющегося наиболее эффективным для гладких подынтегральных функций f .

6.3. Функциональные оценки метода Монте-Карло

Приближения функций вида $L_{(M)}f$ из (6.1) используются также при реализации так называемых дискретно-стохастических численных процедур приближения функций φ , заданных в интегральной форме [37, 56, 57, 80,

86, 101, 102, 103, 106, 107]. Основными примерами таких функций являются интеграл, зависящий от параметра

$$\varphi_1(\mathbf{u}) = \int_V k_1(\mathbf{v}, \mathbf{u}) d\mathbf{v}, \quad \mathbf{u} \in U \subset R^{l_1}, \quad \mathbf{v} \in V \subset R^{l_2}, \quad (6.7)$$

а также решение интегрального уравнения второго рода

$$\varphi_2(\mathbf{u}) = \int_V k_2(\mathbf{v}, \mathbf{u}) d\mathbf{v} + f(\mathbf{u}). \quad (6.8)$$

Дискретно-стохастический метод строится следующим образом. Рассмотрим аппроксимацию функции φ вида (6.1):

$$\varphi(\mathbf{u}) \approx L_{(M)}\varphi(\mathbf{u}) = \sum_{i=0}^M w_i \chi_i(\mathbf{u}).$$

Для простоты рассмотрим случай $w_i = \varphi(\mathbf{x}_i)$. Для оценки значений $\{\varphi(\mathbf{x}_i)\}$ реализуем алгоритмы метода Монте-Карло с числами реализаций $\{n_i\}$:

$$\varphi(\mathbf{x}_i) \approx \tilde{\varphi}_{n_i}(\mathbf{x}_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} \xi_j^{(i)}.$$

Окончательно приближение функции φ имеет вид

$$\varphi(\mathbf{u}) \approx L_{(M)}\tilde{\varphi}(\mathbf{u}) = \sum_{i=0}^M \tilde{\varphi}_{n_i}(\mathbf{x}_i) \chi_i(\mathbf{u}). \quad (6.9)$$

При изучении погрешности $\delta^{(B)} = \rho_B(\varphi, L_{(M)}\tilde{\varphi})$ алгоритма (6.9) возникают проблемы выбора соответствующего нормированного функционального пространства B , а также вероятностного смысла стремления случайной величины $\delta^{(B)}$ к нулю с ростом параметров M и $\bar{n} = \min(n_1, \dots, n_M)$. Достаточно подробно разработаны L_2 -подход, в котором строятся верхние границы величины $\mathbf{E}\delta^{(L_2)}$, и C -подход, в котором величина $\delta^{(C)}$ ограничивается сверху по вероятности [80], [106]. В каждом из этих подходов удается разбить погрешность на два слагаемых; например, для C -подхода

$$\delta^{(C)} \leq \rho_C(\varphi, L_{(M)}\varphi) + \rho_C(L_{(M)}\varphi, L_{(M)}\tilde{\varphi}) = \delta_1^{(C)} + \delta_2^{(C)}. \quad (6.10)$$

Слагаемое $\delta_1^{(C)}$ из (6.10) является неслучайным и оценивается на основании аппроксимационных свойств базиса $\Xi^{(M)}$. В монографиях [37, 80] в качестве аппроксимации $L_{(M)}\varphi$ рассмотрено специальное многомерное линейное приближение, а в работах [102, 103, 106, 107] использована упоминавшаяся выше аппроксимация Стренга-Фикса с базисом (6.4) (при этом верхняя граница для $\delta_1^{(C)}$ получается как следствие леммы 6.1).

Второе (стохастическое) слагаемое $\delta_2^{(C)}$ из (6.10) оценивается на основании свойств устойчивости базиса $\Xi^{(M)}$ (для аппроксимации Стренга-Фикса это свойство отражает лемма 6.2) и соответствующих предельных теорем теории вероятностей. Здесь важным является то, какие оценки $\{\xi^{(i)}\}$ в узлах $\{\mathbf{x}_i\}$ используются. Например, для функции φ_1 из (6.7) можно рассмотреть независимые оценки вида

$$\xi^{(i)}(A) = k_1(\mathbf{x}_i, \boldsymbol{\eta}^{(i)})/p_i(\boldsymbol{\eta}^{(i)}) \quad (6.11)$$

(здесь для каждого узла выбирается своя плотность p_i и случайные векторы $\{\boldsymbol{\eta}^{(i)}\}$ независимы для разных i), а также оценки по методу зависимых испытаний [53]:

$$\xi^{(i)}(B) = k_1(\mathbf{x}_i, \boldsymbol{\eta})/\tilde{p}(\boldsymbol{\eta}); \quad (6.12)$$

здесь плотность \tilde{p} одна и та же для всех узлов i , более того, во всех узлах используются одни и те же случайные числа. Аналогом оценок (6.11) для функции φ_2 из (6.8) являются оценки по методу сопряженных блужданий, а для (6.12) — локальные оценки [14].

При оценке стохастической компоненты погрешности $\delta_2^{(C)}$ алгоритма (6.9) с независимыми оценками в узлах сетки используются предельные соотношения теории порядковых статистик, а для метода зависимых испытаний применяется специальная теория функциональной сходимости последовательностей случайных функций [53, 106].

Для многих используемых в приложениях интегральных уравнений второго рода (в том числе нелинейных и стохастических [37, 80, 86]) можно с успехом использовать аппроксимацию решения по многомерному аналогу метода полигона частот [80, 103]. Здесь стохастические оценки в узлах сетки $\{\xi^{(i)}\}$ являются смещенными, т.е. в соотношениях типа (6.10) появляется дополнительная компонента смещения, для которой требуется строить соответствующую верхнюю границу, согласованную с верхней границей детерминированной компоненты $\delta_1^{(C)}$. Кроме того, оценки $\{\xi^{(i)}\}$ оказываются слабозависимыми и при построении верхней границы стохастической компоненты погрешности $\delta_2^{(C)}$ приходится доказывать специальные аналоги утверждений теории независимых порядковых статистик [103].

Важной также является проблема оптимального (согласованного) выбора параметров алгоритма (6.9): числа узлов M сетки $X^{(M)}$ и числа \bar{n} испытаний $\xi_j^{(i)}$ в узлах сетки. Ставится следующая *задача условной оптимизации*: найти минимум трудоемкости $\min_{M, \bar{n}} S(M, \bar{n})$ при $T^{(B)}(M, \bar{n}) = \tilde{\gamma}$, где $\tilde{\gamma}$ — фиксированное положительное число, а $T^{(B)}$ — верхняя граница погрешности в норме функционального пространства B . Общая схема решения этой задачи такова: из соотношения для $T^{(B)}(M, \bar{n})$ один из параметров (например, \bar{n}) выражается через другой (M) и соответствующее выражение

подставляется в выражение для S , при этом получается функция одного переменного (M), которая и исследуется на минимум. Эта задача решена в работах [80, 102, 103] для L_2 -подхода и C -подхода и для различных (зависимых, независимых, слабозависимых) стохастических оценок $\xi^{(i)}$ в узлах сетки.

6.4. Приложения теории дискретно-стохастических методов

Несмотря на то что основы теории комбинированных алгоритмов разработаны относительно недавно в монографии Г.А. Михайлова [80], число примеров успешного применения этой теории растет год от года. В самой монографии [80] представлены примеры успешного применения методик получения верхних границ погрешности и условно-оптимальных параметров в L_2 -метрике для исследования функциональных алгоритмов теории переноса излучения, а также для глобального решения краевых задач. Достижения последнего десятилетия в этих же областях отражены в [37] (см. также обзор литературы в этой монографии). Аналогичные приложения для C -подхода рассмотрены в работах [56, 57, 101]. Методические разработки теории дискретно-стохастических алгоритмов применимы в случае решения сложных систем, включающих уравнение переноса [105], а также при решении нелинейных задач [86].

Литература

- [1] *Аверина Т.А., Артемьев С.С.* Численное решение неустойчивых стохастических дифференциальных уравнений // Сиб. ЖВМ. 2004 (в печати).
- [2] *Антюфеев В.С., Иванов А.И. и др.* Определение аэрозольных индикатрис рассеяния безоблачной атмосферы в спектральной области 0.55–2.4 мкм // Изв. АН СССР. Сер. ФАО. 1980. Т. 16, № 2. С. 146–155.
- [3] *Артемьев С.С.* Численное решение задачи Коши для систем обыкновенных и стохастических дифференциальных уравнений. Новосибирск: Изд. ВЦ СО РАН, 1993.
- [4] *Артемьев С.С., Михайличенко И.Г., Синицын И.Н.* Статистическое моделирование срочных финансовых операций. Новосибирск: ВЦ СО РАН, 1997.
- [5] *Артемьев С.С., Якунин М.А.* Математическое и статистическое моделирование на фондовых рынках. Новосибирск: ИВМиМГ СО РАН, 2003.
- [6] *Бурмистров А.В., Михайлов Г.А.* Вычисление производных от решения стационарного диффузионного уравнения методом Монте-Карло // ЖВМ и МФ. 2003. Т. 43, № 10. С. 1517–1529.
- [7] *Владимиров В.С.* О применении метода Монте-Карло для отыскания наименьшего характеристического числа и соответствующей собственной функции линейного интегрального уравнения // Теория вероятностей и ее применения. 1956. Т. 1, № 1. С. 113–130.

- [8] *Войтишек А.В.* Применение аппроксимации Стренга–Фикса при вычислении многократных интегралов методом Монте-Карло // Сибирский журнал вычислит. математики РАН. Сиб. отделение. 1999. Т. 2, № 2. С. 111–122.
- [9] *Гихман И.И., Скороход А.В.* Стохастические дифференциальные уравнения. Киев: Наукова думка, 1968.
- [10] *Данфорд Н., Шварц Дж.Т.* Линейные операторы (общая теория). М.: ИЛ, 1962.
- [11] *Джсеттыбаев Е.О., Каргин Б.А.* О наблюдении внутренних волн в океане // Иссл. Земли из космоса. 1983. № 1. С. 84–87.
- [12] *Елепов Б.С., Кронберг А.А., Михайлов Г.А., Сабельфельд К.К.* Решение краевых задач методом Монте-Карло. Новосибирск: Наука, 1980.
- [13] *Ермаков С.М., Михайлов Г.А.* Курс статистического моделирования. М.: Наука, 1976.
- [14] *Ермаков С.М., Михайлов Г.А.* Статистическое моделирование. М.: Наука, 1982.
- [15] *Зуев В.Е., Титов Г.А.* Оптика атмосферы и климат. Томск: Изд-во "Спектр", 1996.
- [16] *Калашникова Н.И., Огородников В.А.* Стохастическая интерполяция однородных случайных полей // Методы и алгоритмы статистического моделирования. Новосибирск: ВЦ СО РАН, 1995. С. 40–51.
- [17] *Кантер Р.Р., Пригарин С.М.* Численное моделирование морского ветрового волнения для исследования поля отраженного оптического излучения // Новосибирск: ВЦ, 1989. (Препринт / АН СССР. Сиб. отделение. ВЦ; 829).
- [18] *Каргин Б.А.* О статистическом моделировании взаимодействия оптического излучения с взволнованной поверхностью моря // Моделирование волновых полей (математические проблемы геофизики). Новосибирск: ВЦ СО АН СССР, 1983. С. 25–40.
- [19] *Каргин Б.А.* Статистическое моделирование поля солнечной радиации в атмосфере. Новосибирск: ВЦ СО АН СССР, 1984.
- [20] *Каргин Б.А.* Статистическое моделирование поля оптического излучения в атмосфере и океане // Дис... докт. физ.-мат. наук. Новосибирск: ВЦ СО АН СССР, 1985.
- [21] *Каргин Б.А.* Численная модель поля солнечной радиации в аэрозольной атмосфере // Оптика атмосферы и океана. 1994. Т. 7, № 8. С. 1112–1119.
- [22] *Каргин Б.А., Лаврентьев А.Е., Пригарин С.М.* Система автоматизации статистического моделирования радиационного поля атмосферы // Оптика атмосферы и океана. 1999. Т. 12, № 3. С. 238–245.
- [23] *Кертисс Д.* Методы Монте-Карло для итерации линейных операторов // УМН. 1957. № 5. С. 149–174.
- [24] *Курбанмурадов О.А., Симонов Н.А., Сабельфельд К.К.* Алгоритмы случайного блуждания по границе. Новосибирск: Наука, 1989.
- [25] *Марченко М.А., Михайлов Г.А.* Весовые алгоритмы статистического моделирования диффузионных процессов // Журн. вычисл. матем. и мат. физики. 2003. Т. 43, № 4. С. 571–584.
- [26] *Марчук Г.И., Агошков В.И.* Введение в проекционно-сеточные методы. М.: Наука, 1981.
- [27] *Марчук Г.И., Лебедев В.И.* Численные методы в теории переноса нейтронов. М.: Атомиздат, 1971.

- [28] *Марчук Г.И.* Уравнение для ценности информации с метеорологических спутников и постановка обратных задач // *Космические исследования.* 1964. Т. 2, № 3. С. 462–477.
- [29] *Марчук Г.И., Михайлов Г.А.* О решении задач атмосферной оптики методом Монте-Карло // *Изв. АН СССР. Физика атмосферы и океана.* 1967. Т. 3, № 3. С. 258–273.
- [30] *Марчук Г.И., Михайлов Г.А.* Результаты решения некоторых задач атмосферной оптики методом Монте-Карло // *Изв. АН СССР. Физика атмосферы и океана.* 1967. Т. 3, № 4. С. 394–401.
- [31] *Марчук Г.И., Михайлов Г.А., Назаралиев М.А., Дарбинян Р.А.* Решение прямых и некоторых обратных задач атмосферной оптики методом Монте-Карло. Новосибирск: Наука, 1968.
- [32] *Марчук Г.И., Михайлов Г.А., Назаралиев М.А., Дарбинян Р.А., Каргин Б.А., Елепов Б.С.* Метод Монте-Карло в атмосферной оптике. Новосибирск: Наука, 1976. (Engl. transl.: Springer Verlag, 1980).
- [33] *Михайлов Г.А.* Численное построение случайного поля с заданной спектральной плотностью // *Докл. АН СССР.* 1978. Т. 283, № 3. С. 793–795.
- [34] *Михайлов Г.А.* Моделирование случайных процессов и полей на основе точечных потоков Пальма // *Докл. АН СССР.* 1982. Т. 262, № 3. С. 531–535.
- [35] *Михайлов Г.А.* Приближенные модели случайных процессов и полей // *ЖВМ и МФ.* 1983. Т. 23, № 3. С. 558–566.
- [36] *Михайлов Г.А.* Оптимизация весовых методов Монте-Карло. М: Наука, 1986. (Engl. transl.: Springer-Verlag, 1992.)
- [37] *Михайлов Г.А.* Весовые методы Монте-Карло. Новосибирск: Изд-во СО РАН, 2000.
- [38] *Михайлов Г.А.* Весовые алгоритмы статистического моделирования. Новосибирск: Изд. ИВМиМГ СО РАН, 2003.
- [39] *Михайлов Г.А., Лотова Г.З.* Метод Монте-Карло и простая вероятностная модель для исследования флуктуаций критических значений параметров процесса переноса // *Доклады РАН.* 2004. Т. 395, № 1. С. 22–25.
- [40] *Михайлов Г.А., Медведев И.Н.* Оптимизация весовых методов Монте-Карло по вспомогательным переменным // *Сиб. матем. журн.* 2004. Т. 45, № 2. С. 399–409.
- [41] *Михайлов Г.А., Рогачинский С.В.* Весовые методы Монте-Карло для решения многочастичных задач, связанных с уравнением Больцмана // *Доклады РАН.* 2002. Т. 383, № 3. С. 731–734.
- [42] *Михайлов Г.А., Середняков А.С.* Численные и асимптотические оценки влияния размерности модели стохастической среды на оценки переноса излучения // *Оптика атмосферы и океана.* 1997. Т. 10, № 3. С. 201–210.
- [43] *Михайлов Г.А., Тройников В.С.* Моделирование случайных полей при решении стохастических задач методом Монте-Карло (свойства реализаций) // *Актуальные проблемы вычислит. и прикладной математики.* Новосибирск: ВЦ СО АН СССР, 1983. С. 122–127.
- [44] *Назаралиев М.А.* Статистическое моделирование радиационных процессов в атмосфере. Новосибирск: Наука, 1990.
- [45] *Огородников В.А.* О динамико-вероятностном прогнозе // *Изв. АН СССР. Сер. ФАО.* 1975. Т. 11, № 8. С. 851–853.

- [46] *Огородников В.А.* Моделирование трехмерных полей геопотенциала с заданной статистической структурой // *Методы Монте-Карло в вычислительной математике и математической физике.* Новосибирск: ВЦ СО АН СССР, 1979. С. 73–78.
- [47] *Огородников В.А.* Некоторые свойства оценок пороговых уровней длительных похолоданий // *Методы статистического моделирования.* Новосибирск: ВЦ СО АН СССР, 1986. С. 25–34.
- [48] *Пригарин С.М., Журавлева Т.Б., Воликова П.В.* Пуассоновская модель многослойной разорванной облачности // *Оптика атмосферы и океана.* 2002. Т. 15, № 10. С. 917–924.
- [49] *Протасов А.В., Огородников В.А.* Динамико-вероятностная модель идентификации климатического местоположения источников загрязнения в атмосфере // *Тр. междунар. конференции ENVIRONMENTIS-2002.* Т. 2. Томск: Изд-во ГУ "Томский ЦНТИ", 2002. С. 310–315.
- [50] *Сабельфельд К.К.* Векторные алгоритмы метода Монте-Карло для решения систем эллиптических уравнений второго порядка и уравнения Ламе // *Доклады АН СССР.* 1982. Т. 262, № 5. С. 1076–1080.
- [51] *Сабельфельд К.К., Михайлов Г.А.* О численном моделировании диффузии частиц в случайных полях скоростей // *Известия АН СССР. Физика атмосферы и океана.* 1980. Т. 16, № 3. С. 229–235.
- [52] *Стечкин С.Б., Субботин Ю.Н.* Сплайны в вычислительной математике. М.: Наука, 1976.
- [53] *Фролов А.С., Ченцов Н.Н.* О вычислении методом Монте-Карло определенных интегралов, зависящих от параметра // *ЖВМ и МФ.* 1962. Т. 2, № 4. С. 714–717.
- [54] *Фролов А.С., Ченцов Н.Н.* Решение трех типичных задач теории переноса методом Монте-Карло // *Метод Монте-Карло в проблеме переноса излучений.* М.: Атомиздат, 1967.
- [55] *Хасъминский Р.З.* Устойчивость систем дифференциальных уравнений при случайных возмущениях их параметров. М.: Наука, 1969.
- [56] *Шкарупа Е.В.* Оптимизация метода полигона частот с оценками по пробегу для глобального решения уравнения переноса // *ЖВМ и МФ.* 2003. Т. 43, № 3. С. 440–452.
- [57] *Шкарупа Е.В.* Оценка погрешности и оптимизация функциональных алгоритмов блуждания по решетке решения задачи Дирихле для уравнения Гельмгольца // *Сибирский мат. журнал.* 2003. Т. 44, № 5. С. 1163–1182.
- [58] *Artemiev S.S., Averina T.A.* Numerical Analysis of Systems of Ordinary and Stochastic Differential Equation. Utrecht: VSP, 1997.
- [59] *Black F., Scholes M.* The pricing of options and corporate liabilities // *J. of Political Economy.* 1973. V. 81. P. 637–659.
- [60] *Elliott F. W., Horntrop D. J., Majda A. J.* A Fourier-wavelet Monte Carlo method for fractal random fields // *J. Comp. Phys.* 1997. V. 132, № 2. P. 384–408.
- [61] *Engelhardt Ch., Kurbanmuradov O., Sabelfeld K., Sukhodolov A.* Numerical study of the statistical characteristics of the mixing processes in rivers. WIAS Preprint № 705. Berlin: WIAS, 2001.
- [62] *Ermakov S.M., Nekrutkin W.W., Sipin A.S.* Random Processes for Classical Equations of Mathematical Physics. Dodrecht: Kluwer, 1989.

- [63] *Halton J.H.* A retrospective and prospective survey of the Monte Carlo methods // *SIAM Rev.* 1970. V. 12. P. 1–63.
- [64] *Hotelling H.* Analysis of a complex of statistical variables into principal components // *J. Educ. Psych.* 1933. № 24.
- [65] *Kalos M.H. and Whitlock P.A.* Monte Carlo Methods (V.1: Basics). N.Y.: John Wiley and Sons, 1986.
- [66] *Kargin B.A., Prigarin S.M.* Numerical modelling of random fields in stochastic problems of atmosphere–ocean optics // *Russian J. Numer. Anal. Math. Modell.* 1996. V. 11, № 4. P. 323–341.
- [67] *Kargin B.A., Rakimgulov K.B.* A weighting Monte Carlo method for modelling the optical radiation field in the ocean–atmosphere system // *Russ. J. Numer. Anal. Math. Modelling.* 1992. V. 7, № 3. P. 221–240.
- [68] *Kloeden P. E., Platen E.* Numerical Solutions of Stochastic Differential Equations. Berlin: Springer, 1992.
- [69] *Kolodko A., Sabelfeld K.* Stochastic particle methods for Smoluchowski coagulation equation: variance reduction and error estimations // *Monte Carlo methods and applications.* 2003. V. 9, № 4.
- [70] *Kolodko A., Sabelfeld K., Wagner W.* A stochastic method for solving Smoluchowski's coagulation equation // *Math. Comput. Simulation.* 1999. V. 49, № 1–2. P. 57–79.
- [71] *Kraichnan R. H.* Diffusion by a random velocity field, physics of fluids. 1970. V. 13. P. 22–31.
- [72] *Kurbanmuradov O., Levykin A., Rannik U., Sabelfeld K.K., Vesala T.* Stochastic Lagrangian footprint calculations over a surface with an abrupt change of roughness height // *Monte Carlo methods and applications.* 2003. V. 9, № 2. P. 167–188.
- [73] *Kurbanmuradov O.A., Orszag O.A., Sabelfeld K.K., Yeung P.K.* Analysis of relative dispersion of two particles by Lagrangian stochastic models and DNS methods // *Monte Carlo methods and applications.* 2001. V. 7, № 3–4. P. 245–264.
- [74] *Kurbanmuradov O.A., Sabelfeld K.K.* Lagrangian stochastic models for turbulent dispersion in the atmospheric boundary layer // *Boundary-Layer Meteorology.* 2000. V. 97, № 2. P. 191–218.
- [75] *Kurbanmuradov O., Sabelfeld K., Koluhin D.* Stochastic Lagrangian models for two-particle motion in turbulent flow. Numerical Results // *Monte Carlo methods and applications.* 1997. V. 3, № 3. P. 199–224.
- [76] *Kurbanmuradov O., Sabelfeld K., Smidts O., Vereecken H.* A Lagrangian stochastic model for transport in statistically homogeneous porous media // *Monte Carlo methods and applications.* 2003. V. 9, № 4.
- [77] *Lukin V.L., Mikhailov G.A.* The probabilistic representation and the Monte Carlo methods for solving the first boundary value problem to the polyharmonic equation // *Russ. J. Numer. Anal. Math. Modelling.* 2004. V. 19, № 5. P. 501–510.
- [78] *Mascagni M., Simonov N.A.* The random walk on the boundary method for calculating capacitance // *J. of Comput. Physics.* 2004. V. 195. P. 465–473.
- [79] *Mikhailov G.A.* Optimization of Weighted Monte Carlo Methods. Berlin: Springer, 1992.

- [80] *Mikhailov G.A.* Minimization of computational costs of non-analogue Monte Carlo methods // Series of Soviet and East European mathematics. 1991. V. 5.
- [81] *Muller M.E.* Some continuous Monte Carlo method for the Dirichlet problem // Ann. Math. Stat. 1956. V. 27, № 3. P. 569–589.
- [82] *Nadler B., Singer A.* Ion flow through protein channels: connecting langevin simulations to continuum. IV IMACS seminar on Monte Carlo methods. Berlin, 2003.
- [83] *Ogorodnikov V.A., Prigarin S.M.* Numerical Modelling of Random Processes and Fields: Algorithms and Applications. Utrecht: VSP, 1996.
- [84] *V.A. Ogorodnikov and A.V. Protasov* Dynamic probabilistic model of atmospheric processes and variational methods of data assimilation // Russ. J. Numer. Anal. Math. Modelling. 1997. V. 12, № 5. P. 461–479.
- [85] *Onishuk A.A., Stasio S.Di, Levykin A.I., Strunin V.P., Karasev V.V., Sabelfeld K.K., Panfilov V.N.* The role of dipole interactions in coagulation of silicon agglomerates // J. of Aerosol Science. 2000. V. 31. P. 950–951.
- [86] *Plotnikov M.Yu., Shkarupa E.V.* Error estimation and optimization in C -space of Monte Carlo iterative solution of nonlinear integral equations // Monte Carlo Methods and Applications. 1998. V. 4, № 1. P. 53–70.
- [87] *Pope S. B.* Turbulent Flows. Cambridge: Cambridge Univ. Press, 2000.
- [88] *Prigarin S.M.* Spectral Models of Random Fields in Monte Carlo Methods. Utrecht: VSP, 2001.
- [89] *Rannik U., Aubinet M., Kurbanmuradov O., Sabelfeld K., Markkanen T., Vesala T.* Footprint analysis for measurements over a heterogeneous forest // Boundary-Layer Meteorology. 2000. V. 97, № 1. P. 137–166.
- [90] *Sabelfeld K.K.* Monte Carlo Methods in Boundary Value Problems. N.Y.: Springer, 1991.
- [91] *Sabelfeld K.K.* Stochastic models for coagulation of aerosol particles in intermittent turbulent flows // Mathematics and Computers in Simulation. 1998. V. 47, № 2–5. P. 85–102.
- [92] *Sabelfeld K.K., Kolodko A.A.* Stochastic lagrangian models and algorithms for spatially inhomogeneous Smoluchowski equation // Mathematics and Computers in Simulation. 2003. V. 61. P. 115–137.
- [93] *Sabelfeld K., Kolyukhin D.* Stochastic Eulerian model for the flow simulation in porous media // Monte Carlo Methods and Applications. 2003. V. 9, № 3. P. 271–290.
- [94] *Sabelfeld K.K., Kurbanmuradov O.A.* Stochastic lagrangian models for two-particle motion in turbulent flows // Monte Carlo Methods and Applications. 1997. № 3. P. 53–72.
- [95] *Sabelfeld K.K., Kurbanmuradov O.A.* Coagulation of aerosol particles in intermittent turbulent flows // Monte Carlo Methods and Applications. 2000. V. 6, № 3. P. 211–261.
- [96] *Sabelfeld K.K., Levykin A.I., Shalimova I.A.* Spherical integral equations and discrete stochastic iterative methods // Monte Carlo Methods and Applications. 2005. V. 11, № 1 (в печати).
- [97] *Sabelfeld K.K., Shalimova I.A.* Spherical Means for PDEs. Utrecht: VSP, 1997.
- [98] *Sabelfeld K.K., Shalimova I.A.* Random walk methods for static elasticity problems // Monte Carlo Methods and Applications. 2002. V. 8, № 2.

- [99] *Sabelfeld K.K., Shalimova I.A.* Discrete random walk over large spherical grids generated by spherical means for PDEs // Monte Carlo Methods and Applications. 2004. V. 10, № 3–4.
- [100] *Sabelfeld K.K., Simonov N.A.* Random Walks on Boundary for Solving PDEs. Utrecht: VSP, 1994.
- [101] *Sabelfeld K.K., Shkarupa E.V.* Functional random walk on spheres algorithm for biharmonic equation: optimization and error estimation // Monte Carlo Methods and Applications. 2003. V. 9, № 1. P. 51–65.
- [102] *Shkarupa E.V., Voytishkek A.V.* Optimization of discretely stochastic procedures for globally estimating the solution of an integral equation of the second kind // Russ. J. Numer. Anal. Math. Modelling. 1997. V. 12, № 6. P. 525–546.
- [103] *Shkarupa E.V., Voytishkek A.V.* Convergence of discrete-stochastic numerical procedures with independent or weakly dependent estimators at grid nodes // J. Stat. Plann. Infer. 2000. V. 85. P. 199–211.
- [104] *Thomson D.J.* Criteria for the selection of stochastic models of particle trajectories in turbulent flows // J. Fluid. Mech. 1987. V. 180. P. 529–556.
- [105] *Voytishkek A.V.* Statistical estimation of radiative flows in numerical solution of radiative-conductive heat transfer problems // Russ. J. Numer. Anal. Math. Modelling. 1992. V. 7, № 4. P. 343–369.
- [106] *Voytishkek A.V.* On the errors of discretely stochastic procedures in estimating globally the solution of an integral equation of the second kind // Russ. J. Numer. Anal. Math. Modelling. 1996. V. 11, № 1. P. 71–92.
- [107] *Voytishkek A.V.* Convergence and optimization of smooth discretely stochastic procedures for globally estimating the solution of an integral equation of the second kind // Russ. J. Numer. Anal. Math. Modelling. 1999. V. 14, № 2. P. 177–191.
- [108] *Voytishkek A.V., Kablukova E.G.* Using the approximation functional bases in Monte Carlo methods // Russ. J. Numer. Anal. Math. Modelling. 2003. V. 18, № 6. P. 521–542.

Параллельные вычисления: новые концепции в науке и образовании

В.В. Воеводин, Вл.В. Воеводин

Содержание

1. Теоретические основы	220
2. Вопросы образования	231
3. Распределенная обработка данных	246
Литература	255

1. Теоретические основы

Начало 80-х годов прошлого столетия. Только что Г.И. Марчук организовал в Академии наук Институт вычислительной математики (тогда он еще назывался Отделом) — небольшой мобильный коллектив высококвалифицированных научных сотрудников, способных решать различные большие и сложные задачи, как теоретические, так и прикладные. Ясно, что такие задачи неизбежно должны были ориентироваться на использование передовой вычислительной техники.

В то время вычислительная техника переживала очередной романтический взлет, обещающий фантастические скорости решения задач. А такие вычислительные системы как ILLIAC IV [14, 17], Cray-1 [24, 25] и т.п., уже внедряемые в практику, с особой силой подогревали интерес к перспективам ее развития. Характерной чертой новой техники было одновременное или, как говорят, параллельное использование большого и даже очень большого числа работающих независимо друг от друга отдельных процессоров. Возникло много принципиально новых проблем, связанных не только с конструированием подобной техники [15, 16, 18, 23]. Необходимо было разрабатывать новые языки программирования и новые типы компиляторов. Было неясно, нужно ли создавать новые численные методы или каким-либо способом можно адаптировать уже существующие и т.п. Совершенно

непонятно было, как же все это связано между собой и с теми наработками, которые появились за предшествующие десятилетия развития вычислительной техники. И главное, совсем не виден был ответ на очень важный вопрос о том, что можно ожидать в будущем и с чем может столкнуться математик-вычислитель, решающий и исследующий большие задачи.

Всю совокупность разнообразных проблем, так или иначе относящихся к особенностям конструирования и использования многопроцессорной вычислительной техники, стали называть параллельными или высокопроизводительными вычислениями. С самого начала работы нового института они были объявлены Г.И. Марчуком приоритетным направлением научных исследований.

Поскольку знания в новой предметной области хотелось получить как можно быстрее, казалось естественным обратиться к существовавшей на тот момент научной литературе. Первое "открытие" было неожиданным. Очень скоро выяснилось, что никаких учебников и даже учебных пособий, более или менее систематически описывающих хотя бы основные положения параллельных вычислений, не было ни в нашей стране, ни за рубежом. Более того, стало очевидно, что они не могли появиться и в ближайшее время просто потому, что основы знаний в новой предметной области не были тогда еще сформированы, как, впрочем, не была очерчена и сама предметная область. Стало также ясно, что если мы хотим активно работать в области параллельных вычислений, надо самим начинать создавать научный фундамент и четко понимать, что мы изучаем, как и зачем.

Работа в столь неопределенной области требовала формулирования некоторой целевой проблемы, по которой можно сверять как выбор нужного направления исследований, так и уровень получаемых результатов. Такую проблему предложил Г.И. Марчук. Она стала называться "Отображение задач вычислительной математики на архитектуру вычислительных систем". Отдельные ее аспекты описаны в [11], а сама проблема в общих чертах сводится к следующему.

Известно, что универсальная вычислительная техника работает не очень эффективно на многих задачах именно из-за своей ориентации на универсальность. Поэтому для некоторых частных, но часто используемых задач выгоднее создавать специализированную технику. Она всегда ориентируется на очень узкий класс задач, например, обработку сигналов, реализацию быстрого дискретного преобразования Фурье, выполнение матричных операций и т.п. Однако решение любой задачи из выделенного класса теперь можно осуществить максимально эффективно. Допустим далее, что процесс решения крупной конкретной задачи представлен как последовательность решений каких-то относительно простых задач, эффективно реализуемых на спецпроцессорах. Тогда для заданной задачи можно попытаться построить вычислительную систему как совокупность

взаимосвязанных спецпроцессоров, ориентированных на решение как раз тех простых задач, на которые разложена исходная задача. В этом случае можно надеяться, что с помощью построенной таким способом вычислительной системы можно решать конкретную целевую задачу наиболее эффективно. Говоря другими словами, в проблеме "Отображение задач вычислительной математики на архитектуру вычислительных систем" по существу речь шла о разработке математической модели процесса построения заказных вычислительных систем, ориентированных на крупные вычислительные приложения.

Четверть века назад, когда проблема отображения впервые была сформулирована, она казалась неподъемной, хотя бы потому, что даже неясно было, с какого конца за нее браться. Однако постепенно формировалось убеждение, что ключевым ее моментом должен стать анализ информационной структуры алгоритмов, т.е. структуры, показывающей обмена данными в процессе решения задачи, причем как на макро-, так и на микроуровне, вплоть до обменов между отдельными элементарными операциями. Источником выявления информационной структуры могли быть только программы на алгоритмических языках и математические соотношения, поскольку именно они позволяют описывать алгоритмы наиболее точно.

Так была сформирована и сформулирована фундаментальная дочерняя проблема — анализ информационной структуры алгоритмов и программ. Она стала не только центральной для проблемы отображения, но и, как было установлено в дальнейшем, связующей для большинства областей, относящихся к параллельным вычислениям: разработка языков параллельного программирования, параллелизирующих компиляторов, переносимого программного обеспечения, параллельных численных методов и многих других. Более того, выяснилось, что сведения об информационной структуре алгоритмов могут быть эффективно использованы при проведении исследований в ряде областей, формально никак не относящихся к параллельным вычислениям, например, при изучении быстрых алгоритмов и ошибок округления. Складывается впечатление, что значение сведений об информационной структуре алгоритмов выходит далеко за рамки параллельных вычислений и, возможно, эти сведения являются или должны являться стержнем большинства исследований в области алгоритмов, реализуемых на вычислительной технике.

В это действительно можно поверить, если учесть, что информационная структура описывается графом, порождаемым самим алгоритмом. Множество вершин графа задается множеством выполняемых операций алгоритма, а множество дуг символизирует передачи в процессе реализации алгоритма результатов выполнения одних операций к другим операциям в качестве их аргументов. Данный граф принято называть графом алгорит-

ма. Он представляет информационное ядро алгоритма, очищенное от всех наслоений, вызванных используемыми формами записи алгоритма.

Граф алгоритма имеет очень прозрачный смысл. Поэтому его легко применять в теоретических исследованиях. Однако чтобы этот граф использовать в реальных приложениях, он должен быть явно задан в какой-либо форме, приемлемой для таких целей. На практике он никогда не бывает известен в нужном виде и граф алгоритма приходится находить с помощью специальных методик из описывающих сам алгоритм программ или математических соотношений. Эти методики чрезвычайно сложны, и на их разработку ушло много лет. В настоящее время методики реализованы на персональном компьютере в виде программной системы, получившей название V-Ray system. Система позволяет эффективно исследовать информационную структуру алгоритмов, описанных на языках Фортран и Си, и адаптировать программы под требования компьютеров параллельной архитектуры.

Необходимость разработки этой системы вызывалась рядом обстоятельств. В настоящее время основные вычислительные мощности достигаются на системах с общим числом отдельных процессоров, исчисляемых сотнями и тысячами. Программировать задачи на таких системах очень трудно. Технологии и языки программирования, широко используемые на последовательных компьютерах и векторных суперкомпьютерах с малым числом процессоров, на больших системах могут применяться лишь частично. Особенно трудно переносить на многопроцессорные системы большие программные комплексы, создававшиеся разными людьми в течение долгого времени.

Широкое распространение кластерных систем, а также использование локальных и глобальных сетей для связи отдельных компьютеров между собой привлекает к процессам вычислений большое число пользователей, не имеющих в области параллельных вычислений достаточных знаний. Поэтому для эффективной постановки задач на многопроцессорные системы необходимо иметь различные программные инструментальные средства, существенно облегчающие процесс постановки. Однако для их разработки требуется предварительное проведение весьма тонких теоретических исследований, касающихся информационной структуры алгоритмов и программ, о чем уже говорилось.

Работы по созданию инструментальных средств, поддерживающих процесс написания программ для вычислительных систем параллельной архитектуры, начались почти одновременно с появлением подобных систем. Одними из первых были разработаны в США инструментальные системы Paraphrase, KAP, Forge [22] (см. также <http://www.apri.com>) и другие. Они были ориентированы на преобразование программ, написанных на языке фортран, и для вычислительных систем того времени в какой-то мере вы-

полняли возложенные на них функции. По мере усложнения архитектуры вычислительных систем стали появляться специальные языки параллельного программирования. При этом довольно четко прослеживается общая тенденция: чем сложнее архитектура используемой вычислительной системы, тем более детальную информацию о структуре применяемых алгоритмов должен указать пользователь в своих программах и тем меньшую часть нужной информации ему удастся получить с помощью существующих программных инструментальных средств. Сейчас имеется немало популярных систем и языков параллельного программирования: MPI, Open MP, HPF, DVM, PVM, mp C (<http://www.ispras.ru/mpc>; <http://www.keldysh.ru/dvm>; <http://www.keldysh.ru/norma>; <http://www.mpiforum.org>; <http://www.open-mp.org>.) и т.п. Все они требуют предоставления дополнительных сведений о параллельных свойствах программ. Для многопроцессорных вычислительных систем преодоление всех трудностей по разработке параллельных программ, включая заботу об эффективности их реализации, сейчас возложено на пользователя. Получение нужных дополнительных сведений также возложено на него.

Анализ существующих программных инструментальных средств, облегчающих процессы написания параллельных программ, показывает, что трудности их разработки во многом объясняются трудностями получения и исследования сведений о параллельной структуре реализуемых алгоритмов. Исследование структуры программ на основе анализа их текстов имеет давнюю историю и началось одновременно с созданием первых компиляторов. Различные теоретические результаты получены в России. У их истоков лежат работы А.П.Ершова [10]. Первые результаты по выявлению параллельной структуры программ получены в США [20, 21]. Они были с успехом внедрены в указанные выше инструментальные системы и компиляторы для векторных систем. Сфера применения разработанных методов анализа программ была ограниченной и не позволяла получать сведения, необходимые для использования многопроцессорных систем [26].

Параллельная структура программ является составной частью информационной структуры. Развивая работы по изучению информационной структуры алгоритмов и программ на основе различных форм их записей, мы, безусловно, нацеливались и на автоматизацию процесса получения необходимых сведений о параллельной структуре программ, написанных на традиционно используемых последовательных языках программирования. При этом важным является то обстоятельство, что для выявления требуемых сведений о структуре не привлекаются никакие дополнительные сведения от пользователя. Нужны только тексты программ. На основе разработанной нами теории совместно с Научно-исследовательским вычислительным центром Московского государственного университета и была создана инструментальная программная система V-Ray

для выявления различного рода скрытого параллелизма в алгоритмах и программах.

В связи с графом алгоритма отметим одну проблему, не очень заметную на первый взгляд, которая, тем не менее, на этапе поисковых исследований вызвала огромнейшие трудности. В графе столько вершин, а по порядку и столько же дуг, сколько выполняется элементарных машинных операций за время реализации алгоритма. Для задач, решаемых на компьютерах часами и днями, их число настолько велико, что для прямого описания всех вершин и дуг соответствующего графа алгоритма не хватит памяти самого большого компьютера в мире. Более того, граф не только нужно описать, его нужно и анализировать. В частности, необходимо находить различные ориентированные разрезы, критический путь, какие-то подграфы и т.п. Все такие задачи имеют полиномиальную или даже экспоненциальную сложность. Следовательно, при прямом задании графа любой его анализ, скорее всего, займет гораздо больше времени, чем решение исходной задачи. Такой анализ практически бесполезен. Граф алгоритма приходится находить либо по текстам программ, либо из математических соотношений. Подобные формы записи почти всегда содержат параметры, например, размеры массивов, которые на момент исследования записей не определены. Следовательно, вся работа по нахождению графа и его исследованию неизбежно должна сводиться к решению параметризованных задач. Это очень серьезная трудность.

Тем не менее проблема построения и исследования параметризованного графа алгоритма и других аналогичных ему графов была успешно решена. Удалось выделить представительный класс алгоритмов и доказать, что все их информационные графы могут быть описаны небольшими наборами кусочно-линейных параметризованных функций. Разработаны также быстрые методы вычисления таких функций. Во многом они-то и позволили сделать упомянутую систему V-Ray достаточно эффективной и приемлемой для практического использования.

Примерно в то же время, когда развернулись работы по проблеме отображения, по всем ведущим компьютерным странам мира прокатилась волна конференций, посвященных обсуждению необходимости и возможности построения вычислительных систем нового типа, так называемых систолических массивов [19]. Необходимость объяснялась сложностью, дороговизной и трудностью использования суперкомпьютеров традиционной архитектуры. Возможность опиралась на потрясающие успехи микроэлектроники, приведшие к тому, что элементарные процессоры при массовом производстве могли стать очень дешевыми и миниатюрными.

Теперь представим следующую ситуацию. Допустим, что процессоры, реализующие отдельные операции, конструктивно выполнены в виде правильных многогранников одного размера, на грани которых выводятся ре-

зультаты и подаются аргументы выполняемых операций. Будем строить любые геометрические фигуры из этих многогранников, состыковывая их по граням без зазоров и наложений. Предположим также, что передача информации от одного процессора к другому может осуществляться только через соприкасающиеся грани. При некоторых условиях такие конструкции можно заставить работать, получив тем самым специализированную многопроцессорную систему. Подавая на свободные входы процессоров входные данные, мы будем получать какие-то результаты со свободных выходов. Вычислительные системы подобного типа и являются систолическими массивами, а отдельные процессоры-многогранники принято называть систолическими ячейками.

Сразу видны несомненные достоинства систолических массивов. Главное из них — отсутствие длинных межпроцессорных связей. Поэтому в них нет значительных потерь при передачах данных. Нет в этих системах и общей памяти, что также приводит к снижению коммуникационных затрат. Однако в течение долгого времени оставался открытым вопрос о том, какие же виды алгоритмов можно реализовывать на системах такого типа. Получить на него полный ответ смогли только мы, привлекая для этого различные элементы информационной структуры, включая граф алгоритма. Систолические массивы показались нам интересными, так как они хорошо вписывались в общую концепцию проблемы отображения. При разложении исходной задачи на простые можно попробовать соответствующие простые задачи реализовывать как раз на систолических массивах.

Систолические массивы явились лишь поводом для того, чтобы серьезно заняться решением более сложной задачи. Разлагая исходную задачу на простые, трудно заранее сказать, что будут представлять собой эти простые задачи. Но если какая-то из них будет решаться очень часто, неизбежно возникает желание построить для нее эффективный спецпроцессор, и тогда встанет вопрос о том, как его делать. Поэтому начали строить и изучать математические модели процессов работы системы взаимосвязанных функциональных устройств.

Снова ключевым элементом оказался граф алгоритма. Поместим в каждую вершину графа функциональное устройство, реализующее соответствующую операцию, и будем считать дуги графа линиями связи, по которым результаты работы одних устройств передаются на входы других. Установим на дугах времена передачи информации, а на функциональных устройствах — времена их срабатывания. При некоторых дополнительных условиях такую систему можно заставить работать. Назовем ее граф-машиной. Она вполне подходит для исследования различных реализаций одного и того же алгоритма.

Несомненным достоинством граф-машины является то, что на ней можно реализовать все теоретически мыслимые как последовательные, так

и параллельные режимы, в том числе наискорейшие. Но сразу видны и ее недостатки: каждое из устройств срабатывает только один раз, по каждой линии связи информация передается только один раз и т.п. Поэтому необходимо было разработать какую-либо процедуру преобразования граф-машины, при которой число устройств и связей уменьшалось бы, но сохранялась возможность наискорейшей реализации алгоритма. Такая процедура была построена на основе использования операций простого гомоморфизма графа, т.е. последовательного слияния пар вершин и исключения возникающих кратных дуг и петель. После этих операций получается другая модельная, вообще говоря, параллельная вычислительная система, у которой в зависимости от способа слияния вершин граф может иметь контуры, быть сильно связанным и даже содержать только одну вершину. Назовем ее конвейерным вычислителем.

Процесс функционирования конвейерного вычислителя был изучен весьма основательно для произвольного ориентированного графа связей. Получены различные интересные факты, дано математическое обоснование таким понятиям как загруженность, ускорение и производительность, ряд законов функционирования параллельных систем получил теоретическое подтверждение, некоторые из них были квалифицированы как неверные. Было также выяснено, при каких условиях на граф конвейерного вычислителя алгоритм может быть реализован за теоретически наискорейшее время.

При построении спецпроцессоров, содержащих большое число процессоров, время передачи информации между процессорами обычно намного меньше, чем время выполнения процессорами отдельных операций. Это хорошо видно как раз на примере систолических массивов. Поэтому при проведении исследований нередко считается, что все передачи осуществляются мгновенно. В такой ситуации доказано, что если при гомоморфной свертке граф-машины сливаются лишь вершины, находящиеся на одном пути графа, то на полученном конвейерном вычислителе реализуется весь спектр временных режимов граф-машины, включая наискорейшие. Простейшим способом реализации подобной свертки является проектирование графа на гиперплоскость. Следовательно, однократное или многократное проектирование графа алгоритма есть эффективный способ построения графа систолического массива, реализующего заданный алгоритм. И это дает ответ на вопрос, какие алгоритмы можно реализовать на систолических массивах.

Становление работ по проблеме "Отображение задач вычислительной математики на архитектуру вычислительных систем" заняло несколько лет. Это был трудный, но интересный период, так как приходилось не только разрабатывать методологию решения проблемы, но одновременно формировать и многократно уточнять саму проблему. За время ее станов-

ления было сделано достаточно много, но еще не было выработано четкой позиции по многим вопросам. Размышления об этом периоде нашли свое отражение в книге [2]. В ней дан обзор существовавших на тот момент материалов по параллельным вычислениям, приведен систематизированный список литературы, насчитывающий более 1000 работ.

С тех пор прошло много времени и получено много различных результатов, однако и сейчас проблема отображения остается не менее трудной. В первый период основные сложности были связаны с тем, чтобы найти те ниточки, потянув за которые можно начать вытягивать всю проблему. Теперь же главная трудность состоит в том, чтобы осознать и систематизировать постоянно расширяющееся многообразие открывающихся связей проблемы отображения со смежными областями.

В настоящее время нет другого реального пути создавать вычислительную технику предельно высокой производительности, кроме как выжимать все возможное из идеи параллелизма. Поэтому проблема отображения оказалась самым тесным образом связанной с вычислительными системами параллельной архитектуры. Уровень развития вычислительной техники зависит не только от технических достижений, но и от уровня развития всех связанных с ней областей фундаментальных и прикладных знаний. А в них и сейчас имеется довольно большая неразбериха, что приводит к различным последствиям.

С одной стороны, открываются большие возможности получать совершенно новые знания, что уже отмечалось в связи с информационной структурой алгоритмов и программ. Но эти знания исключительно трудно доводить до практического применения. Потому что, с другой стороны, в параллельных вычислениях давно установились какие-то традиции и эти традиции очень медленно и трудно реагируют на всякие новшества, тем более радикальные. Мы уже говорили о том, что любой язык параллельного программирования требует от пользователя предоставления дополнительной информации об используемых алгоритмах, которая зависит от архитектуры вычислительной системы и языка. Пользователь знает эту информацию далеко не всегда. Отсутствие внятной политики в развитии архитектур вычислительных систем и языков параллельного программирования привело к тому, что программы пользователя часто оказываются не переносимыми с одной системы на другую. Методы исследования информационной структуры могли бы помочь в разрешении данной ситуации. Однако для этого надо, как минимум, объединять усилия различных специалистов в создании необходимых инструментальных средств, договариваться об унификации тех или иных требований и т.п. Пока эта работа идет с большим трудом.

Заметим, что общий характер трудностей, сопровождающих развитие параллельных вычислений, в целом выглядит таким же, каким он был и во

времена последовательных вычислений. Только для параллельных вычислений все трудности проявляются в значительно более острой форме. Во многом из-за большей сложности самой предметной области. Но, возможно, главным образом вследствие того, что к началу активного внедрения вычислительных систем параллельной архитектуры в практику решения больших прикладных задач, в отличие от последовательных машин, не был подготовлен нужный теоретический фундамент.

В самом деле, последовательные вычисления развивались не одну сотню лет. За это время пришло довольно четкое понимание того, что такое последовательный алгоритм. Вокруг данного понятия сформировался большой раздел математики, называемый теорией алгоритмов и изучающий общие свойства последовательных вычислений. Уточненное понятие алгоритма в терминах идеализированных вычислительных машин привело к очень важному понятию машины Тьюринга. По существу этот идеализированный автомат стал теоретическим прообразом первых ЭВМ. Присоединение к машине Тьюринга памяти сделало ее весьма полезным и даже приближенным к реальности инструментом исследований. Но в основе всего лежали последовательные действия. Не удивительно поэтому, что и ЭВМ в течение длительного периода также развивались по пути выполнения именно последовательных действий. Описание заданий для ЭВМ осуществляется с помощью тех или иных языков программирования. Программа для последовательных ЭВМ однозначно определяет последовательность вполне конкретных операций, которые ЭВМ должна выполнить, чтобы реализовать задание. В этом смысле такие программы принципиально ничем не отличаются от программ машины Тьюринга с памятью. Поэтому и реальные языки последовательного программирования за редким исключением оказались похожими друг на друга. Имевшиеся различия чаще всего отражали особенности ЭВМ и области ее применения.

Таким образом, в отношении последовательных вычислений имевшийся теоретический задел во многом определял как структуру первых ЭВМ, так и структуру программного обеспечения. Конечно, развитие и того, и другого проходило не гладко. Постоянно возникавшие различия между ЭВМ и языками программирования приводили к многочисленным трудностям в работе пользователей, особенно в обеспечении эффективности и преемственности создаваемого программного продукта. Однако долгое время эти трудности не казались принципиальными и сохранялась надежда на их устранение, в первую очередь, за счет введения различного рода стандартов в вычислительном деле.

Тем не менее незаметно для многих в вопросах разработки вычислительной техники назревали революционные изменения, приведшие, в конце концов, к радикальному пересмотру всех представлений о вычислениях.

И определяет эти изменения параллелизм, — параллелизм во имя повышения производительности. Пока параллельное выполнение в компьютере любых операций, передач информации и обращений к памяти не приводило к принципиальным изменениям в языках программирования, у пользователей не было особых причин думать о параллелизме. Но в определенный момент параллелизма в компьютере стало столь много, что его присутствие уже нельзя было прикрыть техническими решениями. И тогда от пользователя стали требовать предоставления дополнительной информации о структуре используемых им алгоритмов, требовать как раз для того, чтобы эффективно использовать заложенный в компьютер параллелизм. Как уже отмечалось, пользователь знает эту информацию далеко не всегда, и добывается она с большим трудом.

К началу массового внедрения вычислительных систем параллельной архитектуры сложилась в целом совсем иная ситуация, чем она была в аналогичное время для последовательных ЭВМ. Не было никакой целостной теории параллельных алгоритмов, аналогичной теории алгоритмов для последовательных вычислений. Существовали лишь отдельные разрозненные результаты. Не было даже сколько-нибудь ясного описания, что же нужно понимать под параллельным алгоритмом. И, конечно, не было никакого формального аппарата, который можно было бы назвать параллельным аналогом машины Тьюринга. Отсутствие теоретического фундамента приводило к серьезным издержкам в практической деятельности и не позволяло планировать работу на перспективу.

В заключение отметим два факта: один, касающийся методологии изучения параллельных вычислений, другой, возможно, очень важный в отношении будущих практических разработок.

Мы уже отмечали значение машины Тьюринга для теории и практики последовательных вычислений. Этот формальный автомат позволяет реализовать любой последовательный алгоритм и в определенном смысле является прообразом любой однопроцессорной ЭВМ. Но любая многопроцессорная ЭВМ также может реализовать если не любой, то почти любой алгоритм. Тем не менее в развитии таких ЭВМ и всего того, что их окружает, совсем не ощущается необходимости введения формального автомата, который позволял бы реализовать любой параллельный алгоритм. Однако появилась острая потребность в некоторой идеализированной машине, на которой можно было бы реализовать любой как параллельный, так и последовательный режим для конкретного алгоритма. С возложенными на нее задачами вполне справилась граф-машина. Вообще говоря, не очень понятно, почему же не появляется необходимость введения идеализированной параллельной машины, на которой можно было бы реализовывать любые алгоритмы в любом режиме. Может быть, просто потому, что развитие параллельной вычислительной техники до сих пор осуществляется

в значительной мере стихийно, и пока еще не сформировались устойчивые принципы ее конструирования?

Теперь несколько слов о другом. С помощью системы V-Ray были проведены массовые эксперименты по выявлению структуры информационных графов и, в частности, графов алгоритмов конкретных численных методов. Была обнаружена удивительная закономерность: большое разнообразие существующих методов не приводит к такому же разнообразию их информационных структур. Точнее, многие графы формально совершенно различных алгоритмов оказались изоморфными, отличаясь друг от друга только содержанием вершин и дуг. Поэтому была выдвинута гипотеза о том, что в конкретных вычислительных областях типовых информационных структур немного. Пока практика подтверждает эту гипотезу. Но если гипотеза действительно верна, то возникает интересная перспектива в развитии проблемы отображения. Для конкретных прикладных областей типовые информационные структуры можно определить заранее. Следовательно, можно заранее предвидеть, какую структуру должны будут иметь спецпроцессоры, если для конкретных областей придется создавать заказную вычислительную технику. Но даже если ее не создавать, можно заранее решить многие вопросы распараллеливания алгоритмов и создания типового параллельного программного обеспечения. В любом случае за выдвинутой гипотезой тянется много интересных следствий. Насколько гипотеза верна в реальности, — покажет будущее. Но что-то интересное в ней все-таки есть.

Многое из того, что сделано в области параллельных вычислений, опубликовано в большом числе статей и в нескольких книгах, из которых книга [4] является итоговой. Она охватывает большой временной период, затрагивает все основные стороны параллельных вычислений и является продуктом многочисленных обсуждений состояния и перспектив развития данной области.

2. Вопросы образования

Характерной чертой современного состояния научно-технического прогресса является использование высокопроизводительных систем для решения самых разнообразных и сложнейших задач науки и производства. Расчет летательных аппаратов, анализ изменений климата, изучение белковых соединений, расшифровка генома человека — это лишь отдельные примеры задач, решить которые невозможно без использования современной техники. Список важнейших задач, требующих применения больших и супербольших компьютеров, огромен. Большие и супербольшие компьютеры имеют много особенностей. Они очень сложны и не удобны в использовании. У них нет того дружественного интер-

фейса, который характерен для персональных компьютеров. Для эффективного их применения от пользователей требуется предоставление такой информации о структуре задачи, которая не всегда бывает известна. Несмотря на все это, большие и супербольшие компьютеры остаются сейчас и, по-видимому, останутся в ближайшем будущем единственным техническим средством решения сложных задач науки, техники и производства.

Возрастающий поток сложных задач требует привлечения большого числа новых кадров. Их подготовка в области параллельных вычислений ведется во многих вузах и университетах России. Институт вычислительной математики РАН с первых же лет своего существования активно включился в эту деятельность. Основная его работа в области образования велась раньше и ведется теперь на кафедре математического моделирования физических процессов Московского физико-технического института и на факультете вычислительной математики и кибернетики Московского государственного университета при активном участии Научно-исследовательского вычислительного центра МГУ. Процесс образования в области параллельных вычислений не может быть качественным без использования самых современных вычислительных средств. Поэтому ИВМ РАН и НИВЦ МГУ всегда уделяли большое внимание созданию полноценных учебных классов, оборудованных первоклассной техникой, обеспечению доступа, в том числе удаленного, к большим компьютерам, оснащению своих организаций новейшими высокопроизводительными системами. Например, ИВМ РАН в настоящее время оборудован кластером, состоящим из 16 узлов (2 x Itanium-2, 1.3 GHz, 2GB RAM, сеть Myrinet 2000/Gigabit Ethernet/Fast Ethernet) и имеющим пиковую производительность 83 Gflops.

Но в образовании в области параллельных вычислений не только техническое оборудование является предметом первоочередной заботы. В настоящее время ощущается острая нехватка учебных материалов по данной проблематике. Писать учебники и учебные пособия в области высокопроизводительных вычислений трудно по многим причинам. Главная из них состоит в том, что проблема решения сложных задач на сложной вычислительной технике является комплексной и охватывает такие разные области как изучение архитектуры компьютеров, языков и систем программирования, численных методов, информационной структуры алгоритмов и программ и многое другое. Ни научных, ни тем более учебных материалов, ориентированных на комплексное рассмотрение этой сложной и чрезвычайно запутанной проблемы, до недавнего времени не было. Имелись лишь отдельные публикации, посвященные описанию отдельных фрагментов проблемы. В таком же стиле были представлены и все немногочисленные отечественные и зарубежные образовательные средства по данной проблеме.

Параллельные вычисления представляют относительно новый, но бурно развивающийся раздел прикладной математики и информатики. Еще несколько десятилетий назад к нему имело отношение лишь небольшое число узких специалистов. Однако сейчас, когда суперкомпьютеры, кластеры и распределенные вычислительные системы стали необходимым элементом решения сложных прикладных задач в самых разных областях, с параллельными вычислениями вынуждены знакомиться очень многие. Остро встает вопрос о формах преподавания данного раздела и его содержания. Ситуация меняется настолько быстро, что сложности возникают даже на стадии выбора базовой части знаний. Многое из того, что необходимо было знать еще три-пять лет назад, в настоящее время устаревает и перестает быть актуальным.

В высокопроизводительных вычислениях установившаяся базовая часть знаний относительно невелика. Однако новые знания здесь появляются в огромном количестве. Это во многом определяется их зависимостью от уровня развития вычислительной техники, которая, в свою очередь, изменяется чрезвычайно быстро. По мере осмысления перспектив использования техники какая-то часть новых знаний переходит в базовые. В целом же базовые знания в данной области пока не вышли из стадии формирования и, скорее всего, не выйдут из этой стадии еще очень долго. Как же в таких условиях организовывать образовательные процессы, чтобы, с одной стороны, не заниматься изучением устаревших знаний и, с другой стороны, не упустить зарождающиеся перспективные знания? Выход видится в разумном сочетании использования традиционных образовательных средств в форме учебников и учебных пособий на бумажных носителях и доступных из сети Интернет электронных образовательных средств, оперативно отражающих текущую ситуацию. Именно по этому принципу создавался весь образовательный комплекс, включая и необходимые вычислительные средства. Такой подход согласуется с требованиями федеральной целевой программы "Развитие единой образовательной информационной среды" и межвузовской научно-технической программы Министерства образования РФ "Создание системы открытого образования" [12, 13].

Структурно собственно образовательный комплекс состоит из трех частей. Первую часть составляют учебные пособия периода 1986–1997 гг. Они посвящены четырем основным темам: архитектуре больших вычислительных систем и их классификации, параллельным расширениям языка программирования фортран, математическому аппарату изучения параллельных процессов и методам преобразования программ под требования целевых вычислительных машин. Эта часть может быть по форме отнесена к традиционным образовательным средствам, хотя и в нетрадиционной предметной области. Вторая часть нетрадиционна как по форме, так и по содержанию. Она представляет информационно-аналитический центр в сети

Интернет для оперативного обеспечения пользователей учебной, научной и организационной информацией по самому широкому спектру вопросов, относящихся к высокопроизводительным вычислениям. И, наконец, третья часть — это итоговое фундаментальное учебное пособие "Параллельные вычисления".

Остановимся подробнее на каждой части. С какой бы стороны ни рассматривать высокопроизводительную технику, главным стимулом ее развития было и остается повышение эффективности процессов решения больших и очень больших задач. Эффективность зависит от производительности компьютеров, размеров и структуры их памяти, пропускной способности каналов связи. Все это необходимо знать любому человеку, желающему быстро решить большую задачу. Знакомству со всеми этими вопросами посвящено учебное пособие [6] "Методы описания и классификация архитектур вычислительных систем". В нем в доступной форме рассказывается о конвейерном и параллельном способе организации вычислений, дается описание конкретных больших компьютеров, собраны вместе и детально обсуждаются различные классификации типов вычислительных систем и их архитектур, делается сравнительный анализ высокопроизводительной вычислительной техники.

Эффективность процессов решения задач зависит не только от совершенства техники. Не в меньшей, если не в большей, степени она определяется уровнем развития языков программирования. По мере совершенствования вычислительной техники и усложнения решаемых задач развиваются и языки программирования. Появление параллельных вычислительных систем поставило перед пользователями целый ряд совершенно не характерных для традиционного программирования проблем. Выявление и описание параллелизма вычислений, распределение работы между различными функциональными устройствами и отдельными процессорами вычислительной системы, распределение данных, синхронизация работы параллельно выполняемых частей программы. Эти и другие подобные проблемы надо как-то решать, поэтому возникла необходимость расширения языка добавлением соответствующих средств описания параллелизма. Язык фортран до настоящего времени является одним из самых распространенных языков программирования. Именно поэтому возникло много его расширений применительно к параллельным вычислительным системам и суперкомпьютерам различной архитектуры. Учебное пособие [9] "Современный фортран на перспективных супер-ЭВМ" посвящено описанию всех основных расширений языка фортран и изложению различных особенностей их применения в практическом использовании.

Одним из самых трудных моментов использования высокопроизводительной вычислительной техники и написания программ, учитывающих особенности их архитектуры, является обнаружение и предоставление той

дополнительной информации о структуре задачи или программы, которая обязательно требуется для организации процесса вычислений. С точки зрения процессов образования положение здесь особенно напряженное. С одной стороны, все без исключения языки и системы параллельного программирования заставляют в принудительном порядке предоставлять дополнительную информацию о структуре вычислительных потоков в задачах. Без этой информации компиляторы на больших вычислительных системах либо не работают совсем, либо работают очень не эффективно. С другой стороны, ни в каких курсах, в том числе посвященных устройству компьютеров, численным методам или программированию, ничего не говорится о том, как и откуда получать подобную информацию. Ничего не говорится прежде всего потому, что до недавнего времени соответствующие сведения можно было найти только в каких-то научных публикациях в форме описаний отдельных фактов, причем нередко описаний неполных, противоречивых и даже ошибочных. Учебное пособие [3] "Математические основы параллельных вычислений" ликвидирует данный пробел. В нем разрозненный до этого времени материал систематизирован, приобрел логическую ясность и стал излагаться на базе строгих математических построений.

Ключевым аспектом высокопроизводительных вычислений является эффективность. Эффективность — понятие многоплановое. Это удобство использования техники и программного обеспечения, наличие необходимого интерфейса, простота доступа и многое другое. Но главное — это достижение близкой к пиковой производительности компьютеров. Данный фактор настолько важен, что всю историю развития вычислительной техники и связанных с ней областей можно описать как историю погони за наивысшей эффективностью решения задач.

Эффективность процессов решения задач можно рассматривать, к тому же, как показатель того, насколько хорошо используется далеко не дешевая высокопроизводительная техника в реальной практической деятельности. Эффективность написания программ — понятие очень тонкое. Иногда достаточно переставить в программе два оператора или слегка изменить описание какого-нибудь цикла, и время реализации программы может измениться в десятки, сотни и даже тысячи раз. Обнаружить, что и как надо изменить в программе, чтобы улучшить ее реальную эффективность, — задача исключительно сложная. По трудности, значению и используемому аппарату она сравнима с задачей поиска дополнительной информации о структуре программ и численных методов, предоставляемой компиляторам высокопроизводительной вычислительной техники. Тем не менее с различными преобразованиями программ с целью повышения эффективности их реализации постоянно приходится иметь дело всем пользователям больших вычислительных систем. Ни на практике, ни в теории выполнения таких преобразований в вузах не учат по тем же причинам, о которых говорилось

выше: нет сформировавшихся учебных материалов. Учебное пособие [1] "Информационная структура алгоритмов" — первое пособие, посвященное важнейшей теме преобразования программ.

Вторая часть образовательного комплекса представляет Информационно-аналитический центр в сети Интернет для оперативного обеспечения пользователей учебной, научной и организационной информацией по самому широкому спектру вопросов, касающихся высокопроизводительных вычислений. Она не традиционна как по форме представления материала, так и по его содержанию. Поэтому на этой части остановимся несколько подробнее.

Как уже отмечалось выше, высокопроизводительные вычисления являются не только сложной и запутанной, но и быстро развивающейся областью знаний. Много в ней меняется очень динамично, и уследить за всеми изменениями не так просто. Поэтому необходимо создавать какую-то постоянно действующую информационную службу, которая смогла бы помочь узнавать последние новости и знакомиться с последними разработками. Принимая во внимание широкое развитие телекоммуникационных средств связи, почти очевидно, что подобная служба должна создаваться на основе использования сети Интернет. Ясно, что такая система нужна профессионалам в области высокопроизводительных вычислений. Но она не менее нужна и для совершенствования процессов образования. Подготовка высококвалифицированных кадров, особенно на этапах обучения на старших курсах и в аспирантуре, требует, чтобы будущие специалисты как можно быстрее погружались в профессиональную среду. В противном случае образование оказывается оторванным от потребностей практики, со всеми вытекающими отсюда негативными последствиями. Поэтому исключительно важно, чтобы будущие специалисты уже на стадии обучения начинали пользоваться теми же источниками знаний и теми же инструментариями, что и профессионалы, получая при этом необходимые для образования сведения.

Исходя из указанных соображений, в Научно-исследовательском вычислительном центре МГУ был задуман, разработан и с 1998 года сделан открытым для широкого использования Информационно-аналитический центр по параллельным вычислениям в сети Интернет. Представленная в нем информация доступна круглосуточно и бесплатно по адресу <http://www.parallel.ru>. Информационно-аналитический центр ориентирован на решение трех основных задач в области высокопроизводительных вычислений: обеспечение пользователей оперативной информацией, поддержка учебного процесса и организация взаимодействия всех лиц, интересующихся данной проблематикой. Зайдя на сервер, можно получить свежую информацию о самых мощных компьютерах мира и России, найти более детальную информацию о конкретных компьютерах, познакомиться

ся с последними библиографическими новинками и разработками в области программного обеспечения, узнать о планируемых заседаниях научных семинаров и готовящихся научных конференциях и молодежных школах, получить сведения о ведущих специалистах в области высокопроизводительных вычислений и сделать многое другое.

Особая задача Информационно-аналитического центра связана с поддержкой учебного процесса. На его компьютерах не только решаются научные и производственные задачи, но и проходят практикум студенты различных факультетов. На сервере Информационно-аналитического центра можно найти всю необходимую информацию: состав компьютеров и правила доступа к ним, описание установленного программного обеспечения и систем программирования, списки консультантов по различным вопросам и т.п. Там же имеются материалы к ряду курсов лекций, описания самых популярных языков программирования, материалы по истории компьютеров и истории построения кластеров. Можно также познакомиться с интересными прикладными задачами, требующими применения больших вычислительных систем, с перспективными языками и технологиями программирования. По некоторым курсам на сервере можно получить учебные задания, проверить свои знания и получить оценку.

Информационно-аналитический центр в сети Интернет по высокопроизводительным вычислениям очень быстро стал популярным и по существу превратился в общероссийский центр, объединяющий лиц, интересующихся большими компьютерами и большими задачами. В Центре есть дискуссионный клуб, в котором можно по сети Интернет завязать обсуждение любого вопроса. Организован консультационный совет, где можно получить консультацию от ведущих специалистов по любым проблемам, относящимся к прикладной математике и информатике. Для желающих регулярно получать свежую поступающую информацию организована система подписки. Подписка бесплатная. Подписчиком может стать любой человек или организация, независимо от уровня профессиональной подготовки. Необходимо лишь зарегистрироваться в Информационно-аналитическом центре. Сейчас в нем зафиксировано более 2500 подписчиков, среди которых есть представители практически всех крупнейших университетов и вузов России.

Высокопроизводительные вычисления долгое время не были на острие внимания научной и преподавательской общественности. Мощности однопроцессорных и, в том числе, персональных компьютеров росли достаточно быстро, обеспечивая основные потребности большинства областей в решении текущих задач и поспевая в какой-то мере за ростом их сложности. Интерес к высокопроизводительным вычислениям сохранялся лишь там, где большие задачи исторически возникли давно. Но, в конце концов, случилось то, что и должно было случиться — начался резкий подъем интереса

к большим задачам едва ли не всюду. И тут стали обнаруживаться настораживающие обстоятельства: специалистов в области высокопроизводительных вычислений катастрофически не хватает; в основной своей массе вузы не готовят кадры, необходимые для постановки и решения больших задач; в области высокопроизводительных вычислений нет учебников, учебных пособий и т.д. Главная причина возникновения создавшегося положения связана с тем, что своевременно не был замечен и правильно оценен назревающий революционный с точки зрения технологии решения задач переход к использованию больших многопроцессорных вычислительных систем. В конечном счете, это привело к тому, что к моменту широкого внедрения таких систем в практику использования математические науки оказались без нужного багажа знаний, особенно в части, касающейся структуры алгоритмов. Без нужного багажа знаний оказались и смежные науки, связанные с разработкой языков программирования, компиляторов и архитектуры вычислительных систем. Очевидно, что в этих условиях невозможно было ожидать появления хороших учебных материалов.

Значительные усилия были затрачены на систематизацию и накопление нужных знаний, на устранение в них узких мест, на создание математического аппарата проведения исследований, на формирование логически ясной структуры материала. Об этом подробно говорится в книге [2]. Она не писалась как учебное пособие, но долгое время использовалось как таковое просто в силу отсутствия других приемлемых материалов. В основе каждого учебного пособия лежит предварительное проведение обширных фундаментальных исследований. Например, одно из них [3], посвященное научным основам, было создано только через 10 лет после первых публикаций. В нем изложение различных аспектов высокопроизводительных вычислений поставлено на строгую математическую основу. Собственно говоря, именно с него высокопроизводительные вычисления стали приобретать черты строгости построения, и их стало возможным преподавать как науку, а не как набор не очень ясных рецептов и рекомендаций. Столь же тщательное научное обоснование лежит в основе каждого учебного пособия.

Как описанные учебные пособия, так и Информационно-аналитический центр по параллельным вычислениям в сети Интернет используются в качестве основы для организации учебного процесса в области высокопроизводительных вычислений во многих вузах России. Широкую апробацию они прошли в Московском государственном университете в высшей компьютерной школе, на факультете вычислительной математики и кибернетики, на физическом факультете, на кафедре математического моделирования физических процессов Московского физико-технического института, в Международном государственном университете г. Дубна. С начала 90-х годов по работам цикла читаются специальные и основные курсы

лекций, работают студенческие и научные семинары, защищено 10 кандидатских диссертаций и одна докторская. Несколько диссертаций готовятся к защите.

На многих научных конференциях, семинарах и молодежных школах неоднократно и достаточно остро поднимался вопрос о написании расширенного учебного пособия, охватывающего разные стороны высокопроизводительных вычислений. По следам особенно острых выступлений на одной из таких молодежных школ было принято решение о написании фундаментального учебного пособия [4] "Параллельные вычисления".

Учебное пособие "Параллельные вычисления" является итогом многолетней научной и педагогической деятельности. Оно вобрало в себя как все ранее написанные учебные пособия, так и результаты выполнения многих практических разработок в области высокопроизводительных вычислений. Естественно, был максимально учтен опыт преподавания и подготовки кадров высшей квалификации в различных учебных заведениях России. Для работы с учебным пособием "Параллельные вычисления" нужны минимальные начальные знания. Некоторые разделы доступны даже школьникам, владеющим навыками работы с компьютером. Вместе с тем в учебном пособии приводятся сведения очень высокой степени сложности, достойные внимания самых серьезных исследователей. Каждый читатель может найти здесь материал по своим интересам и своему уровню знаний.

Учебное пособие "Параллельные вычисления" востребовано научным и преподавательским сообществом высшей школы. Оно получило высокую оценку ведущих специалистов в области высокопроизводительных вычислений и стало первым российским учебным пособием в данной области, рекомендованным Министерством образования РФ для студентов вузов, обучающихся по направлению "Прикладная математика и информатика". За создание цикла научно-образовательных изданий для высшей школы "Высокопроизводительные вычисления" авторам этого учебного пособия была присуждена в 2003 г. премия Правительства Российской Федерации в области образования.

Опыт преподавания параллельных вычислений, как и ряда других дисциплин, и не только математических, показывает, что в российском образовании в целом назревают революционные перемены. Оно вступает на качественно новый уровень: решается задача массового использования компьютерных и информационных технологий в общем и профессиональном образовании. Рассматривается проблема создания в России единой для всех образовательных учреждений информационной среды [12, 13]. Много будет зависеть от того, чем наполняется эта среда. Информационная среда имеет много составляющих: компьютеры, линии связи, обслуживающее программное обеспечение, правила и умение работать в среде и многое другое. Но, возможно, самая главная составляющая, если иметь в виду

именно образование, — это электронное представление собственно учебного материала.

Пока работа в России по его созданию идет достаточно вяло. Это проявляется, в частности, в том, что среди общедоступных электронных образовательных средств (учебники, задачки, справочники, энциклопедии и т.п.) доля российских средств по сравнению с зарубежными невелика. Если предположить, что техническое и концептуальное обеспечение информатизации в России будет в ближайшее время сформировано, но ситуация с собственными электронными образовательными средствами радикально не изменится, то российское образование неизбежно окажется под сильным влиянием зарубежных образовательных средств. Зарубежные образовательные средства чаще всего носят справочно-рецептурный характер. Такие средства не дают возможности изучать предмет глубоко и, следовательно, не очень полезны для подготовки высококлассных специалистов. К тому же, если материал не содержит очень большого числа рисунков, схем, таблиц, фотографий и другого изобразительного материала, то характер электронного изложения оказывается мало чем отличающимся от книжного. Все недостатки подобного подхода к построению электронных образовательных средств хорошо видны на примере математических наук, изучаемых в высших учебных заведениях. Причем в тех вузах, где математическое образование является профилирующим.

Математическое профессиональное образование — одно из наиболее ярких достижений российской высшей школы. Однако в современных условиях приходится прилагать немало усилий, чтобы сохранить завоеванные позиции. Как и образование в целом, математика оказалась под жестким давлением времени. Развитие науки заставляет вводить в процесс обучения новые дисциплины. Из-за этого сокращаются и делаются более поверхностными основные математические курсы. Сама математика развивается сейчас настолько бурно, что основные ее достижения не удастся отразить даже в специальных курсах. При этом образовательные средства в математике практически не меняются в течение многих десятилетий: те же лекции, семинары, экзамены, контрольные и т.п. По существу основными носителями знаний по-прежнему являются преподаватель и книга. Естественно возникает вопрос о том, как сделать доступ к профессиональным математическим знаниям более эффективным.

Проиллюстрируем некоторые проблемы создания электронных образовательных средств на примере линейной алгебры и параллельных вычислений, которые представляют в некотором смысле полярные разделы математики. Линейная алгебра является одним из самых устоявшихся ее разделов, по крайней мере, в базовой части. За последние три-четыре десятилетия содержание курсов по линейной алгебре меняется очень мало. Если проанализировать изданные за этот период учебники и учебные по-

собия, то нетрудно заметить, что они мало чем отличаются друг от друга. Несколько меняется терминология, меняется порядок изложения материала, меняются акценты. Но в целом объем излагаемого материала остается почти одним и тем же. Различные варианты учебников и учебных пособий скорее отражают различия во взглядах авторов на характер изложения материала, чем различия в объемах описываемых знаний. Ясно, что на пути подобных модификаций курсов нельзя существенно улучшить уровень усвоения материала. Многолетний опыт работы с молодыми специалистами говорит о том, что даже выпускники элитных вузов, в которых математика является профилирующей областью знаний, имеют весьма посредственные знания по линейной алгебре. Но ведь линейная алгебра является основой всей прикладной математики, и без хорошего ее знания трудно решить качественно любую более или менее сложную прикладную задачу.

Сравнение линейной алгебры и параллельных вычислений, двух разделов математики, показывает следующее. В линейной алгебре базовая часть знаний установилась и является весьма обширной. Новые знания появляются в большом количестве, но они практически никак не затрагивают базовую часть. В параллельных вычислениях установившаяся базовая часть относительно невелика. Новые знания здесь появляются еще в большом количестве, что во многом определяется их зависимостью от уровня развития вычислительной техники. По мере осмысления перспектив использования техники какая-то часть новых знаний переходит в базовые. В целом же базовые знания в параллельных вычислениях пока не вышли из стадии формирования.

Мы уже отмечали выше, что многое зависит от того, как конкретно будут устроены конечные образовательные средства, по которым, собственно говоря, и должно проходить обучение. Подчеркнем также еще раз, что состав и структура именно этих средств мало зависят от всего сопутствующего окружения. На наш взгляд, поиск ответов на многочисленные вопросы, касающиеся электронных образовательных средств, надо начинать совсем с другого вопроса: "А чем конкретно не устраивает книга как источник знаний?" Без обстоятельного ответа на данный вопрос невозможно построить эффективные образовательные средства нового типа. К сожалению, многочисленные публикации и обсуждения этой темы почти всегда сводятся к констатации лишь каких-то достаточно общих методологических положений и оценок. Они содержат мало конкретных предложений. А это, в свою очередь, не позволяет разрабатывать новые формы представления знаний. Во всяком случае, пытаясь создать такие формы для линейной алгебры и параллельных вычислений, мы не смогли обнаружить подходящие конструктивные идеи. Тем не менее на примерах линейной алгебры и параллельных вычислений попробуем понять, чем может не устраивать

книга как источник знаний и чем может оказаться полезным компьютер при выборе форм представления знаний.

Безусловно, использование компьютера очень эффективно в организации различных видов поиска. Но не очень ясно, что стоит искать, если иметь в виду тот материал, который дается в книге. С помощью компьютера легко показывать различные рисунки, фотографии, схемы, таблицы. Но, например, линейная алгебра является в высшей степени абстрактной наукой. Многомерность ее объектов не позволяет в полной мере использовать иллюстративный материал для сопровождения процесса обучения. Конечно, на начальном этапе освоения линейной алгебры какие-то ее положения можно иллюстрировать соответствующими рисунками из аналитической геометрии. Но это только на начальном этапе. И, наконец, компьютер незаменим при организации всякого рода счета. Но в теоретическом курсе линейной алгебры его практически нет. Возможно, что именно отсутствие достаточного числа иллюстративных элементов, не позволяющее видеть предмет целиком как некоторую совокупность взаимосвязанных образов, делает затруднительным изучение абстрактных дисциплин. О попытках найти такие образы в линейной алгебре стоит сказать несколько подробнее. Тем более, что возможность использования найденных образов далеко выходит за рамки линейной алгебры.

Линейную алгебру изучать действительно очень трудно. С целью облегчения процесса ее освоения в свое время был написан справочник [7]. В нем содержалось много самого разного материала, в том числе приводились весьма подробные сведения по теоретическому курсу линейной алгебры. Эти сведения представляют собой большое число утверждений, содержащих описание отдельных фактов и определений. Никаких доказательств не приводилось. Система утверждений была столь детальной и систематизирована столь тщательно и глубоко, что по ней оказалось достаточно легко находить почти любой факт и понятие из общеобразовательного курса линейной алгебры. Возможно, именно этим обстоятельством объясняется тот факт, что справочник оказался популярным в среде студентов и аспирантов, особенно в период подготовки к экзаменам.

Собственно говоря, последнее обстоятельство и стало главным импульсом в принятии решения создать некоторую электронную систему, поддерживающую процесс освоения нового материала, причем создать на новой методологической и технической основе [5]. Выбор линейной алгебры в качестве предметной области оказался для нас совершенно естественным. Этот курс в той или иной мере читается практически в каждом вузе с естественнонаучным уклоном в образовании. Кроме этого, по данному курсу у нас был накоплен достаточно большой как педагогический опыт, так и общий багаж знаний. Важным фактором, определя-

ющим принятое решение, было и то, что у нас имелся квалифицированный коллектив, владеющий современными программными и сетевыми технологиями.

Допустим, что теоретический курс разбит на конкретные утверждения, которые надо освоить в процессе обучения. Пусть они представляют определения, понятия и факты, связывающие какие-то понятия между собой. Построим ориентированный граф. В качестве вершин возьмем отдельные утверждения. Дуги будем проводить следующим образом. Пусть вершина соответствует некоторому понятию. Каждое новое понятие всегда возникает как следствие совместного рассмотрения нескольких уже введенных ранее понятий. Проведем дуги из вершин, иницирующих новое понятие, в соответствующую ему вершину. Каждый новый факт связывает или использует в своей формулировке также несколько введенных ранее понятий. Проведем соответствующие этим понятиям дуги. Кроме этого, в процессе доказательства справедливости конкретного факта могут быть использованы какие-то другие, ранее установленные факты. И это использование отметим дугами. Назовем построенный граф информационным графом предметной области.

Каждый лектор при изложении своего курса неявно следует какому-то информационному графу: выбирает необходимую совокупность понятий и фактов (вершины графа), выстраивает их в логически связную последовательность (дуги графа) и проводит некоторые доказательства. Каждый конкретный информационный граф всегда является ациклическим. Правда, при этом отдельные утверждения могут совпадать. Если проанализировать различные курсы по одной и той же дисциплине, особенно традиционной, то можно заметить следующее. Конкретные совокупности понятий и фактов всегда выбираются из одной и той же известной совокупности и очень редко "разбавляются" чем-то новым. Почти всегда при изложении понятий и фактов сохраняется одно и то же отношение предшествования. Как правило, меняются местами только те сведения, которые на самом деле не связаны друг с другом.

Все сказанное привело нас к такой идее. Построим объединенный информационный граф. Для этого соберем вместе всю совокупность значимых понятий и фактов, используемых в различных курсах по одной и той же дисциплине, и возьмем их в качестве вершин. Установим описанным выше способом дуги. Если окажется, что в какую-то вершину входит большое число дуг, разобьем соответствующее понятие или факт на более мелкие. Будем добиваться таким разбиением того, чтобы в каждую вершину входило лишь очень небольшое число дуг. И, конечно, будем стремиться сделать граф ациклическим. Далее, мы можем пометить вершины графа, указав, например, уровень их сложности, степень общности и т.п. Мы можем присоединить к графу дополнительные вершины, соответствующие

иллюстрирующим примерам, комментариям, доказательствам и т.п., поместив их надлежащим образом.

Имея размеченный информационный граф предметной области, тексты утверждений с различными пояснениями, а также алфавитный, систематизированный и некоторые другие каталоги, характеризующие структуру всей совокупности утверждений, можно решать большое число самых разных задач, полезных как для лекторов курсов, так и для лиц, изучающих эти курсы. По каталогам можно выбрать совокупность утверждений, связанных с заданным набором понятий и фактов. Вместе с дугами это дает информационный граф заданного курса. Сразу видно, какие доказательства в данном курсе можно проводить полностью, а в каких неизбежно будут пропуски. Параллельная форма графа показывает, какие темы не связаны между собой и насколько их изложение можно разносить друг от друга. Сам по себе информационный граф показывает, на что опирается каждое конкретное утверждение и что опирается на него. Это как раз та информация, которую очень трудно получить из обычного книжного текста. Имея информационный граф, структуру предметной области уже легко визуализировать. Для этого компьютерные технологии подходят как нельзя лучше. Работа с информационным графом и каталогами очень многогранна и ее разнообразие ограничивается только фантазией разработчиков соответствующих систем.

На этих идеях сконструирована и реализована базовая электронная энциклопедия ЛИНЕАЛ, предназначенная для получения теоретических сведений в области линейной алгебры. Она рассчитана на широкий круг пользователей — от школьника до научного работника и преподавателя. Включенный в систему материал содержит сведения, заведомо превышающие то, что дается в традиционных курсах. Систему ЛИНЕАЛ легко использовать для изучения линейной алгебры. Все сведения получаются последовательно с самых азов вплоть до очень серьезных результатов. Для работы с энциклопедией необходимо уметь пользоваться персональным компьютером и иметь начальные математические знания. Точнее, надо понимать действия с вещественными числами и представлять, что такое тригонометрические функции. Ничего другого для теоретического освоения линейной алгебры не требуется. Система ЛИНЕАЛ создана в двух вариантах. Она реализована как автономная программная система для персонального компьютера и доступна в сети Интернет по адресу <http://lineal.guru.ru>.

Безусловно, составление информационного графа, текстов утверждений и каталогов является трудной работой, которая под силу лишь профессионалам самой высокой квалификации в конкретной предметной области. Особенно тщательно приходится формировать разбиение на отдельные утверждения. Например, первоначально предполагалось, что в качестве возможного разбиения курса линейной алгебры можно без изменения

взять разбиение, представленное в первой половине материала из книги [7]. Более того, не предполагалось сопровождать факты доказательствами их правильности, как и было сделано в [7]. Однако более детальное рассмотрение предметной области и методов работы с соответствующей ей базой данных заставили внести изменения в первоначальные планы. Поскольку причины внесения изменений не связаны конкретно с линейной алгеброй, а носят общий характер, на них стоит остановиться.

Удобство пользования базой данных во многом определяется простотой структуры причинно-следственных связей. Чтобы ее достичь, как минимум необходимо иметь небольшое число опорных связей для каждого из утверждений. Это означает, что введение в базу данных любого нового факта или определения должно опираться на наличие в базе лишь малого числа утверждений. Материал из книги [7] не удовлетворяет полностью этим требованиям. Попытка построить по нему информационный граф показала, что имеется достаточно много вершин, в которые входит слишком много дуг. Другими словами, доказательства многих утверждений оказываются весьма сложными в силу отсутствия в [7] необходимых промежуточных утверждений. По этой причине материал из [7] пришлось переписать заново. В новом варианте большинство из утверждений имеют 5–8 опорных связей, из которых больше половины относятся к перечислению используемых понятий.

В процессе написания новой совокупности утверждений для курса линейной алгебры стала понятной необходимость введения в базу данных доказательств правильности имеющихся фактов. Включение доказательств было продиктовано несколькими соображениями. Во-первых, они представляют ценный методологический материал, используемый в процессе освоения предметной области. Во-вторых, не все они просты даже для специалистов. И, наконец, наличие доказательств делает возможным использование системы ЛИНЕАЛ в качестве полноценного учебника по любому курсу, соответствующему заданному набору утверждений.

Идеи, заложенные в систему ЛИНЕАЛ, достаточно универсальны. Они применимы к любой предметной области, как естественнонаучной, так и гуманитарной, которую можно представить как совокупность объектов, объединенных логическими связями. Например, так устроен любой математический курс, особенно устоявшийся. Создавать программные оболочки для систем, подобных ЛИНЕАЛ'у, тоже трудно, и для этого тоже нужны высококвалифицированные специалисты, владеющие передовыми программными и сетевыми технологиями. Чтобы исключить ненужное дублирование работ, программная оболочка системы ЛИНЕАЛ сделана не зависящей от предметной области.

Линейная алгебра, ко всему прочему, оказалась хорошим полигоном, на котором удалось не только отработать принципы создания программной

оболочки, но и накопить достаточный опыт работы с подобными системами. Объектная независимость оболочки позволила приступить к созданию аналогичной системы в области параллельных вычислений.

Конечно, в основу структурирования этой предметной области и построения ее информационного графа было положено учебное пособие [4]. Но в параллельных вычислениях содержание вершин и дуг определено не столь четко, как в линейной алгебре, поскольку между отдельными утверждениями теперь не всегда существует точная причинно-следственная связь. Скорее устанавливаемые связи в параллельных вычислениях отражают какие-то предпочтения в порядке изложения или изучения материала. И здесь появилась еще одна сфера использования рассматриваемых информационных систем — построение различных путеводителей по трудно систематизируемым или плохо формализованным областям.

Учебное пособие [4] содержит фактический материал огромного объема. Невозможно было выделить в нем все сколько-нибудь значимое, так как для этого нужно создавать в книге иерархически очень сложную текстовую структуру. Материал с такой структурой читается на бумажном носителе с большим трудом. Поэтому пришлось представить его традиционным образом. При этом приходится сознавать, что большое число каких-то полезных идей, акцентов и даже отдельных фактов может просто затеряться где-то между строк. Однако ничто не мешает, используя уже разработанную программную среду, создать над учебным пособием информационную надстройку-путеводитель, где весь материал будет отражен со всеми необходимыми деталями и акцентами.

Создание Информационно-аналитического центра, базовой электронной энциклопедии по линейной алгебре и путеводителя над учебным пособием по параллельным вычислениям [4] показывает, насколько велики возможности в создании нетрадиционных электронных образовательных средств.

3. Распределенная обработка данных

Удивительно быстро развивается вычислительная техника. Еще недавно производительность в один гигафлоп прочно ассоциировалась с суперкомпьютерами, а сегодня это обычная характеристика обычной "персоналки". Для того чтобы иметь представление о современных компьютерах с максимальной производительностью, два раза в год формируется список пятисот самых мощных компьютеров мира. В ноябре 2003 года опубликована 22-я редакция списка, определившая лидеров нашего времени. На первом месте находится японский компьютер Earth Simulator производства фирмы NEC с производительностью на тесте Linpack 35.86 Тфлопс. Вторую позицию занимает компьютер ASCI Q производства Hewlett-Packard

с производительностью на тесте Linpack 13.88 Тфлопс. На третьем месте в новом списке оказался кластер X, состоящий из 2200 процессоров Apple G5 2.0 GHz, соединенных коммуникационной сетью InfiniBand. Данный кластер стал третьим компьютером, преодолевшим рубеж 10 Тфлопс, показав на тесте Linpack производительность 10.28 Тфлопс. На данный момент в списке находится 131 система с производительностью на тесте Linpack более 1 Тфлопс (в предыдущей редакции списка было лишь 59 систем). Суммарная производительность 500 компьютеров из списка возросла за полгода с 375 Тфлопс до 528 Тфлопс. Для того чтобы попасть в список, на данный момент требуется производительность более 403.4 Тфлопс (по сравнению с 245.1 Тфлопс в июне 2003 года). Фантастика . . .

Интересно, что лидер нынешнего списка компьютер Earth Simulator держится на своей позиции уже второй год. В нем все уникально: архитектура, характеристики, возможности. Его создание, безусловно, стало возможным только в результате выполнения проекта действительно общенационального масштаба. Производство компьютера было закончено в феврале 2002 года, после чего он был установлен в японском центре морских наук и технологий (Japan Marine Science and Technology Center). Сам компьютер и все его технологическое окружение (системы электропитания, кондиционирования, освещения, сейсмозащиты и т.п.) занимают здание размером 50 x 65 x 17 метров.

Earth Simulator содержит 640 процессорных узлов, соединенных между собой через высокоскоростной переключатель. В состав узла входят 8 векторных арифметических процессоров, работающих над общей для каждого узла оперативной памятью, коммуникационный процессор и процессор для операций ввода/вывода. Оперативная память каждого узла разделена на 2048 банков и имеет объем 16 Гбайт. Пиковая производительность одного арифметического процессора равна 8 Гфлопс, поэтому пиковая производительность всего компьютера, объединяющего 5120 процессоров, равна 40 Тфлопс. Попутно заметим, что производительность на тесте Linpack составляет 89,6% от теоретического максимума — это также рекордный показатель среди участников списка Top500.

И, наконец, еще одной отличительной особенностью компьютера является то, что его архитектура объединяет все три базовых принципа построения высокопроизводительных систем. Если рассматривать его в целом, то Earth Simulator является массивно-параллельным компьютером с распределенной памятью. Вместе с тем каждый процессорный узел построен на принципах SMP-архитектуры, причем основа каждого процессора это векторно-конвейерная обработка. Все лучшие качества основных классов современных высокопроизводительных систем были использованы для создания безусловного лидера — суперкомпьютера Earth Simulator.

Но лидер всегда один и уникален, а что же сегодня чаще всего используется на практике? В настоящее время самый динамичный сегмент принадлежит кластерным системам. Общее число кластеров в 22-й редакции списка Top500 продолжает быстро увеличиваться и составляет уже 208 систем (149 в июне 2003 года), причем в первую десятку наиболее мощных компьютеров сейчас входит 7 кластеров.

Если попытаться определить коротко, то вычислительный кластер — это совокупность компьютеров, объединенных в рамках некоторой сети для решения одной задачи. В качестве вычислительных узлов обычно используются доступные на рынке однопроцессорные компьютеры, двух- или четырехпроцессорные SMP-серверы. Каждый узел работает под управлением своей копии операционной системы, в качестве которой чаще всего используются стандартные ОС: Linux, NT, Solaris и т.п. Состав и мощность узлов может меняться в рамках даже одного кластера, давая возможность создавать неоднородные системы. Выбор конкретной коммуникационной среды определяется многими факторами: особенностями класса решаемых задач, доступностью финансирования, необходимостью последующего расширения кластера и некоторыми другими. Возможно включение в конфигурацию специализированных компьютеров, например, файл-сервера, и, как правило, предоставляется возможность удаленного доступа на кластер через Интернет.

Простор для творчества при проектировании кластеров огромен. Рассматривая крайние точки, кластером можно считать как пару ПК, связанных локальной Ethernet-сетью, так и упоминавшийся выше кластер X, занимающий третье место в списке Top500.

По кластерному пути развития вычислительной техники пошли и в Научно-исследовательском вычислительном центре МГУ в 1999 году. Именно тогда было принято решение сделать основную ставку на кластерные системы, и время показало, что это решение оказалось абсолютно правильным. Первым построили кластер на основе 24-х процессоров Intel Pentium-III/500 MHz с использованием сети SCI в качестве вычислительной коммуникационной среды. Через полгода решили провести модернизацию и добавить еще шесть узлов (12 процессоров). И вот здесь, казалось бы, на столь тривиальном действии, нас поджидала неожиданность — узлов на базе использованных процессоров и нужных нам материнских плат в продаже уже не было. Найдя вариант на базе Intel Pentium-III/550 MHz, проблему модернизации кластера решили введением легкой неоднородности, но одновременно появилось четкое представление о серьезной задаче на ближайшее будущее: процессоры быстро совершенствуются, значит, неоднородность кластеров НИВЦ МГУ будет столь же быстро расти. Можно ли будет в таких условиях использовать весь потенциал вычислительных ресурсов суперкомпьютерного центра или мы будем вынуждены для каждой

задачи выделять лишь некоторое число процессоров какого-либо одного кластера?

Задача оказалась исключительно актуальной. Ниже показана современная структура вычислительного комплекса НИВЦ МГУ:

- Кластер SCI, 36 CPUs, 18 узлов: 2 x Intel Pentium III/500 MHz, 1GB, HDD 3.2 GB, Fast Ethernet + SCI, 18 Gflops,
- Кластер SKY, 40 CPUs, 20 узлов: 2 x Intel Pentium III/850 MHz, 1GB, 2 x HDD 15 GB, Fast Ethernet, 34 Gflops,
- Кластер AQUA, 40 CPUs, 24 узла: 2 x Intel Pentium III/1 GHz, 1GB, HDD 20 GB, Fast Ethernet, 48 Gflops,
- Кластер LEO, 32 CPUs, 16 узлов: 2 x Intel Xeon 2.6 GHz, 2GB, 2 x HDD 40 GB, Fast Ethernet + SCI (D335), 166 Gflops,
- Кластер NEW, 160 CPUs, 80 узлов: 2 x Opteron 248 2.2 GHz, 3GB, HDD 36 GB, InfiniBand+Gigabit Ethernet, 704 Gflops.

Заметим, что кластеры SKY и AQUA могут работать как порознь, так и в режиме объединенного кластера.

Как обеспечить преємственность вычислительной техники для тех пользователей, которым необходимы максимально возможные ресурсы? Как сделать так, чтобы появление новых систем не приводило к забвению старых, а просто давало пользователю приращение мощности за счет добавления новых процессоров к уже имеющемуся парку компьютеров?

Все первые параллельные вычислительные системы были уникальны: оригинальная архитектура, последние достижения микроэлектроники, специальное программное обеспечение, специализированные сервис и поддержка. Но уникальность означает высокую цену и недоступность. Сегодня ситуация резко изменилась. Вычислительный кластер можно собрать в любой лаборатории, отталкиваясь от реальных потребностей в вычислительной мощности и доступного бюджета. Для класса задач, где не предполагается тесного взаимодействия между параллельными процессами, кластерное решение даже на основе обычных персональных компьютеров и сети Fast Ethernet будет достаточно эффективным. А чем такое решение принципиально отличается от локальной сети современной организации? С точки зрения прикладного программиста — почти ничем. Если у него появится возможность использования подобной сети для решения своих задач, то для него такая конфигурация и будет параллельным компьютером.

Одним из ключевых факторов, определяющих развитие технологий параллельной и распределенной обработки данных, является распространение сетевых технологий и сети Интернет. Интернет можно рассматривать как самый большой параллельный компьютер, состоящий из множества компьютеров сети. Своего рода метакомпьютер, к которому всегда был и

будет особый интерес, поскольку никакая отдельная вычислительная система не сравнится по своей мощности с потенциальными возможностями Глобальной сети. Главное — это научиться эффективно использовать ее потенциал.

Конструктивные идеи использования распределенных вычислительных ресурсов для решения сложных задач появились относительно недавно. Первые прототипы реальных систем метакомпьютинга стали доступными с середины 90-х годов. Некоторые претендовали на универсальность, часть систем была сразу ориентирована на решение конкретных задач, где-то ставка делалась на использование выделенных высокопроизводительных сетей и специальных сетевых протоколов, а где-то за основу брались обычные каналы и протокол HTTP.

Проанализировав работы в данной области, можно констатировать, что реальная работа по созданию и апробации систем метакомпьютинга сегодня активно идет по трем направлениям [4]. Первое направление — это создание универсальных метакомпьютерных сред. Практически все основные производители программного обеспечения (Oracle, IBM, HP, Sun и др.) работают в данном направлении. Многие берут в качестве стандарта Globus, создавая программную инфраструктуру для своих платформ. На основе этого же пакета формируются глобальные полигоны, объединяющие высокоскоростными сетями значительные распределенные вычислительные ресурсы. Второе направление получается из первого, если универсальность среды заменить четкой ориентацией на конкретные задачи. Речь идет о создании специализированных метакомпьютерных сред для решения небольшого набора многократно используемых "тяжелых" вычислительных задач. Такая постановка намного более реалистичная, поскольку специфика задачи известна заранее, что помогает спроектировать эффективную среду для ее решения. Третье направление состоит в разработке инструментария для организации распределенных вычислений.

Безусловно, универсальные среды являются перспективным направлением, но они появятся не скоро. В настоящее время не ясны даже принципы их использования. Globus Toolkit, стандарт де-факто, слишком тяжел в установке и сложен в использовании. А что делать, если 2000 компьютеров организации Вам могут отдать лишь на ночь или на два выходных дня? А если администраторы не хотят устанавливать ничего лишнего на свои компьютеры? Нужен простой инструментарий, который помог бы быстро создавать распределенные приложения и использовать доступные вычислительные ресурсы. По такому пути пошли несколько лет назад в НИВЦ МГУ, отрабатывая различные технологии организации и проведения распределенных вычислительных экспериментов. Для их осуществления была разработана специальная система, которая получила название X-Com [8].

Основные требования к проектируемой системе X-Com были сформулированы следующим образом:

- ориентация на вычислительные задачи;
- работа через Интернет, возможность использования всех доступных в сети вычислительных ресурсов различной мощности;
- минимум дополнительных действий и системного вмешательства на используемых ресурсах;
- масштабируемость системы, устойчивость к неоднородности и изменению конфигурации вычислительной среды;
- простота адаптации прикладных программ.

К настоящему времени система прошла апробацию в ходе нескольких масштабных экспериментов с использованием широкого спектра ресурсов: от простых домашних компьютеров до мощных параллельных вычислительных систем, от использования Глобальной сети до работы на выделенном гигабитном полигоне Московского университета. В каждом случае на подготовку и разворачивание распределенного эксперимента требовалось не более 1 часа.

Совместно с группой специалистов из Центра "Биоинженерия" РАН в распределенном режиме решалась задача определения скрытой периодичности в генетических последовательностях. Основную сложность в данном случае представляли высокая вычислительная сложность алгоритмов и огромный объем входных данных. По самым скромным подсчетам для обработки материала, имеющего реальное научное значение, на одном процессоре потребовалось бы обработать несколько гигабайт входной информации и многие месяцы, а то и годы непрерывных вычислений. На вход программы подавались описания структуры генетических последовательностей, закодированные стандартным набором символов a-t-g-c, а на выходе появлялась информация о найденных во входных последовательностях повторах тех или иных фрагментов.

Все входные данные были взяты из специализированного банка данных — центрального хранилища актуальной информации о генетической структуре различных организмов. Во всех расчетах были задействованы географически удаленные вычислительные ресурсы, объединенные с помощью системы X-Com.

Для решения исходной задачи в среде X-Com все входные последовательности разбивались на порции, содержащие от нескольких десятков до нескольких сотен тысяч символов. Часть порций нарезалась с "перехлестом". В этом случае конец предыдущей порции совпадал с началом последующей, что было необходимо для нахождения повторов в самой зоне разрезания. Количеством символов в порции можно было регулировать объем вычислений, необходимый для ее обработки в разумное время. В

проведенных экспериментах время обработки одной порции менялось в пределах нескольких десятков минут. Для сокращения сетевого трафика выполнялась online-компрессия всех передаваемых данных. Общий объем написанного для интеграции с системой X-Com дополнительного кода составил менее 200 строк.

Проводя серию экспериментов, одновременно с решением основной прикладной задачи нами преследовались две дополнительные цели: проверить потенциальные возможности системы X-Com и определить особенности ее работы при пиковых нагрузках. Остановимся подробнее на описании двух наиболее интересных экспериментов.

Цель первого эксперимента — географический размах: 14 удаленных вычислительных систем, принадлежащих 10 организациям, расположенным в 8 городах и 6 различных часовых поясах. В разные моменты времени к эксперименту подключалось 407 процессоров, причем максимальное число одновременно работающих процессоров составило 385, а в среднем по эксперименту эта величина оказалась равной 296. Общее время эксперимента составило около 64 часов. Один средний компьютер справился бы с подобной задачей лишь за два года непрерывной работы. Взаимодействие между компьютерами шло по обычным каналам связи, и суммарно было передано 9 Гбайт данных. Максимальная загрузка исходящего канала данных от центрального сервера составила около 7 Мбит/с. Никаких специальных действий, направленных на обеспечение гарантированной полосы пропускания или увеличение стабильности связи, не предпринималось.

В таблице 1 показаны вычислительные системы, работавшие в этом эксперименте. Число выполненных каждой системой заданий определялось как их мощностью, так и длительностью участия в эксперименте. Так, кластеры Дубны и Переславля-Залесского работали все время, а системы из Твери и Уфы подключались лишь на несколько часов.

Несмотря на значительную географическую удаленность использованных вычислительных систем, эффективность работы системы в целом превысила 95% (отношение суммарного времени полезных вычислений на узлах к общему времени использования данных узлов в эксперименте).

Задача второго эксперимента — работа в условиях максимальной неоднородности. Организаторы эксперимента: НИВЦ МГУ и Межведомственный суперкомпьютерный центр (МСЦ). В расчете участвовали четыре типа вычислительных ресурсов: отдельные компьютеры, учебный класс, традиционные вычислительные кластеры и суперкомпьютер МВС-1000М. Часть компьютеров работала под управлением Windows, а часть — под Linux (как правило, Red Hat). На одной части компьютеров использовались процессоры Intel, а на другой — Alpha. Одновременно в экс-

Таблица 1. Вычислительные ресурсы "географического" эксперимента

Организация и кластер	Город	Платформа	Число процессоров	Выполнение заданий
ЛИТ ОИЯИ. DN1	Дубна	Linux / Intel	16	3824
ЛИТ ОИЯИ. DN2	Дубна	Linux / Intel	16	3712
ИММ УрО РАН	Екатеринбург	Linux / Intel	16	2108
НИИЯФ МГУ	Москва	Linux / Intel	44	4270
НИВЦ МГУ. AQUA	Москва	Linux / Intel	40	10464
НИВЦ МГУ. SCI	Москва	Linux / Intel	32	3366
НИВЦ МГУ. SKY	Москва	Linux / Intel	40	7404
Офисные компьютеры	Нью-Йорк	Windows / Intel	5	834
ИПС РАН	Переславль-Залесский	Linux / Intel	32	11904
Физфак МГУ	Москва	Linux / Intel	10	488
ИМПБ РАН	Пущино	Linux / Intel	40	3656
ГТУ	Тверь	Linux / Intel	16	45
УГАТУ	Уфа	Linux / Intel	64	490
ВМиК МГУ	Москва	Linux / Intel	36	3067

перименте применялись три различных режима использования ресурсов (табл. 2):

- МОНО — монопольное использование, при котором вычислительная система полностью отдавалась под проведение данного эксперимента;
- ОЧЕР — работа через стандартные системы управления очередями заданий. Данный режим не требует менять политику администрирования и использования вычислительной системы при ее подключении и после ее выхода из эксперимента;
- ЗАН — режим "по занятости" позволяет использовать свободные ресурсы вычислительной системы.

С помощью специальных алгоритмов система X-Com определяет, что узел в данный момент не используется, и запускает на нем прикладную задачу. Как только любая другая программа на узле проявляет активность, X-Com приостанавливает свою работу.

Эксперимент продолжался 14,5 часов. Среднее число одновременно работающих процессоров составило 125 (максимум зафиксирован на значении 150). Всего к эксперименту подключалось 437 различных процессоров. Казалось бы, слишком много, учитывая среднее число 125. Но объясняется

Таблица 2. Вычислительные ресурсы "неоднородного" эксперимента

Кластер	Платформа	Режим	Число выданных заданий	Число обработанных заданий	Число оборванных сессий	Доля оборванных сессий, %
МСЦ. КАРРА	Linux / Intel	МОНО	2485	2476	9	0.36
МСЦ. МВС-1000М	Linux / Alpha	ОЧЕР	3203	2796	407	12.70
НИВЦ. AQUA	Linux / Intel	МОНО	9689	9645	44	0.45
НИВЦ. SCI	Linux / Intel	ЗАН	501	467	34	6.78
НИВЦ. SKY	Linux / Intel	ЗАН	5233	5184	49	0.93
НИВЦ. Уч.классы	Windows / Intel	МОНО	1642	1599	43	2.61

это очень просто. Суперкомпьютер МВС-1000М объединяет 768 процессоров, и система управления заданиями при распределении программы на реальные процессоры не обязана учитывать, куда были распределены предыдущие варианты этой же программы.

Всего за время этого эксперимента было передано 26,8 Гбайт данных. Время обработки одной порции данных менялось от 1 секунды до 10 минут. Несмотря на значительную коммуникационную нагрузку и специально созданную неоднородность, эффективность работы всей системы составила 92,9%. В таблице 2 приводятся характеристики вычислительных систем, режимов их использования и агрегированные данные хода вычислений.

Естественно, что наиболее эффективно используются системы в монопольном режиме. Небольшие накладные расходы, возникшие при использовании режима "по занятости", во многом объясняются малым изменением вычислительной нагрузки на соответствующих кластерах во время данного эксперимента. Самый большой процент оборванных сессий оказался при использовании стандартных систем управления очередями. Объясняется это тем, что при постановке в очередь система заказывает для программы некоторое время. После того как программу распределяют на реальные узлы, она начинает штатную обработку порций данных и продолжает это до того момента, пока не исчерпает свой лимит по времени, после чего всегда происходит разрыв сессии между узлом и сервером.

Проведенные эксперименты убедительно показали, насколько эффективно можно решать большие задачи с использованием существующих рас-

пределенных вычислительных ресурсов. Конечно же, X-Com — не универсальна, да и не может быть универсальным решением, одинаково хорошо соответствующим любой задаче и любой вычислительной среде. Но то, что она является одним из эффективных и реально применимых на практике подходов — это доказано. Что использовать в каждом случае, определяется, прежде всего, самим пользователем и стоящими перед ним задачами.

Сейчас важно осознать, что в неоднородных распределенных вычислительных средах заложен колоссальный потенциал, и очень важно не опоздать с его освоением.

Литература

- [1] *Воеводин В.В.* Информационная структура алгоритмов. М.: Изд-во МГУ, 1997.
- [2] *Воеводин В.В.* Математические модели и методы в параллельных процессах. М.: Наука, 1986.
- [3] *Воеводин В.В.* Математические основы параллельных вычислений. М.: Изд-во МГУ, 1991.
- [4] *Воеводин В.В., Воеводин Вл.В.* Параллельные вычисления. СПб.: БХВ-Петербург, 2002.
- [5] *Воеводин В.В., Воеводин Вл.В.* ЛИНЕАЛ: Электронные образовательные средства: новые идеи // Вычислительные методы и программирование. 2003. Т. 4. С. 1–6.
- [6] *Воеводин Вл.В., Капитонова А.П.* Методы описания и классификации архитектур вычислительных систем. М.: Изд-во МГУ, 1994.
- [7] *Воеводин В.В., Кузнецов Ю.А.* Матрицы и вычисления. М.: Наука, 1984.
- [8] *Воеводин Вл.В., Филамофитский М.П.* Суперкомпьютер на выходные // Открытые системы. 2003. №5. С. 43–48.
- [9] *Воеводин Вл.В., Церетели П.А.* Современный фортран на перспективных супер-ЭВМ. М.: Изд-во МГУ, 1995.
- [10] *Ершов А.П.* Современное состояние теории схем программ // Проблемы кибернетики. 1973. №27. С. 87–110.
- [11] *Марчук Г.И., Котов В.Е.* Проблемы вычислительной техники и фундаментальные исследования // Автомат. и вычисл. техн. 1979. №2. С. 3–14.
- [12] Международная конференция "Информационные технологии в открытом образовании" // Материалы конференции. М.: МЭСИ, 2001.
- [13] Федеральная целевая программа "Развитие единой образовательной информационной среды". Основные направления развития образовательных электронных изданий и ресурсов // Материалы научно-практической конференции. М.: РМЦ, 2002.
- [14] *Feierbach G., Stevenson D.* The ILLIAC IV // Supercomputers. 1979. V. 2. P. 77–92.
- [15] *Ginsberg M.* Some observations on supercomputer computational environments // Parallel and large-scale computers: performance, architecture, applications. Amsterdam: North-Holland, 1983. P. 173–184.

- [16] *Gottlieb A., Schwartz J.T.* Networks and algorithms for very-large-scale parallel computation // *Computer*. 1982. V. 15, №1. P. 27–36.
- [17] *Kuck D.J.* ILLIAC IV software and application programming // *IEEE Trans. Comput.* 1968. V. C-17, №8. P. 758–770.
- [18] *Kuck D.J. et al.* The effects of program restructuring, algorithm change and architecture choice on program performance // *Proc. Int. Conf. Parallel Process.* N.Y.: IEEE, 1984. P. 129–138.
- [19] *Kung H.T.* Why systolic architecture? // *Computer*. 1982. 15, №1. P. 37–46.
- [20] *Lampert L.* The coordinate method for parallel execution of DO loops // *Proc. 1973 Samagore Comput. Conf. Parallel Process.* N.Y.: IEEE, 1973. P. 1–12.
- [21] *Lampert L.* The hyperplane method for an array computer // *Proc. 1974 Samagore Comput. Conf. Parallel Process. Lect. Notes Comput. Sci.* 24. Berlin: Springer-Verlag, 1975. P. 113–131.
- [22] *Maske T. et al.* The KAP/ST-100: A Fortran translator for the ST-100 attached processor // *IEEE Proc. of ICPP.* N.Y.: IEEE, 1986. P. 171–175.
- [23] *Owens J.L.* The influence of machine organization on algorithms // *Complexity of sequential and parallel numerical algorithms.* N.Y.: Acad. Press, 1973. P. 111–130.
- [24] *Russell R.M.* The CRAY-1 computer system // *Commun. ACM.* 1978. V. 21, №1. P. 63–72.
- [25] *Sites R.L.* An analysis of the CRAY-1 computer // *Proc. 5th Annu. Symp. Comput. Archit.* N.Y.: IEEE, 1978. P. 101–106.
- [26] *Zhiyu Shen, Zhiyan Li, Pen-Chung Yew.* An empirical study of Fortran programs for parallelizing compilers // *IEEE Trans. on Parallel and Distributed Systems.* N.Y.: IEEE, 1990. P. 350–364.

Метод сопряженных уравнений и анализ сложных систем

*В.И. Агошков, В.С. Владимиров, И.В. Волович,
В.П. Дымников, В.П. Шутяев*

Содержание

Введение	258
1. Сопряженные операторы и сопряженные уравнения . . .	262
1.1. Линейные сопряженные пространства и операторы	262
1.2. Сопряженные уравнения и теория разрешимости линейных уравнений в банаховых пространствах	266
1.3. Подходы к определению сопряженных операторов в нелиней- ных задачах	268
1.4. Свойства операторов, сопряженных к нелинейным	278
2. Метод сопряженных уравнений в теории возмущений . .	284
2.1. Алгоритмы возмущений для линейных задач	285
2.2. Сопряженные уравнения и теория возмущений для линейных функционалов	288
2.3. Сопряженные уравнения и алгоритмы теории возмущений в нелинейных задачах	293
2.4. Сопряженные уравнения в теории чувствительности	299
3. Сопряженные уравнения и законы сохранения в сложных системах	303
3.1. Локальные и нелокальные токи и ассоциированные системы	304
3.2. Сохраняющиеся токи для нелинейных систем	305
3.3. Законы сохранения для двумерных вполне интегрируемых систем	309
3.4. Сопряженные уравнения и интегральные законы сохранения	315
3.5. Сопряженные уравнения для систем гидродинамического типа	316
17 Современные проблемы..., т. 1	257

4. Метод сопряженных уравнений в задачах управления и математического моделирования сложных систем	319
4.1. Задачи оптимального управления	320
4.2. Задача нечувствительного оптимального управления	322
4.3. Задачи вариационного усвоения данных наблюдений	323
4.4. Задачи охраны окружающей среды	327
4.5. Задачи теории ядерных реакторов	330
4.6. Задачи анализа чувствительности климатической системы	333
5. Некоторые нерешенные проблемы теории сопряженных уравнений и их приложений в анализе сложных систем	337
Литература	338

Введение

Метод сопряженных уравнений все более активно проникает в различные области математики и ее приложений. Его основная идея заключается в том, что при исследовании некоторой проблемы для исходной системы уравнений строится сопряженная система и поставленная проблема изучается с использованием решений сопряженной задачи. Развитие этого метода в значительной степени основано на трудах Гурия Ивановича Марчука в течение его почти 60-летней научной деятельности. В этой статье мы даем обзор метода сопряженных уравнений и его применений к анализу различных сложных систем. В последние годы в связи с исследованиями многих прикладных проблем, в частности глобальных изменений на планете Земля и ядерной энергетики, возникли новые постановки проблем, требующие глубокого анализа. К числу таких проблем прежде всего можно отнести проблемы глобальных изменений климата нашей планеты, состояния и защиты от загрязнений окружающей среды, атомных реакторов, иммунологии, сохранения биосферы в условиях резкого увеличения народонаселения, интенсивного развития промышленного производства и многие другие. Объединяются в общую сложную систему проблемы локальных и глобальных возмущений и чувствительности моделей ко входным данным задач. Все это потребовало комплексного изучения больших систем с применением теории сопряженных уравнений. Именно эти сложные системы и теория чувствительности в применении к их исследованию дают новый импульс развитию теории сопряженных уравнений. Теория чувствительности на основе сопряженных уравнений может быть применена не только к ретроспективному изучению процессов, описываемых имита-

ционными моделями, но и к исследованию самих моделей, реализуемых с помощью современных вычислительных технологий. Если к этому добавить проблему четырехмерного анализа данных наблюдений и решение обратных задач, то становится ясным важное значение сопряженных задач в науке.

Первоначально определенные Лагранжем сопряженные операторы нашли глубокое теоретическое обоснование и широкое применение при решении многих задач математической физики. Но истинное значение теории сопряженных уравнений, пожалуй, было впервые оценено физиками при развитии квантовой механики. Уравнение Шредингера потребовало развития аппарата самосопряженных расширений симметрических операторов. Здесь самосопряженные уравнения становятся необходимым математическим аппаратом для формулирования теории малых возмущений в спектральных проблемах.

Один из этапов активного интереса к сопряженным уравнениям и теории малых возмущений относится к теории ядерных реакторов, где сложные задачи теории переноса нейтронов с замедлением их в средах потребовали формулирования соответствующих сопряженных уравнений и решения основных и сопряженных уравнений применительно к оценке первого собственного числа спектральной задачи, связанной с реализацией стационарной цепной реакции. Созданная здесь на базе основных и сопряженных уравнений реактора теория малых возмущений сыграла большую роль в развитии физики реакторов и в реализации проектов атомных электростанций [36].

В дальнейшем было дано развитие теории сопряженных задач по отношению к заданным функционалам для некоторых классов задач математической физики. Оно оказалось плодотворным и для многих других направлений науки. В результате появились общие подходы к исследованию сложных систем и математических моделей. Эти подходы — основное содержание многолетних исследований Г.И. Марчука и его сотрудников и последователей в различных областях математики и ее приложениях к проблемам ядерной энергетики, моделям охраны окружающей среды, теории климата и его изменений, математическим проблемам обработки информации со спутников, математическим моделям в иммунологии и др. Вместе с развитием метода сопряженных уравнений формировался рациональный подход к решению обратных задач и к планированию математического эксперимента.

Г.И. Марчуком был выполнен ряд исследований, оказавших значительное влияние на теорию сопряженных уравнений и методов возмущений. Дальнейшее развитие метода сопряженных уравнений проводилось в работах Ж.-Л. Лионса, В.П. Дымникова, В.И. Агошкова, В.П. Шутяева, В.С. Владимирова, И.В. Воловича, Г.А. Михайлова, С.М. Ермакова и др.

В статье Г.И. Марчука и В.В. Орлова "К теории сопряженных функций" (1961) был в общем виде сформулирован метод получения сопряженных функций (функций ценности) для широкого класса неоднородных линейных уравнений и получены формулы теории возмущений для линейных функционалов, которые впоследствии нашли применение при исследовании сложных систем.

На основе метода сопряженных уравнений и теории возмущений Г.И. Марчуком была разработана общая постановка обратных задач, связанных с восстановлением параметров оператора исходной задачи [33]. С помощью формул малых возмущений Г.И. Марчук предложил алгоритм определения искомых поправок к известным коэффициентам невозмущенной задачи. В результате обратная задача была сведена к задачам линейной алгебры. Было дано развитие теории и на случай, когда невозмущенное состояние значительно отличается от реального. На основе сформулированных постановок обратных задач Г.И. Марчук предложил пути решения проблемы планирования сложного эксперимента, позволяющие среди всевозможных (практически реализуемых) наборов измерений выбрать те, которые оказываются наиболее информативными с точки зрения решения обратной задачи [33]–[42].

Конкретным приложением данных общих постановок обратных задач была постановка задач атмосферной оптики, сформулированная в статье Г.И. Марчука "Уравнение для ценности информации с метеорологических спутников и постановка обратных задач" (1964). Эта статья и другие работы Г.И. Марчука по применению метода статистических испытаний (метод Монте-Карло), выполненные совместно с его учениками, стали основополагающими в атмосферной оптике и теории обратных задач в ней.

Г.И. Марчуком были выполнены важные исследования по теории и приложениям сопряженных уравнений в связи с решением актуальных задач моделирования окружающей среды и оптимизации (с экологической точки зрения) территориального размещения промышленных объектов [38].

В последние годы нелинейные задачи становятся объектом широкого фронта исследований. И, естественно, при этом возникают те или иные обобщения теории сопряженных уравнений, приспособленные для новых классов задач. Большой интерес здесь вызывают подходы, сформулированные в работах С. Ямамура [77], М.М. Вайнберга [13], В.П. Маслова [1], В.С. Владимирова и И.В. Воловича [19], Г.И. Марчука и В.И. Агошкова [43], В.А. Треногина [74]. Впоследствии были предложены и исследованы новые подходы к построению сопряженных операторов в нелинейных задачах математической физики на основе применения групп преобразований, законов сохранения, общих теорем вариационного исчисления (см. работы

Г.И. Марчука, В.И. Агошкова, В.П. Шутяева [46], В.С. Владимирова, Г.И. Марчука [21], В.П. Дымникова [22]).

В настоящей статье рассматривается ряд положений метода сопряженных уравнений в линейных и нелинейных задачах. Рассматриваются различные приложения сопряженных уравнений в теории возмущений, теории чувствительности, законах сохранения и др., а также формулируются нерешенные проблемы, исследование которых представляет значительный интерес. Авторы дают обзор лишь некоторых аспектов теории сопряженных уравнений и делают попытку привлечь внимание к этим новым подходам широкого круга научных сотрудников и инженеров-исследователей, которые, возможно, получают определенный импульс к созданию технологии проектирования или планирования экспериментов при решении прикладных задач. Более подробное изложение обсуждаемых подходов и их математическое обоснование можно найти в специальной литературе, ссылки на которую читатель найдет в конце статьи. В этом списке можно найти как известные монографии по теории сопряженных уравнений, так и оригинальные работы, в которых изучаются специальные положения этой теории и многие приложения теории сопряженных операторов и уравнений, не нашедших полного отражения в настоящей статье.

Структура статьи состоит в следующем. Статья включает 5 разделов, каждый из которых отражает ту или иную роль метода сопряженных уравнений в анализе сложных систем.

В разделе 1 приводятся основные положения метода сопряженных уравнений в линейных и нелинейных задачах, формулируются различные принципы построения сопряженных операторов, а также рассматриваются вопросы разрешимости основных и сопряженных уравнений,

В разделе 2 изложены основные положения метода регулярных возмущений применительно к неоднородным задачам, задачам на собственные значения, к вычислению функционалов, к решению нелинейных задач математической физики, а также к исследованию чувствительности сложных систем. Одновременно показана та значительная роль, которую играют в данных методах сопряженные уравнения и их решения.

В разделе 3 обсуждаются некоторые общие подходы к построению интегральных законов сохранения на базе метода сопряженных уравнений для нелинейных задач.

В разделе 4 представлен ряд важных приложений метода сопряженных уравнений и алгоритмов теории возмущений (проблемы оптимального управления, задачи усвоения данных наблюдений, задачи охраны окружающей среды, проблемы расчета ядерных реакторов).

В разделе 5 сформулированы некоторые нерешенные проблемы теории сопряженных уравнений и их приложений в математической физике и математическом моделировании сложных процессов.

1. Сопряженные операторы и сопряженные уравнения

Теория сопряженных уравнений играет важную и незаменимую роль при анализе многих сложных систем. Математический аппарат сопряженных уравнений в своей классической форме хорошо развит для линейных задач. Однако при рассмотрении нелинейных уравнений даже само понятие сопряженного оператора не является общепринятым. В настоящее время существуют несколько различных подходов к определению сопряженного оператора в нелинейных задачах. Эти подходы приводят к различным сопряженным операторам и сопряженным уравнениям.

В данном разделе мы приводим основные положения теории сопряженных уравнений в линейных и нелинейных задачах, формулируем различные принципы построения сопряженных операторов, а также рассматриваем вопросы разрешимости основных и сопряженных уравнений.

1.1. Линейные сопряженные пространства и операторы

В теории дифференциальных уравнений под сопряженным оператором часто понимается формально сопряженный оператор, который зачастую описывает лишь последовательность операций дифференцирования. Определение сопряженного оператора с позиций функционального анализа глубже и существенно затрагивает граничные условия и др. Но чтобы ввести сопряженный оператор и соответствующее сопряженное уравнение, надо сначала договориться о том, что понимать под сопряженным пространством. Оказывается, что и здесь также имеется несколько возможностей. И чтобы в дальнейшем было понятно, о каких сопряженных операторах и уравнениях идет речь, введем определение сопряженного пространства и сделаем это сначала для случая банаховых пространств.

Сопряженные пространства. Пусть X есть банахово пространство с нормой $\|\cdot\|_X$ и элементами g, f, l, \dots . Рассмотрим множество непрерывных функционалов g^*, f^*, l^*, \dots , заданных над элементами из X , т.е. при определенных значениях $g^*(f), \dots$. Часто для $g^*(f)$ пользуются также обозначением $\langle f, g^* \rangle_X$, и выражение $\langle f, g^* \rangle_X$ называют *отношением двойственности* между f и g^* . В дальнейшем подразумевается, что $g^*(f) \equiv \langle f, g^* \rangle_X$.

Теперь мы можем наложить на функционалы ограничение линейности, т.е. требование выполнения свойства

$$g^*(\alpha f + \beta l) = \alpha g^*(f) + \beta g^*(l), \quad (1.1)$$

где α, β , вообще говоря, есть комплексные числа, или антилинейности (по-

лулинейности, сопряженной линейности):

$$g^*(\alpha f + \beta l) = \bar{\alpha}g^*(f) + \bar{\beta}g^*(l). \quad (1.2)$$

Выбор этого ограничения будет влиять на определение сопряженного пространства X^* . Поэтому условимся в дальнейшем рассматривать *линейные функционалы* [53], хотя можно было бы выбрать и антилинейные функционалы [26].

Над элементами g^*, h^*, \dots определим операции сложения и умножения на числа. Сделать это можно следующим образом:

$$(\alpha g^* + \beta h^*)(f) = \alpha g^*(f) + \beta h^*(f), \quad (1.3)$$

либо эти операции можно определить так:

$$(\alpha g^* + \beta h^*)(f) = \bar{\alpha}g^*(f) + \bar{\beta}h^*(f). \quad (1.4)$$

В дальнейшем мы выбираем определение упомянутых операций согласно (1.3). В результате множество линейных непрерывных функционалов мы превращаем в линейное пространство, которое обозначим X^* и назовем *сопряженным к X* . (Нулевым элементом в X^* будет функционал, тождественно равный нулю на всем X .) Если на X^* ввести норму вида

$$\|g^*\|_{X^*} = \sup_{0 \neq f \in X} \frac{|g^*(f)|}{\|f\|_X} = \sup_{f \in X, \|f\|_X=1} |g^*(f)|, \quad (1.5)$$

то X^* становится нормированным пространством (более того, банаховым). Из такого определения нормы в X^* имеем (обобщенное) неравенство Коши-Буняковского

$$| \langle f, g^* \rangle_X | \leq \|f\|_X \|g^*\|_{X^*}. \quad (1.6)$$

Наряду с определением нормы $\|g^*\|_{X^*}$ согласно (1.5) отметим, что если $f \in X$, то

$$\|f\|_X = \sup_{0 \neq g^* \in X^*} \frac{|g^*(f)|}{\|g^*\|_{X^*}} = \sup_{\|g^*\|_{X^*}=1} |g^*(f)|. \quad (1.7)$$

Сопряженные операторы в банаховых пространствах. После введения сопряженного пространства дадим определение сопряженного оператора в банаховых пространствах. Оно является общепринятым. Пусть X, Y есть два банахова пространства, а A — линейный оператор с областью определения $D(A) \subset X$, плотной в X . Оператор A считается действующим из X в Y (т.е. его область значений $R(A)$ принадлежит Y). Пусть X^*, Y^* — пространства, сопряженные к X, Y , соответственно. Рассмотрим произвольный линейный непрерывный функционал $g^* \in Y^*$. Для него справедливо неравенство $|g^*(Af)| \leq c\|Af\|_Y$ при любом элементе $f \in D(A)$.

Этот функционал мы можем рассмотреть и как функционал над X (фиксируя g^* и изменяя $f \in D(A) \subset X$). Но в этом случае он может оказаться неограниченным, т.е. неравенство $|g^*(Af)| \leq c\|f\|_X$ может и не иметь места. Поэтому рассмотрим лишь те функционалы $g^* \in Y^*$, для которых имеем $|g^*(Af)| \leq c\|f\|_X$ для всех $f \in D(A)$, и обозначим множество таких функционалов через $D(A^*)$ (множество $D(A^*)$ не пусто, так как $g^* \equiv 0$ заведомо принадлежит $D(A^*)$). Тогда при фиксированном $g^* \in D(A^*)$ получим, что $g^*(Af)$ является линейным ограниченным функционалом над X , и обозначим его через $h^*(f)$, т.е.

$$g^*(Af) = h^*(f), \quad h^* \in X^*. \quad (1.8)$$

Таким образом, каждому $g^* \in D(A^*)$ ставится в соответствие элемент $h^* \in X^*$. В силу плотности $D(A)$ в X функционал h^* определяется единственным образом. Однозначное соответствие $g^* \rightarrow h^*$ задает оператор A^* , который называется *сопряженным* к A . Он имеет область определения $D(A^*) \subset Y^*$ и область значений $R(A^*)$ из X^* . Итак,

$$A^*g^* = h^*, \quad g^* \in D(A^*), \quad h^* \in R(A^*), \quad A^* : Y^* \rightarrow X^*. \quad (1.9)$$

Если $f \in D(A)$ и $g^* \in D(A^*)$, то для этих элементов (в силу определения A^*) справедливо соотношение $g^*(Af) = h^*(f)$, т.е.

$$g^*(Af) = A^*g^*(f), \quad (1.10)$$

или (в терминах отношений двойственности в пространствах X и Y)

$$\langle Af, g^* \rangle_Y = \langle f, A^*g^* \rangle_X. \quad (1.11)$$

Отметим также, что если A^*, B^* есть два сопряженных оператора, введенных по сформулированному определению и соответствующих некоторым A и B , то на множестве $D(A^*) \cap D(B^*)$ будем иметь $(\alpha A + \beta B)^* = \alpha A^* + \beta B^*$, и, таким образом, множество сопряженных операторов оказывается линейным.

Сопряженные операторы в гильбертовых пространствах. Пусть теперь X и Y являются гильбертовыми пространствами соответственно со скалярными произведениями $(\cdot, \cdot)_X, (\cdot, \cdot)_Y$ и нормами $\|\cdot\|_X = (\cdot, \cdot)_X^{1/2}$, $\|\cdot\|_Y = (\cdot, \cdot)_Y^{1/2}$. Тогда все изложенное выше остается справедливым. Обозначим через X^*, Y^* сопряженные пространства, соответствующие X, Y . Отмечаем, что при рассмотрении гильбертовых пространств пространства X и X^* (так же, как Y и Y^*) можно отождествить, т.е. считать, что $X = X^*$ ($Y = Y^*$), и это обстоятельство часто используется в исследованиях. В основе этого лежит теорема Рисса о представлении линейного ограниченного функционала.

Теорема 1.1. (Рисс) Каждому элементу g^* пространства X^* , сопряженного к X , соответствует единственный элемент $g \in X$ такой, что $g^*(f) = (f, g)_X$ для всех $f \in X$, причем $\|g\|_X = \|g^*\|_{X^*}$.

Согласно этой теореме между X^* и X существует изометрия $J_X g^* = g$, обладающая свойством антилинейности:

$$J_X(\alpha g^* + \beta h^*) = \bar{\alpha} J_X(g^*) + \bar{\beta} J_X(h^*). \quad (1.12)$$

Это свойство является следствием свойств скалярного произведения, которые приводят к тому, что элементу $(\alpha g^* + \beta h^*)$ соответствует элемент $\bar{\alpha}g + \bar{\beta}h$ при наличии соответствий $g^* \rightarrow g$, $h^* \rightarrow h$. Тогда значение $g^*(f)$ можно представить следующими способами:

$$g^*(f) = (f, g)_X = (f, J_X g^*)_X = J_X^{-1} g(f) \quad (1.13)$$

и справедливо равенство $\|g^*\|_{X^*} = \|g\|_X$. Все это позволяет отождествить X^* с X как абстрактные множества (однако нельзя отождествлять X^* , X как линейные пространства), и именно данное отождествление имеют в виду (что и мы часто делаем в дальнейшем), когда пишут $X \equiv X^*$. Если принимают $X \equiv X^*$, то отношение двойственности $\langle f, g \rangle$ считают совпадающим со скалярным произведением $(f, g)_X$ при $f, g \in X$. Гильбертово пространство в таком случае называют также *основным*.

Обратимся к определению сопряженного оператора в случае гильбертовых пространств X, Y , не отождествляя их пока с сопряженными. Пусть A есть линейный оператор из X в Y с плотной в X областью определения $D(A)$. Если следовать определению сопряженного оператора в банаховых пространствах, то придем к оператору A^* с областью определения $D(A^*)$, действующему из Y^* в X^* и удовлетворяющему соотношению

$$\langle Af, g^* \rangle_Y = \langle f, A^* g^* \rangle_X, \quad (1.14)$$

где $f \in D(A)$, $g^* \in D(A^*)$.

Если же X, Y отождествлять с X^*, Y^* , соответственно, то сопряженный к A оператор будет отличаться от A^* . Сформулируем его определение в следующей форме. Пусть A — линейный оператор из $X \equiv X^*$ в $Y \equiv Y^*$ с плотной в X областью определения $D(A)$. Обозначим через $D(\tilde{A}^*) \subset Y$ множество элементов из Y таких, что для каждого $g \in D(\tilde{A}^*)$ существует единственный элемент $h \in X$, удовлетворяющий равенству $(Af, g) = (f, h)_X$ при всех $f \in D(A) \subset Y$. Пусть A^* есть оператор, имеющий область определения $D(\tilde{A}^*)$ и такой, что $\tilde{A}^* g = h$, или, что эквивалентно

$$(Af, g)_Y = (f, \tilde{A}^* g)_X \quad (1.15)$$

при всех $f \in D(A)$, $g \in D(\tilde{A}^*)$. Данный оператор \tilde{A}^* называется *гильбертово сопряженным* к A .

Однако отметим, что это определение сопряженного оператора в гильбертовых пространствах наиболее удобно, когда $X = Y$, т.е. оператор A действует в одном и том же пространстве. Например, в этом случае можно сравнить A, \tilde{A}^* и ввести определения симметричного и самосопряженного операторов. Так, если A, \tilde{A}^* действуют в одном пространстве Y и $A = \tilde{A}^*$ (что подразумевает также равенство $D(A) = D(\tilde{A}^*)$), то оператор A называется *самосопряженным*. Если же $D(A) \subseteq D(\tilde{A}^*)$ и на $D(A)$ имеем $(Af, g)_Y = (f, Ag)_Y, f, g \in D(A)$, то оператор A *симметричный* и здесь $A \subseteq \tilde{A}^*$.

Свойства линейных сопряженных операторов. Пусть X, Y — банаховы пространства, а $A: X \rightarrow Y$ — линейный оператор с областью определения $D(A)$, плотный в X . Рассмотрим сопряженный оператор $A^*: Y^* \rightarrow X$, определенный равенством (1.11), с областью определения $D(A^*) \subset Y^*$. Лишь в редких случаях можно полно описать структуру $D(A^*)$ и свойства $D(A^*)$ (плотность в Y^* и т.п.). Нередко можно построить *сужение* \tilde{A}^* оператора A^* на заранее выбранное множество $\tilde{D} \subseteq D(A^*)$, для которых имеет место равенство (1.11) при $f \in D(A), g^* \in \tilde{D}$. Оператор \tilde{A}^* часто называют также *формально сопряженным (сопряженным по Лагранжу, ассоциированным)*.

Теорема 1.2. 1) Если Y рефлексивно, то оператор A^* , сопряженный к замкнутому оператору A с плотной областью определения, имеет также плотную в Y^* область определения. 2) A^* — замкнутый линейный оператор. 3) Равенство $D(A^*) = Y^*$ имеет место тогда и только тогда, когда A ограничен на $D(A)$. В этом случае $A^* \in \mathcal{L}(X^*, Y^*), \|A^*\| = \|A\|$.

1.2. Сопряженные уравнения и теория разрешимости линейных уравнений в банаховых пространствах

Математические модели, описывающие физические процессы, включают, как правило, различные дифференциальные и интегральные уравнения, которые удобно исследовать, рассматривая их как уравнения (*прямые уравнения*) в банаховых пространствах. Сопряженные уравнения играют важную роль при исследовании разрешимости прямых уравнений и многих задач математической физики. В настоящем подразделе мы приводим классические основы теории прямых и сопряженных линейных уравнений в банаховых пространствах. Изложение данного подраздела базируется на известной монографии С.Г.Крейна [27].

Пусть на линейном многообразии $D(A)$ банахова пространства E определен линейный оператор A , отображающий $D(A)$ в некоторое банахово пространство F . Множество $D(A)$ называется *областью определения* оператора A .

Рассмотрим линейное уравнение

$$Ax = y, \quad (1.16)$$

где y — заданный элемент пространства F , а x — искомый элемент $D(A)$. Совокупность всех $y \in F$, для которых уравнение (1.16) разрешимо, является линейным многообразием в F и называется *областью значений* $R(A)$ оператора A . Совокупность всех решений соответствующего однородного уравнения $Ax = 0$ является линейным многообразием в E и называется *нуль-пространством* или *ядром* $N(A)$ оператора A .

Уравнение (1.16) называется *однозначно разрешимым* (о.р.) на $R(A)$, если однородное уравнение $Ax = 0$ имеет только нулевое решение, т.е. если $N(A) = 0$. В этом случае каждому $y \in R(A)$ отвечает только одно решение уравнения $y = Ax$. Оператор A имеет на $R(A)$ обратный оператор A^{-1} : $A^{-1}y = x$ ($y \in R(A)$).

Уравнение (1.16) называется *корректно разрешимым* (к.р.) на множестве $R(A)$, если при $x \in D(A)$ справедливо неравенство $\|x\|_E \leq k\|Ax\|_F$, где $k > 0$ и не зависит от x . Из корректной разрешимости следует однозначная разрешимость. В конечномерном случае оба условия эквивалентны.

В случае, когда уравнение (1.16) корректно разрешимо, оператор A имеет на $R(A)$ ограниченный обратный.

Уравнение (1.16) *нормально разрешимо* (н.р.), если $R(A)$ представляет собой подпространство (замкнутое!) пространства F : $R(A) = \overline{R(A)}$.

Уравнение (1.16) *плотно разрешимо* (п.р.), если $R(A)$ плотно в F : $\overline{R(A)} = F$.

Уравнение (1.16) *везде разрешимо* (в.р.), если $R(A) = F$.

Линейный оператор A называется *замкнутым*, если из того, что $x_n \rightarrow x$ и $Ax_n \rightarrow y$ следует, что $x \in D(A)$ и $Ax = y$.

Пусть A^* — сопряженный к A оператор, определенный равенством (1.11). Для *сопряженного уравнения*

$$A^*g = f \quad (g \in D(A^*) \subset F^*, f \in E^*) \quad (1.17)$$

вводятся понятия однозначной, корректной, плотной и везде разрешимости, так же как и для уравнения (1.16).

Теорема 1.3. Для того чтобы уравнение (1.16) было плотно разрешимо ($\overline{R(A)} = F$), необходимо и достаточно, чтобы уравнение (1.17) было однозначно разрешимо ($N(A^*) = 0$).

Теорема 1.4. Для нормальной разрешимости уравнения (1.16) необходимо и достаточно, чтобы выполнялось условие $R(A) \perp N(A^*)$, т.е. чтобы уравнение (1.17) было разрешимо для тех и только тех правых частей, которые ортогональны всем решениям сопряженного однородного уравнения.

Теорема 1.5. Для того чтобы уравнение (1.16) было однозначно разрешимо на $R(A)$, достаточно, чтобы уравнение (1.17) было плотно разрешимо.

Теорема 1.6. Уравнение (1.17) везде разрешимо тогда и только тогда, когда уравнение (1.16) корректно разрешимо на $R(A)$.

Теорема 1.7. Для того чтобы уравнение (1.16) было везде разрешимым, необходимо, чтобы уравнение (1.17) было корректно разрешимым на $R(A^*)$.

Пусть оператор A замкнут. Тогда основные факты о разрешимости прямых и сопряженных линейных уравнений даются следующей таблицей [27]. Уравнение с замкнутым оператором с плотной областью определения:

(A)		(A*)
о.р.	\Leftarrow	п.р.
п.р.	\Leftrightarrow	о.р.
к.р.	\Leftrightarrow	в.р.
н.р.	\Leftrightarrow	н.р.
в.р.	\Leftrightarrow	к.р.

Заметим, что первую стрелку в общем случае даже для ограниченного оператора нельзя обратить. Однако в случае рефлексивных пространств это можно сделать. Так справедлива

Теорема 1.8. Если пространство E рефлексивно, то для замкнутого оператора A с плотной областью определения из однозначной разрешимости уравнения (1.16) следует плотная разрешимость уравнения (1.17).

Доказательства приведенных выше утверждений и другие результаты о разрешимости прямых и сопряженных уравнений приведены в [27].

1.3. Подходы к определению сопряженных операторов в нелинейных задачах

В данном разделе мы приведем обзор различных подходов к определению сопряженных операторов в нелинейных задачах, наметившихся в настоящее время в научной литературе.

Построение сопряженных операторов на основе тождества Лагранжа. Пусть X, Y, H_0 — пространства из некоторой шкалы гильбертовых пространств, среди которых лишь H_0 отождествляется со своим сопряженным, т.е. является основным пространством.

Скалярные произведения, нормы и отношения двойственности при рассмотрении пространств X, Y будем обозначать соответственно как

$$(f, g)_X, \|f\|_X = (f, f)_X^{1/2}, \langle f, g^* \rangle_X \equiv (f, g^*)_{H_0}, \quad f, g \in X, \quad g^* \in X^* \equiv X^{-1},$$

$$(f, g)_Y, \|f\|_Y = (f, f)_Y^{1/2}, \langle f, g^* \rangle_Y \equiv (f, g^*)_{H_0}, \quad f, g \in Y, \quad g^* \in Y^* \equiv Y^{-1}.$$

Предполагается, что X вложено в Y плотно и непрерывно.

Пусть F есть нелинейный оператор, действующий из X в Y . Область определения $D(F) \subset X$ этого оператора предполагается линейным множеством, плотным в X . Пусть также выполнено условие $F(0) = 0$. Определим сопряженный оператор, соответствующий F , на основе тождества Лагранжа [35], [46].

Предположим, что элемент $F(u)$ можно представить в виде

$$F(u) = A(u)u, \quad (1.18)$$

где $A(u)$ есть некоторый линейный оператор, зависящий от u , с областью определения $D(A) \supset D(F)$. Фиксируя элемент $u \in D(F)$, обычным образом, как это делается в теории линейных операторов (см. 1.1), вводим (гильбертов) сопряженный оператор $A^*(u)$:

$$(A(u)w, v)_{H_0} = (w, A^*(u)v)_{H_0}, \quad \forall w \in D(A), \quad v \in D(A^*), \quad (1.19)$$

где

$$D(A^*) = \{g \in Y^* \mid \exists c = \text{const} > 0 : |(A(u)f, g)_{H_0}| \leq c\|f\|_X \quad \forall f \in D(A)\}.$$

Определение 1.1. Оператор $A^*(u)$ с областью определения $D(A^*) \subset Y^*$ и областью значений из X^* называется *сопряженным оператором, соответствующим F* .

Из (1.19) следует, что $A^*(u)$ удовлетворяет тождеству Лагранжа

$$(F(u), v)_{H_0} = (u, A^*(u)v)_{H_0} \quad (1.20)$$

при любых $u \in D(F)$, $v \in D(A^*)$.

Замечание 1.1. Подчеркнем, что в определении 1.1 оператор $A^*(u)$ называется "сопряженным оператором, соответствующим F ", а не "сопряженным к F ". Сужение этого оператора на некоторое множество из $D(A)$ называют также ассоциированным оператором [18].

Заметим, что в случае нелинейного оператора F может оказаться несколько операторов $A^*(u)$, удовлетворяющих (1.20), т.е. определение 1.1 не приводит к однозначно определенному $A^*(u)$. Так, если в представлении (1.18) возможно несколько операторов $A(u) = A_i(u)$ ($i = 1, 2, \dots$), то, вводя согласно (1.19) сопряженные операторы $A_i^*(u)$, получаем несколько сопряженных операторов, соответствующих F .

Пример 1.1. Пусть $(t, x) \in \Omega = (0, 1) \times (0, 1)$, $X = X^* = Y = Y^* = L_2(\Omega)$ — пространство вещественных периодических по t и по x функций с периодом, равным единице по каждой из переменных (и, естественно, определенных при любых $(t, x) \in R^2$). Норма в $L_2(\Omega)$ имеет обычный вид:

$$\|u\|_{L_2} = \left(\int_{\Omega} |u(t, x)|^2 dt dx \right)^{1/2}.$$

Рассмотрим периодическую задачу вида

$$F(u) \equiv \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + au = f, \quad a = \text{const} > 0, \quad (1.21)$$

где оператор F имеет область определения $D(F) = C^{(1)}(\Omega) \subset L_2(\Omega)$. Представим $F(u)$ следующими способами:

$$\begin{aligned} F(u) &= A_1(u)u, \quad \text{где} \quad A_1(u)v = \frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + av, \\ F(u) &= A_2(u)u, \quad \text{где} \quad A_2(u)v = \frac{\partial v}{\partial t} + \left(a + \frac{\partial u}{\partial x}\right)v, \\ F(u) &= A_3(u)u, \quad \text{где} \quad A_3(u)v = \frac{\partial v}{\partial t} + \frac{1}{2} \frac{\partial(uv)}{\partial x} + av. \end{aligned} \quad (1.22)$$

На основе операторов $\{A_i(u)\}$ и равенства (1.22) строим соответствующие сопряженные операторы $\{A_i^*(u)\}$ (для этого можно применить обычную методику интегрирования по частям):

$$\begin{aligned} A_1^*(u)w &= -\frac{\partial w}{\partial t} - \frac{\partial(uw)}{\partial x} + aw = -\frac{\partial w}{\partial t} - u \frac{\partial w}{\partial x} + \left(a - \frac{\partial u}{\partial x}\right)w, \\ A_2^*(u)w &= -\frac{\partial w}{\partial t} + \left(a + \frac{\partial u}{\partial x}\right)w, \\ A_3^*(u)w &= -\frac{\partial w}{\partial t} - \frac{u}{2} \frac{\partial w}{\partial x} + aw. \end{aligned} \quad (1.23)$$

Отметим, что каждой из областей $\{D(A_i^*)\}$ операторов $\{A_i^*\}$ принадлежит пространство периодических функций $W_2^1(\Omega)$. Поэтому по крайней мере на $W_2^1(\Omega)$ операторы $\{A_i^*(u)\}$ имеют вид, представленный в (1.23). Таким образом, неоднозначность представления $F(u)$ в виде $A(u)u$ привела к неоднозначному определению сопряженного оператора, соответствующего F .

Определение сопряженных операторов класса \mathcal{D} . Пусть F — исходный нелинейный оператор, действующий из X в Y с областью определения $D(F)$, плотной в X , где X, Y — пространства из шкалы гильбертовых пространств, введенные выше, причем $Y = X^*$, $F(0) = 0$. Будем теперь предполагать, что F дифференцируем по Гато в любой точке $u \in D(F)$, т.е. в точке $u \in D(F)$ существует предел

$$\lim_{t \rightarrow 0} \frac{F(u + tv) - F(u)}{t} = F'(u)v, \quad (1.24)$$

для всех $v \in D(F)$, где $F'(u)$ — линейный оператор (действующий на v) из X в Y с областью определения $D(F') \supseteq D(F)$. Предел здесь понимается в "сильном" смысле, т.е.

$$\lim_{t \rightarrow 0} \left\| \frac{1}{t} (F(u + tv) - F(u) - F'(u)v) \right\|_Y = 0.$$

Предположим далее, что $F'(u)$ — ограниченный оператор и непрерывный по u для всех $u \in D(F)$. Последнее означает, что $F(u)$ дифференцируем по Фреше, т.е. при фиксированном $u \in D(F)$ и всех $v \in D(F')$ имеем:

$$F(u + v) - F(u) = F'(u)v + w(u, v), \quad (1.25)$$

где $F'(u)$ — линейный непрерывный по v оператор, а элемент $w(u, v) \in Y$ такой, что $\lim_{v \rightarrow 0} w(u, v)/\|v\|_X = 0$. При этом можно считать, что $D(F') = X$.

Будем считать далее также, что $D(F) = X$. Из множества операторов $\{F\}$ с указанными выше свойствами выделим класс \mathcal{D} таких, что каждому $F \in \mathcal{D}$ отвечает оператор G из этого множества, что

$$(w, F'(u)v)_{H_0} = (v, G'(u)w)_{H_0} \quad (1.26)$$

для всех $u, v, w \in X$, где $F'(u), G'(u)$ — производные Гато от F и G , соответственно.

Определение 1.2 [77]. Оператор $G \in \mathcal{D}$, удовлетворяющий (1.26) при любых $u, v, w \in D(F) = X$, называется *сопряженным с F оператором* и обозначается $F^* = G$.

Согласно этому определению оператор, сопряженный к нелинейному оператору F , существует лишь только в том случае, когда F принадлежит классу \mathcal{D} , и если он существует, то является единственным [77, 13].

К сожалению, определением 1.2 не всегда можно воспользоваться, поскольку в целом ряде практических задач возникают нелинейные операторы, не принадлежащие классу \mathcal{D} . Например, как показано в [45], оператор F из примера 1.1 не принадлежит классу \mathcal{D} .

Перейдем теперь к следующему определению сопряженного оператора, которое применимо для более широкого класса нелинейных операторов.

Определение сопряженных операторов с помощью формулы Тейлора. Пусть, по-прежнему, F — нелинейный оператор, действующий из X в Y с областью определения $D(F)$, плотной в X , и $F(0) = 0$. Будем предполагать, что F дифференцируем по Гато в смысле (1.24) и $F'(u)$ (линейный оператор, действующий из X в Y с областью определения $D(F') \supseteq D(F)$) непрерывен по u , т.е. $\|F'(u) - F'(v)\|_{X \rightarrow Y} \rightarrow 0$ при $\|u - v\|_X \rightarrow 0$, $u, v \in D(F)$. В этом случае для F справедлива формула Тейлора с остаточным членом в интегральной форме [43], [68]:

$$F(u) = \int_0^1 F'(tu) dt u, \quad u \in D(F). \quad (1.27)$$

Используя формулу Тейлора (1.27), представим $F(u)$ в виде

$$F(u) = A(u)u, \quad (1.28)$$

где $A(u) = \int_0^1 F'(tu)du$ — линейный оператор, действующий из X в Y с областью определения $D(A) = D(F)$.

Формулу (1.27) можно положить в основу определения сопряженного оператора. Так, фиксируя элемент $u \in D(F)$, обычным образом (согласно (1.19)), вводим сопряженный к $A(u)$ оператор $A^*(u)$, действующий из Y^* в X^* :

$$A^*(u)u^* \in X^*, \quad A^*(u) : Y^* \longrightarrow X^*, \quad u \in D(F) \quad (1.29)$$

с областью определения $D(A^*)$.

Определение 1.3 [68]. Оператор $A^*(u) = \left(\int_0^1 F'(tu)dt \right)^*$ называется *сопряженным оператором, соответствующим нелинейному оператору F* .

Из единственности производной Гаю, единственности интеграла от оператор-функции $F'(tu)$ и из единственности сопряженного оператора к линейному, область определения которого плотна во всем пространстве, следует однозначность $A^*(u)$ в смысле определения 1.3.

Далее легко заметить, что для двух элементов $u \in D(A)$ и $u^* \in D(A^*)$ справедливо тождество $(F(u), u^*)_{H_0} = (u, A^*(u)u^*)_{H_0}$ или

$$(A(u)u, u^*)_{H_0} = (u, A^*(u)u^*)_{H_0}, \quad (1.30)$$

то есть тождество Лагранжа. Если к тому же элемент u есть решение прямого (основного) уравнения

$$F(u) \equiv A(u)u = f, \quad f \in R(F) \subset Y, \quad (1.31)$$

а u^* является решением сопряженного уравнения

$$A^*(u)u^* = g, \quad g \in R(A^*) \subset X^* \quad (1.32)$$

(при фиксированном u — решении (1.31)), то из (1.30) следует *соотношение сопряженности*

$$(f, u^*)_{H_0} = (u, g)_{H_0}, \quad (1.33)$$

которое, как мы увидим ниже, играет важную роль при вычислении функционалов от решений нелинейных задач.

Замечание 1.2. Если F — линейный оператор: $F(u) = Fu$, то определение 1.3 приводит к сопряженному оператору, принятому в теории линейных операторов в гильбертовых пространствах (см. (1.19)), и здесь имеем $A^*(u) \equiv A^* = F^*$, а соотношение (1.30) есть $(Au, u^*)_{H_0} = (u, A^*u^*)_{H_0}$.

Используя операторы $A(u) = \int_0^1 F'(tu)dt$, $A^*(u)$, можно ввести определение оператора, сопряженного к F , и определения симметричного и кососимметричного операторов.

Определение 1.4 [68]. Если $D(F) \subset D(A^*)$, то оператор F^* , определенный на $D(F)$ и задаваемый соотношением

$$F^*(u) = A^*(u)u = \left(\int_0^1 F'(tu) dt \right)^* u, \quad (1.34)$$

называется *оператором, сопряженным к F* .

Определение 1.5. Оператор F , отображающий X в $Y \equiv X^*$, называется *симметричным*, если $A(u) = A^*(u)$ на $D(F)$, и *кососимметричным*, если $A(u) = -A^*(u)$ на $D(F)$ (и тогда $F(u) = -F^*(u)$) при каждом $u \in D(F)$.

Условия, при которых оператор F является симметричным или кососимметричным, сформулированы в [46].

В [46] показано, что если оператор F принадлежит классу \mathcal{D} , введенному выше, то сопряженный оператор, построенный по определению 1.4, совпадает с сопряженным оператором из определения 1.2.

Таким образом, определение 1.4 применимо для более широкого класса операторов, которые дифференцируемы по Гато, но не обязательно принадлежат классу \mathcal{D} .

Пример 1.2. Рассмотрим оператор F и пространства из примера 1.1. Для данного оператора имеем

$$F'(u)v = \frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + \left(a + \frac{\partial u}{\partial x} \right) v. \quad (1.35)$$

Найдем сопряженный оператор $(F'(u))^*$. Поскольку предполагается, что $X = X^* = Y^* = Y = L_2(\Omega)$, то интегрированием по частям получаем

$$(F'(u)v, u^*)_{L_2(\Omega)} = \left(v, -\frac{\partial u^*}{\partial t} - u \frac{\partial u^*}{\partial x} + au^* \right)_{L_2(\Omega)}. \quad (1.36)$$

Таким образом, имеем $W_2^1(\Omega) \subset D(A^*)$, и сужение оператора $(F'(u))^*$ на множество $W_2^1(\Omega)$ есть

$$(F'(u))^* = -\frac{\partial u^*}{\partial t} - u \frac{\partial u^*}{\partial x} + au^*. \quad (1.37)$$

И если ограничиться рассмотрением сопряженного оператора $A^*(u)$ лишь на $W_2^1(\Omega)$, то $A^*(u)$ имеет вид

$$A^*(u)u^* = \left(\int_0^1 F'(tu) dt \right)^* u^* = -\frac{\partial u^*}{\partial t} - \frac{u}{2} \frac{\partial u^*}{\partial x} + au^*. \quad (1.38)$$

Таким образом, сопряженный оператор в смысле определения 1.1 здесь совпадает на $W_2^1(\Omega)$ лишь с одним из сопряженных операторов из примера 1.1 (а именно, с $A_3^*(u)$) в смысле определения 1.3.

Определение 1.4 можно также обобщить, на m -раз дифференцируемые нелинейные операторы, используя формулу Тейлора в общей ее форме записи [43, 46, 68].

Ассоциированные операторы в нелинейных задачах. При изучении дифференциальных уравнений важную роль играют законы сохранения (сохраняющиеся токи) (см. [17]–[20]). Как правило, в приложениях используются локальные токи, однако представляет интерес нахождение и нелокальных сохраняющихся токов.

Метод построения сохраняющихся токов для нелинейных интегро-дифференциальных уравнений, основанный на выборе ассоциированной линейной системы как сопряженной к исходной системе, развит в [17]–[19]. Для дифференциальных уравнений этот метод редуцируется к рассмотренному ранее в [20].

Пусть дана нелинейная система уравнений вида

$$F(u) = A(u)u \equiv \left(\sum_{|\alpha| \leq m} P_\alpha[u](x) \partial^\alpha \right) u(x) = 0, \quad (1.39)$$

где $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ и $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ — мультииндекс, $u \in M(s, r)$ — пространству $(s \times r)$ — матриц-функций от x ; (нелинейные) операторы $u \rightarrow P_\alpha[u]$ отображают $M(s, r)$ в $M(q, s)$.

При фиксированном $u \in M(s, r)$ обычным образом (как это делается в теории линейных дифференциальных операторов) рассмотрим оператор $A^*(u)$ — (формально) сопряженный по отношению к $A(u)$:

$$A^*(u)v \equiv \sum_{|\alpha| \leq m} (-1)^{|\alpha|} \partial^\alpha (v(x) P_\alpha[u](x)),$$

где $v \in M(r, q)$. Отметим, что $A(u)$ — линейный дифференциальный оператор, зависящий от u , при этом оператор $A^*(u)$ был назван в [17]–[19] *ассоциированным* по отношению к исходному нелинейному оператору $F(u) \equiv A(u)u$.

Системы вида $A^*(u)v = 0$ и их решение $v = v(u)$ могут быть использованы для построения сохраняющихся токов в нелинейных задачах [17]–[19] (см. также раздел 3 настоящей статьи).

Сопряженный оператор как класс эквивалентности линейных операторов. Как мы видели выше, сопряженный оператор в ряде случаев определяется неоднозначно в нелинейных задачах. Здесь мы покажем, что его можно вывести из класса эквивалентности линейных операторов,

соответствующих данному нелинейному оператору. Этим достигается единственность сопряженного оператора и выделяется класс нелинейных операторов, для которых сопряженный оператор существует. В линейном же случае класс эквивалентности состоит из единственного оператора — классического сопряженного оператора.

Пусть \mathcal{D} — линейное множество, плотное в гильбертовом пространстве \mathcal{H} , и каждому элементу $\varphi \in \mathcal{D}$ соответствует линейный оператор $A_\varphi \equiv A(\varphi)$, действующий в \mathcal{H} ,

$$\{\psi \rightarrow A_\varphi \psi, \quad \psi \in \mathcal{D}(A_\varphi)\} \quad (1.40)$$

с областью определения $\mathcal{D}(A_\varphi)$, содержащей \mathcal{D} . Пусть, далее, $\mathcal{D}(A_\varphi^*)$ также содержит \mathcal{D} , так что $\mathcal{D}(A_\varphi)$ и $\mathcal{D}(A_\varphi^*)$ плотны в \mathcal{H} . Введем (нелинейные) операторы

$$\varphi \rightarrow A_\varphi^* \varphi = A_\varphi^* \psi|_{\psi=\varphi}, \quad \varphi \in \mathcal{D}, \quad (1.41)$$

с областью определения \mathcal{D} .

Определение 1.6. Назовем множества линейных операторов $\{A_\varphi, \varphi \in \mathcal{D}\}$ и $\{B_\varphi, \varphi \in \mathcal{D}\}$ эквивалентными, $A_\varphi \sim B_\varphi$, если выполнено равенство

$$A_\varphi^* \varphi = B_\varphi^* \varphi, \quad \varphi \in \mathcal{D}. \quad (1.42)$$

Введенная эквивалентность обладает свойствами рефлексивности $A_\varphi \sim A_\varphi$, симметричности $B_\varphi \sim A_\varphi$, если $A_\varphi \sim B_\varphi$, транзитивности $A_\varphi \sim C_\varphi$, если $A_\varphi \sim B_\varphi$ и $B_\varphi \sim C_\varphi$.

Поэтому множества линейных операторов вида (1.40) распадаются на непересекающиеся классы эквивалентности.

Пусть $L(\varphi)$ — (нелинейный) оператор с линейной областью определения $\mathcal{D}(L)$, плотной в \mathcal{H} . По аналогии с линейным случаем определим сопряженный оператор L_φ^* к оператору L_φ .

Определение 1.7 Сопряженным оператором L_φ^* к оператору L_φ назовем класс эквивалентности линейных операторов $\{\psi \rightarrow L_\varphi^* \psi, \varphi \in \mathcal{D}(L)\}$ таких, что область определения $\mathcal{D}(L_\varphi^*)$ оператора L_φ^* содержит $\mathcal{D}(L)$ и выполнено равенство

$$\langle L(\varphi), \psi \rangle = \langle \varphi, L_\varphi^* \psi \rangle, \quad \psi \in \mathcal{D}(L_\varphi^*), \quad \varphi \in \mathcal{D}(L). \quad (1.43)$$

Из этого определения вытекают такие утверждения [21].

Если L — линейный оператор, то класс эквивалентности сопряженного оператора к L состоит из единственного оператора L^ , который не зависит от φ .*

Сопряженный оператор L_φ^ к оператору L_φ единственный (в том смысле, что единственный класс эквивалентности $\{L_\varphi^*, \varphi \in \mathcal{D}(L)\}$).*

Необходимым условием существования сопряженного оператора L_φ^* к оператору L_φ является его представление

$$L(\varphi) = (L_\varphi^*)^* \varphi, \quad \varphi \in \mathcal{D}(L), \quad (1.44)$$

т.е. оператор $L(\varphi)$ есть суперпозиция (нелинейного) оператора $\varphi \rightarrow (L_\varphi^*)^* = L_\varphi^{**}$ и линейного оператора $\psi \rightarrow L_\varphi^{**} \psi$ при каждом $\varphi \in \mathcal{D}(L)$.

Равенство (1.44) следует из (1.43), поскольку предполагается, что $\mathcal{D}(L_\varphi^*)$ плотно в \mathcal{H} . С другой стороны, если оператор $L(\varphi)$ допускает представление

$$L(\varphi) = A(\varphi)\varphi, \quad \varphi \in \mathcal{D}(L), \quad (1.45)$$

где $\varphi \rightarrow A(\varphi)$ — (нелинейный) оператор с линейной областью определения $\mathcal{D}(L)$, плотной в \mathcal{H} , и при каждом $\varphi \in \mathcal{D}(L)$ $\psi \rightarrow A(\varphi)\psi$ есть линейный оператор, причем области определения $\mathcal{D}(A(\varphi))$ и $\mathcal{D}(A^*(\varphi))$ содержат $\mathcal{D}(L)$, то сопряженный оператор L_φ^* существует и эквивалентен оператору $A^*(\varphi)$:

$$L_\varphi^* \sim A^*(\varphi). \quad (1.46)$$

Эквивалентность (1.46) следует из (1.43) и (1.45):

$$\langle A(\varphi)\varphi, \psi \rangle = \langle \varphi, A^*(\varphi)\psi \rangle = \langle \varphi, L_\varphi^* \psi \rangle, \quad \varphi, \psi \in \mathcal{D}(L), \quad (1.47)$$

так что семейство линейных операторов $\{L_\varphi^*, \varphi \in \mathcal{D}(L)\}$ существует и принадлежит тому же классу, что и семейство $\{A^*(\varphi), \varphi \in \mathcal{D}(L)\}$. Таким образом, справедлива

Теорема 1.9. *Для того чтобы для оператора $L(\varphi)$ существовал сопряженный оператор L_φ^* , необходимо и достаточно, чтобы он допускал представление (1.45); при этом справедливо соотношение (1.46).*

Изложенное определение сопряженного оператора как класса эквивалентности допускает распространение и на случай банахова пространства. Однако заметим, что этот класс в общем случае может не включать сопряженные операторы, построенные по другим изложенным ранее принципам. Это следует, в частности, из свойств линейных сопряженных операторов и определения сопряженного оператора в нелинейных задачах как класса эквивалентности, по которым это определение охватывает класс операторов $F(u) \equiv A_i(u)u$, $i = 1, 2, \dots$ с замкнутыми операторами $\{A_i(u)\}$.

Определение сопряженных операторов, действующих в пространствах нелинейных функционалов. До сих пор мы рассматривали сопряженные операторы, действующие в классических сопряженных пространствах — пространствах линейных ограниченных функционалов.

В работе [74] было введено определение сопряженного оператора, действующего в пространствах нелинейных функционалов.

Пусть X — некоторое непустое множество и на нем задана неотрицательная функция двух переменных $\rho(x_1, x_2)$, удовлетворяющая аксиомам тождества и симметрии. Тогда пару $X_\rho = (X, \rho)$ называют слабо метрическим пространством, а ρ — слабой метрикой на множестве X .

Введем пространство нелинейных функционалов, сопряженное к слабо метрическому пространству X_ρ . Фиксируем точку $x_* \in X_\rho$ и через X_ρ^* обозначим множество всех комплекснозначных функционалов φ , определенных всюду на X таких, что $\varphi(x_*) = 0$ и

$$\|\varphi\| = \sup_{x_1 \neq x_2} |\varphi(x_1) - \varphi(x_2)| \rho^{-1}(x_1, x_2) < \infty. \quad (1.48)$$

Функционалы из X_ρ^* являются ρ -липшиц-непрерывными и ρ -ограниченными в следующем смысле $|\varphi(x_1) - \varphi(x_2)| \leq \|\varphi\| \rho(x_1, x_2)$ и $|\varphi(x)| \leq \|\varphi\| \rho(x, x_*)$.

С естественными операциями сложения функционалов и умножения функционалов на числа и с нормой (1.48) X_ρ^* является линейным нормированным пространством. Пространство X_ρ^* называется сопряженным к слабо метрическому пространству X_ρ .

Дадим теперь определение сопряженного оператора (см. [74]). Пусть X_ρ и Y_r — слабо метрические пространства со слабыми метриками $\rho(x_1, x_2)$ и $r(y_1, y_2)$, соответственно. Рассмотрим нелинейный оператор $A: D \subset X_\rho \rightarrow Y_r$ с плотной в X_ρ областью определения D . Фиксируем $x_* \in D$ и положим $y_* = Ax_*$. Введем сопряженные пространства X_ρ^* и Y_r^* так, чтобы принадлежащие им функционалы удовлетворяли условиям $\varphi(x_*) = 0$ и $f(y_*) = 0$, соответственно.

Теорема 1.10 [74]. *Существует единственный однозначный оператор $A^*: D^* \subset Y_r^* \rightarrow X_\rho^*$, удовлетворяющий для всех $x \in D$ и всех $f \in D^*$ условию*

$$f(Ax) = (A^*f)(x). \quad (1.49)$$

Область определения D^ оператора A^* состоит из тех $f \in Y_r^*$, для которых функционал $\tilde{\varphi} = f \circ A$ (суперпозиция) является ρ -ограниченным на D . При этом A^* является линейным оператором.*

Определение 1.8. *Оператор A^* называется сопряженным оператором к оператору A .*

Определение сопряженного оператора A^* в виде (1.49) является довольно узким, однако оно позволило исследовать в ряде случаев вопросы глобальной разрешимости нелинейных уравнений [74].

О других подходах к построению сопряженных операторов. Остановимся еще на одном подходе к построению сопряженных операторов, суть которого состоит [1] в сведении нелинейного уравнения к линейному, если это возможно. Дело в том, что некоторые нелинейные уравнения математической физики можно подходящей заменой свести к линейным.

Если такой замены не существует, то в ряде случаев можно ввести в рассматриваемых линейных функциональных пространствах новые операции сложения и умножения такие, что нелинейное уравнение превращается в линейное относительно этих операций. Для преобразованных уравнений уже применима линейная теория. В частности, можно ввести сопряженное уравнение, провести соответствующие рассуждения, получить нужные формулы, а после этого, используя замену, вернуться к исходной нелинейной задаче. Такой подход к построению сопряженных операторов был использован В.П.Масловым для исследования некоторых нелинейных задач математической физики.

Однако таким методом построения сопряженного оператора не всегда можно воспользоваться. Здесь, во-первых, не всегда удается нелинейную задачу свести к линейной. Во-вторых, часто бывает так, что исходное уравнение нелинейно, а рассматриваемый функционал является линейным, и при соответствующей замене уравнение перейдет в линейное, а функционал окажется нелинейным. Последнее обстоятельство не позволяет нам воспользоваться сопряженным уравнением для вычисления значения функционала. Тем не менее этот метод построения сопряженного оператора нашел свое применение при исследовании ряда нелинейных задач математической физики (см. [1]). Особое внимание в работе [1] уделяется задачам, для которых в рассматриваемых функциональных пространствах можно ввести новые операции сложения и умножения на число и превратить исследуемое нелинейное уравнение в линейное относительно этих операций.

В работе [54] вводится определение сопряженного оператора в нелинейных задачах на основе формулы конечных приращений. Это определение применимо для еще более широкого класса нелинейных операторов, чем приведенные ранее определения. Однако следует отметить, что сопряженные операторы, задаваемые по этому определению, вообще говоря, могут быть введены неоднозначно.

В настоящем разделе мы рассмотрели различные подходы к определению сопряженных операторов в нелинейных задачах. Каждый из этих подходов может быть использован (в условиях его применимости) при исследовании тех или иных нелинейных задач математической физики. Многообразие таких подходов позволяет выбирать тот или иной способ построения сопряженного оператора в зависимости от целей исследования (построение законов сохранения, вычисление функционала, формулировка алгоритма возмущений, исследование разрешимости нелинейной задачи и т.п.).

1.4. Свойства операторов, сопряженных к нелинейным

Как мы видели выше, для построения сопряженных операторов в нелинейных задачах могут быть использованы различные принципы. Есте-

ственно ожидать, что и полученные при этом сопряженные операторы будут различными. Такое разнообразие имеет как отрицательные (возможность неоднозначной трактовки тех или иных алгоритмов), так и свои положительные стороны. Исследователь может попытаться выбрать именно тот сопряженный оператор, который больше подходит для его задачи — имеет лучшие вычислительные или какие-либо другие характеристики. В качестве таких характеристик могут выбираться свойства: положительной определенности A^* , везде разрешимости сопряженного уравнения $A^*(u)w = g$ при $u \in D(F)$, корректной разрешимости этого уравнения, симметричности A^* , кососимметричности $A(u)$ и др.

В настоящем разделе мы приводим некоторые свойства сопряженных операторов, построенных по различным принципам.

Общие свойства основных и сопряженных операторов. Пусть $X, Y, H_0 \equiv H_0^*$ — пространства из некоторой шкалы гильбертовых пространств. Пусть $\Phi : \rightarrow Y$ есть нелинейный оператор с областью определения $D(\Phi) \subset X$, являющейся выпуклым множеством. Считаем Φ непрерывно дифференцируемым по Гато. Тогда

$$\Phi(U) = \Phi(U_0) + \int_0^1 \Phi'(U_0 + tu) dt u, \quad u = U - U_0, \quad U, U_0 \in D(\Phi), \quad (1.50)$$

или

$$F(u) \equiv A(u)u = \Phi(U) - \Phi(U_0), \quad (1.51)$$

где $A(u)$ есть оператор вида

$$A(u) = \int_0^1 \Phi'(U_0 + tu) dt = \int_0^1 \Phi'(tu) dt \quad (1.52)$$

с областью определения

$$D(A) \equiv D(F) = \{v \in X : v + U_0 \in D(\Phi)\}, \quad (1.53)$$

которая предполагается линейным множеством. Пусть $u \in D(A)$ есть любой фиксированный элемент. Введем оператор $A^*(u)$ на основе соотношения

$$(A(u)v, g)_{H_0} = (v, A^*(u)g)_{H_0}, \quad (1.54)$$

где $u, v \in D(A)$, $g \in D(A^*)$. Оператор $A^*(u)$ удовлетворяет равенству

$$(F(u), g)_{H_0} = (u, A^*(u)g)_{H_0} \quad \forall u \in D(F), \quad \forall g \in D(A^*), \quad (1.55)$$

что вытекает из (1.54) при $v = u$.

Итак, определяя оператор $A(u)$ по формуле (1.52), можно однозначно ввести соответствующий ему сопряженный оператор $A^*(u)$. Однако данный оператор является одним из операторов, для которых имеет место представление вида $F(u) = A(u)u$. В данном разделе через $A(u)$ всегда обозначается оператор, вводимый согласно (1.52). Любой другой оператор, построенный согласно другим принципам, будем обозначать через $\tilde{A}(u)$. Таким образом, пусть $\tilde{A}(u)$ есть другой оператор (отличный от $A(u)$) с областью определения $D(\tilde{A})$, действующий из X в Y , для которого справедливо равенство

$$F(u) = \tilde{A}(u)u, \quad \forall u \in D(\tilde{A}). \quad (1.56)$$

Для оператора $\tilde{A}(u)$ также введем сопряженный оператор $\tilde{A}^*(u)$:

$$(\tilde{A}(u)v, g)_{H_0} = (v, \tilde{A}^*(u)g)_{H_0} \quad (1.57)$$

при $\forall u \in D(F)$, $v \in D(\tilde{A})$, $g \in D(\tilde{A}^*)$. Для данного оператора имеет место равенство типа (1.55):

$$(F(u), g)_{H_0} = (u, \tilde{A}^*(u)g)_{H_0} \quad \forall u \in D(F), \forall g \in D(\tilde{A}^*).$$

Из представлений $F(u) = A(u)u$, $F(u) = \tilde{A}(u)u$ заключаем, что $D(A)$, $D(\tilde{A})$ имеют ненулевое пересечение. Поэтому при сравнении операторов $A(u)$, $\tilde{A}(u)$ будем считать, что они определены на $D(A) \cap D(\tilde{A})$ и что это пересечение вновь обозначено через $D(A)$. Данное предположение вводится и при сравнении операторов A^* , \tilde{A}^* .

Пусть u, v — произвольные фиксированные элементы из $D(A)$. Рассмотрим элемент $u + th$, $h \in D(A)$. Если существует билинейный по h и v оператор $B = B(u)$ такой, что

$$(B(u)h)v = \lim_{t \rightarrow 0} \frac{(A(u + th) - A(u))v}{t},$$

то этот оператор будем называть производной от $A(u)$ и обозначать $A_u(u; \cdot) = \partial A(u; \cdot) / \partial u$. Таким образом,

$$\frac{\partial A}{\partial u}(u; h)v = \lim_{t \rightarrow 0} \frac{(A(u + th) - A(u))v}{t}, \quad (1.58)$$

т.е. оператор $A(u)$ дифференцируем. Для оператора $A_u(u; h)$ имеем равенства

$$\frac{\partial A}{\partial u}(u; h) = F'(u)h - A(u)h, \quad \frac{\partial A}{\partial u}(u; h) = \Phi'(U_0 + u)h - A(u)h \quad (1.59)$$

при $\forall u \in D(A)$, $h \in D(F')$.

При изучении свойств оператора $A(u)$ можно воспользоваться тем обстоятельством, что некоторые из них следуют из свойств оператора F или его производной $F'(u) = \Phi'(U_0 + u)$.

Лемма 1.1. *Если при каждом $u \in \bar{S}_R(0) = \{u \in D(F) : \|u\|_X \leq R\}$ оператор $F'(u) = \Phi'(U_0 + u)$ имеет обратный, то оператор $A(u)$ также имеет обратный при $u \in \bar{S}_R(0)$.*

Лемма 1.2. *Если при каждом $u \in \bar{S}_R(0)$ для оператора $F'(u)$ справедлива оценка*

$$\|F'(u)v\|_Y \geq m\|v\|_X, \quad (1.60)$$

где v — произвольный элемент из $D(F)$, а m — положительная постоянная, не зависящая от v , то и для оператора $A(u)$ при любом $u \in \bar{S}_R(0)$ справедлива оценка

$$\|A(u)v\|_Y \geq m\|v\|_X \quad (1.61)$$

с той же постоянной m .

Как известно из теории сопряженных линейных операторов, эти операторы всегда замкнуты. Естественно, что это свойство имеет место и для $A^*(u), \tilde{A}^*(u)$.

Лемма 1.3. *Пусть $Y = X^*$ и оператор F сильно монотонный, т.е.*

$$\langle u - v, F(u) - F(v) \rangle_X \geq m\|u - v\|_X^2, \quad m = \text{const} > 0;$$

тогда оператор $A(u)$ положительно определен на $D(F)$, т.е.

$$\langle v, A(u)v \rangle_X \geq m\|v\|_X^2. \quad (1.62)$$

Обратимся теперь к свойствам оператора $\tilde{A}(u)$, который возникает в представлении $F(u)$:

$$F(u) = \tilde{A}(u)u, \quad u \in D(F). \quad (1.63)$$

При этом вид оператора $\tilde{A}(u)$ (или способ его построения) не конкретизируется. На основе этого оператора можно ввести $\tilde{A}^*(u)$ и соответствующее сопряженное уравнение $\tilde{A}^*(u)w = p$.

Итак, пусть $F(u) = \tilde{A}(u)u$, где $\tilde{A}(u)$ есть некоторый линейный оператор, причем принимаем $D(\tilde{A}(u)) = D(F)$ для $\forall u \in D(F)$.

Лемма 1.4. *Если $\tilde{A}(u)$ непрерывен в нуле в смысле*

$$\lim_{t \rightarrow 0} \|\tilde{A}(tv)v - \tilde{A}(0)v\|_Y = 0 \quad (1.64)$$

для $\forall v \in D(F)$, то

$$\tilde{A}(0) = F'(0) = \Phi'(U_0) = A(0).$$

В силу произвольности линейного оператора $\tilde{A}(u)$ в представлении $F(u) = \tilde{A}(u)u$ равенства $\tilde{A}(0) = F'(0) = A(0)$ будут выполняться вне зависимости от принципа построения $\tilde{A}(u)$, если только выполнено условие (1.64).

Пусть теперь оператор $\tilde{A}(u)$ дифференцируем по Гато в точке u в смысле (1.58) и $\partial\tilde{A}/\partial u \equiv \tilde{A}'$ — его производная.

Лемма 1.5. *Если $\tilde{A}(u)$ дифференцируем, то*

$$\tilde{A}(u; v)u = F'(u)v - \tilde{A}(u)v \quad (1.65)$$

для $\forall v \in D(F)$.

Лемма 1.6. *Если оператор $\tilde{A}(u)$ дифференцируем в нуле и если существует $F''(0, g)$, то*

$$\tilde{A}'(0; g)h + \tilde{A}'(0; h)g = F''(0; g)h, \quad \tilde{A}'(0; u)u = \frac{1}{2}F''(0; u)u \quad \forall u \in D(F) \quad (1.66)$$

для $\forall g, u, h \in D(F)$.

Свойства операторов из класса \mathcal{D} . Отметим, что класс \mathcal{D} содержит все непрерывно дифференцируемые по Гато операторы, обращающиеся в нуле в нуль и имеющие сопряженные операторы $F_{\mathcal{D}}^*$, обладающие такими же свойствами. Следовательно, формулировка свойств операторов из \mathcal{D} в то же самое время будет означать выявление свойств операторов, сопряженных к операторам из \mathcal{D} .

Итак, рассмотрим некоторые свойства сопряженных операторов $F_{\mathcal{D}}^*$ (в смысле определения 1.2), установленных в [13], [77].

Лемма 1.7. *Если F имеет сопряженный оператор, то:*

1) $F_{\mathcal{D}}^*$ также имеет сопряженный оператор и $F^{**} = (F_{\mathcal{D}}^*)^* = F$;

2) $F_{\mathcal{D}}^*(x) = \int_0^1 (F'(tx))^* x dt, \quad x \in X$;

3) $\langle x, F(x) \rangle_X = \langle x, F_{\mathcal{D}}^*(x) \rangle_X$;

4) если G также имеет сопряженный оператор, то $(aF + bG)^* = aF_{\mathcal{D}}^* + bG_{\mathcal{D}}^*$ для любых вещественных a, b .

Лемма 1.8. *Если $F \in \mathcal{D}$, то оператор $F + F_{\mathcal{D}}^*$ симметричный, а $F - F_{\mathcal{D}}^*$ кососимметричный.*

Теорема 1.11. *Если $F \in \mathcal{D}$, то:*

1) он симметричный тогда и только тогда, когда он сильно потенциальный (т.е. если его потенциал дифференцируем по Фреше);

2) он кососимметричный тогда и только тогда, когда он линейный и для любого $x \in X$ выполнено равенство $\langle x, F(x) \rangle_X = 0$.

Теорема 1.12. Для того чтобы оператор $F \in \mathcal{D}$ (т.е. F допускал существование сопряженного), необходимо и достаточно существование линейного кососимметричного оператора $A \in \mathcal{L}(X, X^*)$ такого, что разность $F - A$ есть потенциальный оператор.

Таким образом, множество \mathcal{D} является прямой суммой всех симметричных и кососимметричных операторов, т.е. $F = F_S + A_F$, где $F_S = \frac{1}{2}(F + F_{\mathcal{D}}^*)$, $A_F = \frac{1}{2}(F - F_{\mathcal{D}}^*)$, причем оператор F_S является потенциальным, а A_F линейным.

Свойства сопряженных операторов, построенных на основе формулы Тейлора. Рассмотрим свойства операторов $A^*(u), F^*$, введенных согласно определениям 1.3, 1.4, т.е. операторов

$$\begin{aligned} A^*(u) &= \left(\int_0^1 F'(tu) dt \right)^*, \quad u \in D(F), \\ F^*(u) &= \left(\int_0^1 F'(tu) dt \right)^* u, \quad u \in D(F) \subset D(A^*). \end{aligned} \quad (1.67)$$

Лемма 1.9. Пусть $(F'(u))^*$ есть сопряженный оператор к $F'(u)$ (при фиксированном элементе $u \in D(F)$). Тогда на $D(F)$ имеет место равенство

$$A^*(u) = \left(\int_0^1 F'(tu) dt \right)^* = \int_0^1 (F'(tu))^* dt. \quad (1.68)$$

Соотношение между операторами $F^*, F_{\mathcal{D}}^*$ устанавливает следующая

Лемма 1.10. Пусть $Y \equiv X^*$, а оператор F принадлежит множеству \mathcal{D} . Тогда на $D(F)$ операторы F^* и $F_{\mathcal{D}}^*$ совпадают, т.е. $F^*(u) = F_{\mathcal{D}}^*(u)$, $u \in D(F)$.

Лемма 1.10 позволяет перенести на $F^*(u)$ ряд свойств, сформулированных для оператора $F_{\mathcal{D}}^*$. Однако отмечаем, что подобный подход к изучению свойств оператора F^* справедлив при предположении, что $F \in \mathcal{D}$. Если же отказаться от него, то, чтобы ответить, например, на вопрос о симметричности F , необходимо, по-видимому, провести дополнительные исследования. Так, в частности, без этого предположения неочевидна симметричность оператора $\Phi = F + F^*$. Чтобы доказать это свойство, надо доказать дифференцируемость оператора F^* (что даст дифференцируемость Φ) и получить представление $\Phi(u) = B(u)u$, где $B(u) = \int_0^1 \Phi'(tu) dt$, а затем доказать, что $B(u) = B^*(u)$. Аналогичное замечание относится и к изучению кососимметричности оператора $F - F^*$. В ряде случаев преодолеть отмеченные трудности позволяет следующая

Теорема 1.13. Пусть оператор F с областью определения $D(F)$ действует из X в $Y \equiv X^*$ и выполнены условия:

- 1) F — дважды непрерывно дифференцируем по Гато;
- 2) множество $D(F)$ принадлежит области определения каждого из операторов $(F'(u, v))^*$ ($u, v \in D(F)$);
- 3) при любых $u, v, w, h \in D(F)$ справедливо равенство $(F''(u, v)h, w)_{H_0} = (h, F''(u, v)w)_{H_0}$. Тогда оператор F^* , определяемый формулой (1.67), дифференцируем по Гато и удовлетворяет равенствам

$$((F^*(u))'h, w)_{H_0} = ((F'(u))^*h, w)_{H_0} = (h, F'(u)w)_{H_0}. \quad (1.69)$$

Следствие 1. Если выполнены условия теоремы 1.13, то на множестве $D(F)$ операторы F^* , $F_{\mathcal{D}}^*$ совпадают, и оператор F^* представим в виде суммы потенциального и линейного кососимметричного операторов.

Следствие 2. Если выполнены условия теоремы 1.13, то $F + F^*$ — симметричный оператор, а $F - F^*$ — кососимметричный оператор.

Ряд других свойств сопряженных операторов, построенных по различным принципам, в нелинейных задачах приводятся в [13, 19, 43, 46, 77]. Одним из важных приложений этих свойств является установление свойств разрешимости и единственности решений прямых уравнений $A(u)v = f$ и сопряженных уравнений $A^*(u)q = g$. Здесь также привлекаются результаты по разрешимости линейных уравнений в банаховых пространствах, приведенные в п. 1.2. В свою очередь на основе этих свойств может быть сформулирован ряд условий разрешимости исходного нелинейного уравнения $F(u) = f$ (см., например, [46]), а также осуществлены приложения к формулировке и обоснованию алгоритмов возмущений, численных методов решения уравнения $F(u) = f$, построению законов сохранения, к решению задач оптимального управления и обратных задач с уравнением состояния $F(u) = f$. Ряд таких приложений свойств сопряженных операторов и сопряженных уравнений осуществляется в последующих разделах.

2. Метод сопряженных уравнений в теории возмущений

Сопряженные уравнения играют важную роль в создании эффективных алгоритмов возмущений и все больше и больше находят применение как в теоретических исследованиях, так и в практических расчетах.

В настоящем разделе излагаются основные положения методов регуляризованных возмущений применительно к неоднородным задачам, к вычислению функционалов, к решению нелинейных задач математической физики, а также к исследованию чувствительности сложных систем. Одновременно будет показана та значительная роль, которую играют в данных методах сопряженные уравнения и их решения.

2.1. Алгоритмы возмущений для линейных задач

Алгоритмы возмущений возникли в XIX веке в связи с решением задач небесной механики. Математическая теория возмущения была сформулирована в работах А. Пуанкаре и А. М. Ляпунова. В том виде, в котором алгоритмы возмущений применяются в задачах на собственные значения, они были разработаны в трудах Релея и Шредингера. Математически строгая теория возмущений, по-видимому, начинается с работ Ф. Реллиха. Дальнейшее развитие математическая теория возмущений получила в работах К.О. Фридрихса, Т. Като, Н.Н. Боголюбова и Ю.А. Митропольского, А.Н. Колмогорова, В.И. Арнольда, Ю. Мозера, А.Б. Васильевой и В.Ф. Бутузова, М.И. Вишика и Л.А. Люстерника, Г.И. Марчука, Ж.-Л. Лионса, С.А. Ломова, Н.Н. Моисеева, В.П. Маслова, Н.А. Дмитриева, В.А. Треногина, Р. Беллмана, А.Н. Филатова, М.Д. ван Дейка и многих других. Эти работы продолжили пути развития теории возмущений в применении к широкому классу задач математической физики.

Среди методов, развиваемых в этих работах, мы выделим совокупность методов, которые объединяются в настоящее время в рамках *теории регулярных возмущений*. Ряд положений этой теории и приложений в ней сопряженных уравнений рассматриваются в настоящем подразделе.

Пусть X, Y — гильбертовы пространства соответственно со скалярными произведениями и нормами $(\cdot, \cdot)_X, \|\cdot\|_X = (\cdot, \cdot)_X^{1/2}, (\cdot, \cdot)_Y, \|\cdot\|_Y = (\cdot, \cdot)_Y^{1/2}$. Пусть A — линейный замкнутый оператор с областью определения $D(A)$ (плотной в X), действующий из X в Y . Рассмотрим неоднородное уравнение

$$Au = f, \quad (2.1)$$

которое будем предполагать однозначно разрешимым при любом элементе f из области значений $R(A)$ оператора A . Другими словами, предполагается существование обратного оператора A^{-1} , определенного на $R(A)$.

Уравнение (2.1) будем называть *невозмущенным основным уравнением*. Наряду с (2.1) введем *возмущенное уравнение*

$$A_\varepsilon u_\varepsilon = f_\varepsilon, \quad (2.2)$$

где $A_\varepsilon = A + \varepsilon \delta A$; δA — возмущающий оператор из X в Y с областью определения, совпадающей с $D(A)$ (т.е. $D(A_\varepsilon) = D(A)$), и областью значений $R(\delta A) \subset R(A)$; ε — числовой параметр; $f_\varepsilon = f + \varepsilon \delta f$; $\delta f \in R(A)$.

Отметим, что часто параметр ε в возмущенное уравнение вводится формально, а после проведения исследований он полагается равным необходимому значению. Однако во многих задачах в качестве параметра ε можно выбрать тот или иной параметр, который фактически присутствует в уравнении и имеет физический смысл.

Итак, ставится задача отыскания решения (2.2) с помощью алгоритма возмущений. Формальная схема этого алгоритма состоит в следующем. Предположим, что решение уравнения (2.2) существует и представляется в виде ряда по ε :

$$u_\varepsilon = u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + \dots, \quad (2.3)$$

где $\{u_i\}_{i=0}^\infty \subset D(A)$. Подставляя (2.3) в (2.2), имеем

$$(A + \varepsilon \delta A)(u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + \dots) = f + \varepsilon \delta f.$$

Приравнивая в обеих частях этого равенства коэффициенты при одинаковых степенях ε , получаем следующее уравнение для u_i ($i = 0, 1, \dots$):

$$\begin{aligned} Au_0 &= f, \\ Au_1 &= \delta f - \delta Au_0, \\ Au_i &= -\delta Au_{i-1}, \quad i = 2, 3, \dots \end{aligned} \quad (2.4)$$

Если предположить, что эти уравнения разрешимы и их решения — функции u_i — найдены, то мы можем построить функцию u_ε по формуле (2.3), которую назовем *формальным решением* уравнения (2.2). Если мы ограничимся вычислением u_0, u_1, \dots, u_N , то функцию

$$u_\varepsilon^{(N)} = u_0 + \varepsilon u_1 + \dots + \varepsilon^N u_N \quad (2.5)$$

называют *приближением N -го порядка к u_ε* .

Уже из изложенного выше следует ряд вопросов, которые требуют своего ответа в алгоритмах возмущений:

- 1) разрешимо ли уравнение (2.2);
- 2) представимо ли решение возмущенной задачи в виде ряда (2.3);
- 3) разрешима ли система (2.4) (здесь вопрос не только в обратимости оператора A , но и в принадлежности к $R(A)$ правых частей уравнений системы);
- 4) сходится ли ряд (2.3), и если сходится, то будет ли функция u_ε решением уравнения (2.2);
- 5) какова скорость сходимости ряда (2.3) и какова оценка погрешности $\|u_\varepsilon - u_\varepsilon^{(N)}\|$;
- 6) если требуется построить приближенное решение уравнения (2.2) с заданной точностью, то как это проще сделать: решая непосредственно уравнение (2.2) или решая последовательно необходимое число уравнений системы (2.4); какова при этом эффективность второго пути приближенного решения задачи (*эффективность алгоритмов возмущений*).

Если ответы на поставленные вопросы (и, возможно, на ряд других) не даются, то изложение и применение в практических расчетах алгоритмов возмущений в той или иной степени следует считать формальным. Можно выделить два основных подхода к обоснованию алгоритмов регулярных возмущений и получению ответов на вопросы 1)–6) (см., например, [46]). Суть первого из них состоит в доказательстве аналитической зависимости решения возмущенного уравнения (2.2) от параметра ε , после чего системы уравнений (2.4) являются одним из простых следствий из теории аналитических функций. Во втором подходе сначала формально получают систему (2.4), а затем исследуют разрешимость этих уравнений, получают априорные оценки для $u_\varepsilon^{(N)}$ и устанавливают условия сходимости $u_\varepsilon^{(N)}$ при $N \rightarrow \infty$ к некоторой функции u_ε и последующем доказательстве факта, что u_ε есть решение уравнения (2.2). В обоих подходах существенное значение могут играть дополнительные ограничения на операторы A, A_ε и условия той или иной разрешимости уравнений (2.1), (2.2): наличие априорных оценок типа $\|u\|_X \leq m\|Au\|_Y$, $m = \text{const} > 0$, присутствие свойства корректной разрешимости уравнения (2.1), замкнутости оператора A , ограничения на порядок возмущений и др.

При наличии этих ограничений часто удается провести достаточно полное изучение того или иного алгоритма возмущений и дать его обоснование. Отсутствие одного из них, как правило, в значительной степени усложняет исследование.

Рассмотрим теперь сопряженное по отношению к (2.1) уравнение

$$A^* u^* = g, \quad (2.6)$$

где A^* — сопряженный к A оператор, действующий из Y^* в X^* , а $g \in X^*$. При рассмотрении сопряженных уравнений и их приложений в алгоритмах возмущений мы введем дополнительное ограничение, предполагая, что

$$R(A) = Y, \quad (2.7)$$

т.е. исходное основное уравнение (2.1) везде разрешимо в Y . Это ограничение позволяет нам гарантировать корректную разрешимость сопряженного уравнения (2.6) (см. [27]), т.е. мы будем иметь соотношение

$$\|u^*\|_{Y^*} \leq m\|A^* u^*\|_{X^*}, \quad u^* \in D(A^*), \quad m = \text{const} > 0. \quad (2.8)$$

Тогда возмущенное сопряженное уравнение

$$(A^* + \varepsilon \delta A^*) u_\varepsilon^* = g + \varepsilon \delta g, \quad (2.9)$$

где $D(\delta A^*) = D(A^*)$, $R(\delta A^*) \subseteq R(A^*)$, $\delta g \in R(A^*)$, при выполнении ограничения

$$m |\varepsilon| \|\delta A^*\|_{Y^* \rightarrow X^*} \equiv q^* < 1 \quad (2.10)$$

будет также корректно разрешимым. Повторяя теперь рассуждения, проведенные при рассмотрении прямых задач, можно доказать, что решение уравнения (2.9) представимо сходящимся в Y^* рядом

$$u_\varepsilon^* = u_0^* + \varepsilon u_1^* + \varepsilon^2 u_2^* + \dots, \quad (2.11)$$

где $u_i^* \in D(A^*)$, и u_i^* могут быть определены путем последовательного решения уравнений

$$\begin{aligned} A^* u_0^* &= g, & A^* u_1^* &= \delta g - \delta A^* u_0^*, \\ A^* u_i^* &= -\delta A^* u_{i-1}^*, & i &= 2, 3, \dots \end{aligned} \quad (2.12)$$

Если найти лишь функции $\{u_i^*\}_{i=1}^N$, то функцию

$$u_*^{(N)} = u_0^* + \varepsilon u_1^* + \dots + \varepsilon^N u_N^* \quad (2.13)$$

можно принять за приближение N -го порядка к решению уравнения (2.9).

Обоснование алгоритма возмущений в сопряженных задачах можно осуществить так же, как и в прямых.

2.2. Сопряженные уравнения и теория возмущений для линейных функционалов

В настоящее время математические модели в различных отраслях науки становятся все более сложными и содержательными. Поэтому они требуют предварительного глубокого анализа и, в первую очередь, исследования чувствительности решения и его функционалов в зависимости от вариаций параметров задачи и входных данных. И именно здесь мы сталкиваемся с необходимостью привлечения в исследованиях сопряженных уравнений в зависимости от тех или иных функционалов задач. Сопряженные уравнения в этом случае позволяют оценить вариации функционала в зависимости от вариации входных параметров исходных задач. Именно на это свойство обращается особое внимание в ряде основополагающих работ Г. И. Марчука по теории и приложениям сопряженных уравнений. В данном разделе мы изложим некоторые из методов теории возмущений для линейных функционалов. Причем базироваться они будут на решениях как основных уравнений, так и уравнений с сопряженными операторами.

Линейные функционалы и соотношение сопряженности. Пусть A — линейный оператор, действующий в гильбертовом пространстве X , с областью определения $D(A) \subset X$, плотной в X . Предполагаем, что в X заданы некоторое скалярное произведение (\cdot, \cdot) и норма $\|\cdot\|$: $\|\varphi\| = (\varphi, \varphi)^{1/2}$, $\varphi \in X$.

Рассмотрим линейное уравнение

$$A\varphi = f, \quad (2.14)$$

где f — заданный элемент пространства X , а φ — искомый элемент из $D(A)$. Пусть уравнение (2.14) корректно разрешимо, т.е. оператор A имеет на $R(A)$ ограниченный обратный. Тогда для всякого $f \in R(A)$ существует единственный элемент $\varphi \in D(A)$, являющийся решением уравнения (2.14), причем $\|\varphi\| \leq k\|f\|$, где $k > 0$ не зависит от φ и f .

При решении тех или иных физических задач обычно нужно получить в результате значение некоторой величины, являющейся функционалом от φ . Мы будем рассматривать линейные ограниченные функционалы от φ . Согласно теореме Рисса [53], всякий линейный ограниченный функционал в гильбертовом пространстве X записывается в виде скалярного произведения:

$$J_p[\varphi] = (\varphi, p), \quad (2.15)$$

где p — фиксированный элемент из X .

Введем вместе с оператором A сопряженный к нему оператор A^* с областью определения $D(A^*)$, удовлетворяющий тождеству Лагранжа:

$$(Ah, g) = (h, A^*g) \quad (2.16)$$

для любых элементов $h \in D(A)$, $g \in D(A^*)$. Наряду с уравнением (2.14), которое будем называть *основным*, рассмотрим *сопряженное* неоднородное уравнение

$$A^*\varphi_p^* = p, \quad (2.17)$$

где p — элемент из X , определяющий нужный нам функционал $J_p[\varphi]$ из (2.15).

Поскольку уравнение (2.14) корректно разрешимо, то сопряженное уравнение (2.17) всюду разрешимо [27], т.е. для любого $p \in X$ уравнение (2.17) имеет решение φ_p^* . Подставляя в формулу (2.16) вместо элементов h, g решения уравнений (2.14) и (2.17) φ и φ_p^* , получим

$$(A\varphi, \varphi_p^*) = (\varphi, A^*\varphi_p^*), \quad (2.18)$$

или, воспользовавшись уравнениями (2.14) и (2.17), приходим к *соотношению сопряженности* вида

$$(f, \varphi_p^*) = (\varphi, p). \quad (2.19)$$

Отсюда и из (2.15) следует, что

$$J_p[\varphi] = (f, \varphi_p^*). \quad (2.20)$$

Таким образом, если нам нужно найти значение функционала $J_p[\varphi]$, мы можем получить его двумя способами: либо решить уравнение (2.14) и определить эту величину по формуле (2.15), либо решить уравнение (2.17) и определить ту же величину по формуле (2.20).

Следовательно, каждому линейному функционалу $J_p[\varphi] = (\varphi, p)$ может быть поставлен в соответствие элемент φ_p^* , удовлетворяющий уравнению (2.17), причем в качестве свободного члена этого уравнения следует использовать именно элемент $p \in X$, определяющий интересующий нас функционал $J_p[\varphi]$.

Решение φ_p^* (называемое часто *сопряженной функцией* или *функцией ценности*) имеет глубокий физический смысл и играет важную роль при изучении физических процессов. В работе [40] выяснен смысл функции φ_p^* для широкого класса задач математической физики.

Рассмотренный подход к определению линейных функционалов через сопряженную функцию φ_p^* развивался в ядерной физике [39]. В наиболее общем виде он был сформулирован в работе [48]. Дальнейшее развитие и применение эти идеи получили в работах [34]–[42], [43]–[46] и др. О применении такой методики к решению различных задач математической физики более подробно будет рассказано в следующих подразделах. Мы же остановимся здесь на использовании этого подхода при формулировке теории возмущений для линейных функционалов.

Формулы теории возмущений. Пусть свойства оператора A основной задачи (2.14) как-то изменились и он переходит в возмущенный оператор \tilde{A} :

$$\tilde{A} = A + \delta A, \quad (2.21)$$

где δA — некоторый линейный оператор, действующий в X , с областью определения $D(\delta A)$. Для простоты будем считать, что $D(\delta A) = D(A)$.

Решение φ и значение функционала $J_p[\varphi]$ также изменяется в этом случае: вместо φ имеем $\tilde{\varphi}$, а вместо $J_p[\varphi]$ — значение $\tilde{J}_p = \tilde{J}_p + \delta J_p$, где $\tilde{J}_p = J_p[\tilde{\varphi}]$, $\delta J_p = J_p[\tilde{\varphi}] - J_p[\varphi]$. Таким образом, вместо (2.14) приходим к возмущенной задаче

$$\tilde{A}\tilde{\varphi} = f. \quad (2.22)$$

Установим связь между изменением оператора δA и изменением функционала δJ_p . Предположим, что уравнение (2.22) корректно разрешимо; тогда существует единственный элемент $\tilde{\varphi}$ из $D(A)$, являющийся решением уравнения (2.22).

Рассмотрим φ_p^* — решение сопряженного невозмущенного уравнения (2.17), соответствующее функционалу J_p . Из (2.17), (2.22) получаем

$$(\tilde{A}\tilde{\varphi}, \varphi_p^*) - (\tilde{\varphi}, A^*\varphi_p^*) = (f, \varphi_p^*) - (\tilde{\varphi}, p). \quad (2.23)$$

В правой части (2.23) в соответствии с (2.15) и (2.19) получим

$$(f, \varphi_p^*) - (\tilde{\varphi}, p) = J_p[\varphi] - J_p[\tilde{\varphi}] = -\delta J_p. \quad (2.24)$$

Из (2.21), (2.23) и (2.24) получим общее соотношение для вариации функционала:

$$\delta J_p = -(\delta A \tilde{\varphi}, \varphi_p^*). \quad (2.25)$$

Если вместо уравнений (2.22) и (2.17) рассмотрены соответственно возмущенное сопряженное уравнение

$$(A^* + \delta A^*) \tilde{\varphi}_p^* = p \quad (2.26)$$

и невозмущенное основное уравнение (2.14), то аналогичным путем можно получить также соотношение

$$\delta J_p = -(\varphi, \delta A^* \tilde{\varphi}_p^*), \quad (2.27)$$

которое эквивалентно соотношению (2.25).

В [33, 34] отмечена важная особенность применения формул теории возмущений (2.25), (2.27): поскольку эти формулы описывают вариацию функционала, то для вычисления указанных вариаций нет необходимости знать точное решение основной и сопряженной задач, достаточно воспользоваться их приближенными решениями. Так, если возмущение операторов A , A^* столь мало, что оно не сильно искажает решения φ и φ_p^* , то в формулах (2.25), (2.27) можно заменить приближенно $\tilde{\varphi}$ на φ , а $\tilde{\varphi}_p^*$ на φ_p^* . При этом мы получим две эквивалентные друг другу формулы теории малых возмущений (т.е. первого порядка):

$$\delta J_p \simeq -(\delta A \varphi, \varphi_p^*), \quad (2.28)$$

$$\delta J_p \simeq -(\varphi, \delta A^* \varphi_p^*). \quad (2.29)$$

Исходя из соотношений (2.25), (2.27), в случае необходимости можно получить формулы теории возмущений и более высоких порядков, если возмущение оператора A определить с помощью параметра ε . Тогда в предположении, что решение φ возмущенной задачи (2.22) аналитично по ε в окрестности $|\varepsilon| = 0$, т.е. представимо в виде сходящегося ряда по степеням ε :

$$\tilde{\varphi} = \sum_{i=0}^{\infty} \varepsilon^i \varphi^{(i)}, \quad \varphi^{(0)} = \varphi, \quad (2.30)$$

из (2.25) можно получить представление в виде ряда по ε и для δJ_p :

$$\delta J_p = \sum_{i=1}^{\infty} \varepsilon^i J_{p,i}, \quad (2.31)$$

где

$$J_{p,i} = (\varphi^{(i)}, p).$$

Отыскание элементов $\varphi^{(i)}$ ($i = 1, 2, \dots$) можно осуществить алгоритмом возмущений, сформулированным в 2.1. Если $\varphi^{(1)}, \dots, \varphi^{(N)}$ найдены, то, вычислив $\delta J_{p,i} = (\varphi^{(i)}, p)$ ($i = 1, 2, \dots, N$), найдем также $J_p^{(N)}(\tilde{\varphi})$ — приближение N -го порядка к функционалу $J_p(\tilde{\varphi})$:

$$J_p^{(N)}(\tilde{\varphi}) = (\varphi, p) + \sum_{i=1}^N \varepsilon^i \delta J_{p,i}. \quad (2.32)$$

При условии сходимости алгоритма имеем

$$|J_p(\tilde{\varphi}) - J_p^{(N)}(\tilde{\varphi})| \leq O(|\varepsilon|^{N+1}).$$

Сопряженное уравнение. Для нахождения $J_p(\tilde{\varphi})$ можно воспользоваться решением сопряженного уравнения

$$(A^* + \varepsilon \delta A^*) \tilde{\varphi}_p^* = p. \quad (2.33)$$

Представляя $\tilde{\varphi}_p^* = \sum_{i=0}^{\infty} \varepsilon^i \varphi_i^*$ и привлекая для вычисления φ_i^* ($i = 1, 2, \dots$) алгоритм возмущений (см. 2.1), найдем

$$\delta J_{p,i}^* = (f, \varphi_i^*), \quad i = 1, 2, \dots \quad (2.34)$$

Приближение N -го порядка в терминах решений сопряженных уравнений имеет вид

$$J_p^{(N)}(\tilde{\varphi}) = (f, \varphi^*) + \sum_{i=1}^N \varepsilon^i \delta J_{p,i}^*. \quad (2.35)$$

Если $N = 1$, то получаем формулы теории малых возмущений

$$J_p^{(1)}(\tilde{\varphi}) = (\varphi, p) + \varepsilon \delta J_{p,1}, \quad (2.36)$$

$$J_p^{(1)}(\tilde{\varphi}) = (f, \varphi^*) + \varepsilon \delta J_{p,1}^*, \quad (2.37)$$

где $(\varphi, p) = (f, \varphi^*)$. Отмечаем также, что $\delta J_{p,1} = \delta J_{p,1}^* = -(\delta A \varphi, \varphi^*)$, т.е. для отыскания $\delta J_{p,1}$ не нужно вычислять $\varphi^{(1)}$ или φ_1^* . Заметим также, что при $\varepsilon = 1$ формулы (2.36), (2.37) совпадают с (2.28), (2.29).

Полученные формулы теории возмущений (2.25), (2.27) могут быть использованы для оценки различных эффектов и для анализа измерений, а также в теоретическом исследовании сложных систем и в практических расчетах, когда используют метод замены исследуемой сложной системы упрощенной моделью.

2.3. Сопряженные уравнения и алгоритмы теории возмущений в нелинейных задачах

Алгоритмы возмущений находят широкое применение в исследовании и численном решении нелинейных задач [46]. Одним из широко известных классов нелинейных задач является класс задач на собственные значения, когда отыскиваются пары u, λ — решения, такие, что $F(u, \lambda) \equiv Au - \lambda u = 0$, где оператор A может быть даже линейным. Для данных задач по решению их с помощью алгоритмов возмущений получены глубокие и обширные результаты [26, 15]. Изложение этих алгоритмов на базе использования сопряженных уравнений приводится в [46] и ниже мы не рассматриваем обширную область приложений алгоритмов возмущений и сопряженных уравнений, связанную с задачами на собственные значения.

В настоящем подразделе мы рассмотрим общую схему построения алгоритмов возмущений для нелинейных уравнений общего вида и уравнений с сопряженными операторами. Особое внимание уделяется алгоритмам возмущений на основе сопряженных уравнений.

Формулировка алгоритмов теории возмущений. Пусть X и Y — два гильбертовых пространства. Предполагается, что X вложено в Y плотно и непрерывно.

Рассмотрим нелинейный оператор $\Phi(u, \varepsilon)$, действующий из X в Y и зависящий от числового параметра $\varepsilon \in [-\bar{\varepsilon}, \bar{\varepsilon}]$, $\bar{\varepsilon} > 0$. Область определения $D(\Phi)$ этого оператора предполагается линейным множеством, плотным в X . Пусть оператор Φ при любом фиксированном ε имеет непрерывную производную Гато $\Phi'(u, \varepsilon)$ в каждой точке $u \in D(\Phi)$. Мы рассматриваем Φ' как оператор из X в Y . Мы также предполагаем, что область определения $D(\Phi')$ оператора Φ' содержит $D(\Phi)$.

Рассмотрим уравнение

$$\Phi(U, \varepsilon) = 0, \quad (2.38)$$

которое будем называть возмущенной задачей. Зафиксируем элемент $U_0 \in D(\Phi)$, положим $f(\varepsilon) \equiv -\Phi(U_0, \varepsilon)$ и перейдем от (2.38) к уравнению

$$A(u, \varepsilon)u = f(\varepsilon), \quad (2.39)$$

где

$$A(u, \varepsilon) = \int_0^1 \Phi'(U_0 + tu, \varepsilon) dt, \quad u = U - U_0.$$

Оператор $A(u, \varepsilon)$ действует из X в Y с областью определения $D(A) = D(F)$. Сопряженный к нему оператор имеет вид

$$A^*(u, \varepsilon) = \int_0^1 (\Phi'(U_0 + tu, \varepsilon))^* dt, \quad u \in D(F). \quad (2.40)$$

Этот оператор является оператором из Y^* в X^* ; область его определения мы обозначим через $D(A^*)$. Оператор A^* называется (согласно определению 1.3 из раздела 1) *сопряженным оператором, соответствующим* $\Phi(U, \varepsilon)$; он является одним из сопряженных операторов, которые мы можем вводить при рассмотрении уравнения (2.38).

Далее наряду с (2.39) мы рассмотрим сопряженное уравнение

$$A^*(u, \varepsilon)u^* = g(\varepsilon), \quad (2.41)$$

где элемент $g(\varepsilon) \in X^*$ является аналитическим по ε :

$$g(\varepsilon) = \sum_{i=0}^{\infty} \varepsilon^i g_i, \quad g_i = \frac{1}{i!} \frac{d^i g}{d\varepsilon^i} \Big|_{\varepsilon=0}, \quad g_i \in X^*. \quad (2.42)$$

Ряд утверждений о разрешимости уравнений (2.39) и (2.41) может быть сформулирован на основе свойств операторов $A(u, \varepsilon)$, $A^*(u, \varepsilon)$ и общих результатах разрешимости линейных уравнений в банаховых пространствах и приведен в [46].

В дальнейшем мы будем предполагать, что оператор $\Phi(U, \varepsilon)$ аналитичен по всем своим переменным и уравнение (2.39) имеет единственное решение, представляющееся в виде ряда по степеням ε , сходящегося при $|\varepsilon| < \bar{\varepsilon}$, $\bar{\varepsilon} > 0$:

$$u = \sum_{i=0}^{\infty} \varepsilon^i u_i. \quad (2.43)$$

В качестве U_0 выберем решение уравнения

$$\Phi(U_0, 0) = 0, \quad (2.44)$$

которое будем называть невозмущенной задачей. Тогда $u_0 = 0$ и

$$u = \sum_{i=1}^{\infty} \varepsilon^i u_i \quad (2.45)$$

является решением уравнения (2.39).

Алгоритм возмущений для решения задачи (2.39) состоит в последовательном отыскании поправок u_i в разложении (2.45). Для нахождения вида уравнений для u_i ($i = 1, 2, \dots$) разложим $f(\varepsilon)$ из (2.39) в ряд по степеням ε :

$$f(\varepsilon) = \sum_{i=0}^{\infty} \varepsilon^i f_i = \sum_{i=1}^{\infty} \varepsilon^i f_i, \quad (2.46)$$

где

$$f_0 = f(0) = 0, \quad f_i = \frac{1}{i!} \frac{d^i f}{d\varepsilon^i} \Big|_{\varepsilon=0}, \quad i = 1, 2, \dots \quad (2.47)$$

В частности,

$$f_1 = -\frac{\partial}{\partial \varepsilon} \Phi(U_0, \varepsilon)|_{\varepsilon=0}, \quad f_2 = -\frac{1}{2} \frac{\partial^2 \Phi(U_0, \varepsilon)}{\partial \varepsilon^2} |_{\varepsilon=0} \quad (2.48)$$

$\left(\frac{\partial \Phi}{\partial \varepsilon} \text{ и } \frac{\partial^2 \Phi}{\partial \varepsilon^2}\right)$ есть частные производные от $\Phi(U, \varepsilon)$ по ε при фиксированном U . Подставляя (2.45), (2.46) в (2.39) и выполняя сокращение на ε , получаем уравнение

$$A \left(\sum_{i=1}^{\infty} \varepsilon^i u_i, \varepsilon \right) \left(\sum_{i=1}^{\infty} \varepsilon^{i-1} u_i \right) = \sum_{i=1}^{\infty} \varepsilon^{i-1} f_i. \quad (2.49)$$

Отсюда при $\varepsilon = 0$ получается уравнение для u_1 . Затем дифференцируя последовательно уравнение (2.49) по ε и полагая $\varepsilon = 0$, мы в результате получим бесконечную систему уравнений для определения u_i :

$$\begin{aligned} A_0 u_1 &= f_1, \\ A_0 u_2 &= f_2 - A_1(U_0, u_1) u_1, \\ A_0 u_3 &= f_3 - A_1(U_0, u_1) u_2 - \frac{1}{2} A_2(U_0, u_1, u_2) u_1, \\ &\dots \dots \dots \end{aligned} \quad (2.50)$$

где

$$\begin{aligned} A_0 &= A(0, 0) = \Phi'(U_0, 0), \\ A_1(U_0, u_1) &= \frac{d}{d\varepsilon} A\left(\sum_{i=1}^{\infty} \varepsilon^i u_i, \varepsilon\right)|_{\varepsilon=0} = \int_0^1 \Phi''(U_0, t u_1, 0) dt + \frac{\partial}{\partial \varepsilon} \Phi(U_0, 0), \\ A_2(U_0, u_1, u_2) &= \frac{d^2}{d\varepsilon^2} A\left(\sum_{i=1}^{\infty} \varepsilon^i u_i, \varepsilon\right)|_{\varepsilon=0}. \end{aligned}$$

Формулы (2.50) и составляют суть алгоритма возмущений для отыскания поправок u_i . Последовательно решая эти уравнения, мы получаем

$$U = U_0 + \sum_{i=1}^{\infty} \varepsilon^i u_i. \quad (2.51)$$

Элемент вида

$$U_{(N)} = U_0 + \sum_{i=0}^N \varepsilon^i u_i \quad (2.52)$$

называется *приближением N-го порядка* к U .

Аналогично формулируется алгоритм возмущений для решения сопряженного уравнения (2.41). В предположении, что $u^* = \sum_{i=0}^{\infty} \varepsilon^i u_i^*$, уравнения для отыскания поправок u_i^* ($i = 1, 2, \dots$) будут иметь вид:

$$\begin{aligned} A_0^* u_0^* &= g_0, & A_0^* &= (\Phi'(U_0, 0))^*, \\ A_0^* u_1^* &= g_1 - \frac{d}{d\varepsilon} A^*|_{\varepsilon=0} u_0^*, \\ A_0^* u_2^* &= g_2 - \frac{d}{d\varepsilon} A^*|_{\varepsilon=0} u_1^* - \frac{1}{2} \frac{d^2 A^*}{d\varepsilon^2}|_{\varepsilon=0} u_0^*, \\ &\dots\dots\dots \end{aligned} \tag{2.53}$$

Решив первые $N + 1$ уравнений, мы можем найти приближение N -го порядка к u^* по формуле:

$$U_{(N)}^* = u_0^* + \sum_{i=1}^N \varepsilon^i u_i^*. \tag{2.54}$$

Алгоритмы возмущений вида (2.50), (2.53) называются в литературе алгоритмами регулярных возмущений, поскольку они предполагают наличие у решения задачи аналитической зависимости от параметра возмущения. Для обоснования таких алгоритмов возмущений в конкретных задачах могут быть применены два основных подхода, которые в целом те же, что и в линейных задачах. Однако реализация их в нелинейном случае является более сложной проблемой. Изложение ряда результатов по обоснованию алгоритмов возмущений в нелинейных задачах приводится в [46, 54, 68, 69].

Алгоритм теории возмущений для вычисления функционалов. Рассмотрим нелинейную возмущенную задачу

$$\Phi(\varphi, \varepsilon) = g(\varepsilon), \tag{2.55}$$

где $\Phi(\varphi, \varepsilon)$ — нелинейный оператор, действующий из X в Y (X, Y — банаховы пространства) с областью определения $D(\Phi)$, которая предполагается линейным множеством, плотным в X , $g(\varepsilon) \in Y$, ε — параметр возмущения, $\varepsilon \in [0, 1]$.

В приложениях часто требуется найти не само решение φ возмущенной задачи (2.55), а некоторый функционал от него (см. [41]). Мы будем рассматривать линейные ограниченные функционалы вида

$$J(\varphi) = \langle \varphi, p \rangle_X, \quad p \in X^*, \tag{2.56}$$

где X^* — пространство, сопряженное к X , а $\langle \cdot, \cdot \rangle_X$ — отношение двойственности.

Предположим, что нам известно решение $\varphi = \varphi^{(0)}$ невозмущенной по отношению к (2.55) задачи, которая получается из (2.55) при $\varepsilon = 0$:

$$\Phi(\varphi^{(0)}, 0) = g(0). \quad (2.57)$$

Для $J(\varphi)$ мы можем записать

$$J(\varphi) = J(\varphi^{(0)}) + \delta J, \quad (2.58)$$

где $\delta J = \langle u, p \rangle_X$, $u = \varphi - \varphi^{(0)}$. Нас интересует следующий вопрос: как вычислить δJ , не решая нелинейной задачи (2.55).

Одним из методов приближенного вычисления функционала является алгоритм возмущений. Предположим, что уравнение (2.55) имеет единственное решение φ , которое представляется в виде ряда по степеням ε :

$$\varphi = \varphi^{(0)} + \sum_{i=1}^{\infty} \varepsilon^i u_i, \quad u_i \in D(\Phi), \quad (2.59)$$

сходящегося при $\varepsilon < \bar{\varepsilon}$, $\bar{\varepsilon} > 0$, где поправки u_i определяются с помощью алгоритма возмущений вида (2.50). Это означает, что для функционала $J(\varphi)$ из (2.58) справедливо разложение

$$J(\varphi) = J(\varphi^{(0)}) + \sum_{i=1}^{\infty} \varepsilon^i J_i, \quad (2.60)$$

где $J_i = \langle u_i, p \rangle_X$, $\varepsilon < \bar{\varepsilon}$. Таким образом, для вычисления функционала $J(\varphi)$ мы можем воспользоваться алгоритмом возмущений. Если найдены поправки u_i , $i = \overline{1, N}$, то приближение N -го порядка к $J(\varphi)$ определяется по формуле

$$J_{(N)} = J(\varphi^{(0)}) + \sum_{i=1}^N \varepsilon^i J_i. \quad (2.61)$$

Условия сходимости сформулированного алгоритма возмущений для вычисления функционала совпадают с условиями сходимости алгоритма возмущений для решения исходной нелинейной задачи (2.55), при этом скорость сходимости алгоритма возмущений для вычисления функционала не хуже, поскольку

$$|J - J_{(N)}| \leq \|\varphi - \varphi_{(N)}\|_X \|p\|_{X^*}, \quad (2.62)$$

где $\varphi_{(N)} = \varphi^{(0)} + \sum_{i=1}^N \varepsilon^i u_i$. Как следует из раздела 1, для функционала (2.58) справедливо другое представление на основе соотношения сопряженности. Ниже мы покажем, что вычисление функционала с помощью решений сопряженных уравнений может быть в ряде случаев более предпочтительным.

Алгоритмы теории возмущений на основе сопряженных уравнений. Вычтем $\Phi(\varphi^{(0)}, \varepsilon)$ из левой и правой частей уравнения (2.55), в результате придем к равенству

$$F(u, \varepsilon) = f(\varepsilon), \quad (2.63)$$

где $F(u, \varepsilon) = \Phi(\varphi^{(0)} + u, \varepsilon) - \Phi(\varphi^{(0)}, \varepsilon)$, $u = \varphi - \varphi^{(0)}$, $f(\varepsilon) = g(\varepsilon) - \Phi(\varphi^{(0)}, \varepsilon)$. Нелинейный оператор $F(u, \varepsilon)$ действует из X в Y с областью определения $D(F) = D(\Phi)$, причем $F(0, \varepsilon) = 0$.

Введем сопряженный оператор $A^*(u, \varepsilon)$, соответствующий оператору $F(u, \varepsilon)$, по определению 1.1 в предположении, что элемент $F(u, \varepsilon)$ представим в виде

$$F(u, \varepsilon) = A(u, \varepsilon)u, \quad (2.64)$$

где $A(u, \varepsilon)$ — некоторый линейный оператор, зависящий от u и действующий из X в Y с областью определения $D(A) \supset D(F)$.

Рассмотрим сопряженное уравнение

$$A^*(u, \varepsilon)u^* = p, \quad (2.65)$$

где p — элемент из X^* , определяющий функционал (2.56). Решение u^* сопряженной задачи (2.65) зависит, вообще говоря, от решения u исходной нелинейной задачи (2.63). Для функционала δJ из (2.58) справедливо соотношение сопряженности:

$$\delta J = \langle u, p \rangle_X = \langle f(\varepsilon), u^* \rangle_Y. \quad (2.66)$$

Тогда из (2.58) имеем

$$J(\varphi) = J(\varphi^{(0)}) + \langle f(\varepsilon), u^* \rangle_Y. \quad (2.67)$$

Это представление может быть использовано для приближенного отыскания значения $J(\varphi)$. А именно, если мы найдем элемент \tilde{u}^* , близкий в некотором смысле к u^* (в соответствующей метрике), то мы можем приближенно положить

$$J(\varphi) \cong J(\varphi^{(0)}) + \langle f(\varepsilon), \tilde{u}^* \rangle_Y. \quad (2.68)$$

В качестве такого элемента \tilde{u}^* часто можно взять u_0^* — решение задачи

$$A^*(0, \varepsilon)u_0^* = p \quad (2.69)$$

(см. [40, 41, 44]). Тогда для $J(\varphi)$ получаем

$$J(\varphi) \cong J_1, \quad (2.70)$$

где $J_1 = J(\varphi^{(0)}) + \langle f(\varepsilon), u_0^* \rangle_Y$. Однако в каждом конкретном случае не просто доказать, что u^* и u_0^* близки в соответствующей метрике.

Поскольку для непрерывных в нуле операторов $A(u, \varepsilon)$ из представления (2.64) операторы $A(0, \varepsilon)$ всегда совпадают, то сопряженные операторы $A^*(0, \varepsilon)$ также будут совпадать. Таким образом, свойства решения u_0^* сопряженного уравнения (2.69) не будут меняться в зависимости от выбора сопряженного оператора, соответствующего нелинейному оператору F . В связи с этим для обоснования представления (2.70) можно рассматривать любой из сопряженных операторов, соответствующих F (см. определения из раздела 1.1).

Сформулируем теперь основные этапы алгоритма для приближенного вычисления функционала $J(\varphi)$:

- 1) находим u_0^* — решение сопряженной задачи (2.69): $A^*(0, \varepsilon)u_0^* = p$;
- 2) вычисляем поправку $\langle f(\varepsilon), u_0^* \rangle_Y$;
- 3) находим приближенное значение функционала $J(\varphi)$ по формуле:

$$J(\varphi) \cong J_1 = J(\varphi^{(0)}) + \langle f(\varepsilon), u_0^* \rangle_Y. \quad (2.71)$$

Отметим, что для применения и обоснования этого алгоритма не требуется дифференцируемости оператора F по u . Однако если F дифференцируем в смысле Фреше, то, как известно, этот алгоритм совпадает с первой итерацией по методу Ньютона.

2.4. Сопряженные уравнения в теории чувствительности

Основы теории чувствительности на базе сопряженных уравнений были заложены в работах Г.И. Марчука [34]. Исследования в этих направлениях, приобретающие в настоящее время особенно актуальный характер, проводились в работах Д. Какучи [59], Ф. Диме, М. Навона [62], В.В. Пененко [49], Г.А. Бочарова [12] и других.

В настоящем разделе мы покажем, что алгоритмы вычисления функционалов на основе сопряженных уравнений тесно связаны с оценкой чувствительности функционалов к вариациям параметров задачи и входных данных.

Чувствительность линейных функционалов. Рассмотрим нелинейную задачу в виде

$$\Phi(\varphi, \varepsilon) = 0, \quad (2.72)$$

где $\Phi(\varphi, \varepsilon)$ — нелинейный оператор, действующий из X в Y с областью определения $D(\Phi)$, плотной в X , ε — некоторый параметр. Пусть стандартное значение параметра ε задается при $\varepsilon = \varepsilon_0$, ему соответствует решение $\varphi^{(0)}$, удовлетворяющее задаче

$$\Phi(\varphi^{(0)}, \varepsilon_0) = 0. \quad (2.73)$$

Из (2.72)–(2.73) запишем задачу для $u = \varphi - \varphi^{(0)}$:

$$F(u, \varepsilon) = f(\varepsilon), \quad (2.74)$$

где

$$F(u, \varepsilon) = \Phi(\varphi^{(0)} + u, \varepsilon) - \Phi(\varphi^{(0)}, \varepsilon) = A(u, \varepsilon)u,$$

$$f(\varepsilon) = -\Phi(\varphi^{(0)}, \varepsilon) = \sum_{i=0}^{\infty} \varepsilon^i f_i,$$

а в качестве $A(u, \varepsilon)$ можно взять, например, оператор

$$A(u, \varepsilon) = \int_0^1 F'(tu, \varepsilon) dt = \int_0^1 \Phi'(\varphi^{(0)} + tu, \varepsilon) dt,$$

где Φ' — производная Гато от Φ при условии, если она существует.

При $\varepsilon = \varepsilon_0$ задача (2.74) имеет тривиальное решение $u = 0$. Рассмотрим линейный ограниченный функционал J от решения задачи (2.74) вида

$$J(u) = \langle u, p \rangle_X, \quad (2.75)$$

где $p \in X^*$ — заданная функция. Важным с практической точки зрения является вопрос о чувствительности рассматриваемого функционала к вариациям параметров исходной задачи. Пусть таким параметром является параметр ε .

Чувствительность функционала J к изменениям параметра ε определяется величиной

$$I = \left. \frac{dJ}{d\varepsilon} \right|_{\varepsilon=\varepsilon_0}, \quad (2.76)$$

которая есть значение производной функционала J по ε при стандартном значении параметра. Так, если $\left. \frac{dJ}{d\varepsilon} \right|_{\varepsilon=\varepsilon_0} = 0$, то функционал J называется *нечувствительным* [63] к малым изменениям параметра ε .

Такое определение нечувствительности функционала легко объясняется. В самом деле, в условиях сходимости алгоритма возмущений (по ε) для исходной нелинейной задачи (2.74) для u имеет место представление в виде сходящегося ряда

$$u = \sum_{i=0}^{\infty} (\varepsilon - \varepsilon_0)^i u_i, \quad u_0 = 0, \quad u_i \in D(\Phi). \quad (2.77)$$

Тогда для $J(u)$ справедливо разложение

$$J(u) = \sum_{i=0}^{\infty} (\varepsilon - \varepsilon_0)^i J_i, \quad (2.78)$$

где $J_i = \langle u_i, p \rangle_X$, причем $J_0 = 0$ и $I = \left. \frac{dJ}{d\varepsilon} \right|_{\varepsilon=\varepsilon_0} = J_1$, т.е. величина I из (2.76) совпадает с первой поправкой в разложении (2.78) согласно алгоритму возмущений. Именно поэтому при $I = 0$ функционал J называется нечувствительным к малым изменениям параметра ε .

Для практического отыскания величины I из (2.76) ее удобно выразить через решение соответствующей сопряженной задачи. Поскольку $J_1 = \langle u_1, p \rangle_X$, а u_1 — приближение первого порядка по алгоритму возмущений, то из (2.74) следует, что u_1 удовлетворяет уравнению

$$A(0, \varepsilon_0)u_1 = f_1, \quad (2.79)$$

где $A(0, \varepsilon_0) = F'(0, \varepsilon_0) = \Phi'(\varphi^{(0)}, \varepsilon_0)$, $f_1 = \left. \frac{df(\varepsilon)}{d\varepsilon} \right|_{\varepsilon=\varepsilon_0} = -\left. \frac{\partial \Phi}{\partial \varepsilon}(\varphi^{(0)}, \varepsilon) \right|_{\varepsilon=\varepsilon_0}$. Рассмотрим сопряженное уравнение

$$A^*(0, \varepsilon_0)u^* = p, \quad (2.80)$$

где $p \in X^*$ — элемент, определяющий функционал (2.75), $A^*(0, \varepsilon_0) = (\Phi'(\varphi^{(0)}, \varepsilon_0))^*$. Тогда, согласно соотношению сопряженности, получаем, что

$$I = \left. \frac{dJ}{d\varepsilon} \right|_{\varepsilon=\varepsilon_0} = \langle f_1, u^* \rangle_Y = -\left\langle \left. \frac{\partial \Phi}{\partial \varepsilon}(\varphi^{(0)}, \varepsilon) \right|_{\varepsilon=\varepsilon_0}, u^* \right\rangle_Y. \quad (2.81)$$

Таким образом, величина I , характеризующая чувствительность функционала J к малым изменениям параметра ε определяется формулой (2.81) через решение u^* сопряженной задачи (2.80).

В условиях сходимости алгоритма возмущений для исходной нелинейной задачи мы можем записать приближенную формулу для приращений (с точностью до величин второго порядка малости): $\delta J \cong I \delta \varepsilon$, где $\delta J = J - J_0$, $\delta \varepsilon = \varepsilon - \varepsilon_0$, или, с учетом (2.81),

$$\delta J \cong -\left\langle \left. \frac{\partial \Phi}{\partial \varepsilon}(\varphi^{(0)}, \varepsilon) \right|_{\varepsilon=\varepsilon_0}, u^* \right\rangle_Y \delta \varepsilon, \quad (2.82)$$

где u^* — решение сопряженной задачи

$$(\Phi'(\varphi^{(0)}, \varepsilon_0))^* u^* = p. \quad (2.83)$$

Формула (2.82) непосредственно связывает вариацию функционала с вариациями параметра задачи. Чем больше величина $|I|$ из (2.81), тем "более чувствителен" функционал J к изменениям параметра ε . Именно поэтому функцию u^* — решение сопряженной задачи (2.83) — часто называют *функцией ценности* или *функцией влияния* [34, 30, 46].

Случай нелинейного функционала. Предположим, что математическая модель некоторого физического процесса есть уравнение в пространстве $X \equiv Y$:

$$A(u, y) = f, \quad (2.84)$$

где u — решение, y — параметр (ради простоты будем считать его скалярной функцией). Пусть при заданных y, f решение u известно и нас интересует значение функционала

$$J = J(u). \quad (2.85)$$

Относительно J сделаем предположение, что его вариация

$$\delta J(u, h) = \left. \frac{d}{d\varepsilon} J(u + \varepsilon h) \right|_{\varepsilon=0}$$

представима в виде

$$\delta J(u, h) = (Pu, h), \quad (2.86)$$

где P — линейный оператор, действующий в Y .

Рассмотрим теперь уравнение (2.84) с возмущенным параметром δy :

$$A(u + \delta u, y + \delta y) = f, \quad (2.87)$$

из которого получаем уравнение для вариаций

$$A_u(u, y)\delta u + \frac{\partial A}{\partial y}(u, y)\delta y = 0, \quad (2.88)$$

где A_u — производная Гато от $A(u, y)$ в точке (u, y) , а $\partial A/\partial y$ — частная производная по y (оператор, действующий на δy). Считаем, что область определения операторов A_u и $\partial A/\partial y$ плотны в Y . Кроме того, пусть уравнение (2.88) однозначно разрешимо и имеет место оценка

$$\|\delta u\| \leq c\|\delta y\|. \quad (2.89)$$

Нас теперь будет интересовать значение вариации (2.86) при $h = \delta u$. Поскольку

$$|(Pu, \delta u)| \leq c\|\delta u\| \leq c\|\delta y\|,$$

то $\delta J(u, \delta u)$ есть линейный ограниченный функционал над приращениями δy , представимый в виде

$$\delta J(u, \delta u) = (Pu, \delta u) = (U, \delta y), \quad (2.90)$$

где U — некоторый элемент из Y .

Как обсуждалось ранее, U есть функция ценности вклада изменений параметра в изменение интересующего нас функционала J . Поэтому U можно назвать также *функцией чувствительности модели* (2.84) к изменениям параметра δy по отношению к функционалу $J(u)$ (или, проще, *функцией чувствительности модели по отношению к функционалу $J(u)$*). Знание функции U имеет значительный интерес, поскольку она характеризует вклад изменения δy в той или иной части области (в которой решается задача (2.84)) в значение δJ . Найти U можно с помощью рассмотренных ранее методов.

А именно, введем сопряженное уравнение

$$A_u^*(u, y)u^* = -Pu. \quad (2.91)$$

Пользуясь соотношением сопряженности

$$(\delta u, A_u^*(u, y)u^*) = (A_u(u, y)\delta u, u^*),$$

находим

$$\delta J(u, \delta u) = \left(u^*, \frac{\partial A}{\partial y} \delta y \right) = \left(\left(\frac{\partial A}{\partial y} \right)^* u^*, \delta y \right). \quad (2.92)$$

Сравнивая (2.92) и (2.90), а также предполагая плотность множества приращений δy в Y , заключаем, что

$$U = (\partial A / \partial y)^* u^*. \quad (2.93)$$

Итак, чтобы найти функцию чувствительности U при известных u, y , достаточно решить уравнение (2.91) и воспользоваться представлением (2.93). Такой метод исследования чувствительности модели (2.84) с использованием сопряженной задачи (2.91) экономичен и очень эффективен при оценке влияния малых возмущений.

Формулы вида (2.82)–(2.83), (2.90), (2.91) использовались для оценки чувствительности функционалов во многих работах, среди которых отметим [49, 41]. Именно использование сопряженных уравнений вида (2.83) позволило исследовать чувствительность функционалов в задачах взаимодействия атмосферы и океана, теории климата, охраны окружающей среды, иммунологии и др. (см. [31]–[42], [49, 62, 12]).

3. Сопряженные уравнения и законы сохранения в сложных системах

Законы сохранения играют важную роль в анализе разнообразных сложных систем, понимании динамики процессов, описываемых системами уравнений математической физики, конструировании численных методов решения нестационарных систем уравнений в частных производных на больших временных интервалах.

Один из способов построения законов сохранения базируется на известной теореме Нётер и ее обобщениях и связан он с построением групп преобразований, относительно которых рассматриваемая система уравнений инвариантна. Другой метод построения законов сохранения (сохраняющихся токов) основан на выборе *ассоциированной линейной системы, являющейся сопряженной к исходной системе*, и развитый В.С. Владимировым и В.В. Жариновым [20], В.С. Владимировым и И.В. Воловичем [19]. Данный метод применим к построению законов сохранения как в системах линейных дифференциальных, так и в сложных системах, описываемых нелинейными уравнениями (не обязательно дифференциальными).

В данном разделе рассматриваются основные положения подхода к построению законов сохранения, базирующегося на подборе ассоциированных систем. Этот подход применяется также для построения интегральных законов сохранения в нелинейных эволюционных системах. Ряд положений данного метода построения законов сохранения конкретизируются в применении к системам уравнений гидродинамического типа.

3.1. Локальные и нелокальные токи и ассоциированные системы

При изучении дифференциальных уравнений важную роль играют сохраняющиеся токи (законы сохранения) (см. [17]). Обычно в приложениях используются локальные токи.

Однако представляет интерес нахождение и нелокальных токов. Первым нетривиальным примером использования нелокальных интегралов движения является работа Н.Н. Боголюбова 1960 г. [11] по изучению спектра модельного гамильтониана в теории сверхпроводимости.

В данном разделе излагается метод построения сохраняющихся токов (как локальных, так и нелокальных) для весьма общих систем линейных и нелинейных уравнений с многими независимыми переменными с помощью надлежащим образом подобранной ассоциированной (сопряженной) линейной системы дифференциальных уравнений. Изложение данного раздела базируется на работах [17]–[19].

Локальные и нелокальные токи. В дальнейшем мы будем пользоваться следующими определениями. Пусть Ω — область в \mathbf{R}^n ; B, B_1 — банаховы пространства и $C^k(\Omega; B)$ — пространство k раз непрерывно дифференцируемых функций $x \rightarrow u(x)$ в Ω со значениями в B . Пусть $u \rightarrow F[u]$ — (нелинейный) оператор, заданный на $C(\Omega; B)$.

Сохраняющимся током

$$\vec{J} = (J_1, J_2, \dots, J_n), \quad J_i(x) \equiv J_i[u](x), \quad i = 1, 2, \dots, n$$

в области Ω для уравнения $F[u] = 0$ назовем совокупность n операторов $u \rightarrow J_i[u]$, заданных на $C^k(\Omega; B)$ при некотором k со значениями в $C^1(\Omega; B_1)$ и удовлетворяющих закону сохранения

$$\operatorname{div} \vec{J} \equiv \partial_i J_i(x) \equiv \partial_i J_i[u](x) = 0 \quad \text{в } \Omega \quad (3.1)$$

на всех решениях u уравнения $F[u] = 0$. Здесь подразумевается суммирование по $i = 1, 2, \dots, n$.

В силу леммы Пуанкаре условие (3.1) эквивалентно локальной точности в Ω дифференциальной формы $(-1)^{i-1} J_i[u](x) dx_1 \wedge \dots \wedge dx_i \wedge \dots \wedge dx_n$, т.е. $J_i = \partial_j \Phi_{ji}$, $\Phi_{ji} = -\Phi_{ij}$.

Совокупность токов образует линейное пространство. Два тока назовем *эквивалентными*, если они совпадают на всех решениях уравнения $F[u] = 0$.

Ток $\vec{J}[u](x)$ называется *тривиальным*, если (3.1) выполнено при всех $u \in C^k(\Omega; B)$. Остальные токи называются *нетривиальными*.

Ток \vec{J} , называется *локальным* в Ω , если каждый из операторов $J_i[u](x)$ представляет собой функцию $\Phi_i: \Omega \times B \times \dots \times B \rightarrow B_1$ конечного числа аргументов $x, u, \dots, \partial^\alpha u, \dots$, т.е.

$$J_i[u](x) = \Phi_i(x, u(x), \dots, \partial^\alpha u(x), \dots), \quad x \in \Omega. \quad (3.2)$$

Остальные токи называются *нелокальными*.

Предполагая решение u уравнения $F[u] = 0$ достаточно гладким вплоть до границы $\partial\Omega$, с помощью тока \vec{J} получим *интегральный закон сохранения* в односвязной подобласти $\omega \subset \Omega$ в соответствии с формулой Стокса

$$(-1)^{i-1} \int_{\partial\omega} J_i[u] dx_1 \wedge \dots \wedge dx_i \wedge \dots \wedge dx_n = \int_{\omega} \partial_i J_i dx = 0. \quad (3.3)$$

Здесь и в дальнейшем $\partial_i = \partial/\partial x_i$; если $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ — мультииндекс, то $\partial^\alpha = \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n}$, $x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}$, $|\alpha| = \alpha_1 + \dots + \alpha_n$.

3.2. Сохраняющиеся токи для нелинейных систем

Будем рассматривать нелинейную систему уравнений вида

$$A(u)u = \left(\sum_{|\alpha| \leq m} P_\alpha[u](x) \partial^\alpha \right) u(x) = 0 \quad (3.4)$$

или эквивалентно

$$\sum_{k=0}^m P_{i_1, \dots, i_k} [u](x) \partial_{i_1} \dots \partial_{i_k} u(x) = 0. \quad (3.5)$$

Здесь $x = (x_1, \dots, x_n) \in \Omega \subset \mathbf{R}^n$, $u \in \mathcal{M}(s, r)$ — пространству $s \times r$ -матриц-функций, заданных в Ω и достаточно гладких в Ω ; (нелинейные) операторы $u \rightarrow P_\alpha[u]$ отображают $\mathcal{M}(s, r)$ в $\mathcal{M}(q, s)$ (пространство $\mathcal{M}(s, r)$ играет роль пространства $C^k(\Omega; B)$ из п. 3.1.).

Фиксируем $u \in \mathcal{M}(s, r)$. По аналогии с линейным случаем введем *ассоциированные линейные системы дифференциальных уравнений*

$$A(u)w = \sum_{|\alpha| \leq m} P_\alpha[u](x) \partial^\alpha w(x) = 0, \quad (3.6)$$

$$vA^*(u) \equiv \sum_{|\alpha| \leq m} (-1)^{|\alpha|} \partial^\alpha (v(x) P_\alpha[u](x)) \quad (3.7)$$

относительно $w \in \mathcal{M}(s, r)$ и $v \in \mathcal{M}(r, q)$.

Системе (3.2) при фиксированном u соответствует лагранжиан

$$\mathcal{L}(w, v, u) = \text{Tr}(vA(u)w). \quad (3.8)$$

В силу инвариантности лагранжиана (3.8) относительно группы преобразований $w \rightarrow ws, v \rightarrow c^{-1}v$, где c — постоянная невырожденная $r \times r$ -матрица, по теореме Нётер система (3.6), (3.7) обладает сохраняющимся током в Ω на решениях v, w :

$$\begin{aligned} J_i(x) &\equiv J_i[w, v; u](x) = \\ &= \sum_{0 \leq q+k \leq m-1} (-1)^k \partial_{i_1} \dots \partial_{i_k} \left(v P_{i, i_1, \dots, i_k, j_1, \dots, j_q} [u] \right) \partial_{j_1} \dots \partial_{j_q} w, \end{aligned} \quad (3.9)$$

где $i = 1, 2, \dots, n$. Это утверждение следует из формулы

$$\partial_i J_i = (vA^*(u))w - v(A(u)w), \quad (3.10)$$

справедливой для любых допустимых u, v, w .

Из формулы (3.10) видно, что необходимым и достаточным условием сохранения тока (3.9) на решениях u исходной системы (3.4) является равенство

$$(vA^*(u))w - v(A(u)w) = 0. \quad (3.11)$$

В частности, при $w = u$ необходимым и достаточным условием сохранения тока $J_i[u, v; u](x) \equiv J_i[u](x) \equiv J_i(x)$, $i = 1, 2, \dots, n$, где

$$J_i(x) = \sum_{0 \leq q+k \leq m-1} (-1)^k \partial_{i_1} \dots \partial_{i_k} \left(v P_{i, i_1, \dots, i_k, j_1, \dots, j_q} [u] \right) \partial_{j_1} \dots \partial_{j_q} u, \quad (3.12)$$

в Ω на решениях u уравнения (3.4) является равенство

$$(vA^*(u))u = 0. \quad (3.13)$$

Условие (3.13) будет выполнено, если v — решение ассоциированной линейной системы (3.7).

Подчеркнем, что формула (3.9) определяет операторы $u \rightarrow J_i[u]$, $i = 1, \dots, n$ из $\mathcal{M}(s, r)$ в $\mathcal{M}(r, r)$ в следующем смысле. Пусть заданы операторы $u \rightarrow v[u]$, $u \rightarrow w[u]$, преобразующие соответственно $\mathcal{M}(s, r)$ в $\mathcal{M}(r, q)$ и $\mathcal{M}(s, r)$ в $\mathcal{M}(s, r)$, причем $v[u](x)$ и $w[u](x)$ удовлетворяют уравнению (3.6) в Ω при всех решениях u уравнения (3.4). Тогда каждой такой паре операторов соответствует ток (3.9). Сказанное, в частности, относится и к току (3.12).

Возможны и другие более общие способы сведения задачи о построении сохраняющихся токов к линейным дифференциальным уравнениям.

Пусть задано уравнение $F[u] = 0$. Положим

$$J_i[u](x) = \sum_{|\alpha| \leq 1} [(\partial^\alpha u_{ij}(x)) Q_j^\alpha[u](x) + R_j^\alpha[u](x) \partial^\alpha w_{ij}(x)], \quad (3.14)$$

где $u \rightarrow Q_j^\alpha[u]$, $u \rightarrow R_j^\alpha[u]$ — некоторые фиксированные операторы. Тогда, если $v_{ij}(x) \equiv v_{ij}[u](x)$ и $w_{ij}(x) \equiv w_{ij}[u](x)$, $i, j = 1, \dots, n$, удовлетворяют ассоциированному линейному дифференциальному уравнению $\partial_i J_i[u](x) = 0$, т.е.

$$\sum_{|\alpha| \leq 1} \partial_i [(\partial^\alpha v_{ij}) Q_j^\alpha + R_j^\alpha \partial^\alpha w_{ij}] = 0 \quad (3.15)$$

на решениях u уравнения $F[u] = 0$, то выражения (3.14) определяют сохраняющийся ток для уравнения $F[u] = 0$.

Отметим, что в виде (3.14), (3.15) можно представить любой сохраняющийся ток для уравнения $F[u] = 0$.

Покажем, что по любому сохраняющемуся току \vec{J} для уравнения $F[u] = 0$ можно построить бесконечную последовательность сохраняющихся токов $\vec{J}^{(k)}$, $k = 1, 2, \dots$

Пусть $J_i^{(1)}(x) = J_i(x)\Phi(x) + \Phi_i(x)$, $i = 1, 2, \dots, n$, где (Φ, Φ_i) удовлетворяют линейному уравнению $J_i \partial_i \Phi + \partial_i \Phi_i = 0$ на решениях уравнения $F[u] = 0$. Тогда $J_i^{(1)}$, $i = 1, 2, \dots, n$ — сохраняющийся ток. Аналогично по $J_i^{(1)}$ строится ток $J_i^{(2)}$ и т.д.

Итак, мы видим, что для широкого класса нелинейных уравнений можно построить бесконечную последовательность сохраняющихся токов. Ответ на вопрос о том, насколько эффективно можно использовать эти токи, зависит от того, насколько эффективно удастся исследовать линейные уравнения (3.7), (3.15). Для ряда важных уравнений это удастся сделать [17]–[19].

В качестве примера рассмотрим случай $m = 1$. Уравнение (3.4) принимает вид

$$P_i[u] \partial_i + P_0[u] u = 0.$$

Сопряженная система имеет вид

$$\partial_i (v P_i[u]) - v P_0[u] = 0. \quad (3.16)$$

Сохраняющийся ток (3.9):

$$J_i = v P_i[u] u, \quad i = 1, 2, \dots, n. \quad (3.17)$$

В частности, для нелинейного уравнения Дирака

$$[i\gamma^\nu \partial_\nu + i\gamma^\nu A_\nu(x) + f(\bar{\psi}\psi)]\psi = 0,$$

где γ^ν , $\nu = 0, 1, 2, 3$ — матрицы Дирака, ψ — 4-компонентный спинор, $\bar{\psi}$ — дираковски сопряженный спинор, $\bar{\psi} = \psi^* \gamma^0$, A_ν и f — вещественные функции, линейная система (3.16) имеет вид

$$i\partial_\nu(\gamma^\nu) + iv\gamma^\nu A_\nu(x) + vf(\bar{\psi}\psi) = 0.$$

В качестве решения v можно взять $v = \bar{\psi}$. Тогда ток (3.17) становится током электрического заряда $J^\nu = \bar{\psi}\gamma^\nu\psi$, $\nu = 0, 1, 2, 3$.

Аналогичным образом строятся как локальные, так и нелокальные токи для нелинейного уравнения Шредингера, уравнения Навье–Стокса, для уравнений гидродинамики, уравнения Больцмана, Кортевега–де Фриза и других двумерных интегрируемых систем, Янга–Миллса и его суперсимметричного обобщения и других уравнений, см. [19]. Отметим, что нелокальные токи также представляют интерес для физических приложений. В частности, для кирального поля и ферромагнетика Гейзенберга, Кадомцева–Петвиашвили и Кортевега–де Фриза нелокальные токи приведены в [19].

Интегральные законы сохранения. Построенные локальные или нелокальные токи служат для получения интегральных законов сохранения типа (3.3). Обычно при построении интегральных законов сохранения используются краевые условия для решения. Отметим, однако, что, как видно из формулы Стокса, для токов (3.3), полезные интегральные законы сохранения можно получить, не предполагая выполнения краевых условий для решений. В качестве примера рассмотрим простейшее уравнение Бюргера

$$u_t + uu_x = 0$$

в прямоугольнике $\Omega = [(t, x) : 0 \leq t \leq T, a \leq x \leq b]$. Сопряженное уравнение имеет вид:

$$v_t + (uv)_x = 0.$$

Для его решения $v = u_x$ локальные токи имеют вид $J_1 = uu_x$, $J_2 = u^2 u_x$ и формула (3.3) для $\omega = [(\tau, x) : 0 \leq \tau \leq t, a \leq x \leq b]$ имеет вид [76]

$$\frac{1}{2}u^2(t, b) - \frac{1}{2}u^2(t, a) + \int_0^t [u^2(\tau, b)u_x(\tau, b) - u^2(\tau, a)u_x(\tau, a)]d\tau = \text{const},$$

$$0 \leq t \leq T.$$

Заметим, что этот закон сохранения нетривиален, если решение u не удовлетворяет нулевым граничным условиям.

3.3. Законы сохранения для двумерных вполне интегрируемых систем

Обсудим теперь вопрос о сохраняющихся токах для двумерных вполне интегрируемых систем и построении их методом малого параметра.

Под вполне интегрируемой системой будем понимать такую систему нелинейных дифференциальных уравнений, которую можно представить как условие разрешимости некоторой (переопределенной) системы линейных дифференциальных уравнений с параметром. Известно большое число примеров вполне интегрируемых систем. В случае двух независимых переменных вполне интегрируемые системы представляются в виде уравнения

$$\partial_2 U - \partial_1 V + [U, V] = 0, \quad (3.18)$$

которое есть условие совместимости (условие Фробениуса) следующей переопределенной системы.

$$\partial_1 \psi = U\psi, \quad \partial_2 \psi = V\psi, \quad (3.19)$$

где $U = U(x; \lambda)$ и $V = V(x; \lambda)$; $N \times N$ - матрицы-функции переменных $x = (x_1, x_2)$ и λ , мероморфно зависящие от параметра λ .

Пример 3.1. Уравнение Кортевега-де Фриза [19]:

$$u_t - 6uu_x + u_{xxx} = 0, \quad t = x_1, \quad x = x_2, \quad (3.20)$$

$$U(x; \lambda) = 4i\lambda^3 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + 4\lambda^2 \begin{pmatrix} 0 & 1 \\ u & 0 \end{pmatrix} + 2i\lambda \begin{pmatrix} u & 0 \\ u_x & -u \end{pmatrix} + \begin{pmatrix} -u_x & 2u \\ 2u^2 - u_{xx} & u_x \end{pmatrix}, \quad (3.21)$$

$$V(x; \lambda) = i\lambda \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ u & 0 \end{pmatrix}. \quad (3.22)$$

Пример 3.2. Уравнение синус-Гордон [19]:

$$u_{xy} = \sin u, \quad x = x_1, \quad y = x_2, \quad (3.23)$$

$$U(x; \lambda) = i\lambda \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \frac{u}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$V(x; \lambda) = \frac{\cos u}{4i\lambda} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \frac{\sin u}{4\lambda} \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}. \quad (3.24)$$

Для построения сохраняющихся токов для нелинейного уравнения (3.18) достаточно построить сохраняющиеся токи линейной системы (3.19), зависящие от параметра λ . Разлагая эти токи по λ в окрестности точки λ_0

получаем, вообще говоря, бесконечную последовательность сохраняющихся токов, причем при разложении в окрестности регулярной точки λ_0 получаем нелокальные токи, а при разложении в окрестности полюса — λ_0 -локальные токи.

Применим сначала данный метод для построения нелокальных токов. Будем считать что матрицы $U(x; \lambda)$ и $V(x; \lambda)$ принадлежат классу $C^1_{N \times N}(\Omega)$ по x и аналитичны по λ в круге $|\lambda - \lambda_0| \leq \varepsilon_0$, так что, обозначая $\varepsilon = \lambda - \lambda_0$, получим

$$U(x; \lambda) = \sum_{k=0}^{\infty} U_k(x) \varepsilon^k, \quad V(x; \lambda) = \sum_{k=0}^{\infty} V_k(x) \varepsilon^k, \quad (3.25)$$

где $U_k, V_k \in C^1_{N \times N}(\Omega)$ и ряды (3.25) сходятся равномерно по (x, ε) на каждом компакте области $x \in \Omega$, $|\varepsilon| < \varepsilon_0$.

Решение $\psi(x, \lambda)$ системы (3.19), нормированное условием $\psi(x^0; \lambda) = \psi_0$, существует, единственно, принадлежит классу $C^1_{N \times N}(\Omega)$ по x и аналитично по ε в круге $|\varepsilon| < \varepsilon_0$, так что оно разлагается в ряд по ε :

$$\psi(x, \varepsilon) = \sum_{k=0}^{\infty} \psi_k(x) \varepsilon^k, \quad x \in \Omega, \quad |\varepsilon| < \varepsilon_0. \quad (3.26)$$

Система (3.19) в силу (3.18) обладает сохраняющимся (локальным) током вида [19]

$$J_1 = -V\psi, \quad J_2 = U\psi. \quad (3.27)$$

Ток (3.27) — сохраняющийся (нелокальный) для системы (3.18). Разлагая ток (J_1, J_2) в окрестности точки λ_0

$$J_1(x, \lambda) = \sum_{k=0}^{\infty} J_1^{(k)}(x) \varepsilon^k, \quad J_2(x, \lambda) = \sum_{k=0}^{\infty} J_2^{(k)}(x) \varepsilon^k \quad (3.28)$$

и пользуясь рядами (3.25), (3.27), получаем бесконечную цепочку сохраняющихся (матричных) нелокальных токов для уравнения (3.18):

$$J_1^{(k)}(x) = - \sum_{s=0}^k V_s(x) \psi_{k-s}(x), \quad J_2^{(k)}(x) = \sum_{s=0}^k U_s(x) \psi_{k-s}(x), \quad k = 1, \dots, \quad (3.29)$$

где матрицы $\psi_k(x)$ удовлетворяют уравнениям

$$\partial_1 \psi_k = \sum_{s=0}^k U_s(x) \psi_{k-s}, \quad \partial_2 \psi_k = \sum_{s=0}^k V_s(x) \psi_{k-s}, \quad k = 1, \dots \quad (3.30)$$

Из формулы (3.30) видно, что если выполнены калибровочные условия

$$U(x, \lambda_0) = U_0(x) = 0, \quad V(x, \lambda_0) = V_0(x) = 0, \quad (3.31)$$

то формулы (3.29) и (3.30) упрощаются: $\psi_0(x) = \text{const}$ (будем считать $\psi_0 = I$), $J_1^{(0)} = 0$, $J_2^{(0)} = 0$ и

$$\begin{aligned} J_1^{(k)}(x) &= -V_k(x) - \sum_{s=1}^{k-1} V_s(x)\psi_{k-s}(x), \\ J_2^{(k)}(x) &= U_k(x) + \sum_{s=1}^{k-1} U_s(x)\psi_{k-s}(x), \quad k = 1, 2, \dots, \end{aligned} \quad (3.32)$$

где матрицы-функции ψ_k вычисляются рекуррентно:

$$\psi_k(x) = \oint_{x^0}^x \sum_{s=1}^k U_s(x')\psi_{k-s}(x')dx'_1 + V_s(x')\psi_{k-s}(x')dx'_2 + I, \quad k = 1, 2, \dots \quad (3.33)$$

Система (3.19) инвариантна относительно калибровочных преобразований вида

$$\psi \rightarrow \Lambda^{-1}\psi, \quad U \rightarrow \Lambda^{-1}U\Lambda - \Lambda^{-1}\partial_1\Lambda, \quad V \rightarrow \Lambda^{-1}V\Lambda - \Lambda^{-1}\partial_2\Lambda, \quad (3.34)$$

где $\Lambda(x)$ — любая обратимая матрица класса $C^1_{N \times N}(\Omega)$. Если в качестве матрицы Λ взять матрицу ψ_0 , удовлетворяющую системе

$$\partial_1\psi_0 = U_0\psi_0, \quad \partial_2\psi_0 = V_0\psi_0, \quad \psi(x^0) = I, \quad (3.35)$$

то получим систему (3.19), удовлетворяющую калибровочным условиям (3.31). Отметим, что система (3.35) однозначно разрешима в силу условия (3.18) при $\lambda = \lambda_0$.

Теорема 3.1. *Всякая вполне интегрируемая система (3.18) с матрицами U, V класса $C^1_{N \times N}(\Omega)$ по x имеет бесконечную серию нелокальных сохраняющихся токов (3.32).*

Для построения локальных токов будет использована следующая теорема. Пусть матрицы U, V принадлежат классу $C^1_{N \times N}(\Omega)$ по x и λ — регулярная точка; область Ω содержит точки прямой $x_1 = 0$.

Теорема 3.2. *Если ψ — решение уравнения (3.19) представимо в виде*

$$\psi = \Phi e^S, \quad (3.36)$$

где Φ — обратимая и S — диагональная матрицы, $S|_{x_1=0} = 0$, класса $C^2_{N \times N}(\Omega)$, то

$$\begin{aligned} J_1 &= -\text{diag } T, \quad T = \Phi^{-1}V\Phi - \Phi^{-1}\partial_2\Phi, \\ J_2 &= \Lambda, \quad \Lambda = \partial_1S = \Phi^{-1}U\Phi - \Phi^{-1}\partial_1\Phi \end{aligned} \quad (3.37)$$

— сохраняющиеся токи для уравнения (3.18).

Ток (3.37), вообще говоря, нелокальный. Для получения бесконечной цепочки локальных токов произведем разложение в окрестности полюса.

Пусть матрицы $U(x; \lambda)$ и $V(x; \lambda)$ принадлежат классу $C_{N \times N}^\infty(\Omega)$ по x и λ_0 — полюс порядка $p \geq 1$ матрицы U и порядка $q \geq 0$ ($p \geq q$) матрицы V . В уравнениях (3.18) и (3.19) заменим U на $\varepsilon^{-p}U$ и V на $\varepsilon^{-q}V$:

$$\varepsilon^q \partial_2 U - \varepsilon^p \partial_1 V + [U, V] = 0, \quad (3.38)$$

$$\varepsilon^p \partial_1 \psi = U \psi, \quad (3.39)$$

$$\varepsilon^q \partial_2 \psi = V \psi, \quad (3.40)$$

где матрицы U, V обладают разложениями (3.25) с матрицами U_k и V_k класса $C_{N \times N}^\infty(\Omega)$, $k = 0, 1, \dots$. Предположим, что матрица U_0 имеет различные собственные значения $\lambda_1(x), \lambda_2(x), \dots, \lambda_N(x)$ в области Ω . Тогда $\lambda_i \in C^\infty(\Omega)$ и матрица $U_0(x)$ приводится к диагональному виду с помощью неособенной матрицы-функции Φ_0 класса $C_{N \times N}^\infty(\Omega)$:

$$\Phi_0^{-1} U_0 \Phi_0 = \Lambda_0 = [\lambda_1, \lambda_2, \dots, \lambda_N]. \quad (3.41)$$

Построим формальное асимптотическое разложение по ε решения ψ системы (3.39) (при каждом допустимом x_2). Согласно общей теории [8, 19] система (3.39) обладает формальным обратимым решением вида

$$\psi(x; \varepsilon) = \Phi(x; \varepsilon) \exp[\varepsilon^{-p} S(x; \varepsilon)], \quad (3.42)$$

причем матрицы Φ и $\Lambda = \partial_1 S$ разлагаются в формальные асимптотические ряды

$$\Phi(x; \varepsilon) = \sum_{k=0}^{\infty} \Phi_k(x) \varepsilon^k, \quad \Lambda(x; \varepsilon) = \sum_{k=0}^{\infty} \Lambda_k(x) \varepsilon^k. \quad (3.43)$$

При этом коэффициенты Φ_k, Λ_k разложений (3.43) определяются следующим образом. Подставляя выражение (3.42) в уравнение (3.39), для матриц Φ и Λ получаем уравнение

$$\varepsilon^p \partial_1 \Phi = U \Phi - \Lambda \Phi. \quad (3.44)$$

Подставляя ряды (3.43) в уравнение (3.44) и приравнявая нулю коэффициенты при одинаковых степенях ε , для матриц Φ_k и Λ_k получаем рекуррентное соотношение

$$U_0 \Phi_k - \Phi_k \Lambda_0 + U_k \Phi_0 - \Phi_0 \Lambda_k = \partial_1 \Phi_{k-p} + \sum_{s=1}^{k-1} \Phi_{k-s} \Lambda_s - U_s \Phi_{k-s}, \quad k = 1, 2, \dots, \quad (3.45)$$

где матрицы Φ_0, Λ_0 определены равенством (3.41) и $\Phi_s = 0$ при $s < 0$. В качестве решений Φ_k и Λ_k уравнений (3.45) выбираем матрицы

$$\Phi_k = \Phi_0 C_k, \quad \Lambda_k = \text{diag} P_k, \quad k = 1, 2, \dots, \quad (3.46)$$

где матрицы C_k, P_k определяются формулами

$$(C_k)^{ij} = \begin{cases} 0, & i = j \\ \frac{1}{\lambda_j - \lambda_i} (P_k)^{ij}, & i \neq j, \end{cases} \quad (3.47)$$

$$P_k = \Phi_0^{-1} \left[U_k \Phi_0 - \partial_1 \Phi_{k-p} + \sum_{s=1}^{k-1} U_s \Phi_{k-s} - \Phi_{k-s} \text{diag} P_s \right]. \quad (3.48)$$

Таким образом, задача построения формальных асимптотических разложений (3.43) свелась к алгебраическим операциям и операции дифференцирования, т.е. $\Phi_k(x)$ и $\Lambda_k(x)$ локально выражаются через $U_s(x)$, $s < k$.

Чтобы найти ток $J_\nu(x, \lambda)$, $\nu = 1, 2$, воспользуемся теоремой 3.2. Заменяя в формулах (3.36) и (3.37) U на $\varepsilon^{-p}U$, V на $\varepsilon^{-q}V$, S на $\varepsilon^{-p}S$, J_ν на $\varepsilon^{-p}J_\nu$, $\nu = 1, 2$, получаем

$$J_1 = -\text{diag} T, \quad J_2 = \Lambda, \quad (3.49)$$

где J_ν разлагаются в формальные асимптотические ряды (3.28) и матрица T удовлетворяет уравнению

$$\varepsilon^p \partial_2 \Phi = \varepsilon^{p-q} V \Phi - \Phi T. \quad (3.50)$$

Разлагая матрицу T в формальные асимптотические ряды

$$T(x; \varepsilon) = \sum_{k=0}^{\infty} T_k(x) \varepsilon^k, \quad (3.51)$$

из (3.51) получаем рекуррентные соотношения для T_k :

$$\begin{aligned} T_k &= 0, \quad k = 0, 1, \dots, p - q - 1, \\ T_k &= \Phi_0^{-1} \left[V_{k-p+q} \Phi_0 - \partial_2 \Phi_{k-p} + \sum_{s=1}^{k-p+q} V_{k-p+q-s} \Phi_s - \Phi_s T_{k-s} \right], \\ &k = p - q, p - q + 1, \dots \end{aligned} \quad (3.52)$$

Итак, в силу (3.28) и (3.43), (3.49), (3.51) получена следующая бесконечная цепочка локальных сохраняющихся токов для системы (3.18):

$$J_1^{(k)} = \text{diag} T_k, \quad J_2^{(k)} = \text{diag} P_k, \quad k = 0, 1, \dots, \quad (3.53)$$

где матрицы T_k и P_k определяются формулами (3.48) и (3.52), причем $P_0 = \Lambda_0$ и $C_0 = I$. Заменяя в формулах (3.48) и (3.52) Φ_k на $\Phi_0 C_k$, переписываем их в виде ($k = 0, 1, 2, \dots$)

$$\begin{aligned} P_k &= \Phi_0^{-1} U_k \Phi_0 - \Phi_0^{-1} \partial_1 (\Phi_0 C_{k-p}) + \sum_{s=1}^{k-1} \Phi_0^{-1} U_s \Phi_0 C_{k-s} - C_{k-s} \text{diag} P_s, \\ T_k &= \Phi_0^{-1} V_{k-p+q} \Phi_0 - \Phi_0^{-1} \partial_2 (\Phi_0 C_{k-p}) + \sum_{s=p-q}^{k-1} \Phi_0^{-1} V_{k-p+q} \Phi_0 C_{k-s} - C_{k-s} T_s. \end{aligned} \quad (3.54)$$

При $\Phi_0 = I$ формулы (3.54) упрощаются

$$\begin{aligned}
 P_k &= U_k - \partial_1 C_{k-p} + \sum_{s=1}^{k-1} U_s C_{k-s} - C_{k-s} \text{diag} P_s, \\
 T_k &= V_{k-p+q} - \partial_2 C_{k-p} + \sum_{s=p-q}^{k-1} V_{s-p+q} C_{k-s} - C_{k-s} T_s, \quad k = 0, 1, 2, \dots
 \end{aligned}
 \tag{3.55}$$

Для обоснования изложенной схемы приведем две группы достаточных условий, при которых формальные асимптотические разложения (3.43) сходятся асимптотически по ε , равномерно по x в возможно меньшей области $\Omega' \subset \Omega$ и их можно дифференцировать по x и ε любое число раз с сохранением равномерной по x оценки остаточного члена. При этих условиях асимптотические ряды (3.43) будут определять решение (3.42) системы уравнений (3.39). Такими условиями являются следующие.

- I. Матрицы $U(x; \lambda)$ и $V(x; \lambda)$ принадлежат классу $C_{N \times N}^\infty(\Omega)$ и λ_0 — полюс порядка $p \geq 1$ матрицы U и порядка $q \geq 0$ ($p \geq q$) матрицы V .
- II. Матрица $U_0(x) = U(x; \lambda_0)$ обладает различными собственными значениями: $\lambda_1(x), \dots, \lambda_N(x)$, $\lambda_j(x) \neq \lambda_k(x)$; $k \neq j$, $x \in \Omega$.
- IIIa. Матрица $U(x_1, x_2; \lambda)$ аналитична по x_1 при каждом допустимом значении x_2 (таким, что $x = (x_1, x_2) \in \Omega$). Или
- IIIb. Функция $\text{Re} [\lambda_j(x) - \lambda_k(x)]$ не меняет знак при j фиксированном, $k = 1, 2, \dots, N$, $x \in \Omega$.

Теорема 3.3. При выполнении условий I-II-IIIa или I-II-IIIb всякая вполне интегрируемая система (3.18) имеет бесконечную серию локальных сохраняющихся токов, вычисляемых по рекуррентным формулам (3.47), (3.48), (3.53) и (3.54).

Пример 3.3. Уравнение Кортевега-де Фриза (см. пример 3.1): $\varepsilon = 1/\lambda$

$$(\lambda_0 = \infty), \quad p = 3, \quad q = 1, \quad \Phi_0 = I, \quad \Lambda_0 = [4i, -4i] = U_0 = P_0, \quad T_0 = 0;$$

$$P_1 = U_1, \quad T_1 = 0; \quad P_2 = \frac{i}{2} \partial_x U_1, \quad T_2 = V_0;$$

$$P_3 = \begin{pmatrix} 0 & u \\ u^2 - u_{xx} & 0 \end{pmatrix}, \quad T_3 = 0;$$

$$P_4 = \frac{i}{2} \begin{pmatrix} u_{xx} - 3u^2 & -u_x \\ u^2 + u_t & 3u^2 - u_{xx} \end{pmatrix}, \quad T_4 = \frac{i}{2} \begin{pmatrix} -u & 0 \\ 0 & u \end{pmatrix}.$$

Первый нетривиальный ток: $J_1^{(4)} = u$, $J_2^{(4)} = u_{xx} - 3u^2$.

Пример 3.4. Уравнение синус-Гордон (см. пример 3.2): $\varepsilon = 1/\lambda$ ($\lambda_0 = \infty$), $p = 1$, $q = 0$, $\Phi_0 = I$, $\Lambda_0 = [i, -i] = U_0 = P_0$, $T_0 = 0$. Нечетные токи и нулевой ток тривиальны. Низшие нетривиальные токи: $J_1^{(2)} = \cos u$, $J_2^{(2)} = \frac{1}{2} u_x^2$; $J_1^{(4)} = -u_{xx} \sin u$, $J_2^{(4)} = \frac{1}{4} u_x^4 + u_x u_{xxx}$.

3.4. Сопряженные уравнения и интегральные законы сохранения

Ряд подходов к построению законов сохранения, изложенных в предыдущем разделе, могут быть распространены на эволюционные уравнения общего вида. Опишем некоторые из них, следуя работам [22]–[24].

Рассмотрим систему уравнений

$$\frac{\partial u}{\partial t} = F(u), \quad u|_{t=0} = u_0, \quad t \in [0, T], \quad (3.56)$$

где $u = u(t) \in \mathcal{H}$, \mathcal{H} — некоторое вещественное гильбертово пространство со скалярным произведением (\cdot, \cdot) .

Пусть $F(u)$ — нелинейный оператор, для которого можно ввести сопряженный согласно процедурам, изложенным выше. Для этого представим $F(u)$ в виде

$$F(u) = A(u)u, \quad (3.57)$$

где $A(u)$ — нелинейный оператор, зависящий от u , с областью определения, плотной в \mathcal{H} . Такое представление в общем случае неоднозначно. Далее, рассматривая $A(u)$ как линейный оператор, определенный известной функцией $u(t)$, можно стандартным образом построить для него сопряженный оператор, удовлетворяющий тождеству Лагранжа

$$(A(u)\varphi, \psi) = (\varphi, A^*(u)\psi). \quad (3.58)$$

Если удастся построить представление (3.57), такое, что оператор $A(u)$ оказывается кососимметрическим, т.е. $A^*(u) = -A(u)$, то умножая (3.56) скалярно в \mathcal{H} на u , получаем закон сохранения вида:

$$(u, u)(t) = \text{const}. \quad (3.59)$$

Таким образом, если оператор $F(u)$ допускает представление (3.57) с кососимметрическим оператором $A(u)$, то система (3.56) допускает закон сохранения (3.59). Данное утверждение можно обобщить. Так, пусть S — линейный самосопряженный оператор, не зависящий от t , причем $SA(u) = -A^*(u)S$, где $A(u)$ — оператор в (3.57), т.е. представление (3.57) и оператор S такие, что SA — кососимметрический оператор. Тогда, действуя оператором S на уравнение (3.56) и умножая результат скалярно в \mathcal{H} на решения u уравнения (3.56), получаем закон сохранения $(Su, u)(t) = \text{const}$.

Примерами систем (3.56), для которых справедлив закон сохранения $(u, u)(t) = \text{const}$ или $(Su, u)(t) = \text{const}$, являются системы гидродинамического типа, рассматриваемые ниже.

Отмечаем, что приведенные выше законы для системы (3.56) были получены, базируясь на представлении (3.57), и они не использовали решения

сопряженных уравнений. Однако привлечение этих уравнений открывает новые возможности в построении законов сохранения.

Рассмотрим систему (3.56), некоторое представление (3.57) и введем сопряженное уравнения вида

$$\frac{\partial v}{\partial t} = -A^*(u)v + f(u), \quad (3.60)$$

где $f(u)$ — некоторая функция, в общем случае нелинейным образом зависящая от u . Она может также не зависеть от u или равняться нулю. Умножая скалярно уравнение (3.56) на v , уравнение (3.60) на u и складывая, в силу тождества Лагранжа получим

$$\frac{\partial}{\partial t}(u, v) = (f(u), v). \quad (3.61)$$

Если выбрать $f(u)$ ортогональным u , то из (3.61) следует закон сохранения

$$(u, v) = \text{const}. \quad (3.62)$$

Таким образом, можно заключить, что множество представлений $F(u) = A(u)u$ и множество $f(u)$ при условии $(f(u), u) = 0$ определяет множество законов сохранения.

Теорема 3.4. *Если в системе (3.56) имеется закон сохранения $(\Phi(u), u) = \text{const}$ ($\Phi(u)$ — гладкое отображение) и $F(u)$ представимо в виде $F(u) = A(u)u$, то всегда можно построить уравнение вида (3.60), которое имеет решение $v = \Phi(u)$ с условием $(f(u), u) = 0$ так, что $(v, u) = (\Phi(u), u) = \text{const}$.*

Используя решения сопряженного уравнения (3.60) можно предложить следующую методику получения законов сохранения: задавая функции $f(u)$ такие, что $(f(u), u) = 0$ и отыскивая решения уравнения (3.60), получаем закон сохранения $(u, v(u))(t) = \text{const}$. Следуя этой методике, можно строить разнообразные законы сохранения. Конечно, представляют интерес случаи, когда решение $v = v(u)$ уравнения (3.60) может быть построено явно, т.е. когда мы "угадаем" вид $v = v(u)$ при заданном $f(u)$, то найдем явный вид закона $(u, v(u))(t) = \text{const}$. Как показывают примеры из предыдущего подраздела, для ряда интересных уравнений математической физики это можно сделать. Некоторые такие примеры для систем гидродинамического типа приведены ниже.

3.5. Сопряженные уравнения для систем гидродинамического типа

Здесь, следуя [22], мы сформулируем ряд утверждений, касающихся сопряженных уравнений для одного класса уравнений математической физики — систем гидродинамического типа.

Мы будем рассматривать системы гидродинамического типа с вещественными коэффициентами, решения которых принадлежат вещественным конечномерным пространствам. Рассмотрим задачу вида

$$\frac{\partial u}{\partial t} = F(u), \quad t \in [0, T], \quad u = u_0 \text{ при } t = 0, \quad (3.63)$$

где $u = (u^1, u^2, \dots, u^N) \in R^N$, $F(u) = (Q_1(u), Q_2(u), \dots, Q_N(u))$ при $\forall t \in [0, T]$, $u_{(0)} = (u_0^1, u_0^2, \dots, u_0^N) \in R^N$. В покомпонентной форме система (3.63) имеет вид

$$\frac{du^i}{dt} = Q_i(u), \quad u^i|_{t=0} = u_0^i, \quad i = 1, 2, \dots, N. \quad (3.64)$$

Определение 3.1. [51]. Система (3.64) называется системой гидродинамического типа, если она квадратично нелинейна, обладает квадратичным законом сохранения: $u^i u^i = \text{const}$, и лиувиллева: $\frac{\partial Q_i}{\partial u^i} = 0$.

В силу сделанных ограничений на характер нелинейности $\{Q_i(u)\}$ систему (3.64) можно также представить в форме

$$\frac{du^i}{dt} = \Gamma_{ijk} u^j u^k, \quad E \equiv u^i u^i = \text{const}, \quad (3.65)$$

где Γ_{ijk} удовлетворяют условиям симметрии: $\Gamma_{ijk} = \Gamma_{ikj}$. Для того чтобы в (3.65) был закон сохранения, $E = \text{const}$, должно выполняться соотношение цикличности [51]:

$$\Gamma_{ijk} + \Gamma_{kji} + \Gamma_{jki} = 0, \quad (3.66)$$

которое является другой формой условия существования закона сохранения энергии, $E = \text{const}$, а именно, в терминах ограничений на динамический тензор $\{\Gamma_{ijk}\}$.

Линеаризуем систему (3.65) относительно ее решения u . Обозначая через δu отклонения от решения, получим уравнение:

$$\frac{d\delta u^i}{dt} = \Gamma_{ijk} \delta u^j u^k + \Gamma_{ijk} u^j \delta u^k. \quad (3.67)$$

Поскольку для системы (3.65) справедлив закон $u^i u^i = \text{const}$, то отсюда немедленно следует, что система (3.65)–(3.67) допускает закон сохранения $u^i \delta u^i = \text{const}$.

Из существования инварианта $(\delta u, u) = \text{const}$ можно сделать ряд полезных выводов. Например, справедлива

Теорема 3.5. Устойчивое и неустойчивое многообразия решения системы гидродинамического типа ортогональны ее решению, если в начальный момент времени решение линеаризованной системы ортогонально решению исходной системы.

Под устойчивым и неустойчивым многообразием понимается многообразие, соответствующее положительным и отрицательным показателям Ляпунова. Если решение линеаризованной системы в начальный момент времени задается ортогональным решением исходной системы, то оно остается ортогональным и при всех t .

Теорема 3.6. Систему гидродинамического типа всегда можно привести к виду

$$\frac{du}{dt} = A(u)u, \quad (3.68)$$

где $A(u)$ — кососимметрический оператор: $A^*(u) = -A(u)$.

Простейшим примером уравнения вида (3.68) является "уравнение триплета" в канонической форме [51]:

$$\begin{aligned} \frac{du_1}{dt} &= pu_2u_3, & \frac{du_2}{dt} &= qu_1u_3, & \frac{du_3}{dt} &= ru_1u_2, \\ p + q + r &= 0. \end{aligned} \quad (3.69)$$

Представление (3.68) в данном случае является однопараметрическим:

$$\begin{aligned} \frac{d}{dt} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} &= A(u)u \equiv \\ &\equiv \begin{pmatrix} 0 & p\alpha u_3 & (1-\alpha)pu_2 \\ -p\alpha u_3 & 0 & (q+\alpha p)u_1 \\ -(1-\alpha)pu_2 & -(q+\alpha p)u_1 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \end{aligned} \quad (3.70)$$

α — произвольное вещественное число.

Поскольку $A(u)$ — кососимметрическая матрица, то заключаем о справедливости закона $u^i u^i = \text{const}$. (Обратим внимание также на то, что существование этого закона включается в определение системы гидродинамического типа, т.е. в нашем случае его и не надо доказывать на основе представления (3.70)!)

Однако для триплета существует еще один квадратичный закон сохранения [22]:

$$I \equiv (q-r)u_1^2 + (r-p)u_2^2 + (p-q)u_3^2 = \text{const}. \quad (3.71)$$

Можно построить сопряженное уравнение для (3.69), следствием которого будет закон сохранения (3.71).

Если рассматривать функционал I как скалярное произведение (u, u^*) , то для выполнения (3.71) необходимо, чтобы сопряженное к (3.69) уравнение имело решение $u^* = ((q-r)u_1, (r-p)u_2, (p-q)u_3)^T$. Не приводя подробного исследования этой проблемы, отметим, что в данном случае мы

также имеем однопараметрическое семейство представлений для триплета, частным случаем которого является представление

$$\frac{d}{dt} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 0 & pu_3 & 0 \\ q\beta u_3 & 0 & (1-\beta)qu_1 \\ 0 & ru_1 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \quad (3.72)$$

где $\beta = -\frac{p(q-r)}{q(r-p)}$. Соответствующее сопряженное уравнение имеет вид

$$\frac{d}{dt} \begin{pmatrix} u_1^* \\ u_2^* \\ u_3^* \end{pmatrix} = \begin{pmatrix} 0 & q\beta u_3 & 0 \\ pu_3 & 0 & ru_1 \\ 0 & (1-\beta)qu_1 & 0 \end{pmatrix} \begin{pmatrix} u_1^* \\ u_2^* \\ u_3^* \end{pmatrix}. \quad (3.73)$$

Нетрудно видеть, что достаточным условием существования инварианта $(Su, u) = \text{const}$ является существование такого представления исходной системы гидродинамического типа, решение сопряженной задачи для которого удовлетворяет соотношению $v = Su$.

Если для простоты считать, что S — невырожденный симметрический оператор, то это условие сводится к выполнению следующего соотношения $-S^{-1}A^*S = A$, или $SA + A^*S = 0$, что эквивалентно $SA = K$, где K — кососимметрический оператор.

В случае представления для триплета (3.72) и $S = \text{diag}(q-r, r-p, p-q)$ такой кососимметрический оператор имеет вид

$$K = \begin{pmatrix} 0 & (q-r)pu_3 & 0 \\ -(q-r)u_3 & 0 & r(p-q)u_1 \\ 0 & -r(p-q)u_1 & 0 \end{pmatrix}.$$

В заключение отметим, что техника сопряженных уравнений позволяет находить новые полезные законы сохранения для широкого класса нелинейных задач.

4. Метод сопряженных уравнений в задачах управления и математического моделирования сложных систем

Сопряженные операторы и уравнения находят широкое применение в исследовании и решении различных классов задач управления, в том числе оптимального или нечувствительного управления, обратных задач, задач вариационной ассимиляции данных наблюдений и др., а также при математическом моделировании разнообразных сложных систем, где они стали

в настоящее время неотъемлемой частью теории математического моделирования. В данном разделе будут представлены некоторые из приложений сопряженных операторов и уравнений в ряде областей теории сложных систем и их приложений.

4.1. Задачи оптимального управления

Рассмотрим одну из задач теории управления, встречающуюся в проблемах охраны окружающей среды. Пусть требуется найти функции $U(x, t)$, $V(x)$, являющиеся решением задачи вида

$$\begin{aligned} \frac{\partial U}{\partial t} - \mu \Delta U + \sum_{i=1}^n v_i(x, t) \frac{\partial U}{\partial x_i} + \lambda_0(x, t)U - Q(x, t) &= 0, \\ (x, t) \in \Omega \equiv D \times (0, T), \quad D \subset \mathbf{R}^n, \\ U(x, t) = U_{(\Gamma)}(x, t), \quad x \in \partial D, \quad t \in (0, T), \\ U(x, 0) = V(x), \quad x \in D, \\ J_\gamma(V) = \inf_V J_\gamma(V), \end{aligned} \tag{4.1}$$

где

$$J_\gamma(V) = \sum_{i=1}^N \alpha_i (U_i - C_i)^2 + \gamma \int_D (V(x))^2 dx,$$

$$\gamma = \text{const} \geq 0, \quad \mu = \text{const} > 0, \quad \alpha_i = \text{const} > 0,$$

$$U_i = \int_\Omega p_i(x, t) U(x, t) dx dt,$$

$p_i(x, t) \geq 0$ — "весовая функция" класса $L_\infty(\Omega)$; $V(x)$, $U_{(\Gamma)}(x, t)$, $Q(x, t)$ — заданные функции; $\{C_i\}$ — заданные константы; $\lambda_0 \geq 0$, $\lambda_0 \in L_\infty(\Omega)$. Предполагается, что функции $\{v_i(x, t)\}$ обладают производными $\{\partial v_i / \partial x_i\}$, удовлетворяющими условию

$$\sum_{i=1}^n \frac{\partial v_i}{\partial x_i} = 0, \quad (x, t) \in \Omega.$$

Предположим, что задача (4.1) имеет решение U, V . Рассмотрим тогда (4.1) при $V \rightarrow \tilde{V} = V + V_1$, где V_1 — произвольное "приращение" к V . Тогда для вариации $\delta J_\gamma(V, V_1)$ функционала J_γ справедливо следующее представление:

$$\delta J_\gamma = 2 \sum_{i=1}^N \alpha_i (U_i - C_i) \int_\Omega p_i(x, t) U_1 dx dt + 2\gamma \int_D V(x) V_1(x) dx, \tag{4.2}$$

где U_1 является решением задачи вида

$$\begin{aligned} \frac{\partial U_1}{\partial t} - \mu \Delta U_1 + \sum_{i=1}^n v_i \frac{\partial U_1}{\partial x_i} + \lambda_0 U_1 &= 0, \quad (x, t) \in \Omega, \\ U_1 &= 0, \quad (x, t) \in \partial D \times (0, T), \\ U_1(x, 0) &= V_1(x), \quad x \in D. \end{aligned} \quad (4.3)$$

Рассмотрим сопряженную задачу, задаваемую следующим соотношениями:

$$\begin{aligned} -\frac{\partial u^*}{\partial t} - \mu \Delta u^* - \sum_{i=1}^n v_i \frac{\partial u^*}{\partial x_i} + \lambda_0 u^* &= \sum_{i=1}^N \alpha_i (U_i - C_i) p_i(x, t), \quad (x, t) \in \Omega, \\ u^* &= 0, \quad (x, t) \in \partial D \times (0, T), \\ u^*(x, T) &= 0. \end{aligned} \quad (4.4)$$

Для решения задач (4.3), (4.4) справедливо равенство Лагранжа:

$$\int_{\Omega} \sum_{i=1}^N \alpha_i (U_i - C_i) p_i(x, t) U_1(x, t) dx dt = \int_D u^*(x, 0) V_1(x) dx. \quad (4.5)$$

Из (4.5), (4.2) получаем соотношение

$$\gamma V(x) = -u^*(x, 0). \quad (4.6)$$

Таким образом, решение задачи (4.1) необходимо удовлетворяет уравнениям

$$\begin{aligned} \frac{\partial U}{\partial t} - \mu \Delta U + \sum_{i=1}^n v_i \frac{\partial U}{\partial x_i} + \lambda_0 U - Q &= 0, \quad (x, t) \in \Omega, \\ U &= U_{(\Gamma)}, \quad (x, t) \in \partial D \times (0, T), \\ U(x, 0) &= V(x), \quad x \in D; \end{aligned} \quad (4.7)$$

$$\begin{aligned} -\frac{\partial u^*}{\partial t} - \mu \Delta u^* - \sum_{i=1}^n v_i \frac{\partial u^*}{\partial x_i} + \lambda_0 u^* &= \sum_{i=1}^N \alpha_i (U_i - C_i) p_i(x, t), \quad (x, t) \in \Omega, \\ u^* &= 0, \quad (x, t) \in \partial D \times (0, T), \quad u^*(x, T) = 0, \quad x \in D; \end{aligned} \quad (4.8)$$

$$\gamma V(x) = -u^*(x, 0). \quad (4.9)$$

Решив задачу (4.7)–(4.9) подходящим методом, можно помимо функций U, V, u найти, например, значение функционала

$$J_0 \equiv \sum_{i=1}^N \alpha_i (U_i - C_i)^2 / 2,$$

который часто имеет ясную физическую интерпретацию, и знание его важно для практики при исследовании и решении многих прикладных задач (например, задач охраны окружающей среды [39, 40, 46]) и др.).

4.2. Задача нечувствительного оптимального управления

Рассмотрим задачу

$$\begin{aligned} \frac{\partial U}{\partial t} + AU + f(U) &= 0, \quad (x, t) \in \Omega, \\ U(x, 0) &= V^0(x), \quad x \in D, \end{aligned} \quad (4.10)$$

где

$$AU = -\mu \Delta U + \sum_{i=1}^N v_i \frac{\partial U}{\partial x_i}, \quad (4.11)$$

$$f(U) = \lambda(x, t, U)U - Q(x, t), \quad (4.12)$$

а функции из области определения оператора A считаем удовлетворяющими условию $U = U(\Gamma)$, $(x, t) \in \partial D \times (0, T)$. Предположим, что функции $U(\Gamma)$, V^0 подобраны так, чтобы удовлетворялись некоторые критерии оптимальности (см. п. 4.1).

Рассмотрим еще одну задачу оптимального управления, связанную с (4.10). Так, на практике возможна ситуация, когда $V^0(x)$ может возмущаться посредством составляющей " $\tau V_1^0(x)$ " с известной функцией $V_1^0(x)$, но неизвестным малым параметром $\tau \in \mathbf{R}$ (т.е. $V^0 \rightarrow \tilde{V} = V^0 + \tau V_1^0$). При этом пусть нас интересует, например, значение функционала

$$\Phi_1(U) = \sum_{i=1}^N \alpha_i^{(0)} U_i \equiv \int_{\Omega} U(x, t) p(x, t),$$

где

$$\alpha_i^{(0)} = \text{const} \geq 0, \quad p(x, t) = \sum_{i=1}^N \alpha_i^{(0)} p_i(x, t),$$

и желательно, чтобы это значение было нечувствительно к появлению возмущения τV_1^0 в начальном условии. Поставим вопрос: нельзя ли выбрать управление $v(x, t)$ (которое можно интерпретировать как возмущение функции источников $Q(x, t)$) такое, чтобы функционал $\Phi_1(U)$ от решения задачи

$$\begin{aligned} \frac{\partial U}{\partial t} + AU + f(U) &= v, \quad (x, t) \in \Omega, \\ U(x, 0) &= V^0(x), \quad x \in D, \end{aligned} \quad (4.13)$$

был нечувствителен к появлению возмущения τV_1^0 в начальном условии? (Для простоты мы здесь не вводим дополнительных ограничений на управление v .) Ответ на данный вопрос был дан в работах Ж.-Л. Лионса (в ко-

торых можно найти строгое определение "нечувствительности функционала" [63]). А именно, чтобы $\Phi_1(U)$ был нечувствителен к возмущению вида τV_1^0 , необходимо, чтобы функции U, v были решением задачи

$$\begin{aligned} \frac{\partial U}{\partial t} + AU + f(U) &= v, \quad (x, t) \in \Omega, \\ U(x, 0) &= V^0(x), \quad x \in D; \end{aligned} \quad (4.14)$$

$$\frac{\partial q}{\partial t} - \mu \Delta q - \sum_{i=1}^n v_i \frac{\partial q}{\partial x_i} + (f'(U))^* q = \sum_{i=1}^N \alpha_i^{(0)} p_i(x, t), \quad (x, t) \in \Omega,$$

$$q = 0, \quad (x, t) \in \partial D \times (0, T),$$

$$q(x, T) = 0, \quad x \in D; \quad (4.15)$$

$$\int_D q(x, 0) V_1^0(x) dx = 0, \quad (4.16)$$

где f' есть производная Гато от $f(U)$.

В заключение отметим, что доказательство существования решения систем типа (4.14)–(4.16) в случае нелинейных функций $f(U)$ не является тривиальным. Достаточно полные результаты по данной проблеме получены Ж.-Л. Лионсом для линейных задач. Некоторые результаты для нелинейных задач можно найти в работах [46, 25, 54].

4.3. Задачи вариационного усвоения данных наблюдений

В настоящее время в связи с исследованиями глобальных изменений очень важной является проблема получения и рационального использования данных измерений с целью ретроспективного анализа в различных областях знаний. Математическая модель данной проблемы может быть сформулирована как задача об усвоении и обработке многомерных (включающих зависимость от временной и пространственных переменных) данных, представляющая собой одну из задач оптимального управления.

Начиная с работ Р. Беллмана, Л.С. Понтрягина, Н.Н. Красовского, А.М. Летова, Ж.-Л. Лионса, Р. Гловинского, А. Балакришнана, Г.И. Марчука, Н.Н. Моисеева, Ю.С. Осипова, А.Б. Куржанского, Ю.Г. Евтушенко, постановки и изучение таких задач на основе теории сопряженных уравнений привлекают внимание многих исследователей, занимающихся приложениями методов оптимального управления для практического решения тех или иных проблем. Наибольшее развитие эти методы получили в задачах ядерной энергетики, физики атмосферы и океана, экологии и др. (это работы Г.И. Марчука (1961), В.В. Пененко и Н.Н. Образцова (1976), Г.Р. Контарева (1980), Дж. Льюиса и Дж. Дербера (1985),

И. Навона (1987), Ф. Диме (1982), Ж.-Л. Лионса (1988), П. Куртье и О. Таллаграна (1987), А. Лоренца (1988), В.И. Агошкова (1993), В.И. Агошкова и Г.И. Марчука (1993), Г.И. Марчука и В.Б. Залесного (1993), В.М. Ипатовой (1992), Г.И. Марчука и В.П. Шутяева (1994), В.Б. Залесного и М. Венцеля (1996), и других).

Эти работы тесно примыкают к исследованиям по обратным задачам (см. работы А.Н. Тихонова, М.М. Лаврентьева, В.К. Иванова, А.Л. Бухгейма, Ж.-Л. Лионса и др.)

Оказалось [2, 54, 58], что задачи об усвоении данных эквивалентны, в определенном смысле, некорректно поставленным задачам, что обуславливает трудности решения этих задач и одновременно приводит к необходимости привлечения здесь методов регуляризации, разработанных в трудах А.Н. Тихонова, М.М. Лаврентьева, Ж.-Л. Лионса и др.

Рассмотрим задачу о выборе начального состояния системы, описываемой решением нелинейного эволюционного уравнения. Эта задача в проблемах геофизической гидродинамики известна как "задача об усвоении данных наблюдений".

Пусть H и X — действительные сепарабельные гильбертовы пространства, такие, что X вложено в H плотно и непрерывно; H^* , X^* — пространства, сопряженные к H , X . Введем в рассмотрение пространства $L_2(0, T; H)$, $Y = L_2(0, T; X)$, $L_2(0, T; X^*)$ абстрактных функций $f(t)$ со значениями в H , X , X^* , соответственно, и пространство

$$W = W(0, T) = \left\{ f \in L_2(0, T; X) : \frac{df}{dt} \in L_2(0, T; X^*) \right\},$$

$$\|f\|_W = \left(\left\| \frac{df}{dt} \right\|_{L_2(0, T; X^*)}^2 + \|f\|_{L_2(0, T; X)}^2 \right)^{1/2}.$$

Через $C^0([0, T]; H)$ обозначим банахово пространство непрерывных функций $[0, T] \mapsto H$ с нормой

$$\|f\|_{C^0([0, T]; H)} = \max_{t \in [0, T]} \|f\|_H.$$

Мы предполагаем, что

$$H \equiv H^*, \quad (\cdot, \cdot)_{L_2(0, T; H)} = (\cdot, \cdot), \quad \|\cdot\| = (\cdot, \cdot)^{1/2},$$

$$L_2(0, T; H) \equiv L_2^*(0, T; H) \equiv L_2(0, T; H^*).$$

Пусть $a(t; \varphi, \psi)$ — билинейная форма, определенная для всех $t \in [0, T]$ и для любых $\varphi, \psi \in X$ и удовлетворяющая неравенствам:

$$|a(t; \varphi, \psi)| \leq c_1 \|\varphi\|_X \|\psi\|_X, \quad c_1 = \text{const} > 0, \quad (4.17)$$

$$c_2 \|\varphi\|_X^2 \leq a(t; \varphi, \varphi), \quad c_2 = \text{const} > 0, \quad \forall t \in [0, T], \quad \forall \varphi, \psi \in X. \quad (4.18)$$

Через $A(t) \in \mathcal{L}(Y, Y^*)$ обозначим оператор, порождаемый этой формой:

$$(A(t)\varphi, \psi)_H = a(t; \varphi, \psi) \quad \forall \varphi, \psi \in X. \quad (4.19)$$

Рассмотрим квазилинейную эволюционную задачу вида

$$\begin{cases} \frac{d\varphi}{dt} + A(t)\varphi + \tau F(\varphi) = f(t), & t \in (0, T) \\ \varphi(0) = u, \end{cases} \quad (4.20)$$

где $f \in Y^*$, $u \in H$, $\tau \in [-\tau_0, \tau_0]$ — малый параметр, $\tau_0 \in \mathbf{R}^+$, $F(\varphi)$ — нелинейный оператор, непрерывно дифференцируемый по Фреше, $F: Y \rightarrow Y^*$.

Будем считать, что функция начального условия u в (4.20) неизвестна, и рассмотрим следующую задачу об усвоении данных: найти $u \in H$, $\varphi \in W$ такие, что

$$\begin{cases} \frac{d\varphi}{dt} + A(t)\varphi + \tau F(\varphi) = f, & t \in (0, T) \\ \varphi(0) = u \\ S(u) = \min_u S(u), \end{cases} \quad (4.21)$$

где

$$S(u) = \frac{\alpha}{2} \|u\|_H^2 + \frac{1}{2} \|B\varphi - \hat{\varphi}\|_Z^2,$$

$\alpha, \beta = \text{const} \geq 0$, Z — некоторое гильбертово пространство (пространство наблюдений) со скалярным произведением $(\cdot, \cdot)_Z$ и нормой $\|\cdot\|_Z = (\cdot, \cdot)_Z^{1/2}$, $B: Y \rightarrow Z$ — линейный ограниченный оператор, $\hat{\varphi} \in Z$. Функция $\hat{\varphi}(t)$ задается, как правило, с помощью априорных данных наблюдений.

При сформулированных предположениях существует линейный ограниченный оператор $C: Z \rightarrow Y^*$, для которого справедливо равенство

$$(\theta, B\psi)_Z = (C\theta, \psi) \quad \forall \theta \in Z, \psi \in Y. \quad (4.22)$$

В этом случае необходимое условие оптимальности сводит задачу (4.21) к системе для отыскания функций $\varphi, \varphi^* \in W$, $u \in H$ вида [28]:

$$\frac{d\varphi}{dt} + A(t)\varphi + \tau F(\varphi) = f, \quad t \in (0, T); \quad \varphi(0) = u, \quad (4.23)$$

$$-\frac{d\varphi^*}{dt} + A^*(t)\varphi^* + \tau(F'(\varphi))^*\varphi^* = C\hat{\varphi} - K\varphi, \quad t \in (0, T); \quad \varphi^*(T) = 0, \quad (4.24)$$

$$\alpha u - \varphi^*(0) = 0, \quad (4.25)$$

где $(F'(\varphi))^*: Y \rightarrow Y^*$ — оператор, сопряженный к производной Фреше от F в точке $\varphi \in W$, $A^*(t): Y \rightarrow Y^*$ — сопряженный к $A(t)$ оператор, $K = CB$ — линейный ограниченный оператор, действующий из Y в Y^* ; K является симметричным и неотрицательно определенным.

Рассмотрим задачу (4.23)–(4.25) при $\tau = 0$. Следуя [2, 58, 54], сведем ее к уравнению для управления u . Решения эволюционных задач (4.23), (4.24) при $\tau = 0$, как известно [28, 29], φ, φ^* можно представить в виде:

$$\varphi = G_0 u + G_1 f, \quad \varphi^* = G_1^{(T)}(C\hat{\varphi} - K\varphi),$$

где $G_0 : H \rightarrow W$, $G_1 : Y^* \rightarrow W$, $G_1^{(T)} : Y^* \rightarrow W$ — линейные ограниченные операторы. Тогда, исключая φ, φ^* из (4.23), (4.24) при $\tau = 0$, на основе (4.25) приходим к уравнению для u :

$$Lu = P, \quad (4.26)$$

где оператор управления $L : H \rightarrow H$ и правая часть P определяются по формулам:

$$L = \alpha E + T_0 G_1^{(T)} K G_0, \quad P = T_0 G_1^{(T)} C \hat{\varphi} - T_0 G_1^{(T)} K G_1 f,$$

E — единичный оператор, $T_0 : W \rightarrow H$ — оператор следа: $T_0 \varphi = \varphi|_{t=0}$.

Рассмотрим оператор управления L при $\alpha = 0$, обозначив его \bar{L} . Справедлива

Лемма 4.1. *Оператор $\bar{L} : H \rightarrow H$ непрерывен, самосопряжен и неотрицательно определен:*

$$(\bar{L}v, v)_H \geq 0 \quad \forall v \in H.$$

Если $(K\psi, \psi) > 0 \quad \forall \psi \in Y$, то оператор \bar{L} положителен.

Следствие. *В условиях леммы 4.1. справедливы следующие утверждения:*

I. *Область значений $R(\bar{L})$ оператора \bar{L} плотна в H .*

II. *Уравнение $\bar{L}u = P$ однозначно и плотно в H разрешимо.*

Замечание 4.1. В случае, когда $Z = \mathbf{R}^n$, $n \in \mathbf{N}$, а оператор наблюдений $B : Y \rightarrow Z$ задается по формуле: $B\varphi = ((\varphi, p_1), \dots, (\varphi, p_n))^T$, где $p_i \in Y^*$, $i = 1, \dots, n$, операторы $C : Z \rightarrow Y$ и $K : Y \rightarrow Y^*$ из (4.22), (4.24) определяются следующим образом:

$$C\theta = \sum_{i=1}^n \theta_i p_i, \quad K\varphi = \sum_{i=1}^n (\varphi, p_i) p_i, \quad (4.27)$$

где $\theta = (\theta_1, \dots, \theta_n)^T \in Z$. Тогда $(K\varphi, \psi) = (K\psi, \varphi)$ и $(K\psi, \psi) = \sum_{i=1}^n (\psi, p_i)^2 \geq 0 \quad \forall \varphi, \psi \in Y$.

Условиям леммы 4.1 удовлетворяет, например, оператор $K = E$ (тождественный оператор), который возникает в случае, когда $Z = Y^0$, $B = E$, тогда $C = E$. Именно этот случай был рассмотрен в [54] для симметричного и не зависящего от времени оператора A , где доказано, что оператор

управления \bar{L} является вполне непрерывным. Для несимметричного и зависящего от времени оператора A свойство вполне непрерывности сохраняется для достаточно широкого класса задач, в том числе и для случая, когда оператор наблюдения B вырожден [54].

Из леммы 4.1 вытекает, что при $\alpha > 0$ оператор $L: H \rightarrow H$ положительно определен. Тогда справедлива теорема о разрешимости линейной системы (4.23)–(4.25):

Теорема 4.1. Пусть $f \in Y^*$, $\hat{\varphi} \in Z$. Тогда при $\alpha > 0$ задача (4.23)–(4.25) при $\tau = 0$ имеет единственное решение $\varphi_0 \in W$, $\varphi_0^* \in W$, $u_0 \in H$, для которого справедлива оценка

$$\|\varphi_0\|_W + \|\varphi_0^*\|_W + \|u_0\|_H \leq c_0(\|C\hat{\varphi}\|_{Y^*} + \|f\|_{Y^*}),$$

где $c_0 = \text{const} > 0$.

Из Теоремы 4.1 методом малого параметра доказывается разрешимость нелинейной системы [54]:

Теорема 4.2. Пусть выполнены условия теоремы 4.1 и для некоторого R справедливы неравенства

$$\|F'(\xi)\|_{Y \rightarrow Y^*} \leq k_1, \quad \|F'(\xi) - F'(\eta)\|_{Y \rightarrow Y^*} \leq k_2\|\xi - \eta\|_W \quad (4.28)$$

для любых $\xi, \eta \in B(\varphi_0, R) = \{\varphi \in Y : \|\varphi - \varphi_0\|_W \leq R\}$, где $k_i = k_i(\varphi_0, R) = \text{const} > 0$. Тогда при $|\tau| < \tau_0$, где

$$\tau_0 = 1/c_0[k_1 + k_2(R + \|\varphi_0^*\|_W) + \frac{1}{R}(\|F(\varphi_0)\|_{Y^*} + k_1\|\varphi_0^*\|_W)]^{-1},$$

задача (4.23)–(4.25) имеет единственное решение $(\varphi, \varphi^*, u) \in W \times W \times H$, удовлетворяющее условию

$$\|\varphi - \varphi_0\|_W + \|\varphi^* - \varphi_0^*\|_W + \|u - u_0\|_H \leq R.$$

Если оператор $F(\varphi)$ аналитический, то в условиях теоремы 4.2 для решения (φ, φ^*, u) задачи (4.23)–(4.25) справедливы представления в виде рядов по степеням τ , сходящихся при $|\tau| < \tau_0$.

Нелинейную систему (4.23)–(4.25) можно решать методом последовательных приближений, на каждом шаге которого возникают линейные задачи об усвоении данных. Значительное разнообразие итерационных методов решения задач такого типа обсуждается в работах [58, 71, 72, 54].

4.4. Задачи охраны окружающей среды

Как следует из предыдущих разделов, использование сопряженных функций для представления линейных функционалов в форме $J_p = (f, \varphi_p^*)$ позволяет часто просто оценить вариации функционала J_p (или нескольких

таких функционалов) от вариаций входных параметров исходных задач (например, от вариаций функции f). Это свойство сопряженных функций было положено во многих работах в основу решения важных прикладных задач охраны окружающей среды и позволило предложить конструктивные экономичные подходы к решению этих задач. Изложим идею этих подходов на примере задачи о моделировании экологических ситуаций в акваториях водных бассейнов.

Интенсивное развитие промышленности требует выяснения оптимальных условий размещения новых промышленных предприятий и технологических ограничений на стоки, загрязняющие водные бассейны (моря, озера, заливы и т.д.), с таким расчетом, чтобы загрязнения прибрежных зон было минимальным. Математически эта задача сводится к проблеме минимакса.

Рассмотрим водный бассейн с областью определения D , которую будем считать цилиндром с боковой (береговой) поверхностью S и постоянной глубиной H . Задача диффузии гидрозоля имеет вид

$$\begin{aligned} \left(\frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z} \right) \varphi + \sigma \varphi - \frac{\partial}{\partial z} \nu \frac{\partial}{\partial z} \varphi - \mu \Delta \varphi &= Qw(r - r_0), \\ \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} &= 0, \\ w &= 0, \quad \text{при } z = 0, z = H, \\ \frac{\partial \varphi}{\partial n} &= 0 \quad \text{на } S, \\ \frac{\partial \varphi}{\partial z} &= 0 \quad \text{при } z = 0, z = H, \\ \varphi(r, T) &= \varphi(r, 0), \end{aligned} \tag{4.29}$$

где φ — концентрация загрязняющего гидрозоля, $w(r - r_0)$ — функция источника, $r = (x, y, z)$, $0 \leq t \leq T$, Q — мощность источника, r_0 — точка предполагаемого стока, u, v, w — компоненты вектора скорости \bar{u} . Предполагается, что решение задачи (4.29) имеет годовой ход, т.е. T — интервал времени, равный году.

Задачу (4.29) можно записать в операторной форме:

$$A\varphi = f. \tag{4.30}$$

Для решения задачи о размещении стоков промышленного предприятия в акватории водного бассейна введем в рассмотрение функционал

$$J_{p_i} = \int_0^T dt \int_D p_i(r) \varphi(r, t) dr, \tag{4.31}$$

где

$$p_i(r) = \begin{cases} 1 & \text{при } r \in D_i, \\ 0 & \text{при } r \notin D_i, \end{cases}$$

D_i — область, подлежащая охране от загрязнений. Введем в рассмотрение ограничения по санитарной норме, т.е. потребуем, чтобы r_0 принадлежало такой области Σ_0 , для каждой точки которой выполнялось бы условие

$$J_{p_i}(r_0) \leq c_i, \quad (4.32)$$

где c_i — константа, связанная с санитарной нормой для района D_i . Итак, задача сводится к решению уравнения (4.30) при ограничениях (4.32).

Для решения уже этой задачи введем в рассмотрение сопряженную задачу

$$A^* \varphi_{p_i}^* = p_i. \quad (4.33)$$

Теперь, пользуясь соотношением сопряженности $(A\varphi, \varphi_{p_i}^*) = (\varphi, A^* \varphi_{p_i}^*)$, где $(\cdot, \cdot) \equiv (\cdot, \cdot)_{L_2(D \times [0, T])}$, получаем, что функционал J_{p_i} можно представить следующим образом:

$$J_p = Q \int_0^T dt \int_D \varphi_{p_i}^*(r, t) w(r - r_0) dr. \quad (4.34)$$

Именно на основе представления (4.34) можно решать задачу о размещении стоков. Так, пусть уравнение (4.33) решено. Тогда найдем функционал J_{p_i} , параметрически зависящий от r_0 . По формуле (4.34), меняя значения r_0 , мы можем построить линии одинаковых значений функционала $J_{p_i}(r_0)$, а значит, и определить области, где $J_{p_i} \leq c_i$. Если окажется, что именно этой области принадлежит интересующая нас D_i , то задача выбора допустимых значений Q и r_0 решена. Если же такой области не найдется, то необходимо, например, уменьшить значение интенсивности и решить задачу выбора Q вновь.

Как легко заметить, сформулированная задача легко обобщается на случай, когда мы имеем n экологически охраняемых зон D_i ($i = 1, 2, \dots, n$). Здесь рассматривается n функционалов $J_{p_i}(r_0)$ при ограничениях $J_{p_i}(r_0) \leq c_i$ ($i = 1, 2, \dots, n$). Для практической реализации алгоритма решения задачи необходимо будет один раз решить n сопряженных задач, а затем уже решить следующую задачу минимакса:

$$\max_{i=1,2,\dots,n} B_i(r_0) = \min_{r_0 \in \Sigma_0},$$

где $B_i(r_0) = J_{p_i}(r_0)/c_0$. Заметим, что здесь эта проблема минимакса решается путем явного перебора, который выполняется элементарно. В этом состоит замечательное свойство сопряженных задач, позволяющих проблеме минимакса свести к простейшей реализации. Если бы мы не воспользовались сопряженными задачами, то для решения задачи минимакса нам потребовалось бы решать большое количество задач с различным положением источников промышленных отходов. Такая задача едва ли была бы

разрешима с заданной точностью даже при использовании самых совершенных средств вычислительной техники.

4.5. Задачи теории ядерных реакторов

В 1958 г. вышла в свет монография Г.И. Марчука "Численные методы расчета ядерных реакторов", переизданная в 1961 г. под названием "Методы расчета ядерных реакторов" в существенно переработанном и дополненном виде. Эти монографии вызвали большой научный интерес, а изложенные в них результаты стали основополагающими для развития методов расчета ядерных реакторов в нашей стране.

Одной из первых проблем в теории ядерных реакторов, изученных Г.И. Марчуком с использованием теории сопряженных уравнений и теории малых возмущений, была проблема корректной замены интегро-дифференциального уравнения Больцмана

$$(\Omega, \nabla)\varphi + \sigma\varphi = \int dv' \int W(x, \mu_0, v' \rightarrow v)\varphi(x, \Omega', v')d\Omega'$$

эффективными малогрупповыми моделями вида

$$(\Omega, \nabla)\varphi^j + \sigma^j\varphi^j = \sum_l \int W^{l \rightarrow j}(x, \mu_0)\varphi^l(x, \Omega')d\Omega',$$

где σ^j , $W^{l \rightarrow j}$ — групповые константы. Основным здесь был вопрос выбора данных констант. На основе теории сопряженных уравнений и алгоритмов возмущений Г.И. Марчук разработал методы построения таких моделей. Им были получены выражения для σ^j , $W^{l \rightarrow j}$:

$$\sigma^j = \frac{\int dx \int \varphi^{*j} d\Omega \int_{v_{j-1}}^{v_j} \sigma \varphi dv}{\int dx \int \varphi^{*j} d\Omega \int_{v_{j-1}}^{v_j} \varphi dv},$$

$$W^{l \rightarrow j}(x, \mu_0) = \Delta v_j \sum_{m=0}^{\infty} \frac{2m+1}{2} W_m^{l \rightarrow j} p_m(\mu_0),$$

где

$$W_m^{l \rightarrow j} = \frac{\int \varphi_m^{*j} dx \int_{v_{l-1}}^{v_l} \varphi_m dv \int_{v_{j-1}}^{v_j} W(x, \mu_0, v \rightarrow v') dv'}{\int_{G_n} \varphi_m^{*j} dx \int_{v_{l-1}}^{v_l} \varphi_m dv},$$

G_n — зона реактора, в которой ищутся групповые константы, φ_m^{*j} , φ_m — коэффициенты Фурье в разложении функций φ^{*j} и φ в ряды по сферическим функциям. Здесь $\{\varphi^{*j}\}$ есть решение сопряженных уравнений

$$-(\Omega, \nabla)\varphi^{*j} + \sigma^j \varphi^{*j} = \sum_l \int W^{l \rightarrow j} \varphi^{*j} d\Omega'.$$

Эти выражения для групповых констант в определенном смысле являются оптимальными. Было показано также, что они слабо зависят от функций $\{\varphi_m^{*j}\}$, φ . Это обстоятельство позволило сформулировать практические численные алгоритмы определения групповых констант, близких к оптимальным. Г.И. Марчуком и сотрудниками возглавляемого им отдела в Физико-энергетическом институте в 50-е годы были созданы эффективные малогрупповые модели реактора в различных приближениях метода сферических гармоник. Эти модели широко использовались для серийных расчетов критических масс ядерных реакторов и, в том числе, первой атомной электростанции.

Рассмотренный выше метод построения приближенных математических моделей допускает распространение на ряд других задач теории ядерных реакторов, в частности на задачи переноса излучения для нестационарного интегро-дифференциального уравнения вида

$$\begin{aligned} & \frac{1}{v} \frac{\partial \varphi}{\partial t} + \bar{\Omega} \bar{\nabla} \varphi + \Sigma(\bar{r}, E)\varphi - \\ & - \int d\bar{\Omega}' \int dE' \Sigma(\bar{\Omega}' \rightarrow \bar{\Omega}; E' \rightarrow E)\varphi(\bar{r}, E', \bar{\Omega}', t) = q(\bar{r}, E, \bar{\Omega}, t), \end{aligned} \quad (4.35)$$

где неизвестная функция $\varphi(\bar{r}, E, \bar{\Omega}, t)$ есть поток частиц (например, нейтронов), летящих в направлении $\bar{\Omega}$ с энергией E в точке \bar{r} в момент времени t ; $\Sigma(\bar{r}, E)$ — полное макроскопическое сечение взаимодействия; $\Sigma(\bar{\Omega}' \rightarrow \bar{\Omega}; E' \rightarrow E)$ — дифференциальное сечение перехода из $(E', \bar{\Omega}')$ в $(E, \bar{\Omega})$ при столкновениях (мы не разделяем здесь это сечение на сечение упругого и неупругого рассеяния, деления и т.д.); $q(r, E, \bar{\Omega}, t)$ — распределение источников излучения.

Пусть среда, в которой мы ищем решение уравнения (4.35), занимает объем V , ограниченный поверхностью S . Если вне объема V источники отсутствуют и среда вне V не отражает излучения, то естественным граничным условием для φ является

$$\varphi(\bar{r}_s, E, \bar{\Omega}, t) = 0, \quad \text{при } \bar{\Omega} \bar{n} < 0, \quad (4.36)$$

где \bar{n} — внешняя нормаль к поверхности S в точке \bar{r}_s .

Считая, что источники действуют не бесконечно долго, получим начальное условие:

$$\varphi(\bar{r}, E, \bar{\Omega}, t) = 0, \quad \text{при } t \rightarrow \infty. \quad (4.37)$$

Сопряженное к (4.35) уравнение имеет вид

$$-\frac{1}{v} \frac{\partial \varphi_p^*}{\partial t} + \bar{\Omega} \bar{\nabla} \varphi_p^* + \Sigma \varphi_p^* - \int d\bar{\Omega}' \int dE' \Sigma(\bar{\Omega} \rightarrow \bar{\Omega}'; E \rightarrow E') \varphi_p^*(\bar{r}, E', \bar{\Omega}', t) = p(\bar{r}, E, \bar{\Omega}, t) \quad (4.38)$$

с граничными и начальными условиями

$$\varphi_p^*(\bar{r}_s, E, \bar{\Omega}, t) = 0, \quad \text{при } \bar{\Omega} \bar{n} > 0, \quad (4.39)$$

$$\varphi_p^*(\bar{r}, E, \bar{\Omega}, t) = 0, \quad \text{при } t \rightarrow \infty. \quad (4.40)$$

Отметим, что уравнение (4.38) можно получить, исходя из физического смысла сопряженных функций $\varphi_p^*(\bar{r}, E, \bar{\Omega}, t)$ как ценности частиц (например, нейтронов) по отношению к функционалу

$$J_p[\varphi] = \int_V d\bar{r} \int dE \int d\bar{\Omega} \int_{-\infty}^{+\infty} dt \varphi(\bar{r}, E, \bar{\Omega}, t) p(\bar{r}, E, \bar{\Omega}, t). \quad (4.41)$$

При этом ценность нейтронов $\varphi_p^*(\bar{r}, E, \bar{\Omega}, t)$ равна тому значению, которое принимает функционал $J_p[\varphi]$ при впуске одного нейтрона в точку $x = (\bar{r}, E, \bar{\Omega}, t)$.

Предположим теперь, что различные возмущения в рассматриваемой системе приводят к изменению сечения взаимодействия излучения с веществом:

$$\Sigma \rightarrow \Sigma + \delta\Sigma,$$

$$\Sigma(\bar{\Omega}' \rightarrow \bar{\Omega}; E' \rightarrow E) \rightarrow \Sigma(\bar{\Omega}' \rightarrow \bar{\Omega}; E' \rightarrow E) + \delta\Sigma(\bar{\Omega}' \rightarrow \bar{\Omega}; E' \rightarrow E).$$

При этом изменяется оператор как основного уравнения $A \rightarrow A + \delta A$, так и сопряженного уравнения $A^* \rightarrow A^* + \delta A^*$, причем

$$\delta A = \delta\Sigma - \int d\bar{\Omega}' \int dE' \delta\Sigma(\bar{\Omega}' \rightarrow \bar{\Omega}; E' \rightarrow E),$$

$$\delta A^* = \delta\Sigma - \int d\bar{\Omega}' \int dE' \delta\Sigma(\bar{\Omega} \rightarrow \bar{\Omega}'; E \rightarrow E').$$

Используя соотношения сопряженности, получим формулы теории возмущений в виде

$$\delta J_p = - \int d\bar{r} \int d\bar{\Omega} \int dE \int dt \varphi_p^* [\delta\Sigma \tilde{\varphi} - \int d\bar{\Omega}' \int dE' \delta\Sigma(x' \rightarrow x) \tilde{\varphi}(x')], \quad (4.42)$$

$$\delta J_p = - \int d\bar{r} \int d\bar{\Omega} \int dE \int dt \varphi [\delta\Sigma \tilde{\varphi}_p^* - \int d\bar{\Omega}' \int dE' \delta\Sigma(x' \rightarrow x) \tilde{\varphi}_p^*(x')]. \quad (4.43)$$

Как уже отмечалось, соотношения теории возмущений наряду с их прямым использованием (для оценки различных эффектов и для анализа измерений) могут быть полезными для нахождения упрощенной модели сложной системы, такой, в которой некоторая интересующая нас величина J_p имеет то же значение, что и в истинной системе.

В качестве примера может быть рассмотрена гомогенизация гетерогенной системы или усреднение сечений по энергии. Формулы (4.42) и (4.43) позволяют получить метод усреднения сечений по пространству и по энергии. Действительно, подставляя, например, в соотношение (4.42) $\delta\Sigma = \bar{\Sigma} - \Sigma(\bar{r}, E)$ и полагая $\delta J_p = 0$, получим

$$\bar{\Sigma} = \frac{\int d\bar{r} \int dE \Sigma(\bar{r}, E) \int d\bar{\Omega} \tilde{\varphi} \varphi_p^*}{\int d\bar{r} \int dE \int d\bar{\Omega} \tilde{\varphi} \varphi_p^*}, \quad (4.44)$$

т.е. сечения следует усреднять с весом $\tilde{\varphi} \varphi_p^*$.

К числу таких методов относится и *метод эффективных граничных условий*, заключающийся в замене истинных условий некоторыми упрощенными условиями, но такими, которые приводят к правильному значению потока излучения вдали от границы [36].

4.6. Задачи анализа чувствительности климатической системы

При математическом моделировании изменений климата для различных районов планеты важную роль играет теория чувствительности избранных функционалов по отношению к континентам, Мировому океану, начальным данным, внешним источникам и внутренним параметрам задачи. Проблема чувствительности климата позволяет на основе реальных данных оценить ретроспективно качество моделей и найти новые механизмы, ответственные за формирование климата.

В этом разделе мы обсудим общие подходы к оценке чувствительности на основе простых теоретических моделей. Затем эти результаты могут быть обобщены на более сложные постановки задач.

Рассмотрим термическое взаимодействие атмосферы с Мировым океаном и континентами. Будем рассматривать трехмерную модельную область Ω в сферической системе координат (λ, ψ, z) , где λ — долгота, ψ — широта и z — высота, измеренная от поверхности Земли, которая полагается сферической, как и для океана; h_1 — толщина атмосферного слоя, h_2 — толщина активного слоя океана, h_3 — толщина деятельного слоя почвы; S — поверхность Земли, $S = S_1 \cup S_2 \cup S_3$, S_1 — поверхность континентов, свободная от льда и снега.

Область $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$ состоит из:

- слоя атмосферы $\Omega_1 = \{(\lambda, \psi, z) : (\lambda, \psi) \in S_1, 0 < z < h_1\}$;
- активного слоя океана $\Omega_2 = \{(\lambda, \psi, z) : (\lambda, \psi) \in S_2, -h_2 < z < 0\}$;
- верхнего слоя почвы $\Omega_3 = \{(\lambda, \psi, z) : (\lambda, \psi) \in S_3, -h_3 < z < 0\}$.

Изучим следующую задачу для поля температуры $T(\lambda, \psi, z, t)$ в интервале $\Omega_t = (0, T)$:

$$\alpha \frac{\partial T}{\partial t} + \operatorname{div}(\alpha \mathbf{u} T) - \frac{\partial}{\partial z} \left(\bar{\nu} \frac{\partial T}{\partial z} \right) - \bar{\mu} \delta T = \varepsilon \quad (4.45)$$

с начальными условиями

$$T = T_0(\mathbf{r}), \quad \text{при } t = 0, \quad (4.46)$$

где $\mathbf{r} = (\lambda, \psi, z)^T$, $\bar{\nu} = \alpha \nu$, $\bar{\mu} = \alpha \mu$ и ν, μ — коэффициенты вертикального и горизонтального обмена, соответственно, принятые зависящими от высоты z , поскольку они различны для атмосферы и океана. Функция ν зависит также от горизонтальных координат и времени. Переменные в уравнении (4.45) означают следующее: ε — источник лучистой энергии; T — функция отличная от нуля только в атмосфере, т.е. в области Ω_1 ; $\alpha = c_p \rho(z)$, где c_p — удельная теплоемкость среды, $\rho(z)$ — стандартная плотность атмосферы (при $z > 0$) и океана (при $z < 0$); $\mathbf{u}(\mathbf{r}, t)$ — вектор скорости ветра в атмосфере Ω_1 и течений в океане Ω_2 .

Положим, что значения \mathbf{u} , ν , μ заданы в области определения задачи, причем $\mathbf{u}(\mathbf{r}, t) = 0$, $\mu = 0$ в Ω_3 . Предположим также, что вектор-функция $\mathbf{u}(\mathbf{r}, t)$ удовлетворяет в атмосфере и океане простейшему уравнению неразрывности

$$\operatorname{div}(\rho \mathbf{u}) = 0 \quad (4.47)$$

и что ее нормальная компонента равна нулю на боковых границах Мирового океана так же, как и на поверхностях $z = 0$, $z = h_1$, $z = -h_2$. Естественно, что для атмосферы, если это нужно, уравнение неразрывности может рассматриваться с учетом изменяющихся со временем плотности и массы.

Уравнение (4.45) рассматривается в области $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$, $t \in \Omega_t = (0, \bar{t})$. Граничные условия должны быть сформулированы на границе области Ω при $z = 0$, $z = h_1$, $z = -h_2$, $z = -h_3$ и на поверхностях раздела атмосфера-континент, что обеспечивает существование единственного решения задачи Коши.

Условия равенства температур и скачка в потоках тепла задаются на поверхностях раздела между атмосферой и океаном, атмосферой и континентами ($z = 0$):

$$[T] = 0, \quad \left[\bar{\nu} \frac{\partial T}{\partial z} \right] = F(\lambda, \psi, z), \quad \text{при } z = 0, \quad (\lambda, \psi) \in S_2 \cup S_3, \quad (4.48)$$

где $[f] = f|_{z=-0} - f|_{z=+0}$ — скачок в значениях функции f от z в точке $z = 0$, и $F(\lambda, \psi, z)$ считается функцией, известной по данным наблюдений.

Следующее условие ставится на границе между атмосферой и частью поверхности Земли, покрытой льдом и/или снегом:

$$\bar{\nu} \frac{\partial T}{\partial z} + F = 0 \quad \text{при } z = 0, \quad (\lambda, \psi) \in S_1. \quad (4.49)$$

Условие отсутствия потока тепла принимается на верхней границе атмосферы (при $z = h_1$) и на нижних границах активных слоев в океане и почве:

$$\bar{\nu} \frac{\partial T}{\partial z} = 0 \quad \text{при } z = h_1, \quad z = -h_2, \quad z = -h_3. \quad (4.50)$$

Наконец, примем для простоты, что границы между континентами и Мировым океаном имеют цилиндрическую форму и что

$$\bar{\mu} \frac{\partial T}{\partial n} = 0 \quad \text{при } \partial\Omega_2, \quad (4.51)$$

где $\partial\Omega_2$ — боковая поверхность области Ω_2 , а n — нормаль к $\partial\Omega_2$.

Таким образом, задача (4.45)–(4.51) поставлена полностью, и мы полагаем, что начальные данные задачи определяют ее единственное решение, которое принадлежит гильбертовому пространству $L_2(\Omega \times \Omega_t)$. Примем, что это решение достаточно гладкое, так что дальнейшие преобразования обоснованы.

Рассмотрим вопрос о чувствительности решения задачи (4.45)–(4.51) к различным функционалам. Следующий функционал является самым интересным для нас:

$$J = \int_0^{\bar{t}} dt \int_S F^*(\lambda, \psi, t) T(\lambda, \psi, 0, t) dS, \quad (4.52)$$

где $F^*(\lambda, \psi, t)$ — некая весовая функция, связанная с полем температуры на поверхности $z = 0$. Например, если мы хотим определить среднюю температуру в каком-либо избранном районе континента ω при $z = 0$ в интервале $\bar{t} - \tau \leq t \leq \bar{t}$, то в качестве F^* мы выберем функцию

$$F^*(\lambda, \psi, t) = \begin{cases} 1/(\tau \text{mes } \omega) & \text{если } (\lambda, \psi) \in \omega, \bar{t} - \tau \leq t \leq \bar{t} \\ 0 & \text{в противном случае,} \end{cases} \quad (4.53)$$

где $\text{mes } \omega$ означает, как обычно, площадь района ω . Тогда функционал (4.52) записывается в виде

$$J = \frac{1}{\tau} \int_{\bar{t}-\tau}^{\bar{t}} dt \left(\frac{1}{\tau \text{mes } \omega} \int_{\omega} T(\lambda, \psi, 0, t) dS \right). \quad (4.54)$$

Выражение (4.54) представляет собой среднюю температуру по интервалу $\bar{t} - \tau \leq t \leq \bar{t}$ для избранного района ω . Функционалы такого типа наиболее интересны в теории изменений климата.

Поскольку функционал задачи определен в форме (4.52) или (4.54), мы можем сформулировать сопряженную задачу для (4.45)–(4.51):

$$-\alpha \frac{\partial T^*}{\partial t} - \operatorname{div}(\alpha \mathbf{u} T^*) - \frac{\partial}{\partial z} \left(\bar{\nu} \frac{\partial T^*}{\partial z} \right) - \bar{\mu} \delta T^* = 0 \quad (4.55)$$

с "начальным" условием

$$T^*(\bar{r}, t) = 0, \quad \text{при } t = \bar{t}. \quad (4.56)$$

Уравнение (4.55) рассматривается в области $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$, при $t \in \Omega_t$. Что же касается поверхности $z = 0$, а также других "особых" поверхностей, где необходимое условие дифференцируемости решения T^* и его производных не выполняется, то следует определить дополнительные граничные условия:

$$[T^*] = 0, \quad \left[\bar{\nu} \frac{\partial T^*}{\partial z} \right] = F^*, \quad \text{при } z = 0, \quad (\lambda, \psi) \in S_2 \cup S_3, \quad (4.57)$$

где F^* определяется функционалом (4.52) или, в частности, (4.53).

Следующее условие ставится на границе между атмосферой и частью земной поверхности, покрытой льдом и/или снегом:

$$\bar{\nu} \frac{\partial T^*}{\partial z} + F^* = 0 \quad \text{при } z = 0, \quad (\lambda, \psi) \in S_1. \quad (4.58)$$

На верхней границе атмосферы (при $z = h_1$) и на нижних границах активных слоев в океане и почве ставятся следующие граничные условия:

$$\bar{\nu} \frac{\partial T^*}{\partial z} = 0 \quad \text{при } z = h_1, \quad z = -h_2, \quad z = -h_3. \quad (4.59)$$

На границах между континентами и Мировым океаном мы полагаем

$$\bar{\mu} \frac{\partial T^*}{\partial n} = 0 \quad \text{при } \partial\Omega_2, \quad (4.60)$$

где n — внешняя нормаль к пограничной цилиндрической поверхности "океан-суша".

Теперь изучим другую форму функционала (4.52), выраженную через решение T^* сопряженной задачи. Умножим скалярно уравнение (4.45) на T^* и уравнение (4.55) на T , проинтегрируем по времени в Ω_t по всей области определения решения Ω с учетом начальных данных (4.46), (4.56) и

граничных условий (4.48)–(4.51), (4.57)–(4.60) и вычтем результаты друг из друга. Интегрируя по частям, получим:

$$J = \int_{\Omega} \alpha T^*(\mathbf{r}, 0) T_0(\mathbf{r}) d\Omega + \int_0^{\bar{t}} dt \int_S F(\lambda, \psi, t) T^*(\lambda, \psi, 0, t) dS + \\ + \int_0^{\bar{t}} dt \int_{\Omega_1} \varepsilon T^*(\lambda, \psi, z, t) d\Omega_1. \quad (4.61)$$

Заметим, что левая часть этого соотношения в точности совпадает с функционалом J , определенным формулой (4.52).

Соотношение (4.61) описывает чувствительность функционала J к локализации различных районов и входных параметров задачи, поскольку весовая функция T^* характеризует влияние начальных данных $T_0(\bar{r})$, потоков тепла $F(\lambda, \psi, t)$ на границе $z = 0$ и источников лучистой энергии в атмосфере ε .

Методы сопряженных уравнений нашли применение для практического решения широкого класса задач науки и техники: от проблем глобальных изменений до задач математического моделирования в иммунологии и медицине, включая теорию измерений и планирования эксперимента.

5. Некоторые нерешенные проблемы теории сопряженных уравнений и их приложений в анализе сложных систем

Сформулируем некоторые проблемы, решение которых представляет интерес в теории сопряженных операторов (уравнений) и их приложений в анализе сложных систем.

Проблема 1. В нелинейных задачах дать ответ на вопрос: какой из множества сопряженных операторов является "наилучшим" при рассмотрении конкретной задачи.

Проблема 2. Указать методы отыскания решений "ассоциированных систем" и систем вида (3.11), (3.56) для построения законов сохранения при рассмотрении конкретных классов систем уравнений.

Проблема 3. Разработка и реализация эффективных итерационных алгоритмов при решении задач вариационной ассимиляции данных наблюдений.

Проблема 4. Исследование спектральных свойств операторов управления вида (4.26), возникающих в задачах вариационного усвоения данных наблюдений.

Проблема 5. Развитие методов сопряженных уравнений второго порядка для построения и исследования Гесссианов в задачах оптимального управления.

Проблема 6. Исследование чувствительности оптимальных решений к ошибкам наблюдений в задачах вариационной ассимиляции данных на основе методов сопряженных уравнений.

Проблема 7. Рассматривается задача, решения которой удовлетворяют некоторому набору законов сохранения. Разработать численные методы построения приближенного решения данной задачи, при этом на приближенном решении законы сохранения должны выполняться с заданной точностью.

Проблема 8. Приведенные выше законы сохранения были сформулированы по существу в рамках лагранжева формализма. Вывести законы сохранения в рамках гамильтонова формализма, используя скобки Пуассона.

Проблема 9. Сформулировать интегральные законы сохранения для решений, не обязательно удовлетворяющих нулевым граничным условиям (см. п. 3.3).

Проблема 10. Использовать нелокальные законы сохранения для доказательства существования решений нелинейных уравнений и их исследования.

Проблема 11. Построить нелокальные законы сохранения для уравнения Бюргера

$$u_t + uu_x = \varepsilon u_{xx}, \quad u = u(x, t), \quad \varepsilon > 0$$

на основе решения v ассоциированного уравнения $v_t + (vu)_x = -\varepsilon v_{xx}$.

Представленные в данной работе исследования выполнены при финансовой поддержке РФФИ (гранты № 03-01-00779, № 04-01-00615), грантов Президента РФ (НШ-150.2003.1 и НШ-1542.2003.1), а также программы фундаментальных исследований ОМН РАН.

Литература

- [1] Авдошин С.М., Белов В.В., Маслов В.П. Математические аспекты синтеза вычислительных сред. М.: МИЭМ, 1984.
- [2] Агошков В.И. Методы оптимального управления и сопряженных уравнений в задачах математической физики. М.: ИВМ РАН, 2003.
- [3] Агошков В.И. Сопряженные уравнения в алгоритмах возмущений N-го порядка точности // Сопряженные уравнения и теория возмущений в задачах математической физики. М.: ОВМ АН СССР, 1985. С. 62–85.
- [4] Агошков В.И. Алгоритмы возмущений N-го порядка для функционалов от решений нелинейных задач и оценка их скорости сходимости // Сопряженные уравнения и алгоритмы возмущений в задачах математической физики. М.: ОВМ АН СССР, 1989. С. 3–30.

- [5] *Агошков В.И., Ипатова В.М.* О разрешимости основных и сопряженных уравнений в нелинейных задачах // *Сопряженные уравнения в задачах математической физики.* М.: ОВМ АН СССР, 1990. С. 3–46.
- [6] *Агошков В.И., Дубовский П.Б., Шутяев В.П.* Методы решения задач математической физики. М.: Физматлит, 2002.
- [7] *Агошков В.И., Ипатова В.М.* Об использовании законов сохранения при построении сопряженных операторов. Препринт № 256. М.: ОВМ АН СССР, 1990.
- [8] *Агошков В.И., Мишнева А.П.* О нахождении коэффициента дисперсии в нелинейном параболическом уравнении. Препринт № 200. М.: ОВМ АН СССР, 1988.
- [9] *Агошков В.И., Попыкин А.И., Шихов С.Б.* К теории малых возмущений для уравнения переноса // *Сопряженные уравнения и теория возмущений в задачах математической физики.* М.: ОВМ АН СССР, 1985. С. 76–84.
- [10] *Алоян А.Е., Йорданов Д.Л., Пененко В.В.* Численная модель переноса примесей в пограничном слое атмосферы // *Метеорология и гидрология.* 1981. № 1. С. 32–43.
- [11] *Боголюбов Н.Н.* К вопросу о модельном гамильтониане в теории сверхпроводимости. Избр. труды. Киев: Наукова Думка, 1971. Т. 3. С. 110–173.
- [12] *Бочаров Г.А.* Сопряженные уравнения и анализ чувствительности математических моделей. Деп. в ВИНТИ № 2858–В94. М., 1994.
- [13] *Вайнберг М.М.* Функциональный анализ. М.: Просвещение, 1979.
- [14] *Вишик М.И., Люстерник Л.А.* Некоторые вопросы возмущений краевых задач для дифференциальных уравнений в частных производных // *ДАН СССР.* 1959. Т. 129, № 6. С. 1203–1206.
- [15] *Вишик М.И., Люстерник Л.А.* Возмущение собственных значений и собственных элементов для некоторых несамосопряженных операторов // *ДАН СССР.* 1060. Т. 130, № 2. С. 251–253.
- [16] *Владимиров В.С.* Уравнения математической физики. М.: Наука, 1981.
- [17] *Владимиров В.С., Волович И.В.* Законы сохранения для нелинейных уравнений // *ДАН СССР.* 1984. Т. 279, № 4. С. 843–847.
- [18] *Владимиров В.С., Волович И.В.* Законы сохранения для нелинейных уравнений // *Актуальные проблемы вычислительной математики и математического моделирования.* Новосибирск: Наука. 1985. С. 147–162.
- [19] *Владимиров В.С., Волович И.В.* Локальные и нелокальные токи для нелинейных уравнений. // *Теор. и матем. физика.* Т. 62, № 1984. С. 3–29.
- [20] *Владимиров В.С., Жаринов В.В.* Замкнутые формы, ассоциированные с линейными дифференциальными операторами // *Дифф. уравнения.* 1980. Т.16, № 5. С. 845–867.
- [21] *Владимиров В.С., Марчук Г.И.* Об определении сопряженного оператора для нелинейных задач // *Докл. РАН.* 2000. Т. 372, № 2. С. 165–168.
- [22] *Дымников В.П.* Сопряженные уравнения систем гидродинамического типа // *Изв. РАН, Физика атмосферы и океана,* 2001. Т.37, № 4. С.459–462.
- [23] *Дымников В.П.* Об общих представлениях интегральных законов сохранения нелинейных дифференциальных уравнений // *Докл. РАН.* 2002. Т.382, № 2. С. 173–175.

- [24] *Дымников В.П.* Сопряженные уравнения, интегральные законы сохранения и консервативные разностные схемы для уравнений динамики двумерной идеальной несжимаемой жидкости // ДАН. 2002. Т.386, № 3. С. 320–312.
- [25] *Ипатова В.М.* Задача усвоения данных для модели общей циркуляции океана в квазигеострофическом приближении. Деп. в ВИНТИ № 2333-В92. М., 1992.
- [26] *Като Т.* Теория возмущений линейных операторов. М.: Мир, 1972.
- [27] *Крейн С.Г.* Линейные уравнения в банаховом пространстве. М.: Наука, 1971.
- [28] *Лионс Ж.-Л.* Оптимальное управление системами, описываемыми уравнениями с частными производными. М.: Мир, 1972.
- [29] *Лионс Ж.-Л., Маджсенес Э.* Неоднородные граничные задачи и их приложения. М.: Мир, 1971.
- [30] *Льюис Дэс.* Ценность. Сопряженная функция. М.: Атомиздат, 1972.
- [31] *Марчук Г.И.* Численное решение задач динамики атмосферы и океана. Л.: Гидрометеониздат, 1974.
- [32] *Марчук Г.И.* Методы вычислительной математики. 3-е изд. М.: Наука, 1989.
- [33] *Марчук Г.И.* О постановке обратных задач // ДАН СССР. 1964. Т. 156, № 3. С. 503–506.
- [34] *Марчук Г.И.* Уравнение для ценности информации с метеорологических спутников и постановка обратных задач // Космич. исслед. 1964. Т. 2, вып.3. С. 462–477.
- [35] *Марчук Г.И.* Основные и сопряженные уравнения динамики атмосферы и океана // Метеорология и гидрология. 1974. № 2. С. 9–37.
- [36] *Марчук Г.И.* Численные методы расчета ядерных реакторов. М.: Атомиздат, 1958.
- [37] *Марчук Г.И.* Методы долгосрочного прогноза погоды на основе решения основных и сопряженных задач // Метеорология и гидрология. 1974. № 3. С. 17–34.
- [38] *Марчук Г.И.* Окружающая среда и проблемы оптимизации размещения предприятий // ДАН СССР. 1976. Т. 227, № 5. С. 1056–1059.
- [39] *Марчук Г.И.* Методы расчета ядерных реакторов. М.: Атомиздат, 1961.
- [40] *Марчук Г.И.* Математическое моделирование в проблеме окружающей среды. М.: Наука, 1982.
- [41] *Марчук Г.И.* Сопряженные уравнения и анализ сложных систем. М.: Наука, 1992.
- [42] *Марчук Г.И.* Моделирование изменений климата и проблема долгосрочного прогноза погоды // Метеорология и гидрология. 1979. № 7. С. 25–36.
- [43] *Марчук Г.И., Агошков В.И.* Сопряженные уравнения в нелинейных задачах и их приложения // Функциональные и численные методы математической физики. Киев: Наукова думка, 1988. С. 138–142.
- [44] *Марчук Г.И., Агошков В.И., Шутяев В.П.* Сопряженные уравнения и алгоритмы возмущений в прикладных задачах // Вычислительные процессы и системы. Вып.4. М.: Наука, 1986. С. 5–62.
- [45] *Марчук Г.И., Агошков В.И., Шутяев В.П.* Сопряженные уравнения и алгоритмы возмущений. М.: ОВМ АН СССР, 1986.
- [46] *Марчук Г.И., Агошков В.И., Шутяев В.П.* Сопряженные уравнения и методы возмущений в нелинейных задачах математической физики. М.: Наука, 1993.

- [47] *Марчук Г.И., Лебедев В.И.* Численные методы в теории переноса нейтронов. М.: Атомиздат, 1981.
- [48] *Марчук Г.И., Орлов В.В.* К теории сопряженных функций // *Нейтронная физика*. М.: Госатомиздат, 1976. С. 30–45.
- [49] *Марчук Г.И., Пененко В.В.* Исследование чувствительности дискретных моделей динамики атмосферы и океана // *Изв. АН СССР. Физика атмосферы и океана*. 1979. Т. 15, № 11. С. 1123–1131.
- [50] *Маслов В.П.* Теория возмущений и асимптотические методы. М.: Изд-во МГУ, 1965.
- [51] *Обухов А.М.* Турбулентность и динамика атмосферы. Л.: Гидрометеиздат, 1988.
- [52] *Треногин В.А.* Развитие и приложения асимптотического метода Люстерника–Вишика // *УМН*. 1970. Т. XXV, вып.4. С. 123–156.
- [53] *Треногин В.А.* Функциональный анализ. М.: Наука, 1980.
- [54] *Шутяев В.П.* Операторы управления и итерационные алгоритмы в задачах вариационного усвоения данных. М.: Наука: 2001.
- [55] *Шутяев В.П.* О свойствах решения сопряженного уравнения в одной нелинейной гиперболической задаче // *Дифф. уравнения*. 1992. Т. 28, № 4. С. 706–715.
- [56] *Шутяев В.П.* О вычислении функционала в одной нелинейной задаче с использованием сопряженного уравнения // *ЖВМ и МФ*. 1991. Т. 30, № 9. С. 1278–1288.
- [57] *Шутяев В.П.* О свойствах оператора управления в одной задаче об усвоении данных и алгоритмах ее решения // *Мат. заметки*. 1995. Т. 57, вып. 6. С. 941–944.
- [58] *Agoshkov V.I., Marchuk G.I.* On solvability and numerical solution of data assimilation problems // *Russ. J. Numer. Anal. Math. Modelling*. 1993. V. 8, № 1. P. 1–16.
- [59] *Cacuci D.G.* Sensitivity theory for nonlinear system: II. Extensions to additional classes of responses // *J. Math. Phys*. 1981. V. 22. P. 2803–2812.
- [60] *Dymnikov V.P.* Adjoint equations, integral conservation laws, and conservative difference schemes for nonlinear equations of mathematical physics // *Russ. J. Numer. Anal. Math. Modelling*. 2003. V. 18, № 3. P. 229–242.
- [61] *Le Dimet F.X., Talagrand O.* Variational algorithms for analysis and assimilation of meteorological observation: theoretical aspects // *Tellus A*. 1986. V. 38. P. 97–110.
- [62] *Le Dimet F.X., Navon I.M., Daescu D.N.* Second-order information in data assimilation // *Monthly Weather Review*. 2002. V. 130, № 3. P. 629–648.
- [63] *Lions J.L.* Contrôlabilité Exacte Perturbations et Stabilisations de Systèmes Distribués. Paris: Masson, 1988.
- [64] *Marchuk G.I.* Formulation of the theory of perturbations for complicated models // *Applied Math. and Optimization*. 1975. V. 2, № 1. P. 1–33.
- [65] *Marchuk G.I.* Formulation of the theory of perturbations for complicated models. Part I. The Estimation of the Climate Change // *Geofisica International, Mexico*. 1975. V. 15, № 2. P. 103–156. Part II. Weather Prediction. *Ibid.* P. 169–182.
- [66] *Marchuk G.I.* Mathematical issues of industrial effluent optimization // *J. of the meteorological Society of Japan*. 1982. V. 60, № 1. P. 481–485.

- [67] *Marchuk G.I.* Perturbation theory and the statement of inverse problems // Lect. Notes Comput. Sci. 1973. № 4, Pt.2. P. 159–166.
- [68] *Marchuk G.I., Agoshkov V.I.* Conjugate operators and algorithms of perturbation in non-linear problems: 1. Principles of Construction of conjugate operators // Soviet J. Numer. Anal. Math. Modelling. 1988. № 1. P. 21–46.
- [69] *Marchuk G.I., Agoshkov V.I.* Conjugate operators and algorithms of perturbation in non-linear problems: 2. Perturbation algorithms // Soviet J. Numer. Anal. Math. Modelling. 1988. № 2. P. 115–136.
- [70] *Marchuk G.I., Agoshkov V.I., Shutyaev V.P.* Adjoint Equations and Perturbation algorithms in Non-linear Problems. New York: CRC Press, 1996.
- [71] *Marchuk G.I., Shutyaev V.P.* Iteration methods for solving a data assimilation problem // Russ. J. Numer. Anal. Math. Modelling. 1994. Vol.9, № 3. P. 265–279.
- [72] *Marchuk G.I., Zalesny V.B.* A numerical technique for geophysical data assimilation problem using Pontryagin's principle and splitting-up method // Russ. J. Numer. Anal. Math. Modelling. 1993. V. 8, № 4. P. 311–326.
- [73] *Penenko V., Obratsov N.N.* A variational initialization method for the fields of the meteorological elements // Meteorology and Hidrology (English transl.). 1976. V. 11. P. 1–11.
- [74] *Trenogin V.* Non-linear operators in weak metric spaces and its conjugate operators // Works of Int. Conf. for 175th birthday of P.L.Chebyshev. M.: Moscow Univ. publ., 1996. V. 1. P. 335–337.
- [75] *Van Dyke M.D.* Perturbation Methods in Fluid Mechanics. N.Y.: Academic Press, 1964.
- [76] *Vladimirov V.S.* Conservation laws for differential equations // Functional Analytic and Complex Methods, their Interactions, and Applications to Partial Differential Equations, Proc. of the Intern. Graz Workshop, 12–16 February 2001, Graz, Austria, p. 137–142.
- [77] *Yamamoto S.* The adjoints of differentiable mappings // J. Austral. math. Soc. 1968 V. 8. P. 397–407.

Содержание

К 80-летию академика Г.И. Марчука	5
<i>Н.С. Бахвалов, Г.М. Кобельков, Ю.А. Кузнецов, В.И. Лебедев, И.К. Лифанов, Ю.М. Нечепуренко, В.В. Шайдунов</i> Численные методы решения задач математической физи- ки	9
<i>Ю.В. Василевский, В.П. Ильин, Е.Е. Тыртышников</i> Вычислительные технологии	100
<i>Г.А. Михайлов, Б.А. Каргин, С.М. Пригарин, В.С. Антюфеев, В.А. Огородников, К.К. Сабельфельд, С.С. Артемьев, А.В. Вой-</i> <i>тишек</i> Стохастическое моделирование и метод Монте-Карло . .	149
<i>В.В. Воеводин, Вл.В. Воеводин</i> Параллельные вычисления: новые концепции в науке и образовании	220
<i>В.И. Агошков, В.С. Владимиров, И.В. Волович, В.П. Дымников, В.П. Шутяев</i> Метод сопряженных уравнений и анализ сложных сис- тем	257

Научное издание

**Современные проблемы
вычислительной математики
и математического моделирования**

В двух томах

Том 1

Вычислительная математика

*Утверждено к печати
Ученым советом
Института вычислительной математики
Российской академии наук*

Зав. редакцией *Н.А. Степанова*
Редактор *Л.В. Пеняева*
Художник *Е.А. Шейко*
Художественный редактор *Ю.И. Духовская*

Напечатано с оригинал-макета
Института вычислительной математики РАН

Подписано к печати 21.01.2005
Формат 70 × 100 1/16. Гарнитура Таймс
Печать офсетная
Усл.печ.л. 28,0. Усл.кр.-отг. 28,0. Уч.-изд.л. 21,0
Тираж 570 экз. Тип. зак. 3859

Издательство "Наука". 117997, Москва, Профсоюзная ул., 90

E-mail: secret@naukaran.ru Internet: www.naukaran.ru

Отпечатано с готовых диапозитивов
в ГУП "Типография "Наука"
199034, Санкт-Петербург, 9 линия, 12

ISBN 5-02-033436-7(общ.)



9 785020 334366

ISBN 5-02-033716-1



9 785020 337169

